A novel tool to provide predictable alignment data irrespective of source and image quality acquired on mobile phones: what engineers can offer clinicians

Teng Zhang[1], PhD, MEng, ORCID: 0000-0002-5310-8766

Chuang Zhu[2], PhD, MEng

Qiaoyun Lu[2], MEng

Jun Liu[2], PhD, MEng

Ashish Diwan[3], PhD, FRACS. ORCID: 0000-0003-1037-8421

Jason Pui Yin Cheung[1], MBBS, MMedSc, PDipMDPath, MS, MD, ORCID: 0000-0002-7052-0875

Department of Orthopaedics and Traumatology[1], the University of Hong Kong

Department of Computer Science[2], Beijing University of Posts and Telecommunications

Spine Service & Spine Labs[3], St George and Southerland Clinical School, The University of New South Wales

Correspongding Authors:

Ashish Diwan a.diwan@spine-service.org

Jason Cheung cheungjp@hku.hk

Address:

Division of Spine Surgery

Department of Orthopaedics and Traumatology

The University of Hong Kong

5/F Professorial Block, Queen Mary Hospital, Pokfulam, Hong Kong

Abstract:

Purpose: Existing automated spine alignment are based on original X-rays that are not applicable for teleradiology for spinal deformities patients. We aim to provide a novel automated vertebral segmentation method enabling accurate sagittal alignment detection, with no restrictions imposed by image quality or pathology type.

Methods: 428 optical images of original sagittal X-rays taken by smartphones or screenshots for consecutive patients attending our spine clinic were prospectively collected. Of these, 300 were randomly selected and their vertebrae were labelled with Labelme. The ground truth were specialists measured sagittal alignment parameters. Pre-trained Mask R-CNN was fine-tuned and trained to predict the vertebra level(s) on the remaining 128 testing cases. The sagittal alignment parameters including the thoracic kyphosis (TK), lumbar lordosis (LL) and sacral slope (SS) were auto-detected, based on the segmented vertebra. Dice Similarity Coefficient (DSC) and mean Intersection over Union (mIoU) were calculated to evaluate the accuracy of the predicted vertebra. The detected sagittal alignments were then quantitatively compared with the ground truth.

Results: The DSC was 84.6±3.8% and mIoU was 72.1±4.8% indicating accurate vertebra prediction. The sagittal alignments detected were all strongly correlated with the ground truth (p<0.001). Standard errors of the estimated parameters had a small difference from the specialists' results (3.5°for TK and SS; 3.4° for LL).

Conclusion: This is the first study using fine-tuned Mask R-CNN to predict vertebral locations on optical images of X-rays accurately and automatically. We provide a novel alignment detection method that has a significant application on teleradiology aiding out of hospital consultations.

Keywords: Transfer learning; Mask R-CNN; Spinal deformity; Teleradiology; Out-of-hospital consultation; Automated analysis

Introduction

Spinal deformity is a prevalent spine problem and their symptoms can be disabling[1]. The sagittal alignment parameters (i.e. thoracic kyphosis, lumbar lordosis, sacral slope and pelvic tilt) are critical for clinical diagnosis, follow-up and surgical planning of spinal deformities[2,3]. The measurements of these parameters are commonly performed by trained personnel using the build-in tools of the hospital-based picture archiving and communication system (PACS). The measurement is a manual and time-consuming process, and presents with unavoidable intra- and inter-observer variabilities[4]. Hence there is interest to develop automated spinal alignment analysis tools to improve clinical productivity and the accuracy of the measurements[5]. Previous studies have demonstrated both rule-based (i.e. Surgimap, X Align, Integrated Global Alignment, etc) and artificial intelligence approaches[6-9] to detect vertebrae in the coronal plane. However, the automated segmentation of vertebrae on the sagittal plane has not been studied, which can be valuable in pre-operative planning enabling precise and standardized clinical evaluation between surgeons and improves patient management[2,3,10].

DICOM based images on hospital PACS have been superseded by internet based PACS due to its easily accessible and cost-efficient features, which can also facilitate real-time and out of hospital consultations[11]. Since the majority of spinal surgeons use mobile phones for quick communication[12], capturing optical images of the original radiographs or on the stationary PACS using their mobile phone for discussion with other clinicians is common practice (teleradiology). The convenient optical image of X-rays on a mobile phone can assist in triaging and decision-making. Thus, a combination of automated alignment detection and smartphone acquired images from varying sources will add value in the current clinical practice. However, previous automated spinal image assessments[6-9] were based on original X-rays with high and consistent image quality. This is not easily accessible in teleradiology, and for out of hospital consultation. Especially when optical images are taken by phones, inconsistencies in image quality are expected with various lighting, contrast, rotation and vibration. Therefore, an automated tool for segmenting vertebrae and performing automated alignment analysis using optical spinal images (irrespective of their source) can reduce the burden on the laborious manual alignment analysis and provide an easily accessible tool for teleradiology.

Amongst the engineering and information technology tools available, the well-developed Mask Region-based Convolutional Network (R-CNN)[13] extends Faster R-CNN[15] by adding a branch for predicting segmentation masks on each Region of Interest (RoI). This in parallel with the existing branch for classification and bounding box

regression, provide accurate and fast automated segmentation of objects (i.e. cars, trees, animals, humans, etc.). The mask branch is a small Fully Convolutional Network (FCN)[16] applied to the region of interest that predicts a segmentation mask by a pixel-to-pixel manner. Thus, they can provide accurate segmentation results of natural images, but the transfer learning of the Mask R-CNN on spinal X-rays has not been attempted.

This study aims to provide a robust automated vertebral segmentation method and to eliminate the previous restrictions of automated segmentation created by variable imaging qualities. The study objectives include 1) preparing a diverse dataset of optical images captured from the original sagittal X-rays; 2) tuning the pre-trained Mask R-CNN by natural images based on our collected dataset; and 3) evaluating the accuracy of the predicted segmentation and alignment measurements. We hypothesise that the fine-tuned Mask R-CNN can accurately predict the segmentation of the vertebrae (based on its performance in accurately segmenting natural images) and thus enable measurements of sagittal spinal alignment based on optical images of X-rays simply taken by a smartphone.

Material and methods

Dataset and image pre-processing

This study was approved by the local health regulatory authorities. A prospective consecutive collection of 428 optical images of the original sagittal X-rays of patients with spinal deformity were collected. These patients were referred to scoliosis clinic between 1 January 2019 and 30 April 2019. Images were randomly acquired using smartphones (Figure 1) or screenshot of X-rays displayed on the PACS. The collected images have various unintentional experimental settings such as the noise from devices, illumination, vibration, motion and contrast variations, resulting in different image qualities. Thus, the raw optical images of the X-Ray were of different sizes, ranging from 600×319 to 934×1384 pixels. Additionally, most of the X-Ray images contained the whole spine, but a few captured the entire body. Considering such variations existed in the X-Ray images, we firstly performed pre-process on the images.

Our first preprocessing step was to unify the input images for the subsequent training, we cropped and resized all the remaining samples to 840×360. To deal with the smaller raw images, we upscaled both horizontally and vertically using a bilinear interpolation method; for larger raw images, we cropped and downsized the height or width to ensure each input image was proper for training. Then the average intensity was normalised to reduce the influences

of illumination fluctuations. Figure 2A illustrated the preprocessing procedure using three samples. Figure 2B and Figure 2C demonstrated the region proportion statistics of 60 randomly selected samples.

Augmentation and labelling

Practically, due to the various disturbances in the external environment, an automated segment and detection system was quite sensitive to the quality of X-Ray images. Any small changes in the image acquisition process and light intensities could mislead the auto-detection system and result in poor outcomes. To relieve the effects from outside circumstances and enhance the robustness of our system, we also simulated the real-world situation and generated different types of low-quality images based on our optical images of X-rays after pre-processing. This was achieved by introducing contrast vibration, and scaling on the existing images, as well as randomised horizontal flipping and rotations ranging from 1° to 5°. An open-source software known as Labelme (MIT, Computer Science and Artificial Intelligence Laboratory) was used to indicate the coordinates of the vertebral border to provide ground truth for the segmentation (Figure 3a).

Transfer leaning and network fine-tuning

We exploited transfer learning as it is a sub-field of machine leaning, The Mask R-CNN used in this study was pre-trained on the Common Objects in Context (COCO) dataset[14], and the pre-trained model was fine-tuned with our labelled optical images of the X-rays. The vertebra mask was classified into two classes according to the location of vertebrae enabling feature learning. One class was the vertebra in the thoracic cavity, and the remaining lumbar vertebra was another class. The model was trained by the two stages. In the first 50 epochs (1000 steps per epoch), only the head layers were trained by the learning rate of 0.001, and then all layers were trained by the learning rate of 0.0001. Due to the discrete nature of the vertebra, the non-maximum suppression threshold, which controlled the percent of overlapping of the instances, was adjusted to 0.0001, indicating no overlapping between vertebrae. 300 out of the 428 images were randomly selected to train the model and 128 out of the remained images were used to test the model.

Post-processing

There were some noise masks in some segmentation images (Figure 4A). However, the noise was removed by calculating the slope change of adjacent spines, because the sagittal curvature of the spine was within a small range (0.6°-32° for this dataset). To remove the noise, firstly, the geometric center c1,c2,c3,..., (Figure 4B) was calculated according to the mask. Secondly, we calculated the slope of adjacent spines, $k1=\triangle y1/\triangle x1$, $k2=\triangle y2/\triangle x2$,..., (Figure

4C), and then obtained the slope change, k1-k2,..., . The threshold of slope change was set as 1.7, being a tangent value of 60°, and the masks were removed if the slope change of adjacent spines was more than the threshold. Figure 7D indicated the de-noised segmentation result.

Automated vertebral segmentation and alignment measurements

Key sagittal alignment parameters including the thoracic kyphosis (TK: T5-12), lumbar lordosis (LL: L1-5) and sacral slope (SS) were detected using the following steps. Pelvic tilt and pelvic incidence were not routinely measured due to inadequate exposure of the femoral heads in some images. A binary image was generated from the mask (Figure 5A) and then the outline of the mask was obtained (Figure 5B). The boundary contours were fitted with the smallest circumscribed rectangle (Figure 5C). The upper and lower border of the endplates of specific vertebra could then be detected. The least-squares method was used to get the straight line fitted to the endplates (Figure 5D), and the angle of the line could be detected by calculating the angle between the straight line. The sagittal alignment measurements made by spine specialists were considered the ground truth. The ground truth was measured using the build-in system of PACS as a part of routine clinical practice for deformity diagnosis. The specialists who measured the images were unaware of this study during the reading as they were acquiring the information as part of routine clinical practice.

Testing and statistical considerations

To evaluate the accuracy of the mask generated segmentation, the Dice similarity coefficient (DSC) and Mean Intersection over Union (mIoU) were calculated for each image. The equations are shown below,

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

$$IoU = \frac{A \cap B}{A \cup B},$$

where A is the ground truth and B is the segmentation result, gauging the similarity between A and B.

The comparison of the mean between the ground truth and the automated alignment measurements were conducted by calculating the mean and standard deviation of the absolute value of error for TK, LL and SS. In order to quantitatively compare the results obtained for the testing set with the ground truth, we also conducted linear regression and Bland-Altman analyses by using the Python library StatsModels[14]. For data reporting, R^2 represented

the square of correlation coefficient. Standard errors of the estimate (S) represented the standard deviation of the differences between the prediction results and was reported along with the *p* value.

Results

The sagittal alignment parameters of the dataset had TK ranging from 0.6° to 71.3°, LL from 15.8° to 89.6° and SS from 8.2°-72.2° indicating the diversity in spinal pathologies. With the fine-tuned Mask R-CNN, the prediction of the segmentation was accurate with average DSC of 84.6±3.8% and mIoU of 72.1±4.8%. For key alignment parameters, in comparison with the ground truth, the absolute value of mean error was within an acceptable range of 3.0°-3.3° with a standard deviation of 1.2°-1.3° (Table 1).

The overall validity of the newly proposed transfer learning approach was confirmed by our analysis of the results with small differences between the testing results and the ground truth (Table 2, Figure 6). Those differences were not apparent in the qualitative visual evaluation (Figure 4, 5). The results of the testing were strongly correlated with the ground truth. The R^2 (being the square of correlation coefficient) showed a strong correlation between the prediction result and ground truth ranging from 0.91 to 0.95, with $p<0.001$ in all cases (Table 2). The slope of the regression line was close to the ideal value of 45° (indicating a perfect match being the testing results and the ground truth) in all cases (Table 2: range 42-44°, Figure 6). Furthermore, standard errors of the estimate (S), defined as the standard deviation of the differences between the prediction result and the ground truth were 3.4° to 3.5°. For all sagittal parameters included in this study (TK, LL, and SS), Bland-Altman analysis (Figure 7) showed a minimal and negative mean difference between the ground truth and the prediction result (mean difference -0.2, -0.7 and -0.6 respectively to TK, LL, and SS).

Discussion

This study is the first to combine transfer leaning (fine-tuned pre-trained Mask R-CNN) and teleradiology for fully automated vertebral prediction and sagittal alignment prediction, whereas previous automated methods were focused on coronal alignment with other methods and used original X-ray images. To our knowledge for the first time a relatively large dataset of optical images of original X-rays displayed on PACS was established and finetuned Mask R-CNN was used to predict the vertebral locations. This dataset consists of images with different sizes, contrast, intensity, minor rotations. Moreover, some images did not capture the femoral heads while some are full-body scans.

The diversity in this dataset mimics the real-life clinical scenario, especially when using teleradiology. Potential applications of this method include accelerating spinal deformity screening, and out of hospital consultation when patients cannot have access to certain medical specialists, as well as clinical trials, in order to avoid the interrater disagreement. Furthermore, this may act as a quick data collection tool for research purposes.

Previous automated segmentation studies have demonstrated that automated alignment measurements can be generated based on Cobb angles on the coronal images of patients with idiopathic scoliosis[5-9,15,16]. It is worth noting that one study directly regressed Cobb angles from input images[16]. However, we argue that learning to detect end vertebrae as a recognition/classification task is more stable than learning to estimate Cobb angles directly as a regression task. The application of intermediate supervision can improve the reliability of the predicted results. For example, Horng *et al*[15] performed spine segmentation using a U-Net[17], and they computed Cobb angles directly from segmentation results according to its definition, which is similar to our approach using Mask R-CNN to segment the vertebra and then detect the sagittal alignment parameters. There is one recent study that used both coronal and sagittal X-rays from multiple populations as the training dataset to predict the alignment parameters[18], but this previous work was based on original X-rays archived from the PACS system, which is not applicable for teleradiology and out of hospital consultation.

For this study, pre-trained Mask R-CNN[13] on natural images (COCO dataset) were used and no new machine learning model were developed. However, using a small training dataset with 300 images, the fine-tuned Mask R-CNN significantly saved the time to develop a new machine learning method and was able to accurately predict the vertebra mask (DSC: 84.6±3.8% and mIoU: 72.1±4.8%) on a relatively large testing dataset (128 images). The testing data consists of 30% of this diversified dataset. Previous reported methods on vertebral segmentation had DSC ranging from 77.3% to 91.2%, which is comparable to the predictive accuracy of our method[19-21]. Thus, the vertebra prediction performance (Figure 3) is reliable. This accurate segmentation approach can also assist in future three-dimensional spinal reconstruction from X-ray images.

Additional to the accurate segmentation, the detection accuracy of sagittal alignment parameters including TK, LL and SS[22] is comparable with the clinicians' interrater variance[23-26]. Previous manual assessments done by experienced clinicians demonstrated absolute disagreement of approximately 1.5°[23], 3.9°[24] and 5.8°[26] for mean LL. Previous work reported large variations of TK measurements ranging from 6.2°[26] to 7.3°[27]. However, the variance in TK from previous studies was more likely to be caused by different TK measurement methods. Our

work reported a small absolute error of 3.0°-3.3° for the mean and 3.4°-3.5° for S based on the regression analyses using an automated method on diverse images. This small error indicated a possible clinical application for this method in the future.

It must be acknowledged that key sagittal alignment parameters such as pelvic tilt and pelvic incidence were not detected in this study. It was because during the photo taking process, the femoral head was not identified for some photos or screenshots of the original X-rays. These imaging tasks were performed by different technicians and research assistants. This diversity was an advantage as it mimicked the real-life scenario. However, for future work, standardised image capturing guidance should be developed to obtain improved optical images of the X-rays to study the pelvic parameters as well. Furthermore, it is important to determine whether this method can be implicated into routine clinical practice to facilitate further detection of sagittal alignment parameters in teleradiology.

For future studies, images with spinal instrumentation will be included. Postoperative alignments are equally important, and they provide information regarding surgical outcomes. A new dataset containing pre and post-operative X-rays should be established for future work and to test the methods developed in this study. We suspect further fine-tuning of the Mask R-CNN will be required when images with spinal instrumentations introduced. Nevertheless, a novel automated sagittal alignment detection method for deformity patients was developed based on optical images of X-rays. This is the first study using fine-tuned Mask R-CNN to achieve automated segmentation of vertebrae on optical images of sagittal X-rays. This robust method has significant clinical applications in deformity research and patient management by facilitating accurate and mobile alignment detection.

References

1. Wong AYL, Samartzis D, Cheung PWH, Yin Cheung JP (2019) How Common Is Back Pain and What Biopsychosocial Factors Are Associated With Back Pain in Patients With Adolescent Idiopathic Scoliosis? Clin Orthop Relat Res 477 (4):676-686. doi:10.1097/CORR.0000000000000569
2. Luk KD, Vidyadhara S, Lu DS, Wong YW, Cheung WY, Cheung KM (2010) Coupling between sagittal and frontal plane deformity correction in idiopathic thoracic scoliosis and its relationship with postoperative sagittal alignment. Spine 35 (11):1158-1164. doi:10.1097/BRS.0b013e3181bb49f3
3. Barnes D, Stemper BD, Yogananan N, Baisden JL, Pintar FA (2009) Normal coupling behavior between axial rotation and lateral bending in the lumbar spine - biomed 2009. Biomed Sci Instrum 45:131-136
4. Carman DL, Browne RH, Birch JG (1990) Measurement of scoliosis and kyphosis radiographs. Intraobserver and interobserver variation. J Bone Joint Surg Am 72 (3):328-333
5. Kundu R, Chakrabarti A, Lenka PK (2012) Cobb angle measurement of scoliosis with reduced variability. arXiv preprint arXiv 1211 (5355)
6. Zhang J, Lou E, Le LH, Hill DL, Raso JV, Wang Y (2009) Automatic Cobb measurement of scoliosis based on fuzzy Hough Transform with vertebral shape prior. J Digit Imaging 22 (5):463-472. doi:10.1007/s10278-008-9127-y
7. Zhang J, Lou E, Shi X, Wang Y, Hill DL, Raso JV, Le LH, Lv L (2010) A computer-aided Cobb angle measurement method and its reliability. J Spinal Disord Tech 23 (6):383-387. doi:10.1097/BSD.0b013e3181bb9a3c
8. Sardjono TA, Wilkinson MH, Veldhuizen AG, van Ooijen PM, Purnama KE, Verkerke GJ (2013) Automatic Cobb angle determination from radiographic images. Spine 38 (20):E1256-1262. doi:10.1097/BRS.0b013e3182a0c7c3
9. Safari A, Parsaei H, Zamani A, Pourabbas B (2019) A Semi-Automatic Algorithm for Estimating Cobb Angle. J Biomed Phys Eng 9 (3):317-326. doi:10.31661/jbpe.v9i3Jun.730
10. Eyvazov K, Samartzis D, Cheung JP (2017) The association of lumbar curve magnitude and spinal range of motion in adolescent idiopathic scoliosis: a cross-sectional study. BMC Musculoskelet Disord 18 (1):51. doi:10.1186/s12891-017-1423-6
11. Swinfen R, Swinfen P (2002) Low-cost telemedicine in the developing world. J Telemed Telecare 8 (6):63-65
12. Ozdalga E, Ozdalga A, Ahuja N (2012) The smartphone in medicine: a review of current and potential use among physicians and students. J Med Internet Res 27 (5):e128
13. He K, Gkioxari G, Dollar P, Girshick R (2018) Mask R-CNN. IEEE Trans Pattern Anal Mach Intell. doi:10.1109/TPAMI.2018.2844175
14. Seabold S, Perktold J Statsmodels: econometric and statistical modeling with python. In: The 9th Python in Science Conference, 2010. p 61
15. Horng M-H, Kuok C-P, Fu M-J, Lin C-J, Sun Y-N (2019) Cobb angle measurement of spine from x-ray images using convolutional neural network. Comput Math Method M 2019:18
16. Sun H, Zhen X, Bailey C, Rasoulinejad P, Yin Y, Li S (2017) Direct estimation of spinal cobb angles by structured multi-output regression. Paper presented at the Inf Process Med Imaging,
17. Ronneberger O, Fischer P, T. B (2015) U-net: Convolutional networks for biomedical image segmentation. Paper presented at the International Conference on Med Image Comput Comput Assist Interv
18. Galbusera F, Niemeyer F, Wilke HJ, Bassani T, Casaroli G, Anania C, Costa F, Brayda-Bruno M, Sconfienza LM (2019) Fully automated radiological analysis of spinal disorders and deformities: a deep learning approach. Eur Spine J: official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society 28 (5):951-960. doi:10.1007/s00586-019-05944-z
19. Sungkhun S, Rasmequan S, Chinnasarn K, Rodtuk A (2016) Vertebral body segmentation using aggregate superpixels. Paper presented at the IEEE: Internat Joint Conf Comput Sc Soft Eng
20. Lim PH, Bagci U, Bai L (2013) Introducing Willmore flow into level set segmentation of spinal vertebrae. IEEE Trans Biomed Eng 60 (1):115-122. doi:10.1109/TBME.2012.2225833
21. Shelhamer E, Long J, Darrell T (2017) Fully Convolutional Networks for Semantic Segmentation. IEEE Trans Pattern Anal Mach Intell 39 (4):640-651. doi:10.1109/TPAMI.2016.2572683
22. Vrtovec T, Pernus F, Likar B (2009) A review of methods for quantitative evaluation of spinal curvature. Eur Spine J: official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society 18 (5):593-607. doi:10.1007/s00586-009-0913-0
23. Harrison DE, Harrison DD, Cailliet R, Janik TJ, Holland B (2001) Radiographic analysis of lumbar lordosis: centroid, Cobb, TRALL, and Harrison posterior tangent methods. Spine 26 (11):E235-242. doi:10.1097/00007632-200106010-00003
24. Hong JY, Suh SW, Modi HN, Hur CY, Song HR, Park JH (2010) Reliability analysis for radiographic measures of lumbar lordosis in adult scoliosis: a case-control study comparing 6 methods. Eur Spine J: official publication of

the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society 19 (9):1551-1557. doi:10.1007/s00586-010-1422-x
25. Porto AB, Okazaki VHA (2017) Procedures of assessment on the quantification of thoracic kyphosis and lumbar lordosis by radiography and photogrammetry: A literature review. J Bodyw Mov Ther 21 (4):986-994. doi:10.1016/j.jbmt.2017.01.008
26. Porto AB, Okazaki VHA (2018) Thoracic Kyphosis and Lumbar Lordosis Assessment by Radiography and Photogrammetry: A Review of Normative Values and Reliability. J Manipulative Physiol 41 (8):712-723. doi:10.1016/j.jmpt.2018.03.003
27. Briggs AM, Van Dieen JH, Wrigley TV, Greig AM, Phillips B, Lo SK, Bennell KL (2007) Thoracic kyphosis affects spinal loads and trunk muscle force. Phy ther 87 (5):595-607

Figure legends

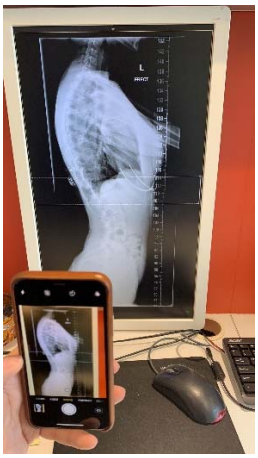
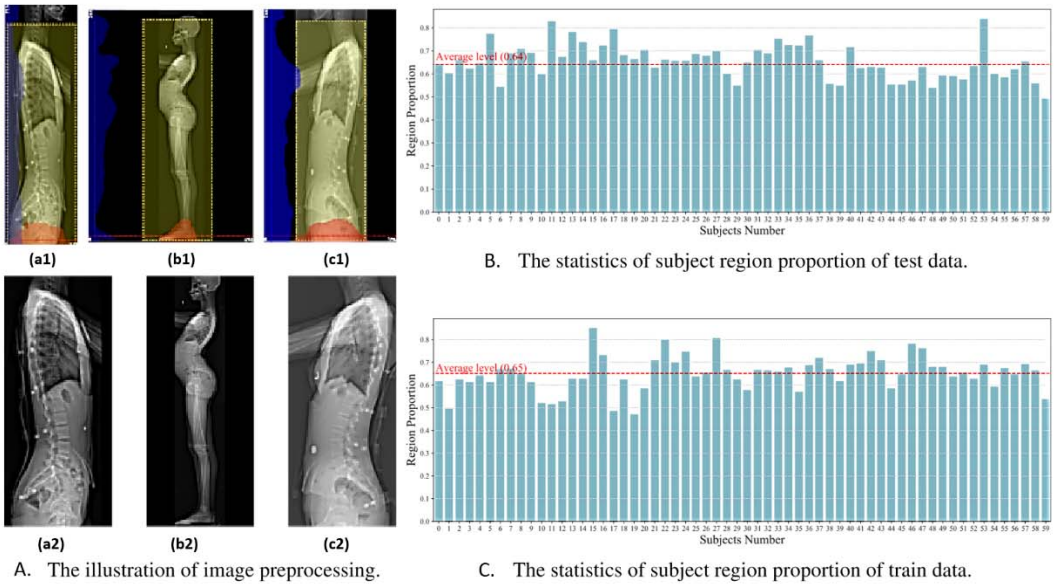


Figure 1. Acquired optical image of X-rays from a stationary PACS.



A. The illustration of image preprocessing.

B. The statistics of subject region proportion of test data.

C. The statistics of subject region proportion of train data.

Figure 2. Preprocessing of the optical images and image statistics. Image A illustrates the preprocessing procedure using three samples. Images a1, b1 and c1 are the optical images of the X-ray with different sizes and orientations. The yellow boxes are crop regions, and a2, b2 and c2 are the processed images. Image B demonstrates the statistics of region proportion of randomly selected 60 test samples after preprocessing, while image C demonstrates the statistics of randomly selected 60 training samples indicated an average level of region proportion being 0.64-0.65.

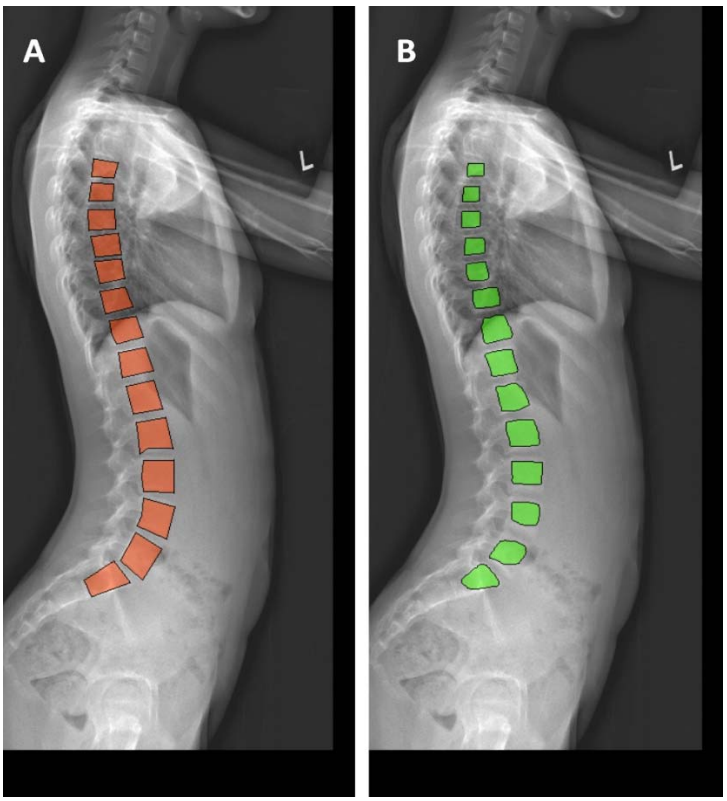


Figure 3. An example of a comparison of the ground truth and the automated segmentation result. The mask of ground truth is shown in orange (A) whereas the mask of the generating segmentation is in green (B).

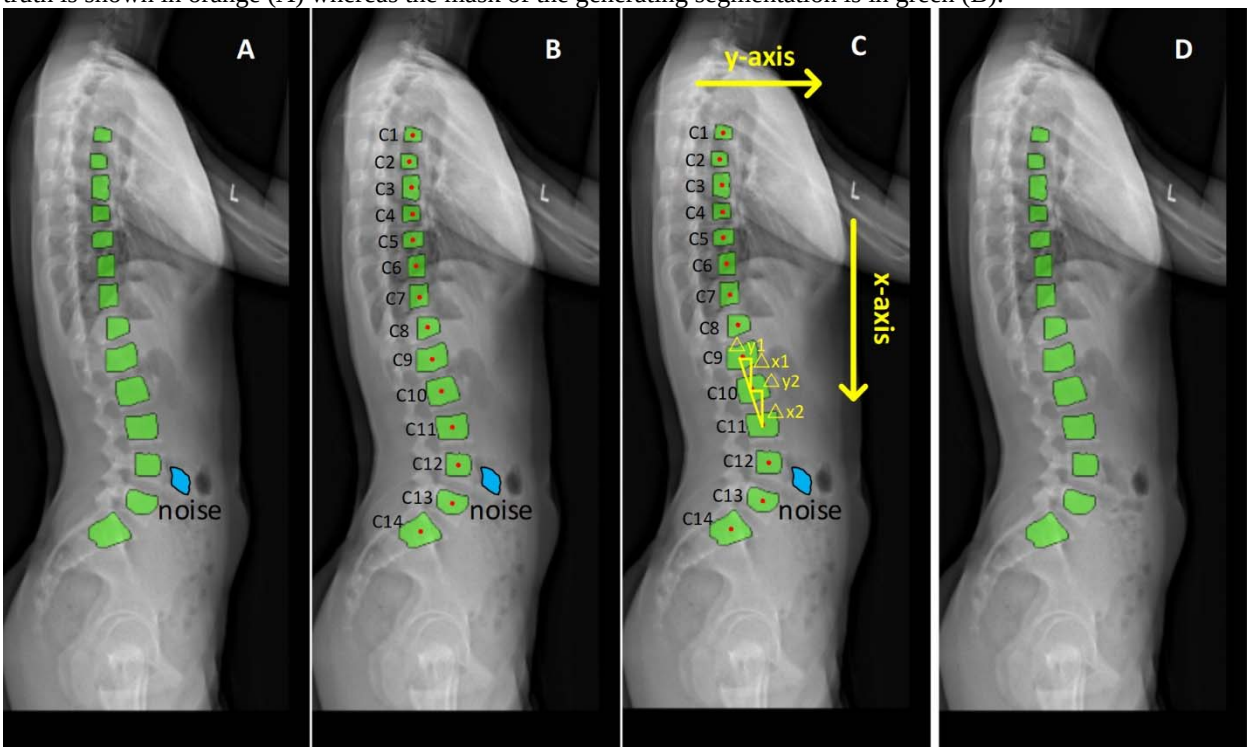


Figure 4. Some noise masks in the segmentation image. Image A illustrates the noise (blue) generated through our fine-tuned network. The centre of the vertebra was detected (B) and the slope of adjacent spine was calculated (C). If the slope was bigger than 60°, the outlier mask was considered as noise and automatically eliminated (D).

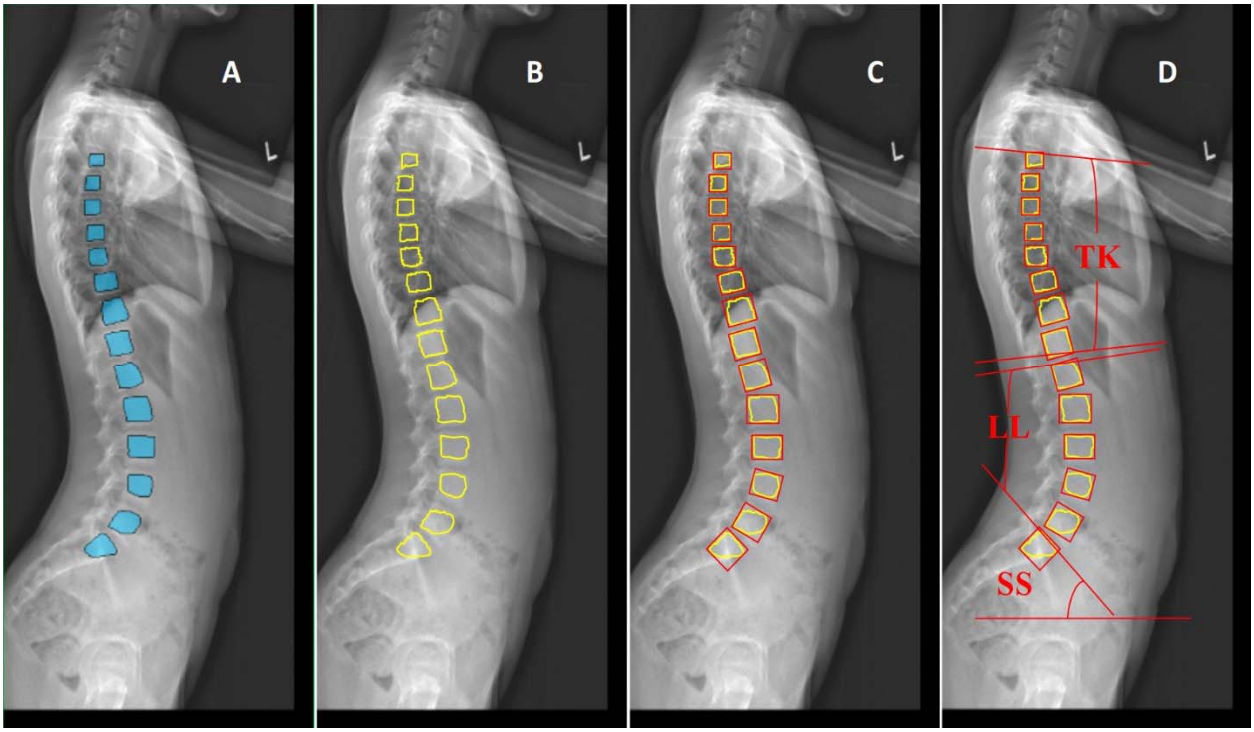


Figure 5. Automated vertebral segmentation and alignment measurements. The mask of the vertebral body was generated automatically (A) and the outline of the mask was obtained (B). The smallest circumscribed rectangle was automatically selected to fit each mask (C) and the endplates of the specific vertebra (T5, T12, L1, L5, and S1) were detected to calculate the thoracic kyphosis (TK), lumbar lordosis (LL) and sacral slope (SS) shown in image D.

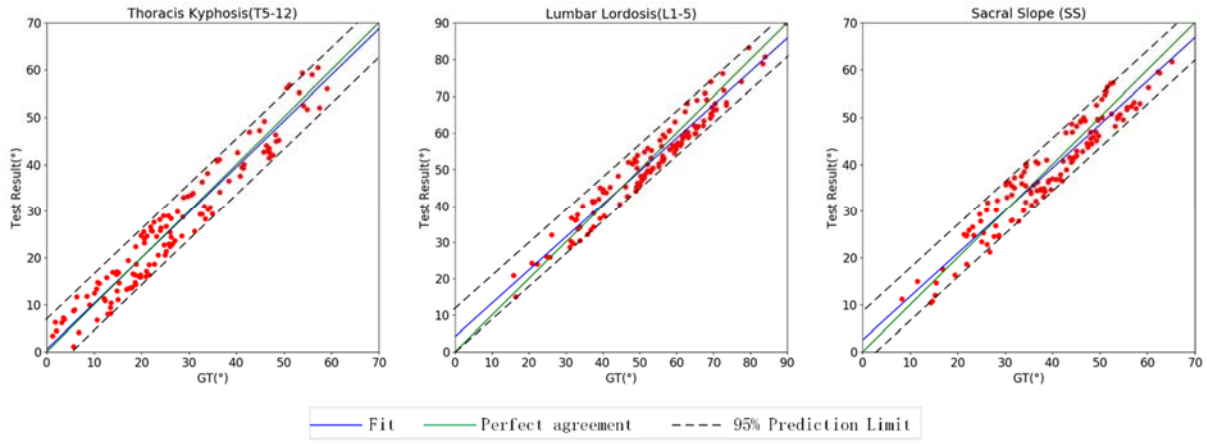


Figure 6. Regression analysis of the alignment parameters predicted by the Mask R-CNN detections (y-axis) versus the ground truth results measured manually by the spine specialists (x-axis). For TK, LL and SS, good agreement between the auto-detected degrees and the ground truth is observed. All units are in degrees.

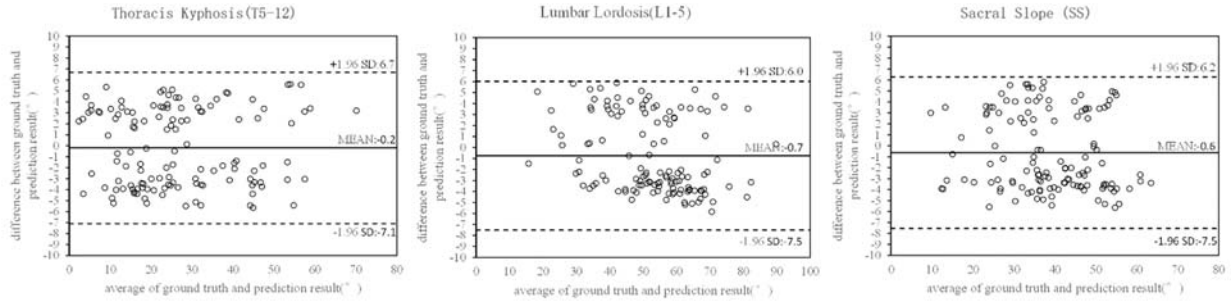


Figure 7. Bland-Altman plots comparing the agreement of sagittal alignment parameters between the Mask R-CNN detections versus the ground truth results measured manually by spine specialists. Y-axis showed the difference between the measurements between automated results and the ground truth. The X-axis represented the average of these measures ((automated results + ground truth)/2). Small mean differences from -0.7° to -0.2° were shown between the auto-detected alignment parameters and the ground truth. All units are in degrees.

Tables

Table 1. Mean and standard deviation of the absolute value of the error between the ground truth and the Mask R-CNN facilitated sagittal alignment prediction.

Parameter	Mean	Standard Deviation
Thoracic Kyphosis (T5-12)	3.0°	1.2°
Lumbar Lordosis (L1-5)	3.2°	1.3°
Sacral Slope (SS)	3.3°	1.3°

Table 2. Regression analysis of the correlation between the values of the anatomical parameters evaluated by spine specialists and those calculated by the Mask R-CNN.

Parameter	R^2	*p* value	Slope of the regression line	Standard error of the measurement(S)
Thoracic Kyphosis (T5-12)	0.944	<0.001	44°	3.5°
Lumbar Lordosis (L1-5)	0.946	<0.001	42°	3.4°
Sacral Slope (SS)	0.914	<0.001	43°	3.5°