

# KSSSS 2026

The 43<sup>rd</sup> International Congress of  
Korean Society of Spine Surgery  
May 20 (Wed) - 22 (Fri), 2026  
Lotte Hotel Seoul (Sogong-dong), Seoul, Korea



# Measurement of Lumbar Lordosis Using a Deep Learning Based Artificial Intelligence Model



**Seong-Ho Oh, M.D** <sup>1,2)</sup> , **Soo-Bin Lee, M.D., Ph.D.** <sup>1)</sup> , **Dong-Sik Chae, M.D., Ph.D.** <sup>1)</sup> ,  
**Kyung-Yil Kang, M.S.** <sup>1,2)</sup> , **Min-Kyu Lee** <sup>3)</sup>

Department of Orthopaedic Surgery, Catholic Kwandong University College of Medicine, International St. Mary's Hospital, Incheon, Korea<sup>1)</sup>

College of Medicine, Catholic Kwandong Graduate School, Gangneung, Korea<sup>2)</sup>

Walk101 Co., Ltd., Incheon, Republic of Korea<sup>3)</sup>

# Introduction



- **Lumbar lordosis (LL)** : Essential for posture, weight distribution, and spinal balance.
- Abnormal LL (Hyper/Hypo) ➔ **Chronic low back pain, sagittal imbalance.**
- Manual measurement of LL: **Time-intensive, inter- and intra-observer variability**

# Introduction



- **Artificial intelligence (AI)** development
  - ➔ Advances medical image analysis with accuracy
- Study objective: Developing a **Deep Learning model for automatic LL measurement** on whole spine X-rays
  - ✓ Improves measurement consistency
  - ✓ Reduces clinician workload,
  - ✓ Enables scalable spinal angle measurements

# Methods



- **Initial dataset** : 2,397 standing whole spine lateral radiographs (2019–2021).
- **Final Dataset** : **888 high-quality radiographs** selected
  - ➔ Divided into training (740), validation (98), and test (50) sets

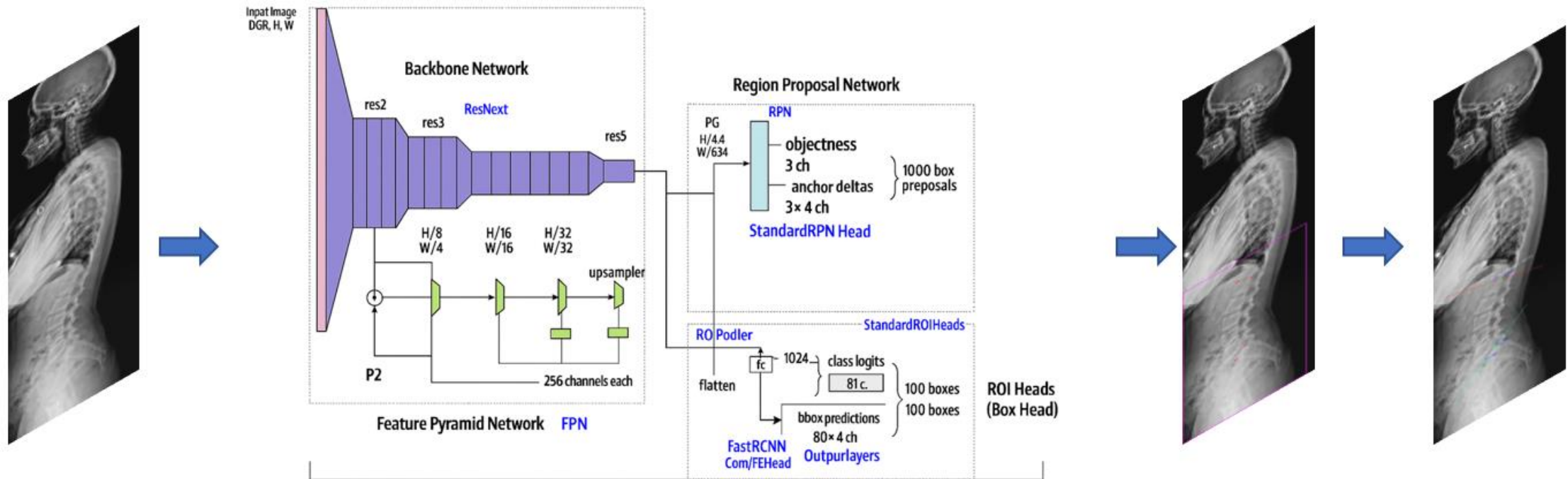
	Training set	Validation set	Test set
X-ray images, n	740	98	50
Age, mean±SD years (range)	48.24±25.36 (3–89)	48.11±25.76 (7–83)	48.7±24.78 (8–86)
Sex ratio, male:female	0.488:0.512	0.459:0.541	0.440:0.560

**Table 1.**  
**Datasets and patient demographics**

SD: standard deviation.

- **Annotation** : **Four anatomical keypoints defined at L1 and S1 vertebral landmarks** (Superior-anterior and superior-posterior)
- **Process** : Labeled by a registered nurse under the supervision of an orthopedic surgeon with 10+ years of experience.

# Methods



**Fig. 1. Overview of the Detectron2 architecture.** It consists of a backbone network (e.g., ResNeXt), a feature pyramid network (FPN) for multi-scale feature extraction, and keypoint detection heads tailored for X-ray image analysis

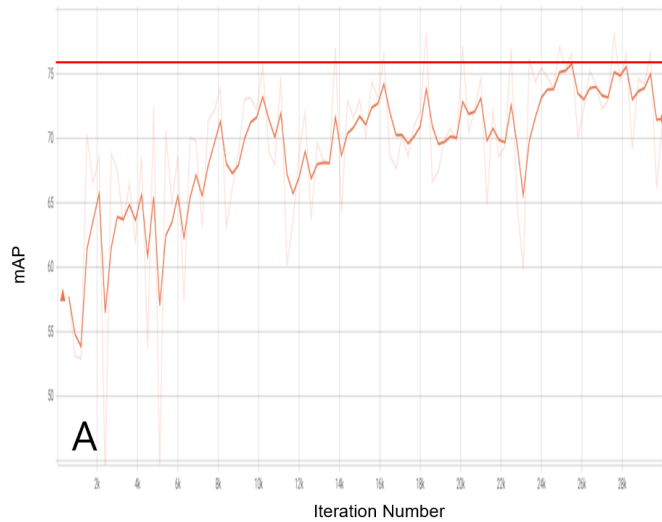
- ❖ Framework: Detectron2 (Facebook AI Research).
- ❖ Architecture: Keypoint R-CNN with Feature Pyramid Network (FPN).
- ❖ Backbone: ResNeXt-101 (32x8d) initialized with ImageNet weights.
- ❖ Training: 30,000 iterations with data augmentation (rotation, brightness, saturation).

# Methods



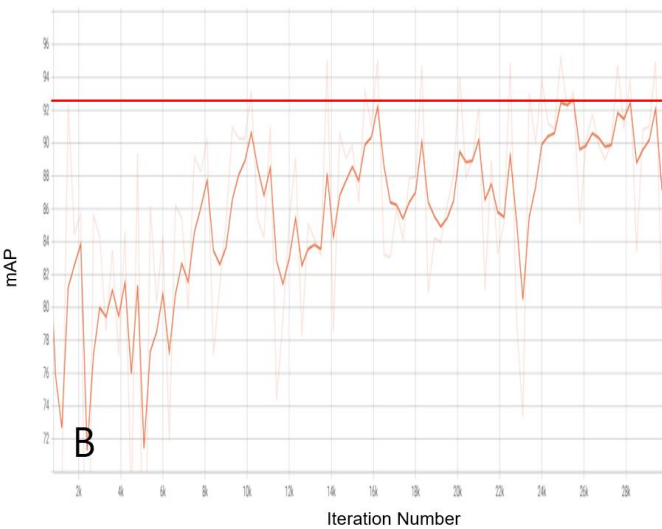
- **Testing** : AI results compared against manual annotations by two expert orthopedic surgeons using **50 X-rays**
- Statistical Metrics
  - Intraclass correlation coefficient (ICC) : Reliability and agreement
  - Pearson's correlation : Strength of correlation
  - Bland-Altman analysis : Systematic bias & limits of agreement.

# Results

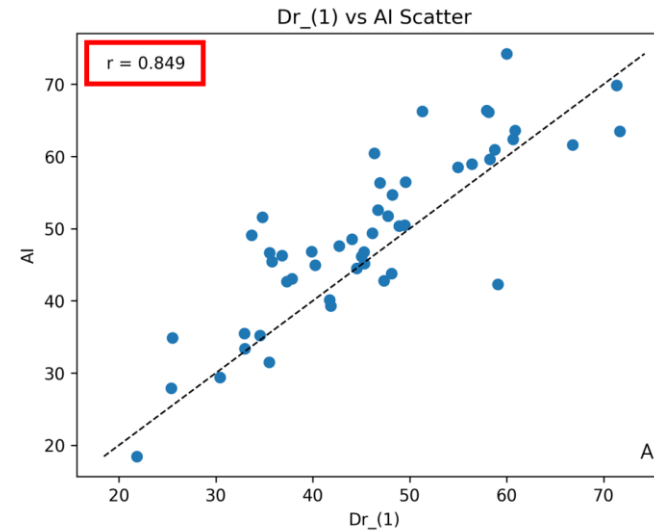


**Fig. 2. Mean Average Precision (mAP) for both bounding boxes and keypoints.**

(A) The mAP for bounding boxes achieved a best score of 75.78.

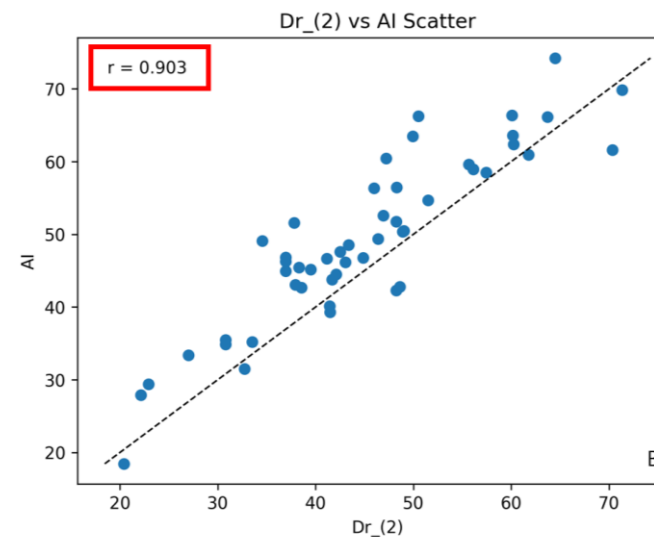


(B) The keypoint mAP, calculated based on Object Keypoint Similarity (OKS), achieved a best score of 92.63.



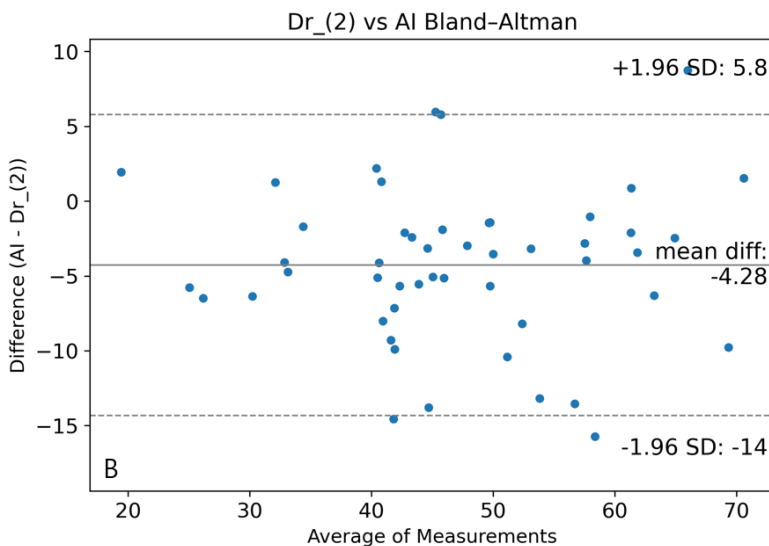
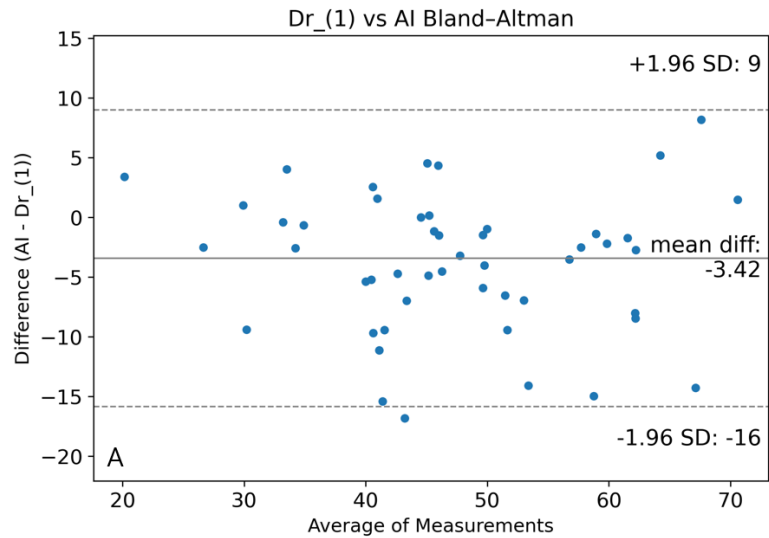
**Fig. 3. Scatterplots comparing AI-inferred and manually measured lumbar lordosis angles.**

(A) AI vs. Orthopedic surgeon (identity line:  $y=x$ ;  $r=0.849$ ,  $p<0.001$ )



(B) AI vs. Orthopedic surgeon 2 (identity line:  $y=x$ ;  $r=0.903$ ,  $p<0.001$ ).

# Results



**Fig. 4. Bland–Altman plots comparing AI-inferred and manually measured lumbar lordosis angles.**

(A) AI vs. Orthopedic surgeon 1 (mean bias:  $-3.42^\circ$ ; 95% limits of agreement:  $-16^\circ$  to  $+9^\circ$ ).

(B) AI vs. Orthopedic surgeon 2 (mean bias:  $-4.28^\circ$ ; 95% limits of agreement:  $-14^\circ$  to  $+5.8^\circ$ ).

→ **Minor systematic underestimation bias was observed, but overall agreement remains clinically acceptable.**

## ICC Results:

❖ AI vs. Surgeon 1: 0.918.

❖ AI vs. Surgeon 2: 0.949.

❖ Combined Agreement: 0.962  
**(Excellent agreement).**

# Conclusion



- AI can reliably and efficiently measure spinal alignment parameters
- Clinical application of AI → reduce interobserver variability, streamline diagnostic processes
- AI-assisted spinal analytics requires ongoing validation to achieve clinical potential