Automatic Extraction of Spinopelvic Parameters Using Deep Learning to Detect Landmarks as Objects

AliAsghar MohammadiNasrabadi[1] ALIASGHARMN@UWATERLOO.CA
William McNally[1] WMCNALLY@UWATERLOO.CA
Gemah Moammer[2] GEMAH.MOAMMER@GRHOSP.ON.CA
John McPhee[1] MCPHEE@UWATERLOO.CA
[1] *Systems Design Engineering, University of Waterloo, Canada*
[2] *Head of Spine Program GRH, SMH and CMH McMaster University, Canada*

Editors: Under Review for MIDL 2022

Abstract

Surgeons measure spinopelvic parameters from X-ray images to evaluate spinopelvic alignment preoperatively for surgical planning. Automatic extraction of these parameters not only saves time but also provides consistent measurements, avoiding human error. In this paper, we introduce a new approach to automatic spinopelvic parameter extraction, which considers landmarks as objects. The landmarks are extracted using a deep learning object detection algorithm that can address the drawbacks of heatmap-based regression. The model is evaluated using two datasets totalling 1000 lateral spinal and pelvic X-ray images. Acceptable accuracy is achieved when comparing the reference manual parameter measurements with those obtained automatically by our prediction model.

Keywords: medical imaging, deep learning, spinopelvic parameters, object detection

1. Introduction

Surgeons use spinopelvic parameters for preoperative evaluation and diagnosis of spinal and pelvic malalignment disorders. For the analyses, they prefer to use X-ray images, which impose lower doses of radiation than CT-scans, taken in standing position to reflect true alignment under the effect of the physiological axial loads on the spine. The most important parameters to be evaluated are the Sacral Slope (SS), Pelvic Tilt (PT), Pelvic Incidence (PI), Lumbar Lordosis (LL), and Sagittal Vertical Axis (SVA). Certain anatomical landmarks are used to extract these parameters from lateral X-ray images. Manual annotation of these landmarks is time-consuming and the accuracy of this process is dependent on the annotator and image quality. To improve consistency, there has been a growing interest in automatic approaches that extract these parameters. Deep Learning (DL) methods have produced the best results to date in terms of landmark-prediction accuracy. Heatmap-based regression, the most common method used for landmark detection in recent studies (Weng et al., 2019; Korez et al., 2020; Yeh et al., 2021), has some inherent drawbacks. We are thus motivated to introduce a new approach with an inspiration taken from keypoint detection (McNally et al., 2021), in which landmarks are considered as objects using bounding boxes. An existing object detector has been adapted for our use. After a brief introduction to our approach, it is validated using two datasets containing manually annotated landmarks. The automatically extracted spinopelvic parameters are then compared with those extracted using the manual annotations.

2. Material and Method

In total, 750 (3-foot standing) lateral spine X-ray images (DS1) were collected from patients who had been referred to the Grand River Hospital (GRH), Kitchener, Canada (2016-2022). We also used another dataset of 250 lateral lumbar spine and pelvic images (DS2), provided by Intellijoint Surgical. The datasets were divided into training (80%), validation (10%), and test (10%) sets. To automatically extract the spinopelvic parameters of interest, we consider landmarks as objects, and calculate the spinopelvic parameters (SS, PT, PI, LL, and SVA). The landmarks are extracted through deep learning object detection using the YOLO algorithm (Redmon et al., 2016), which can address the drawbacks of heatmap-based regression like overlapping heatmap signals and post-processing requirements (McNally et al., 2021). Our landmark object concept is an imaginary object representation in which the coordinates of a landmark are considered at the center of a bounding box (bx, by) with equal width (bw) and height (bh) (McNally et al., 2021). In the labeled dataset, 10 classes of landmark objects have been defined (c_i), which are centers of femoral heads, and anterior/posterior points of S1, L1, C7 superior end plates, and L5 inferior end-plate. Therefore, the resulting label includes the class number and bounding box features $C_i = (c_i, bx_i, by_i, bw_i, bh_i)$. The labels have been manually annotated using Matlab (Figure 1-A). The sizes of the landmark object bounding boxes have been optimized for the best prediction accuracy.

3. Results and Discussion

The model has been evaluated using two datasets, and the results are shown in Table 1. Spinopelvic parameters are calculated from automatically extracted landmarks, and the model's prediction accuracy is evaluated based on manual labels. Considering the limited number of images, especially in DS2, our approach showed good performance. In addition, the model was able to detect required landmarks even when they are covered with an obstacle such as the protective lead shield that often covers the femoral head.

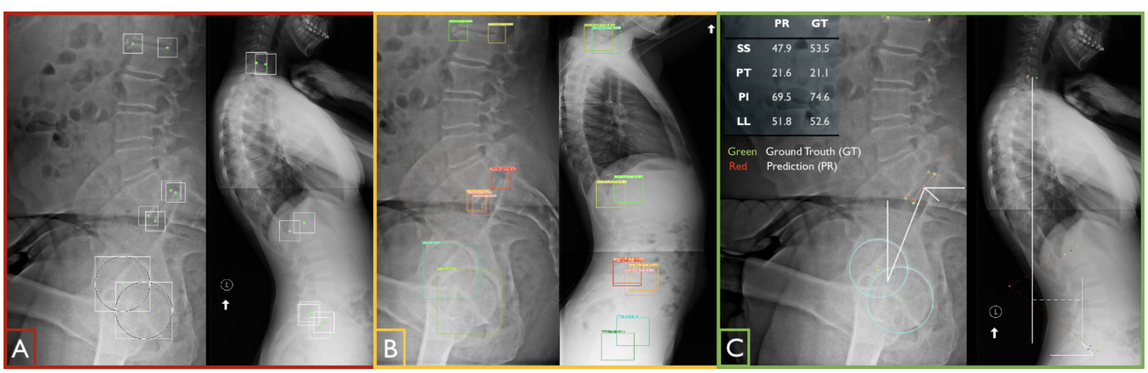

Figure 1: A: Manual labeling of landmarks and specifying bounding boxes. B: Extracted landmarks as objects with different confidences, trained model output, evaluated with the test set. C: choosing the highest confidence for each landmark as the prediction results and calculation of the spinopelvic parameters

Table 1: Statistical comparison of the main values of reference manual parameter measurements (GT) and those obtained automatically by the prediction model (PR) along with standard deviations (SD) of each parameter evaluated in the two datasets (DS1 & DS2) and relative accuracy of prediction (Acc)

	Error		**Mean PR (±SD)**		**Mean GT (±SD)**		***Acc (%)***	
	DS1	DS2	DS1	DS2	DS1	DS2	*DS1*	*DS2*
SS°	5.7	7.9	42.5 (8.7)	35.9 (14.8)	38.9 (10.3)	30.5 (10.4)	*87.1*	*74.1*
PT°	2.7	2.5	18.1 (10.6)	21.7 (21.4)	19.3 (9.4)	22.1 (19.5)	*85.0*	*88.4*
PI°	6.3	8.4	53.6 (10.2)	62.4 (9.1)	50.8 (12.6)	56.7 (13.7)	*88.7*	*85.1*
LL°	9.7	9.5	47.5 (21.4)	33.9 (17.1)	46.2 (16.2)	37.1 (17.8)	*80.3*	*74.3*
$\mathbf{SVA}_{cm}$	0.7	-	4.4 (4.8)	-	3.4 (3.2)	-	*84.7*	-

The outputs of the model are bounding boxes (objects) with different confidences, of which the center of each box is the predicted landmark (Figure 1-B). The predictions with the highest confidence in each class are the actual required outputs needed to calculate the spinopelvic parameters (Figure 1-C). The spinopelvic parameters have been calculated using the extracted landmarks. Table 1 indicates the comparison of the parameters obtained from Ground Truth (GT) labels and those from predicted landmarks.

Acknowledgments

This research was funded by Intellijoint Surgical, the Natural Sciences and Engineering Research Council of Canada, and the Ontario Centres of Excellence.

References

Robert Korez, Michael Putzier, and Tomaž Vrtovec. A deep learning tool for fully automated measurements of sagittal spinopelvic balance from x-ray images: performance evaluation. *European Spine Journal*, 29(9):2295–2305, 2020.

William McNally, Kanav Vats, Alexander Wong, and John McPhee. Rethinking keypoint representations: Modeling keypoints and poses as objects for multi-person human pose estimation. *arXiv preprint arXiv:2111.08557*, 2021.

Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 779–788, 2016.

Chi-Hung Weng, Chih-Li Wang, Yu-Jui Huang, Yu-Cheng Yeh, Chen-Ju Fu, Chao-Yuan Yeh, and Tsung-Ting Tsai. Artificial intelligence for automatic measurement of sagittal vertical axis using resunet framework. *Journal of clinical medicine*, 8(11):1826, 2019.

Yu-Cheng Yeh, Chi-Hung Weng, Yu-Jui Huang, Chen-Ju Fu, Tsung-Ting Tsai, and Chao-Yuan Yeh. Deep learning approach for automatic landmark detection and alignment analysis in whole-spine lateral radiographs. *Scientific reports*, 11(1):1–15, 2021.