



Automated Measurement of Lumbar Lordosis on Radiographs Using Machine Learning and Computer Vision

Global Spine Journal
2020, Vol. 10(5) 611-618
© The Author(s) 2019
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/2192568219868190
journals.sagepub.com/home/gsj



Brian H. Cho, BS^{1*}, Deepak Kaji, BA^{1*}, Zoe B. Cheung, MD¹, Ivan B. Ye, BA¹, Ray Tang, BA¹, Amy Ahn, BS¹, Oscar Carrillo, BS¹, John T. Schwartz, BS¹ , Aly A. Valliani, BA¹, Eric K. Oermann, MD¹, Varun Arvind, BS¹, Daniel Ranti, BS¹, Li Sun, DO¹, Jun S. Kim, MD¹, and Samuel K. Cho, MD¹ 

Abstract

Study Design: Cross sectional database study.

Objective: To develop a fully automated artificial intelligence and computer vision pipeline for assisted evaluation of lumbar lordosis.

Methods: Lateral lumbar radiographs were used to develop a segmentation neural network (n = 629). After synthetic augmentation, 70% of these radiographs were used for network training, while the remaining 30% were used for hyperparameter optimization. A computer vision algorithm was deployed on the segmented radiographs to calculate lumbar lordosis angles. A test set of radiographs was used to evaluate the validity of the entire pipeline (n = 151).

Results: The U-Net segmentation achieved a test dataset dice score of 0.821, an area under the receiver operating curve of 0.914, and an accuracy of 0.862. The computer vision algorithm identified the L1 and S1 vertebrae on 84.1% of the test set with an average speed of 0.14 seconds/radiograph. From the 151 test set radiographs, 50 were randomly chosen for surgeon measurement. When compared with those measurements, our algorithm achieved a mean absolute error of 8.055° and a median absolute error of 6.965° (not statistically significant, P > .05).

Conclusion: This study is the first to use artificial intelligence and computer vision in a combined pipeline to rapidly measure a sagittal spinopelvic parameter without prior manual surgeon input. The pipeline measures angles with no statistically significant differences from manual measurements by surgeons. This pipeline offers clinical utility in an assistive capacity, and future work should focus on improving segmentation network performance.

Keywords

lumbar, lordosis, sagittal balance, artificial intelligence, neural networks, machine learning, radiographic image interpretation, computer-assisted, radiography, spinopelvic parameters, angle measurement

Introduction

Spinal alignment is increasingly being recognized as a key, quantitative assessment of spinal health, and is associated with various spinal disorders such as adolescent idiopathic scoliosis, adult spinal deformity, and degenerative spondylolisthesis.¹⁻⁴ Malalignment, and the resulting compensatory response to maintain upright posture, places additional strain on key spinal, pelvic, and lower extremity structures that can cause arthritis and pain.^{5,6} Restoring proper alignment in the coronal and sagittal plane is therefore essential for improving biomechanical

efficiency and preventing further progression of disease.^{5,7,8} In clinical practice, preoperative radiographic assessment of spinal alignment is conducted by measuring key angles and

¹ Icahn School of Medicine at Mount Sinai, New York, NY, USA

* Brian H. Cho and Deepak Kaji contributed equally to this work.

Corresponding Author:

Samuel K. Cho, Department of Orthopaedic Surgery, Icahn School of Medicine at Mount Sinai, 425 West 59th Street, 5th Floor, New York, NY 10019, USA.
Email: samuel.cho@mountsinai.org



distances using various landmarks and comparing them with established alignment targets.^{3,6}

Poor sagittal alignment in particular has recently been associated with negative health-related quality of life (HRQoL).^{2,9} Sagittal alignment is characterized by 5 key radiographic parameters: cervical lordosis (CL), thoracic kyphosis (TK), pelvic incidence minus lumbar lordosis (PI-LL), sagittal vertical axis (SVA), and pelvic tilt (PT).^{3,6} While surgeons require these manually acquired radiographic measurements for presurgical planning, the process is both time-consuming and prone to rater-dependent error.^{10,11} Furthermore, previous studies have demonstrated significant differences between standing and supine angle measurements.^{4,12} Intraoperative surgical measurements better reflect postoperative sagittal balance because rigid fixation from instrumentation can prevent the passive corrections in spinopelvic parameters that occur in the supine position.¹³ Therefore, the development of automated tools for making rapid, intraoperative calculation of sagittal parameters may be useful for proper evaluation of alignment correction and to improve surgical decision making.

Machine learning, deep learning in particular, is being deployed on medical data to triage patients, automate preoperative planning, and predict outcomes for surgeons.¹⁴⁻¹⁶ Within the field of orthopedics, segmentation using neural networks is a particularly promising technique for automatically identifying bony structures in medical images. Spurred by advances in computer vision techniques, multiple groups have attempted to apply artificial intelligence and other computational methods on radiographs to combat the numerous sources of variability inherent to radiographic angle measurement. While many studies have attempted coronal Cobb angle calculations, they required a large amount of a priori annotation by physicians on the input radiographs. This level of input may help reduce variability but has little intraoperative value and precludes robust spinal curvature assessment at scale.^{17,18} Other groups attempted to work directly on unannotated radiographs to identify landmarks but did not measure any sagittal angles.^{19,20}

The purpose of this study was to develop a novel fully automated machine learning pipeline that reliably measures lumbar lordosis (LL) from radiographic images. Our methods do not require any a priori feature engineering or landmark identification. We report a combined segmentation and computer vision pipeline to measure lumbar lordosis in 0.14 seconds with potential perioperative value.

Materials and Methods

Materials

A total of 780 radiographs were collected from patients who received a lateral lumbar X-ray at our orthopedics department over a 1-year period. All radiographs were standardized and taken by an X-ray technician with the patient standing neutral weightbearing. Only 1 radiograph was selected for each patient to reduce potential bias of training the model on multiple X-rays from a single patient. Any patients with prior spine

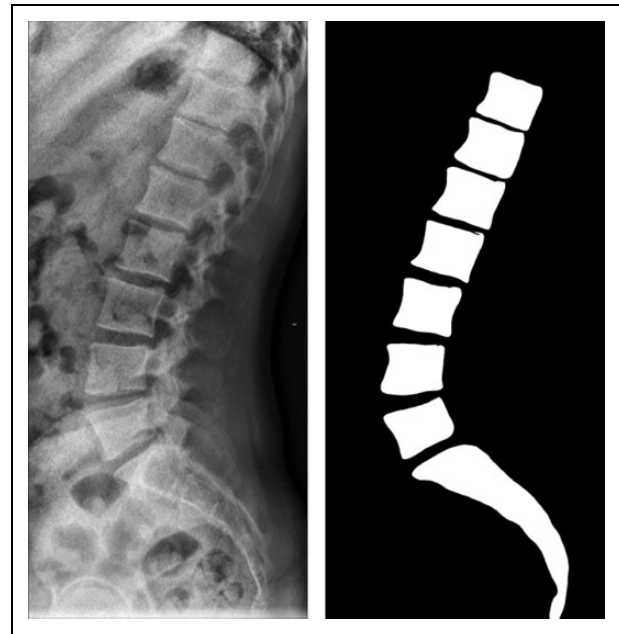


Figure 1. Example of a raw radiograph and its corresponding manually generated binary mask.

surgery or spine instrumentation were excluded. Standing lateral X-rays were used because they are higher quality and better standardized than intraoperative X-rays, increasing the probability of successfully training the model. Binary masks were generated by manually annotating every vertebral body in each radiograph using Photoshop (Adobe Systems, San Jose, CA) (Figure 1). This study was approved by our institutional review board.

Model Training and Optimization

Radiographs were preprocessed using adaptive histogram equalization to improve contrast and normalize signal intensity.²¹ The dataset was split into 629 learning images (80% of total data) and 151 test images (20% of total data) for future performance testing. The learning data was then synthetically augmented using a custom script to 12 580 images and further split into 70% training and 30% validation data (Figure 2). The augmentations included flipping, randomly rotating, and randomly cropping the radiographs to incorporate natural variations inherent to clinical radiographs into the training dataset.

We utilized U-Net, a well-established convolutional neural net (CNN) architecture for segmentation, to generate bone segmentations of the radiographs by optimizing dice similarity coefficient (DSC) loss.^{22,23} DSC is a standard metric for evaluating the accuracy of segmentations by comparing the overlap between CNN generated segmentations and manually generated masks. While a loss function defined by a simple pixel accuracy may be more intuitive, the DSC outperforms these more naive approaches for segmentation based problems.²³ The final model was trained using batch size of 20 on a

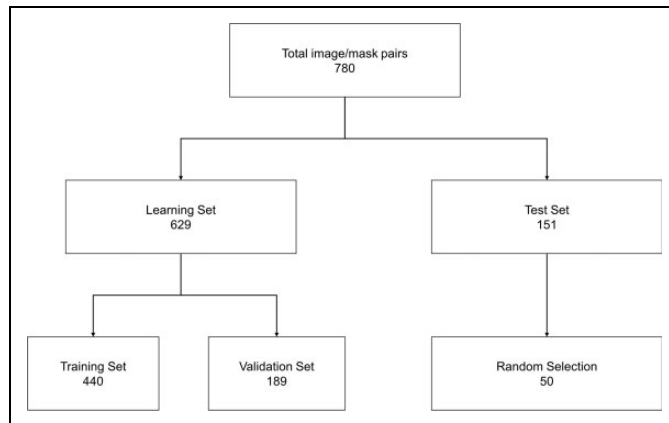


Figure 2. Overview of data workflow for training and testing the U-Net. Augmentation involved flips, random rotations, and random zooms. Each dataset was randomized prior to splitting.

NVIDIA GeForce GTX 1080-TI GPU for 200 epochs over 24.4 hours.

Beginning with raw radiographs, the algorithm first segments the image. Then, it automatically identifies the L1 and S1 vertebra from the segmentation and approximates their superior endplates to calculate the LL angle (Figure 3). Failure to properly identify L1 and S1 was treated as an algorithm failure, and no Cobb angle was measured for these cases. The algorithm was written using Python (version 3.5) and Keras (version 2.2.0).^{24,25}

Statistical Analysis

The segmentation performance was evaluated using DSC and the area under the receiver operating characteristic curve (AUC). The algorithm-generated angles were compared to manual angle measurements from a chief resident (JK), spine fellow (SL), and attending surgeon (SKC) on 50 randomly selected radiographs from the test dataset using Welch's 2-sample *t* test. Each surgeon measured every radiograph twice over a 2-week period to evaluate intra- and interrater reliability using interrater correlation coefficient (ICC) with a 2-way random effects model.²⁶ Measurements from surgeon 3 (SKC) were used as the gold standard for comparison. All statistical analysis was performed with R (version 3.4.4).²⁷

Results

Automatic Segmentation of Vertebral Bodies

The final U-Net achieved a training DSC of 0.966 and validation DSC of 0.923 with an overall training time of approximately 24.4 hours. The U-Net performed well on segmenting the test dataset, with a test DSC of 0.821, AUC of 0.914, and accuracy of 0.862 (Figure 4).

Calculation of Lumbar Lordosis Angle From Segmentations

The algorithm measured LL Cobb angles for 127 of the 151 radiographs in the test dataset, an overall success rate of 84.1%, with an average speed of 0.14 seconds/radiograph. From the 151-image test set, 50 radiographs were randomly chosen for surgeon measurement. On comparison, the algorithm succeeded in measuring angles for 42 images, a success rate of 84%.

Manual measurements among the surgeons demonstrated excellent intra- and interrater reliability, consistent with previous studies evaluating the surgeon reliability of radiographic spine angle measurements, with an overall ICC of 0.958 (95% CI: 0.931-0.976).^{10,11,28-30} The intrarater correlation coefficients (IaCC) for surgeons 1, 2, and 3 were 0.933 (95% CI: 0.878-0.963), 0.984 (95% CI: 0.970-0.991), and 0.970 (95% CI: 0.945-0.984), respectively. The algorithm had higher variability in the mean absolute difference (MAD) from gold standard measurements, with a standard deviation of 12.989°, compared with 3.232° and 3.152° for surgeons 1 and 2 (JK, LS), respectively (Table 1). However, the algorithm still achieved good accuracy, with an overall mean absolute angle difference of 8.055 degrees and was not statistically different from the gold standard measurements ($P = .372$). Compared with the gold standard measurements, the algorithm was even more accurate with a median absolute angle difference of 6.965°, and this was also not statistically different ($P = .161$).

The sorted bar plot of the raw angle differences between the algorithm and gold standard measurements demonstrated higher rates of overestimation than underestimation (Figure 5). A subpopulation analysis revealed much lower variation in absolute angle difference in the images with 6 compared with 7 or more segmented vertebral bodies (Figure 6).

Discussion

Accurate measurement of radiographic parameters is essential for proper assessment of sagittal alignment and surgical planning.^{31,32} The present study demonstrates the first fully automated system for the assessment of sagittal alignment on routine lumbar imaging. The pipeline in this study demonstrates strong segmentation quality (assessed by DSC) and accurate spinopelvic measurement when compared with 3 orthopedic surgeons.

While the advent of digital radiographs and computer-assisted measurement software have simplified the process of acquiring those parameters, they still require extensive manual input from the surgeon—increasing surgeon demand and introducing potential for interrater variability.^{28,30,33} Previous work on automatic extraction of spinal parameters focused mostly on computed tomography and magnetic resonance data, as they contain higher-resolution data with less noise and allow for 3-dimensional reconstruction.³⁴⁻³⁷ These imaging modalities allow for high-fidelity segmentations and measurement of additional parameters such as apical vertebral rotation but are not routinely used for monitoring or intraoperative imaging due

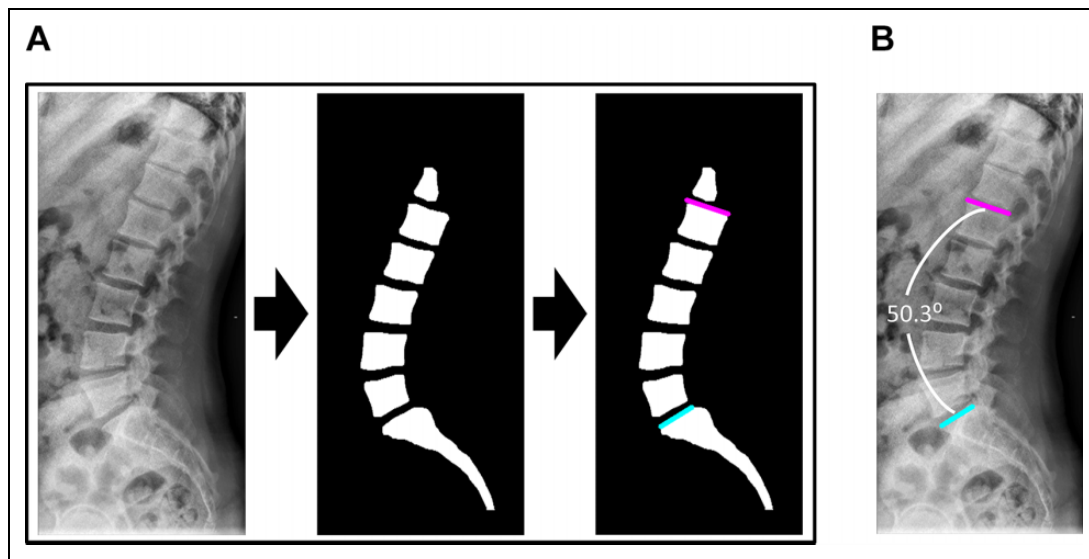


Figure 3. Overview of algorithm workflow for automatic lumbar lordosis angle calculation. (A) The raw radiograph is captured and pre-processed. Bony segmentation is generated from the raw radiograph with the trained U-Net. The LI and SI slopes are identified from the segmented image with a computer vision algorithm. (B) Overlay of the LI and SI slopes on the raw radiograph demonstrates proper slope placement and accurate angle estimation.

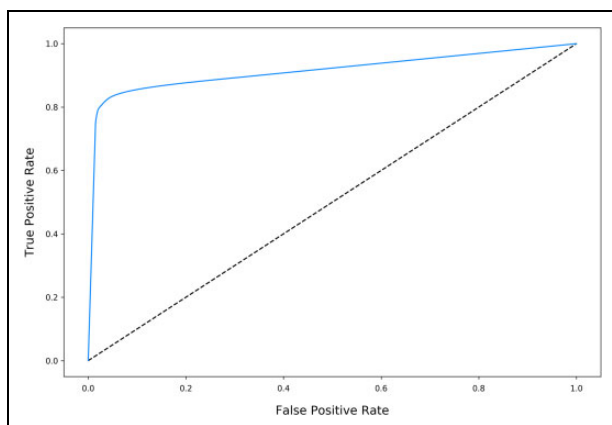


Figure 4. Receiver operating characteristic (ROC) of the U-Net for the test dataset. The overall test area under the ROC curve (AUC) was 0.914 and the overall test accuracy was 0.862. The dotted line denotes AUC = 0.50.

to exposure to radiation and high price, limiting their clinical utility.³⁸⁻⁴⁰ Therefore, this study aimed to evaluate the effectiveness of a fully automated, rapid LL Cobb angle measuring algorithm on lateral lumbar radiographs.

Previous groups have utilized machine learning models such as faster region-based convolutional neural networks (faster R-CNN) as well as traditional computer vision techniques to automatically localize the spine on radiographs.^{17,41-44} However, many of these studies were limited by small sample sizes, due to a lack of widely available source of labeled data, and utilized single vertebra-level segmentation—increasing complexity of the model and thus potential for error.^{17,41} One systematic review of Cobb angle measurement also noted that all previous Cobb angle computerized approaches, even ones that were deemed “automatic,” require landmark identification or another user input to generate each measurement, reducing scalability.⁴⁵ Recent work by Al Arif, et al⁴⁶ demonstrated how a high-performance cervical vertebra segmentation algorithm

Table 1. Absolute Angle Difference Performance Metrics.

| Operator | Minimum | Q1 | Median | Mean | Q3 | Maximum | SD | P ^a |
|---|---------|-------|--------|--------|--------|---------|--------|----------------|
| Relative to gold standard (deg) | | | | | | | | |
| Algorithm | 0.668 | 3.810 | 6.965 | 13.441 | 21.857 | 50.528 | 12.989 | 0.161 |
| Surgeon 1 | 0.300 | 1.762 | 4.050 | 4.474 | 6.650 | 14.000 | 3.232 | 0.224 |
| Surgeon 2 | 0.100 | 1.375 | 3.050 | 3.529 | 4.825 | 18.400 | 3.152 | 0.460 |
| Relative to overall surgeon average (deg) | | | | | | | | |
| Algorithm | 0.187 | 3.815 | 8.055 | 13.069 | 19.834 | 54.395 | 13.126 | 0.372 |

^aP values computed using Welch's 2-sample t test.

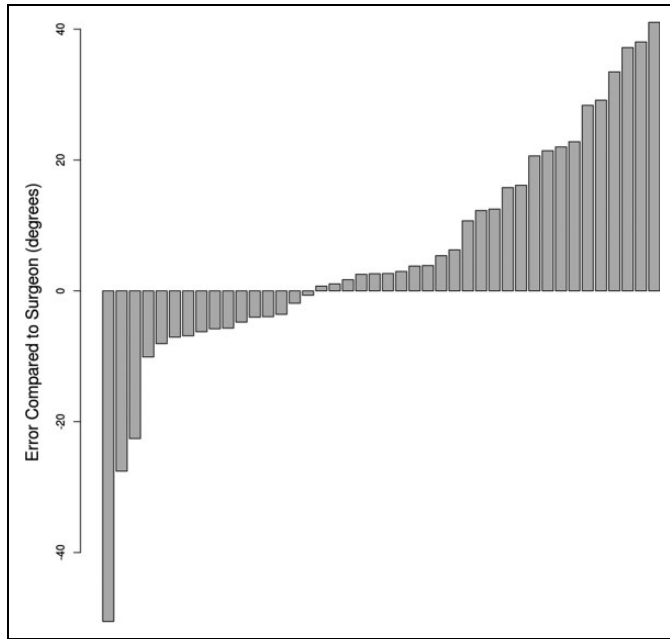


Figure 5. Sorted bar plot of predicted angle error compared to the gold standard measurements ($n = 42$). The algorithm overestimated in 26 radiographs and underestimated in 16 radiographs.

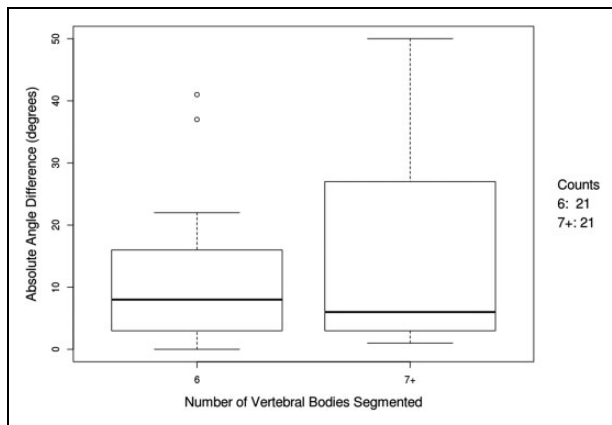


Figure 6. Box-whisker plot of absolute angle difference for radiographs with 6 and 7+ vertebral bodies segmented.

that achieved an average DSC of 0.944 suffered decreases in performance to a DSC of 0.840 when integrated to a fully automatic workflow due to accumulation of error in the multi-model pipeline.⁴⁶ In comparison, our model segmented the entire radiograph at once and still achieved a similar test DSC of 0.852 on lateral lumbar radiographs, which contains additional physiological radiopaque artifacts such as bowel gas and panniculus that increases segmentation difficulty. Most of the algorithm failures were due to segmentation of too few vertebral bodies (<5 lumbar + 1 sacrum) or due to the mis-segmentation of separate vertebral bodies as a single, fused body (Figure 7). The importance of segmentation quality on performance was also demonstrated in our subpopulation analysis (Figure 6), where the variability in absolute angle

difference was much higher in the radiographs with higher than expected number of segmented vertebral bodies (7+ vs 6).

Segmentation allows the algorithm to determine where the vertebral bodies are located in a radiograph, just as spinal surgeons identify the vertebral bodies before identifying the end plates to use for measurement. While this may seem like a trivial task, the large variability in posture and vertebral shape as well as the presence of various radiopaque artifacts make this a difficult problem using traditional computer vision techniques. Computer vision algorithms exhibit robust performance on visual tasks and are largely insensitive to a lack of training data, but they tend to fail on images with high complexity such as radiographs. Deep learning algorithms and CNNs, on the other hand, can tolerate higher complexity but require training on massive datasets in order to identify refined features. While the size of our dataset is significantly larger than those reported in other spine segmentation studies, training a neural network to directly predict the angle from raw radiographs would require an unreasonably large dataset. Segmentation networks are therefore necessary to reduce the complexity of the input image so that robust computer vision algorithms can be utilized. Further improving the robustness of segmentation will therefore be essential for improving the performance of future algorithms, as the computer vision techniques rely on the algorithm-generated segmentation to determine which landmarks to use for measurement.

Our overall median absolute angle difference of 8.055° is larger than the error margins from surgeons (Table 1). While the t test showed that the algorithm measurements were not significantly different from surgeon measurements, demonstrating good accuracy, the standard deviation was much higher for the algorithm, demonstrating lower precision. Considering the lower precision relative to surgeon measurement, this algorithm may find perioperative clinical utility in an assistive capacity. We envision the algorithm could be integrated into manual tools for digital radiograph measurement to provide a visualized default measurement suggestion similar to that seen in Figure 8. The surgeon could then adjust the interactive measurement visualization tools as needed from the automatically generated starting point, reducing surgeon input compared to fully manual measurement. Even in cases of inaccurate measurement suggestion, the algorithm often still succeeds in locating one of the necessary end plates (Figure 8b). Thus, deployment of this algorithm could provide clinical utility despite its lower precision by providing a bridge between manual measurement and fully automated measurement. Before fully automating the measurement of PI – LL with clinically acceptable error, consideration should be made to achieve absolute error and standard deviation that are sufficiently small to avoid affecting management. While there have been attempts at automating the measurement of Cobb angles in adolescent idiopathic scoliosis,^{17,47,48} there is an unfortunate lack of literature on automating the measurement of other radiographic parameters such as pelvic incidence and sagittal alignment that are important for grading and surgical planning.⁴⁹ Further work in providing near real-time measurements

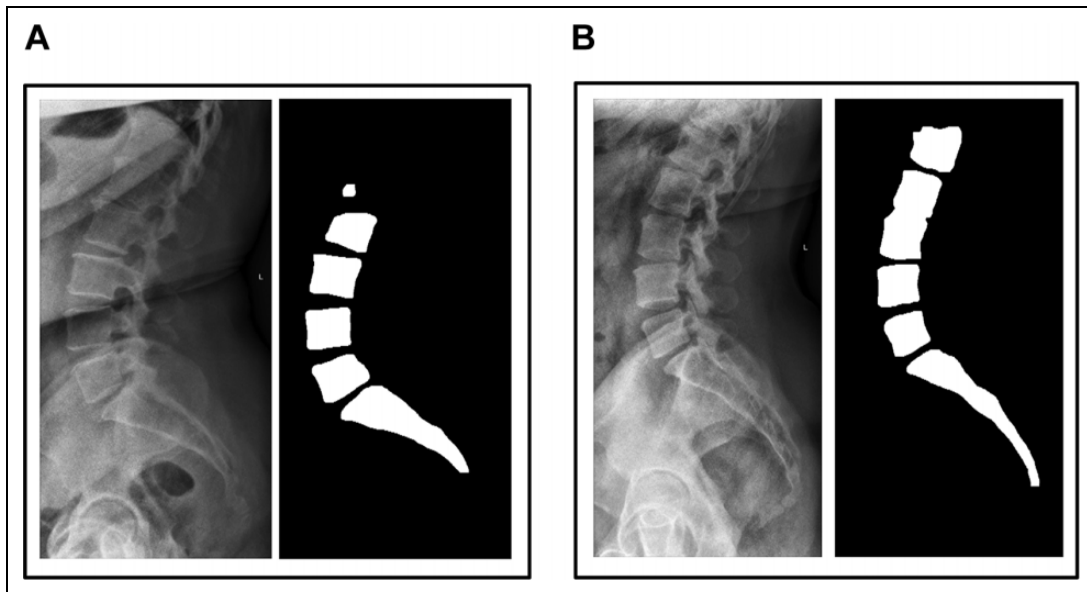


Figure 7. Examples of U-Net segmentation failures. (A) Image and corresponding segmentation characterized by L1 segmentation failure. (B) Image and corresponding segmentation characterized by fused L2 and L3 vertebrae.

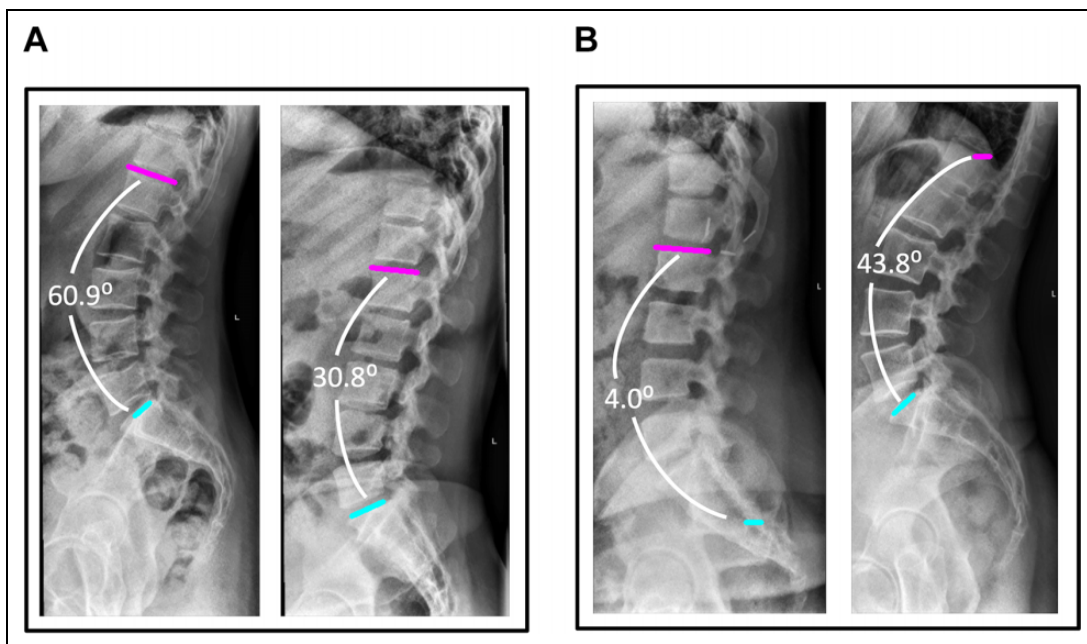


Figure 8. Computer-generated visualizations of end plate location and angle measurement for accurate and inaccurate algorithm measurements. (A) Accurate algorithm measurements corresponding to gold standard measurements of 61.1° (left) and 32.8° (right). (B) Inaccurate algorithm measurements corresponding to gold standard measurements of 37.7° (left) and 66.4° (right).

from radiographs may increase the utility of these measurements in the outpatient as well as pre-, intra-, and postoperative settings, leading to improved screening and/or surgical outcomes.

This study has several limitations. The first limitation is that radiographs with any implants were excluded. This exclusion criteria simplified the process of generating masks, which allowed us to create a large enough database for robust training,

optimization, and testing. However, it limits the utility of our tool in the setting of postoperative follow up or revision, as a U-Net trained solely on radiographs without implants is most likely to include the highly radiopaque implant in the segmentation. A large number of image/mask pairs with advanced augmentation techniques may be necessary to account for the high variability in the types of implants used and the levels at which they are placed.⁴⁶ Another limitation is that all the

radiographic data originated from a single hospital system, which may not reflect the differences in X-ray machines and acquisition techniques used in other institutions that could introduce additional artifacts to the radiographs. Additionally, this study did not stratify the radiographs by severity of ASD, which has been shown to increase the MAD for LL Cobb angle between human observers.¹¹ Severe deformities in the vertebral bodies that are not included in the training data may affect the quality of the segmentation, decreasing the performance of the algorithm. Finally, our comparison analysis excluded radiographs that the algorithms failed to identify the L1 and S1 vertebra on. This may have biased our analysis, as we do not know how the algorithm would have performed on them if the segmentation quality was higher.

Conclusion

This is the first published fully automatic algorithm that measures the LL Cobb angle using lateral lumbar radiographs. Deep learning, in combination with computer vision, is a promising tool in automating the measurement of various radiographic parameters. Our algorithm accurately measures the LL Cobb angle and is not statistically different from manual measurements made by surgeons, suggesting potential clinical utility in an assistive capacity. As techniques in bone segmentation and computer vision improve over time, these types of tools may prove useful in preoperative and intraoperative surgical decision making. Future work should focus on improving the quality of segmentation to increase the reliability of measurements. We envision this to be possible with multilabel segmentation nets that incorporate the femoral heads, which would allow us to calculate important measurements such as pelvic tilt, pelvic incidence, and PI – LL.



Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

John T. Schwartz, BS  <https://orcid.org/0000-0003-3682-5605>
Samuel K. Cho, MD  <https://orcid.org/0000-0001-7511-2486>

References

- Liu H, Li S, Zheng Z, Wang J, Wang H, Li X. Pelvic retroversion is the key protective mechanism of L4-5 degenerative spondylolisthesis. *Eur Spine J*. 2015;24:1204-1211.
- Scheer JK, Tang JA, Smith JS, et al; International Spine Study Group. Cervical spine alignment, sagittal deformity, and clinical implications: a review. *J Neurosurg Spine*. 2013;19:141-159.
- Schwab FJ, Blondel B, Bess S, et al; International Spine Study Group. Radiographical spinopelvic parameters and disability in the setting of adult spinal deformity: a prospective multicenter analysis. *Spine (Phila Pa 1976)*. 2013;38:E803-E812.
- Brink RC, Colo D, Schlösser TPC, et al. Upright, prone, and supine spinal morphology and alignment in adolescent idiopathic scoliosis. *Scoliosis Spinal Disord*. 2017;12:6.
- Lafage R, Liabaud B, Diebo BG, et al. Defining the role of the lower limbs in compensating for sagittal malalignment. *Spine (Phila Pa 1976)*. 2017;42:E1282-E1288.
- Diebo BG, Varghese JJ, Lafage R, Schwab FJ, Lafage V. Sagittal alignment of the spine: what do you need to know? *Clin Neurol Neurosurg*. 2015;139:295-301.
- Jackson RP, McManus AC. Radiographic analysis of sagittal plane alignment and balance in standing volunteers and patients with low back pain matched for age, sex, and size. A prospective controlled clinical study. *Spine (Phila Pa 1976)*. 1994;19:1611-1618.
- Gelb DE, Lenke LG, Bridwell KH, Blanke K, McEnery KW. An analysis of sagittal spinal alignment in 100 asymptomatic middle and older aged volunteers. *Spine (Phila Pa 1976)*. 1995;20:1351-1358.
- Merrill RK, Kim JS, Leven DM, Kim JH, Cho SK. Beyond pelvic incidence-lumbar lordosis mismatch: the importance of assessing the entire spine to achieve global sagittal alignment. *Global Spine J*. 2017;7:536-542.
- Hicks GE, George SZ, Nevitt MA, Cauley JA, Vogt MT. Measurement of lumbar lordosis: inter-rater reliability, minimum detectable change and longitudinal variation. *J Spinal Disord Tech*. 2006;19:501-506.
- Hong JY, Suh SW, Modi HN, Hur CY, Song HR, Park JH. Reliability analysis for radiographic measures of lumbar lordosis in adult scoliosis: a case-control study comparing 6 methods. *Eur Spine J*. 2010;19:1551-1557.
- Salem W, Coomans Y, Brismée JM, Klein P, Sobczak S, Dugailly PM. Sagittal thoracic and lumbar spine profiles in upright standing and lying prone positions among healthy subjects: influence of various biometric features. *Spine (Phila Pa 1976)*. 2015;40:E900-E908.
- Lehman RA Jr, Lenke LG, Helgeson MD, Eckel TT, Keeler KA. Do intraoperative radiographs in scoliosis surgery reflect radiographic result? *Clin Orthop Relat Res*. 2010;468:679-686.
- Titano JJ, Badgeley M, Schefflein J, et al. Automated deep-neural-network surveillance of cranial images for acute neurologic events. *Nat Med*. 2018;24:1337-1341.
- Kim JS, Merrill RK, Arvind V, et al. Examining the ability of artificial neural networks machine learning models to accurately predict complications following posterior lumbar spine fusion. *Spine (Phila Pa 1976)*. 2018;43:853-860.
- Lafage R, Pesenti S, Lafage V, Schwab FJ. Self-learning computers for surgical planning and prediction of postoperative alignment. *Eur Spine J*. 2018;27(suppl 1):123-128.
- Zhang J, Li H, Lv L, Zhang Y. Computer-aided Cobb measurement based on automatic detection of vertebral slopes using deep neural network. *Int J Biomed Imaging*. 2017;2017:9083916.
- Sun H, Zhen X, Bailey C, Rasoulinejad P, Yin Y, Li S. Direct estimation of spinal Cobb angles by structured multi-output regression. In: Niethammer M, Styner M, Aylward S, eds.

- Information Processing in Medical Imaging. Lecture Notes in Computer Science.* Vol 10265. Cham, Switzerland: Springer International; 2017:529-540.
19. Galbusera F, Bassani T, Costa F, Brayda-Bruno M, Zerbi A, Wilke HJ. Artificial neural networks for the recognition of vertebral landmarks in the lumbar spine. *Comput Methods Biomech Biomed Eng Imaging Vis.* 2018;6:447-452.
 20. Wu H, Bailey C, Rasoulinejad P, Li S. Automatic landmark estimation for adolescent idiopathic scoliosis assessment using BoostNet. In: Descoteaux M, Maier-Hein L, Franz A, Jannin P, Collis DL, Duchesne S, eds. *Medical Image Computing and Computer Assisted Intervention—MICCAI 2017.* Cham, Switzerland: Springer; 2017:127-135.
 21. Pizer SM, Philip Amburn E, Austin JD, et al. Adaptive histogram equalization and its variations. *Comput Vision Graphic Image Process.* 1987;39:355-368.
 22. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. arXiv:1505.04597 [cs.CV] 2015. <https://arxiv.org/abs/1505.04597>. Published May 18, 2015. Accessed July 18, 2019.
 23. Zou KH, Warfield SK, Bharatha A, et al. Statistical validation of image segmentation quality based on a spatial overlap index. *Acad Radiol.* 2004;11:178-189.
 24. van Rossum G. *Python Reference Manual.* Amsterdam, Netherlands: CWI (Centre for Mathematics and Computer Science); 1995.
 25. Sato M, Horie K, Hara A, et al. Application of deep learning to the classification of images from colposcopy. *Oncol Lett.* 2018;15:3518-3523.
 26. Gisev N, Bell JS, Chen TF. Interrater agreement and interrater reliability: key concepts, approaches, and applications. *Res Social Adm Pharm.* 2013;9:330-338.
 27. Core Team R. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing; 2013.
 28. Tanure MC, Pinheiro AP, Oliveira AS. Reliability assessment of Cobb angle measurements using manual and digital methods. *Spine J.* 2010;10:769-774.
 29. Zhang J, Lou E, Le LH, Hill D, Raso J, Wang Y. Computer-assisted Cobb angle measurement on posteroanterior radiographs. *Stud Health Technol Inform.* 2008;140:151-156.
 30. Wu W, Liang J, Du Y, et al. Reliability and reproducibility analysis of the Cobb angle and assessing sagittal plane by computer-assisted and manual measurement tools. *BMC Musculoskelet Disord.* 2014;15:33.
 31. Zhao X, Du L, Xie Y, Zhao J. Effect of lumbar lordosis on the adjacent segment in transforaminal lumbar interbody fusion: a finite element analysis. *World Neurosurg.* 2018;114:e114-e120.
 32. Sparrey CJ, Bailey JF, Safaee M, et al. Etiology of lumbar lordosis and its pathophysiology: a review of the evolution of lumbar lordosis, and the mechanics and biology of lumbar degeneration. *Neurosurg Focus.* 2014;36:E1.
 33. Zhang J, Lou E, Shi X, et al. A computer-aided Cobb angle measurement method and its reliability. *J Spinal Disord Tech.* 2010;23:383-387.
 34. Chu C, Belavy DL, Armbrrecht G, Bansmann M, Felsenberg D, Zheng G. Fully automatic localization and segmentation of 3D vertebral bodies from CT/MR images via a learning-based method. *PLoS One.* 2015;10:e0143327.
 35. Forsberg D, Lundström C, Andersson M, Knutsson H. Model-based registration for assessment of spinal deformities in idiopathic scoliosis. *Phys Med Biol.* 2014;59:311-326.
 36. Daenzer S, Freitag S, von Sachsen S, et al. VolHOG: a volumetric object recognition approach based on bivariate histograms of oriented gradients for vertebra detection in cervical spine MRI. *Med Phys.* 2014;41:082305.
 37. Cerveri P, Manzotti A, Marchente M, Confalonieri N, Baroni G. Mean-shifted surface curvature algorithm for automatic bone shape segmentation in orthopedic surgery planning: a sensitivity analysis. *Comput Aided Surg.* 2012;17:128-141.
 38. Ilharreborde B, Dubousset J, Le Huec JC. Use of EOS imaging for the assessment of scoliosis deformities: application to postoperative 3D quantitative analysis of the trunk. *Eur Spine J.* 2014;23(suppl 4):S397-S405.
 39. Kato S, Debaud C, Zeller RD. Three-dimensional EOS analysis of apical vertebral rotation in adolescent idiopathic scoliosis. *J Pediatr Orthop.* 2017;37:e543e547.
 40. Zheng YP, Lee TT, Lai KK, et al. A reliability and validity study for Scolioscan: a radiation-free scoliosis assessment system using 3D ultrasound imaging. *Scoliosis Spinal Disord.* 2016;11:13.
 41. Sa R, Owens W, Wiegand R, Chaudhary V. Fast scale-invariant lateral lumbar vertebrae detection and segmentation in X-ray images. *Conf Proc IEEE Eng Med Biol Soc.* 2016;2016:1054-1057.
 42. Duong L, Cheriet F, Labelle H. Towards an automatic classification of spinal curves from X-ray images. *Stud Health Technol Inform.* 2006;123:419-424.
 43. Ruhan S, Owens W, Wiegand R, et al. Intervertebral disc detection in X-ray images using faster R-CNN. *Conf Proc IEEE Eng Med Biol Soc.* 2017;2017:564-567.
 44. Lecron F, Benjelloun M, Mahmoudi S. Cervical spine mobility analysis on radiographs: a fully automatic approach. *Comput Med Imaging Graph.* 2012;36:634-642.
 45. Langensiepen S, Semler O, Sobottke R, et al. Measuring procedures to determine the Cobb angle in idiopathic scoliosis: a systematic review. *Eur Spine J.* 2013;22:2360-2371.
 46. Al Arif S, Knapp K, Slabaugh G. Fully automatic cervical vertebrae segmentation framework for X-ray images. *Comput Methods Programs Biomed.* 2018;157:95-111.
 47. Zhang J, Lou E, Le LH, Hill DL, Raso JV, Wang Y. Automatic Cobb measurement of scoliosis based on fuzzy Hough transform with vertebral shape prior. *J Digit Imaging.* 2009;22:463-472.
 48. Sardjono TA, Wilkinson MH, Veldhuizen AG, van Ooijen PM, Purnama KE, Verkerke GJ. Automatic Cobb angle determination from radiographic images. *Spine (Phila Pa 1976).* 2013;38:E1256-E1262.
 49. Smith JS, Shaffrey CI, Fu KMG, et al. Clinical and radiographic evaluation of the adult spinal deformity patient. *Neurosurg Clin N Am.* 2013;24:143-156.