Journal of Orthopaedic Surgery and Research

RESEARCH

Open Access

Development of a model for measuring sagittal plane parameters in 10–18-year old adolescents with idiopathic scoliosis based on RTMpose deep learning technology

Zhijie Kang[1†], Guopeng Shi[1†], Yong Zhu[2], Feng Li[3], Xiaohe Li[1*] and Haiyan Wang[1*]

Abstract

Purpose The study aimed to develop a deep learning model for rapid, automated measurement of full-spine X-rays in adolescents with Adolescent Idiopathic Scoliosis (AIS). A significant challenge in this field is the time-consuming nature of manual measurements and the inter-individual variability in these measurements. To address these challenges, we utilized RTMpose deep learning technology to automate the process.

Methods We conducted a retrospective multicenter diagnostic study using 560 full-spine sagittal plane X-ray images from five hospitals in Inner Mongolia. The model was trained and validated using 500 images, with an additional 60 images for independent external validation. We evaluated the consistency of keypoint annotations among different physicians, the accuracy of model-predicted keypoints, and the accuracy of model measurement results compared to manual measurements.

Results The consistency percentages of keypoint annotations among different physicians and the model were 90–97% within the 4-mm range. The model's prediction accuracies for key points were 91–100% within the 4-mm range compared to the reference standards. The model's predictions for 15 anatomical parameters showed high consistency with experienced physicians, with intraclass correlation coefficients ranging from 0.892 to 0.991. The mean absolute error for SVA was 1.16 mm, and for other parameters, it ranged from 0.22° to 3.32°. A significant challenge we faced was the variability in data formats and specifications across different hospitals, which we addressed through data augmentation techniques. The model took an average of 9.27 s to automatically measure the 15 anatomical parameters per X-ray image.

Conclusion The deep learning model based on RTMpose can effectively enhance clinical efficiency by automatically measuring the sagittal plane parameters of the spine in X-rays of patients with AIS. The model's performance was found to be highly consistent with manual measurements by experienced physicians, offering a valuable tool for clinical diagnostics.

Keywords Adolescent idiopathic scoliosis, RTMpose, Deep learning, Sagittal plane

†Zhijie Kang is first author and Guopeng Shi is a co-first author.

*Correspondence:
Xiaohe Li
798242742@qq.com
Haiyan Wang
18647398868@163.com
Full list of author information is available at the end of the article

Introduction

AIS is a common asymmetrical spinal deformity characterized by rib prominence and a pelvic tilt [1]. It is widely recognized that AIS results from the interaction of multiple factors, including environmental, hormonal, and genetic factors, although its exact etiology remains unclear. The prevalence of AIS among adolescents is 2–3%, with a higher incidence in females than in males [2, 3]. School screening studies have reported AIS prevalence rates ranging from 0.47 to 5.2% [4]. The primary diagnostic criterion for AIS is an X-ray showing a coronal plane Cobb angle of ≥ 10° [5]. Based on the Cobb angle, AIS can be classified as mild (Cobb angle < 20°), moderate (Cobb angle < 40°), and severe (Cobb angle > 40°), with mild and moderate cases being the most common and typically managed nonsurgically, unlike the severe cases that require surgical intervention [2].

The normal spine exhibits physiological curvatures in the sagittal plane, including the cervical, thoracic, lumbar, and sacral curvatures, which are crucial for maintaining an upright posture, buffering stress from movement, and facilitating physiological functions [6, 7]. Scoliotic deformity can lead to vertebral rotation and coronal plane deviation, which, in severe cases, may significantly disrupt sagittal balance [8]. Although much research has focused on the coronal plane, the importance of sagittal balance in the treatment of AIS has been relatively neglected, even though coronal parameters alone cannot fully assess the pathological characteristics of patients [9]. Therefore, quantifying sagittal plane anatomical parameters is of significant importance for preoperative assessment, intraoperative guidance, and reduction of postoperative complications [10, 11]. However, manual measurement of anatomical parameters can be time-consuming and vulnerable to significant interindividual variability [12].

Deep learning technology, which can effectively extract image features, has been widely applied in various medical fields, particularly imaging-assisted diagnosis [13]. This technology utilizes computer processing and artificial intelligence to generate optimized weighting factors to ensure accuracy and reduce repetitive tasks [14]. Deep learning models can fit real statistical models of the spine to X-ray images [15]. The U-Net model can segment landmarks on X-rays, effectively measuring the lumbar lordosis (LL) with an average processing time of 0.14 s per X-ray, and its measurement results are consistent with those obtained by clinicians [16]. The ResUNet model can accurately measure the sagittal vertical axis (SVA), thereby facilitating the assessment of spinal sagittal alignment [17]. However, these methods mostly focus on specific regions of the spine, particularly the lumbosacral area and do not comprehensively assess the sagittal balance status of patients with AIS.

Deep learning technology has revolutionized the field of medical imaging, particularly in the diagnosis and assessment of spinal deformities [15–17]. This study introduces a novel approach that employs the RTMpose algorithm to locate key points and assess the sagittal balance status of the spine based on 15 anatomical parameters distributed across the cervical, thoracic, lumbar, and femoral heads. Among the various deep learning algorithms available, we chose to employ the RTMpose algorithm for several compelling reasons. RTMpose is a real-time multi-person pose estimation network that has demonstrated exceptional performance in accurately locating keypoints and assessing posture across multiple body regions [18]. Its ability to handle complex imaging data and provide precise anatomical landmark identification makes it particularly suited for the intricate task of measuring spinal parameters in AIS patients. Furthermore, RTMpose's efficiency in processing full-spine X-ray images aligns with the clinical need for rapid diagnostic tools. Unlike other deep learning models that may focus on specific regions or require extensive computational resources, RTMpose offers a balanced approach, combining accuracy with practicality, It is designed to work with a standard computational setup, making it a viable solution for clinical environments where resources may be limited [18, 19]. Additionally, RTMpose's architecture allows for the integration of various imaging modalities and datasets, which was essential for our multicenter study. Its flexibility in handling diverse data formats and its robust performance across different centers further supported our choice.

Materials and methods

Source of data

This was a retrospective, multicenter diagnostic study using full-spine sagittal plane X-ray images (lateral view) collected from five hospitals in Inner Mongolia: the Affiliated Hospital of Inner Mongolia Medical University, the Second Affiliated Hospital of Inner Mongolia Medical University, Baotou City Central Hospital, the First Affiliated Hospital of Baotou Medical College, and Bayannur Hospital. From January 1, 2021, to June 30, 2023, 500 full-spine sagittal plane X-ray images were collected from the Affiliated Hospital of Inner Mongolia Medical University, the Second Affiliated Hospital of Inner Mongolia Medical University, Baotou City Central Hospital, the First Affiliated Hospital of Baotou Medical College. From July 1, 2023, to September 30, 2023, 60 full-spine sagittal plane X-ray images were independently collected from the Bayannur Hospital for external validation (Fig. 1). The study was approved by the local Institutional Review Board

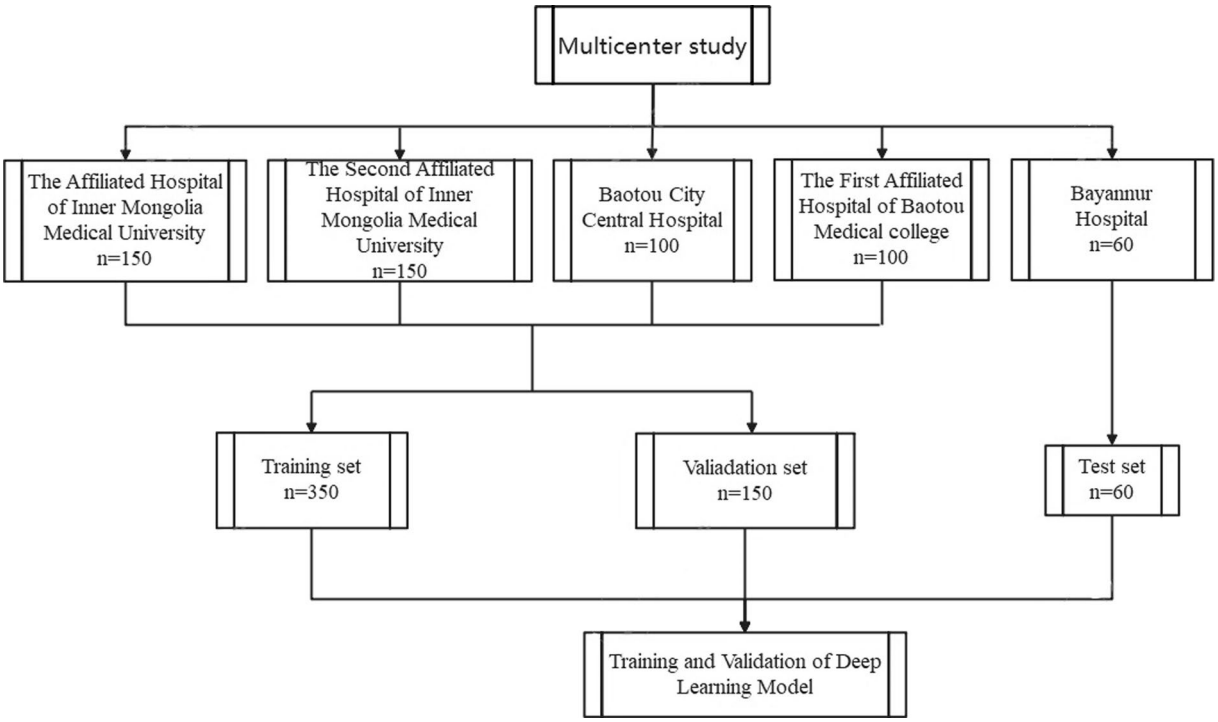

Fig. 1 Schematic diagram of data sources

(IRB) with ethical approval number (YKD2019036), and the requirement for informed consent was waived.

Inclusion and exclusion criteria

Selection criteria of patients with AIS: Patients aged 10–18 years, with a coronal plane Cobb angle of ≥ 10°, without congenital scoliosis, spinal trauma, or other diseases causing spinal deformity, without any other pathological changes, not having received any intervention treatments, and diagnosed with AIS based on full-spine sagittal plane X-ray images.

Image acquisition

Our multicenter approach was designed to enhance the generalizability of our findings by capturing a diverse patient population and accounting for variations in imaging protocols. However, we recognize that differences in imaging techniques across centers could introduce potential biases. To mitigate this, we implemented rigorous standardization protocols for image acquisition at each center. These protocols included uniform patient positioning, standardized exposure settings (120 kV tube voltage and 60 mA tube current), and the use of consistent imaging equipment from reputable manufacturers (Philips DR and Siemens DR systems). Additionally, all images were scaled to a resolution of 1024 × 1024 pixels to ensure consistency in image processing. To further address potential biases, we employed data augmentation techniques, which not only helped in handling variations in data but also simulated the effects of different imaging conditions. This approach allowed our model to learn and adapt to the nuances of various imaging techniques, thereby reducing the impact of technique-related biases on our results. We also conducted a sensitivity analysis to assess the robustness of our model's performance across different centers. The model demonstrated consistent accuracy and reliability in predicting sagittal plane parameters, indicating that any biases introduced by imaging techniques were effectively managed.

Methods for measuring sagittal balance parameters

The sagittal balance parameters were measured as follows:

1. Cervical Cobb angle(CCA): angle between the lower endplates of C2 and C7 (Fig. 2 A-CCA) [11].
2. Thoracic kyphosis (TK): The Cobb angle between the upper endplate of T4 and the lower endplate of T12 (Fig. 2 A-TK) [8].
3. Lumbar lordosis (LL): Cobb angle between the upper endplates of L1 and S1 (Fig. 2 A-LL) [8].
4. Thoracolumbar kyphosis(TLK): Cobb angle between the upper endplate of T10 and the lower endplate of L2 (Fig. 2 C-TLK) [8].

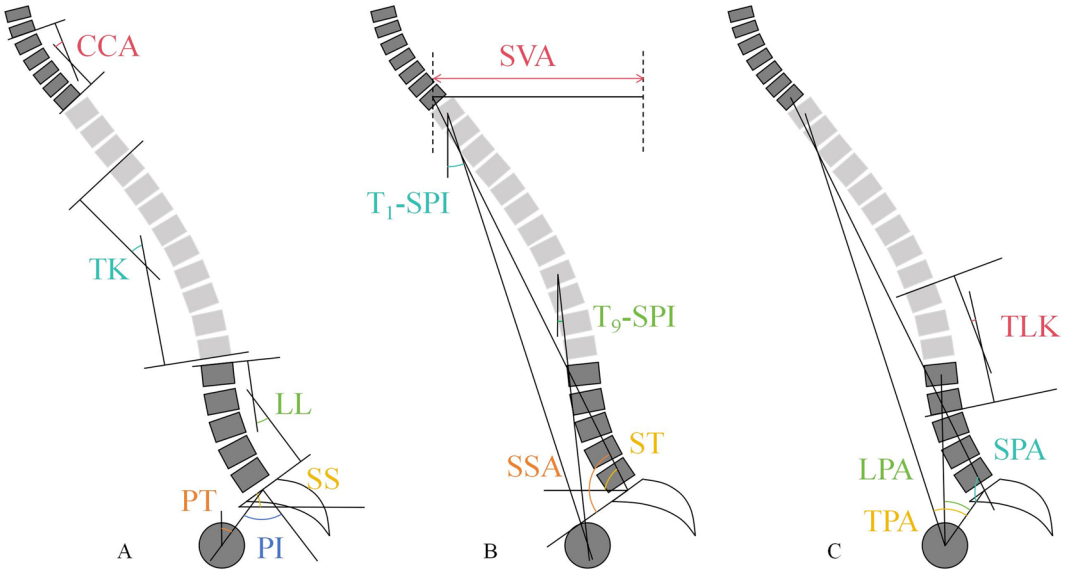

Fig. 2 Schematic diagram of sagittal plane parameter measurements

5. Sagittal vertical axis (SVA): The horizontal distance between the vertical line through the center of C7 and the posterior superior corner of S1, with a positive val-ue if the plumb line is anterior to the superior posterior edge of S1 and a negative value if it is posterior (Fig. 2 B-SVA) [6].
6. Spinal tilt (ST): angle between the line connecting the midpoint of C7 to the midpoint of the upper endplate of S1 and the horizontal line (Fig. 2 B-ST) [20].
7. Spinosacral angle (SSA): angle between the line connecting the center of C7 to the midpoint of the upper endplate of S1 and the upper endplate of S1 (Fig. 2 B-SSA) [20].
8. Spinopelvic angle (SPA): The angle made by the line connecting the midpoint of the upper endplate of S1 with the midpoints of C7 and the femoral head (Fig. 2 C-SPA) [21].
9. T1 spinopelvic inclination (T1-SPI): The angle between the line from the midpoint of T1 to the midpoint of the femoral head and a vertical line through the midpoint of T1, with a positive value if the line inclines forward relative to the vertical line and a negative value if it inclines backward (Fig. 2 B-T_1SPI) [22].
10. T9 spinopelvic inclination (T9-SPI): The angle between the line from the midpoint of T9 to the midpoint of the femoral heads and a vertical line through the midpoint of T9, with a positive value if the line inclines forward relative to the vertical line and a negative value if it inclines backward (Fig. 2 B-T_9SPI) [22].
11. T1 pelvic angle(TPA): The angle formed by the line connecting the midpoint of the femoral head with the midpoint of the upper endplate of S1 and the midpoint of T1 (Fig. 2 C-TPA) [22].
12. Lumbar pelvic angle(LPA): the angle formed by the line connecting the midpoint of the femoral heads to the midpoint of L1 and the midpoint of the upper endplate of S1 (Fig. 2 C-LPA) [22].
13. Sacral slope (SS): The angle between the upper edge of S1 and the horizontal line (Fig. 2 A-SS) [8].
14. Pelvic incidence (PI): The angle between the line from the midpoint of the upper edge of S1 to the center point of the femoral head and a vertical line through the midpoint of the upper edge of S1 (Fig. 2 A-PI) [8].
15. Pelvic tilt (PT): The angle between the line from the midpoint of the upper edge of S1 to the center point of the femoral heads and the vertical line (Fig. 2 A-PT) [8].

Data annotation and key point setting The full-spine sagittal plane X-ray images from the training and validation datasets were annotated by two experienced doctors (including a radiologist and a spine surgeon, both from the Affiliated Hospital of Inner Mongolia Medical University and the Second Affiliated Hospital of Inner Mongolia Medical University, with at least 15 years of experience in diagnosing AIS using X-ray). Annotations

were made using the image annotation software Labelme, marking 38 key points, including the vertices on both sides of the lower endplate of C2, C7, T1, T4, T9, T10, T12, L1, and L2, the two most lateral points on the upper edge of the S1 vertebra, and the center points of both femoral heads. The key points were named in the order of the bone structures mentioned above for ease of subsequent organization and data analysis. The key points were C2LD, C2RD, C7LU, C7RU, C7LD, C7RD, T1LU, T1RU, T1LD, T1RD, T4LU, T4RU, T4LD, T4RD, T9LU, T9RU, T9LD, T9RD, T10LU, T10RU, T10LD, T10RD, T12LU, T12RU, T12LD, T12RD, L1LU, L1RU, L1LD, L1RD, L2LU, L2RU, L2LD, L2RD, S1, S2, U1, and U2. Subsequently, the spinal target frames were outlined to ensure that all the key points were within the outlined target frame (Fig. 3). Determining the midpoint of the vertebral body using the average coordinate method. The test dataset was used to evaluate the final performance of the model, with data annotations independently completed by two doctors from the Affiliated Hospital of Inner Mongolia Medical University and the Second Affiliated Hospital of Inner Mongolia Medical University (one resident with 3 years of experience as a junior doctor and one senior chief doctor with 15 years of experience) and the completed model to ensure the objectivity of the annotations. Before measuring all the anatomical parameters, uniform standard training was conducted to determine the measurement plan. After the first measurement, the image data were shuffled, and 2 weeks later, the two doctors repeated the measurements. The average of the two measurements taken by each person was used as the reference standard [12].

Fig. 3 Schematic diagram of key point annotation

Data processing

Data augmentation

The data used in this study were collected from five collaborating hospitals: the Affiliated Hospital of Inner Mongolia Medical University, the Second Affiliated Hospital of Inner Mongolia Medical University, Baotou Central Hospital, the First Affiliated Hospital of Baotou Medical College, and the Hospital of Bayannur. Throughout the data collection process, owing to differences in equipment, operational methods, environmental conditions, application needs, and technical standards across hospitals, the collected data formats and specifications varied. These differences added complexity to the data integration and model training processes. To address this issue, Python (version 3.8) was used for dataset processing, employing methods, such as color transformation, noise addition, and image blurring, to augment and enrich the dataset. This data augmentation enhanced the diversity of the dataset features, increased the value of the limited dataset, and made it equivalent to larger datasets in terms of information content and application effectiveness.

Data preprocessing

All images were scaled to a resolution of 1024×1024 pixels while maintaining their original resolution for machine learning purposes.

Model construction

This model was developed in collaboration with Shenyang Simo Network Technology Co., Ltd. The construction of the model was based on RTMpose, a deep learning model designed for pose estimation that is officially known as a real-time multi-person pose-estimation network. This model possesses sufficient depth to capture keypoint information within images. We divided the network structure into two parts: the Backbone and the gated attention unit (GAU). The backbone module was pretrained using a heatmap-based method for feature extraction. The GAU divided the horizontal and vertical axes into equally wide bins and discretized the continuous coordinates into integer bin labels. The model was then trained to predict the bin in which a keypoint was located. By utilizing a large number of bins, the quantization error was reduced to the subpixel level [18] (Fig. 4).

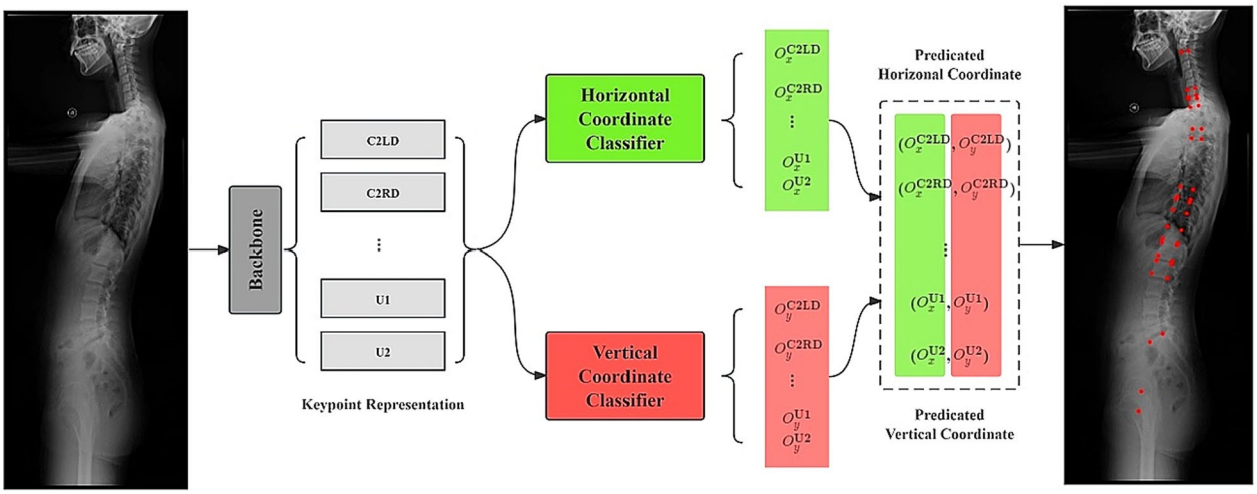

Fig. 4 RTMpose deep learning model construction process

Model performance validation

The dataset used in this study consisted of 560 sets of data divided into training, validation, and test datasets in a 7:2:1 ratio. The training dataset was used to train the model and the validation set was used to determine the parameters within the network model. After each training iteration, the accuracy of the validation set was calculated and compared with the accuracy of the previous training sessions to select the model parameters with the highest accuracy. The test set was used to assess the model's performance. After completing data annotation, the model construction phase began with the establishment of a deep neural network. The training data were input into the model architecture to complete initial model training. Subsequently, the model was calibrated to enhance its performance. After prediction using the model, the 60 test dataset images were resampled to their original pixels for comparison with the annotations made by the radiologists on the original images. Subsequent statistical analysis was conducted to compare the consistency between the two sets of measurements and calculate the model's sensitivity, specificity, and accuracy (Fig. 5).

Statistical analysis

Statistical analysis was performed using Python 3.8 and SPSS 26, with a P-value of < 0.05 considered statistically significant. For quantitative data that followed a normal distribution and showed homogeneity of variance, the mean ± standard deviation was used; otherwise, the median (lower quartile, upper quartile) was used. The

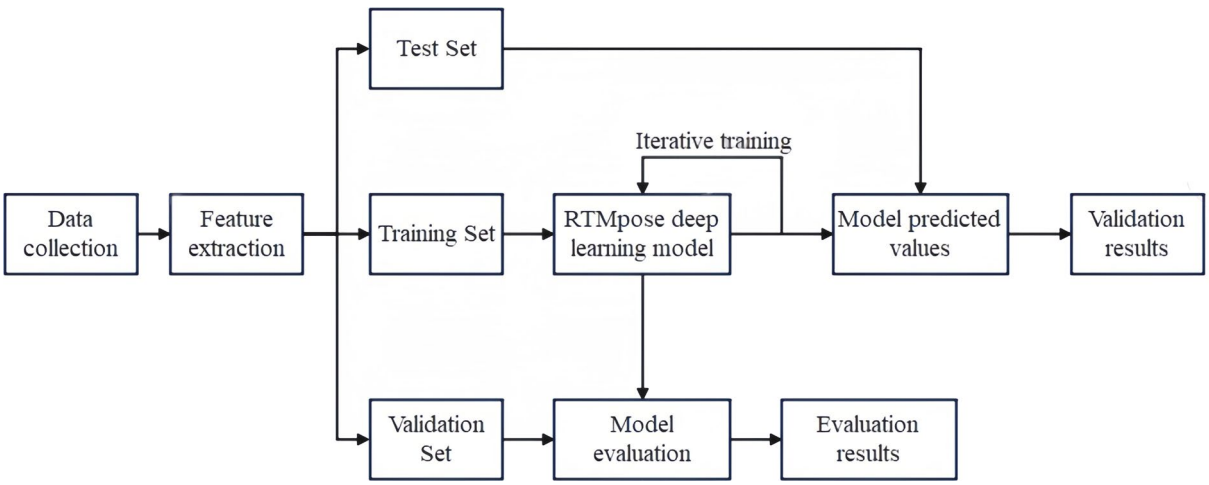

Fig. 5 Deep learning model validation process diagram

Kruskal–Wallis H test was used to compare the time taken for model predictions and measurements by different doctors. The reliability of the model results was evaluated by comparing the percentage range of the model results with the doctor annotations at 1, 2, 3, and 4 mm. The percentage of correct keypoints (PCKs) was used to assess the accuracy of model predictions of the keypoints. The intraclass correlation coefficient (ICC) was used to evaluate the consistency between model results and doctor measurements, with an ICC ≥ 0.75 considered reliable. The mean absolute error (MAE) was used to assess the error between the model and the doctor measurement results.

Results

General information

This study included 560 patients with AIS who underwent full-spine sagittal plane X-ray imaging. Of these, 350, 150, and 60 were allocated to the training, validation, and test datasets, respectively. The male-to-female ratio was 1:5 for the training, validation, and testing datasets. The average age of patients in the training, validation, and test datasets were 13.52 (range 10–18 years), 13.73 (range 10–18 years), and 13.58 years (range 10–18 years) (Table 1).

Consistency of different doctors' annotations

Among the groups, the percentages of keypoint-to-keypoint distances within a 3-mm threshold were 93% (Doctor 1 and Doctor 2), 92% (Doctor 1 and the model), and 83% (Doctor 2 and the model); within a 4-mm threshold, the percentages were 97% (Doctor 1 and Doctor 2), 94% (Doctor 1 and the model), and 90% (Doctor 2 and the model) (Table 2).

Proportion of correct model key point estimations

The model's predictions of key points compared to the reference standard key points within the 1-mm, 2-mm, 3-mm, and 4-mm thresholds were 15–49%, 59–94%, 78–100%, and 91–100%, respectively. The prediction percentage within the 4-mm threshold was above 90% (Table 3) (Fig. 6).

Table 1 Data information [$M(Q_L,Q_U)$]

	Training Set	Validation Set	Test Set
Age (Year)	13 (13,14)	14 (13,14)	14 (13,14)
Sex (Male/Female)	58/292	25/125	11/49
Total	350	150	60

Table 2 Intra-observer and inter-observer image annotation consistency (%)

Threshold	1 mm	2 mm	3 mm	4 mm
Doctor 1 vs. Doctor 2	49	82	93	97
Doctor 1 vs. Model	34	81	92	94
Doctor 2 vs. Model	27	66	83	90

Table 3 Proportion of correct model key point estimations within 1–4-mm thresholds of the reference standard (%)

Threshold (mm)	1	2	3	4
C2LD	33	89	100	100
C2RD	46	84	97	100
C7LU	39	92	97	97
C7RU	37	94	97	98
C7LD	39	89	97	98
C7RD	40	89	95	97
T1LU	39	85	97	98
T1RU	47	94	97	100
T1LD	40	86	95	98
T1RD	37	89	95	97
T4LU	30	78	92	94
T4RU	34	84	94	94
T4LD	40	86	92	92
T4RD	36	79	92	92
T9LU	23	86	94	94
T9RU	39	84	94	94
T9LD	28	82	92	94
T9RD	49	81	91	94
T10LU	39	82	94	94
T10RU	46	88	94	94
T10LD	26	76	88	91
T10RD	39	89	92	94
T12LU	27	85	92	92
T12RU	37	75	92	92
T12LD	27	71	89	92
T12RD	31	82	91	92
L1LU	18	72	91	92
L1RU	34	79	91	94
L1LD	24	75	89	94
L1RD	31	75	92	94
L2LU	31	78	92	94
L2RU	40	84	91	94
L2LD	28	71	86	92
L2RD	34	75	91	94
U1	15	60	84	91
U2	18	59	78	92
S1	21	69	89	91
S2	33	75	88	92

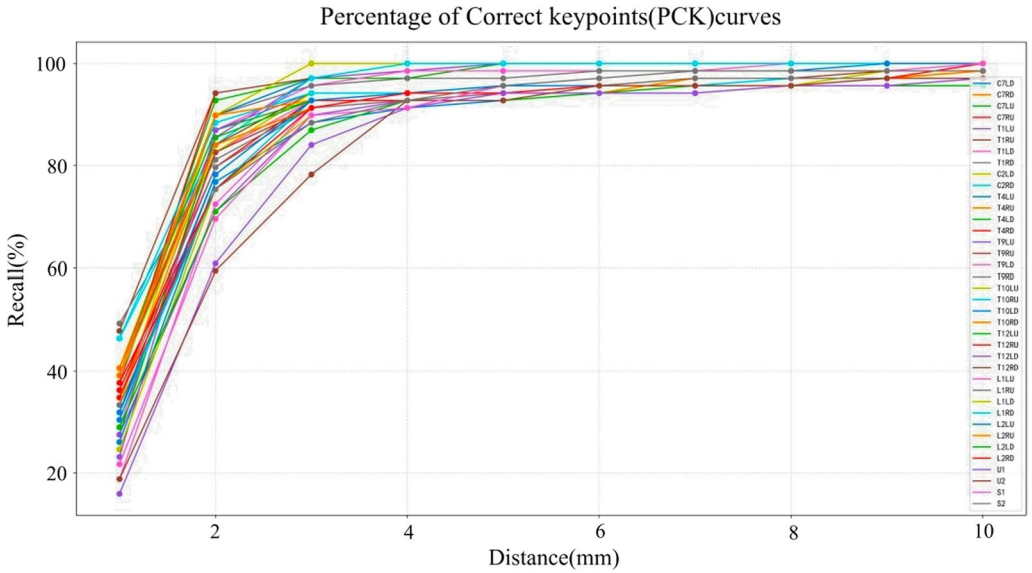

Fig. 6 Key point model prediction performance

Comparison of model predictive outcomes with reference standard results

In this study, we compared the predictive outcomes of our model with the reference standard results. Our model demonstrated excellent performance in predicting anatomical parameters, showing a high degree of consistency with the reference standard, as indicated by the ICC values ranging from 0.892 to 0.991 across different parameters. (Table 4). In terms of MAE, the model's measurements for 15 sagittal parameters yielded MAE values ranging from 0.22 to 3.32 when compared to the reference standard. All of these values fall within an acceptable error range, further validating the effectiveness of the model. The Bland–Altman plot offers a visual method for intuitively assessing the consistency between the manual measurements made by physicians and those obtained from the RTMpose model, instead of relying solely on correlation. This approach facilitates the evaluation of whether the RTMpose measurement method could potentially replace traditional manual measurement methods used by physicians. The results of the visualization indicate that the measurements generated by

Table 4 ICC results for the reference standard, model, senior doctor, and junior doctor

Measurement parameters	Reference standard vs Model	*P*	Reference standard vs Senior doctor	*P*	Reference standard vs Junior doctor	*P*
TK(°)	0.892	0.000	0.800	0.000	0.769	0.000
LL(°)	0.953	0.000	0.758	0.000	0.894	0.000
PT(°)	0.979	0.000	0.898	0.000	0.804	0.000
SS(°)	0.933	0.000	0.916	0.000	0.884	0.000
PI(°)	0.912	0.000	0.887	0.000	0.877	0.000
CCA(°)	0.991	0.000	0.933	0.000	0.920	0.000
TLK(°)	0.892	0.000	0.762	0.000	0.751	0.000
SVA(mm)	0.990	0.000	0.906	0.000	0.843	0.000
SSA(°)	0.942	0.000	0.766	0.000	0.734	0.000
ST(°)	0.941	0.000	0.895	0.000	0.761	0.000
SPA(°)	0.973	0.000	0.869	0.000	0.720	0.000
T1-SPI(°)	0.977	0.000	0.973	0.000	0.962	0.000
T9-SPI(°)	0.984	0.000	0.983	0.000	0.977	0.000
TPA(°)	0.969	0.000	0.850	0.000	0.585	0.000
LPA(°)	0.957	0.000	0.825	0.000	0.764	0.000

the RTMpose model are closer to the manual measurements made by physicians (Table 5) (Fig. 7). Overall, the RTMpose model showed high precision and low error in measuring the sagittal parameters of the spine, with results closer to those of senior physicians, outperforming the measurements of junior physicians.

Time consumption for parameter measurement using deep learning model and doctors of different seniority levels

The time taken to measure 15 sagittal plane anatomical parameters was compared among the model, senior doctors, and junior doctors. The results showed that the time consumption of the model was significantly lower than that of both senior and junior doctors ($P=0.000$), with the model taking an average of 9.27 s per X-ray measurement (Table 6).

Discussion

AIS is the most common spinal curvature disorder, predominantly affecting adolescents aged 10–18 years [23]. Adolescence is the peak period of growth and development of the human body. During this stage, rapid body growth can occur. If spinal deformities are not detected or treated in a timely manner, the condition may progressively worsen, leading to severe physical deformities and potential health issues [24]. Compared to healthy adolescents, patients with AIS exhibit a significant imbalance in the sagittal plane of the spine [25]. The diagnosis of AIS is a complex process that requires comprehensive consideration and assessment from multiple perspectives. The evaluation of spinal sagittal balance is an important criterion for diagnosing AIS [5]. The assessment of spinal sagittal balance is crucial in determining the treatment approach for AIS. In treating AIS, doctors adjust and choose treatment methods based on the specific conditions of the patient's spinal sagittal balance, as long-term brace treatment can affect spinal sagittal balance[10]. If AIS is not detected and treated in its early stages, it often progresses as the patient grows, indicating that the angle of the spinal curvature may increase over time, leading to more severe spinal deformities. In adulthood, patients with untreated or improperly treated AIS may present with significant clinical symptoms[26]. Assessing sagittal balance of the spine in patients with AIS is a highly complex process that involves multiple sagittal spinal parameters collectively reflecting the overall morphology and balance of the spine in the sagittal plane. Precise measurement of these sagittal spinal parameters requires doctors to have a high level of professional knowledge and experience and to rely on high-quality radiological examination results, undoubtedly increasing the workload of clinicians in the diagnostic process. The RTMpose model, a deep-learning-based AI technology, relies on deep learning algorithms to analyze images or videos of individuals and identify and locate key points on the human body, such as the head, arms, and legs, thereby constructing the posture of the human body. It is primarily used for real-time detection and assessment of two-dimensional human posture, with applications in motion analysis, augmented reality, and human–computer interaction [18, 19]. This model feature is beneficial for assessing the overall sagittal balance of the spine on X-rays, offering significant value for a deeper understanding of spinal structure and function in patients with AIS. Therefore, this study built a deep learning model based on the RTMpose to automatically identify the spinal morphology and accurately measure 15 sagittal spinal anatomical parameters, thereby assisting clinicians in assessing the overall sagittal balance of the spine. The use of bounding boxes for the localization of the cervical, thoracic, lumbar spine, pelvis, and femoral heads, including the necessary key points within them, offers the advantage of allowing the model to focus on the spinal morphology, thereby eliminating the influence of unnecessary detection areas and improving the accuracy of key points with geometric shapes.

Artificial intelligence technology, with its high accuracy and short diagnostic time, is widely applied in the medical field. Deep learning technology has great potential for segmenting or reconstructing spinal morphology; however, factors, such as data diversity, image quality, and specialized equipment, limit its development [15, 16]. Therefore, ensuring accuracy is the primary prerequisite for the widespread application of deep

Table 5 MAE results for the reference standard, model, senior doctor, and junior doctor

Measurement parameters	Model	Senior doctor	Junior doctor
TK(°)	3.32	5.59	5.61
LL(°)	2.88	6.56	7.15
PT(°)	0.73	2.04	4.69
SS(°)	2.03	2.68	5.35
PI(°)	2.26	3.70	4.61
CCA(°)	1.19	3.34	4.24
TLK(°)	2.48	3.36	3.75
SVA(mm)	1.16	6.09	9.20
SSA(°)	2.04	4.08	4.98
ST(°)	0.27	0.74	1.50
SPA(°)	0.85	2.60	3.97
T1-SPI(°)	0.22	0.27	0.42
T9-SPI(°)	0.27	0.31	0.46
TPA(°)	0.67	2.04	5.30
LPA(°)	0.61	1.73	2.40

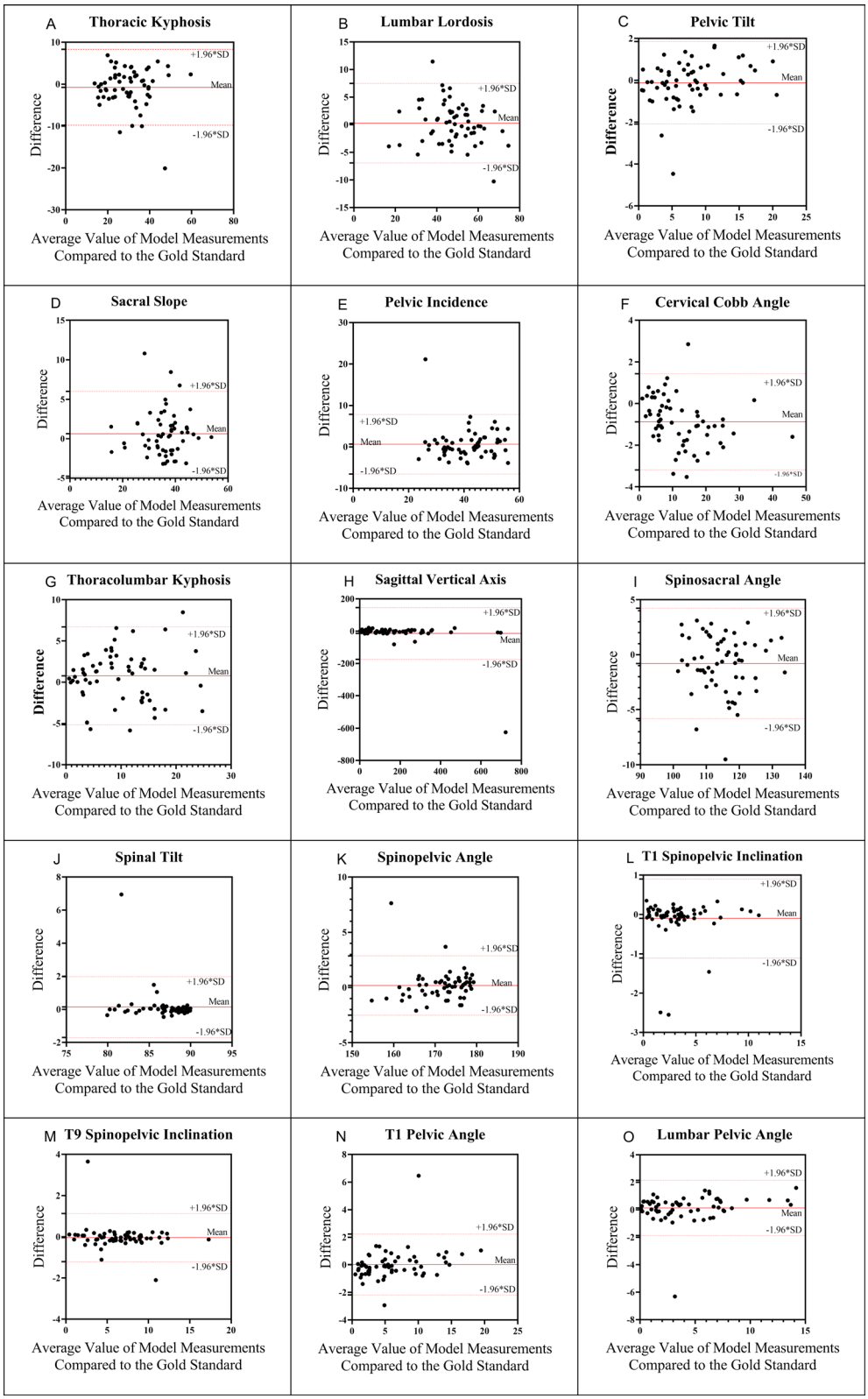

Fig. 7 Bland–Altman plot of model predicted values vs. reference standard anatomical parameters

Table 6 Comparison of time consumption for parameter measurement between deep learning model and doctors of different seniority levels [M(Q_L,Q_U),(S)]

	Model runtime	Time consumption of a senior doctor	Time consumption of a junior doctor	*H*	*P*
Measurement time	9.16 (8.44–10.20)	494.69 (490.04–498.94)	604.60 (583.84–615.72)	159.117	0.000

learning technology in clinical diagnostics and is the key to ensuring patient safety and improving diagnostic efficiency. By contrast, techniques using keypoints for feature extraction can identify and analyze keypoints in images to extract representative information that contains the critical structures and features of the image. This approach is highly robust to minor image variations and can effectively eliminate the impact of factors, such as image quality and differences in imaging equipment, thereby significantly improving the accuracy of the model in medical diagnostics [14]. Chen et al. [27] indicated that when the average distance error of keypoint recognition by deep learning models is 2.977 and the standard deviation is 2.386, the diagnostic results can be used for clinical and biomechanical analyses. In our study, the consistency of keypoint distances within 3 mm was 83–93%, and within 4 mm, it was 90–97%, indicating good consistency between doctors and between doctors and the model, thus demonstrating the reliability of the annotation results. The PCK results showed that the accuracy of the model results within a 3-mm distance range was 78–100%, and within a 4-mm range was 91–100%, indicating that the RTMpose can be applied to automatically measure the sagittal parameters of patients with AIS, assisting clinicians in assessing spinal sagittal balance. Although the model results were good, there were still keypoints with low accuracy; the accuracy of the right femoral head keypoint U2 within the 3-mm range was below 80%. The main reasons for this result could be as follows: (1) the anatomical landmarks of the femoral head center are not obvious, posing a challenge for model training; (2) overlapping phenomena that cause the femoral head to coincide with other bone tissues (such as the hip bone), leading to a decrease in keypoint accuracy; and (3) a small training dataset, resulting in the model's low recognition of this keypoint.

With the continuous development of deep learning technology in the medical field, the development of models that can automatically measure spinal anatomical parameters has become a research hotspot. This technology can analyze medical imaging data and automatically identify and quantify key anatomical features of the spine, thereby providing accurate data support for diagnosis and treatment, improving the precision and efficiency of diagnostics, and promoting the optimization of medical processes and the formulation of personalized treatment plans [28]. When measuring sagittal spinal anatomical parameters, there is a significant deviation in the measurement results among doctors with different professional backgrounds and experience levels, affecting the consistency and accuracy of the diagnostic results [29]. Therefore, improving the efficacy of deep-learning models and reducing resulting errors are key issues to address. Distributed deep learning methods can reduce the deviation between model predictions and doctor measurements [30, 31]. Our study builds a deep learning model based on the RTMpose, first annotating 38 keypoints involving 15 anatomical parameters to assess the overall sagittal balance of the spine and then incorporating bounding boxes to include the keypoints, which can locate the cervical, thoracic, lumbar spine, pelvis, and femoral heads according to spinal morphology, thereby improving model efficacy. Nguyen et al. [30] used a deep learning model built with convolutional neural networks to predict the segmental motion angle between adjacent vertebrae T12-S1, with an average deviation of 1.76° between the model results and physician measurements. In our study, the average error in the angle parameters was 1.42°. Chae et al. [31] constructed a distributed convolutional neural network model and found that the deviations in LL, PT, PI, and SS were 3.17°, 2.64°, 1.45°, and 2.51°, respectively. In our model, the average absolute errors in LL, PT, PI, and SS were 2.88°, 0.73°, 2.26°, and 2.03°, respectively, indicating that the deep learning model based on the RTMpose reduced the assessment error of sagittal parameters, showing its potential for improving the accuracy of measuring spinal sagittal anatomical parameters and providing more accurate data support for clinical diagnosis and treatment. Nguyen et al. [14] built a deep learning model predicting the average absolute errors of anatomical parameters PI, PT, SS, L1I, T1I, C2I, LL, TK, C2-7L, L1S, T1S, C2S sequentially as 2.205°, 1.156°, 3.171°, 2.252°, 5.842°, 3.355°, 3.895°, 5.737°, 6.318°, 2.128°, 6.241°, 3.708°, with correlations > 0.8 for all parameters except C2-7L and T1S. In our study, the model's predicted values compared to the reference standard for the anatomical parameters TK, LL, PT, SS, PI, CCA, TLK, SSA, ST, SPA, T1-SPI, T9-SPI, TPA, and LPA had average absolute errors of 3.32°, 2.28°, 0.73°, 2.03°, 2.26°, 1.19°, 2.48°, 2.04°, 0.27°, 0.85°, 0.22°, 0.27°, 0.67°, and 0.61°, respectively, with MAE ranging from 0.22° to 3.23° and ICC between 0.892°and

0.991°. Compared to the aforementioned studies, our study showed better model efficacy and accurately measured all spinal sagittal parameters. Additionally, comparing the model with the reference standard, high-seniority doctors, and low-seniority doctors, the model's predicted values were closer to the reference standard, showing that the model results were superior to those of high-seniority doctors, who, in turn, were superior to those of low-seniority doctors, aligning with the findings of Zhang [32] and Horng [33].

Our study demonstrates that the RTMpose model can accurately measure sagittal plane parameters in X-rays of patients with AIS. The high ICC and low mean absolute errors MAE indicate the model's potential for clinical use. However, as with any study, there are limitations that deserve further discussion. One significant limitation is the inter-hospital variability in imaging techniques and equipment. Despite our efforts to standardize image acquisition protocols and employ data augmentation to account for these variations, there may still be underlying biases that could affect the model's generalizability. The differences in imaging parameters, patient positioning, and equipment quality across centers could introduce inconsistencies in the data. Future research should aim to control for these variables more rigorously, possibly by conducting a prospective study where imaging protocols can be strictly standardized across all participating centers. Additionally, our study lacks prospective validation, which is a critical step in confirming the clinical applicability of our model. The retrospective nature of our study means that our findings are subject to the biases inherent in historical data. While our model performed well in the test dataset, its performance in a real-time clinical setting, where it would be applied to new, unseen data, remains to be seen. Therefore, we strongly recommend that future work includes a prospective trial to validate the model's accuracy and reliability in a live clinical environment. We acknowledge these limitations and believe that addressing them will be crucial for the further development and refinement of our model. By doing so, we can enhance its clinical utility and contribute to the advancement of AI-assisted diagnostics in orthopedics.

Conclusion

The deep learning model based on the RTM pose can automatically measure the sagittal parameters of the spine on X-rays in patients with AIS. This model can quickly and accurately identify the corresponding keypoints on full-spine sagittal X-rays. The anatomical parameter results measured by the model are highly consistent with those of senior doctors, effectively improving the work efficiency of clinical doctors.

Acknowledgements
We thank Bayannur City Hospital, Inner Mongolia Medical University Affiliated Hospital, Inner Mongolia Medical University Second Affiliated Hospital, Baotou Central Hospital, and the First Affiliated Hospital of Baotou Medical College for their assistance with this study. Special thanks to Shenyang Shengwei Medical Technology Co., Ltd (www.simoai.cn). for their assistance in the development of the deep learning model. We thank Zhang Yunfeng, Jin Feng, Zhang Kai, Xu Yangyang, Da Yifeng, Fang Yuan, Wu Chao, Wang Di and other teachers for their hard work.

Authors' contribution
Zhijie Kang wrote the manuscript. Guopeng Shi is responsible for manuscript and data collation. Haiyan Wang and Xiaohe Li provided directions for writing and submitting the manuscript. Yong Zhu was in charge of study design and data collection. Feng Li was responsible for data analysis and interpretation. All authors read and approved the final manuscript.

Funding
This research was funded by the Inner Mongolia Autonomous Region Health and Health Technology Project(202201188), Inner Mongolia Natural Science Foundation (2021MS08086), Inner Mongolia Medical University "Achievement Transformation" Project (YKD2020CGZH009), 2021 Inner Mongolia Medical University "Zhiyuan Talent" Project (ZY0120017), Inner Mongolia Department of Education Higher Education Innovation Team Development Plan (NMGIRT2227), Inner Mongolia Autonomous Region Key Research, Development, and Achievement Transformation Project (2023YFHH0003), Innovation Team Development Plan for Higher Education Institutions in Inner Mongolia Autonomous Region(NMGIRT2419), Joint project with Inner Mongolia Medical University (YKD2022LH039), Science and Technology Program of the Joint Fund of Scientific Research for the Public Hospitals of Inner Mongolia Academy of Medical Sciences (2023GLLH0225).

Availability of data and materials
Raw data are not publicly available to preserve individuals' privacy under the European General Data Protection Regulation.

Declarations

Ethics approval
The study protocol was approved (YKD2019GG115) by the Ethics Committee of Inner Mongolia Medical University. Written informed consent was not required because of the retrospective nature of the study.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Human Anatomy, Graduate School, Inner Mongolia Medical University, Hohhot 010010, Inner Mongolia, China. [2]Tumor Hospital, Affiliated to Inner Mongolia Medical University, Inner Mongolia Medical University, Hohhot 010000, Inner Mongolia, China. [3]Department of Spinal Surgery, The Second Affiliated Hospital of Inner Mongolia Medical University, Hohhot 010000, Inner Mongolia, China.

Received: 27 August 2024 Accepted: 3 December 2024
Published online: 11 January 2025

References

1. Feustel A, Konradi J, Wolf C, Huthwelker J, Westphal R, Chow D, Hülstrunk C, Drees P, Betz UJB. Influence of lateral sitting wedges on the rasterstereographically measured scoliosis angle in patients aged 10–18 years with adolescent idiopathic scoliosis. Bioengineering. 2023;10(9):1086.
2. Negrini S, Donzelli S, Aulisa AG, Czaprowski D, Schreiber S, de Mauroy JC, Diers H, Grivas TB, Knott P, Kotwicki TJS, et al. 2016 SOSORT guidelines:

orthopaedic and rehabilitation treatment of idiopathic scoliosis during growth. Scoliosis Spinal Disord. 2018;13:1–48.
3. Cortés-Pérez I, Salamanca-Montilla L, Gámiz-Bermúdez F, Obrero-Gaitán E, Ibáñez-Vera AJ, Lomas-Vega RJC. Vestibular morphological alterations in adolescent idiopathic scoliosis: a systematic review of observational studies. Children. 2022;10(1):35.
4. Konieczny MR, Senyurt H, Krauspe RJ. Epidemiology of adolescent idiopathic scoliosis. J Child Orthop. 2013;7(1):3–9.
5. Mitsiaki I, Thirios A, Panagouli E, Bacopoulou F, Pasparakis D, Psaltopoulou T, Sergentanis TN, Tsitsika AJC. Adolescent idiopathic scoliosis and mental health disorders: a narrative review of the literature. Children. 2022;9(5):597.
6. Young E, Regan C, Currier BL, Yaszemski MJ, Larson ANJ. At mean 30-year follow-up, cervical spine disease is common and associated with thoracic hypokyphosis after pediatric treatment of adolescent idiopathic scoliosis. J Clin Med. 2022;11(20):6064.
7. Gardner A, Berryman F, Pynsent PJJA. Statistical modelling of how the sagittal alignment of the cervical spine is affected by adolescent idiopathic scoliosis and how scoliosis surgery changes that. J Anat. 2022;241(2):437–46.
8. Zhang Z, Wang L, Li JC, Liu LM, Song YM, Yang X. Characteristics of sagittal alignment in patients with severe and rigid scoliosis. Orthop Surg. 2023;15(6):1607–16.
9. Ghorbani F, Ranjbar H, Kamyab M, Babaee T, Kamali M, Razavi H, Sharifi P, Janani L. Effect of brace treatment on craniovertebral to lumbopelvic sagittal parameters in adolescents with idiopathic scoliosis: a systematic review. Asian Spine J. 2023;17(2):401.
10. Zhang Z, Ma X, Yin J, Shu L, Gao R, Ma J, Zhou XJO, Surgery T. Alterations of sagittal alignment and thoracic cage parameters after long-term bracing in adolescents with idiopathic scoliosis. Orthop Traumatol Surg Res. 2020;106(7):1257–62.
11. Le Huec J-C, Thompson W, Mohsinaly Y, Barrey C, Faundez A. Sagittal balance of the spine. Eur Spine J. 2019;28:1889–905.
12. Wu C, Meng G, Lian J, Xu J, Gao M, Huang C, Zhang S, Zhang Y, Yu Y, Wang H. A multi-stage ensemble network system to diagnose adolescent idiopathic scoliosis. Eur Radiol. 2022;32(9):5880–9.
13. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, Van Der Laak JA, Van Ginneken B, Sánchez CI. A survey on deep learning in medical image analysis. Med Imaging Anal. 2017;42:60–88.
14. Nguyen TP, Jung JW, Yoo YJ, Choi SH, Yoon J. Intelligent evaluation of global spinal alignment by a decentralized convolutional neural network. J Digit Imaging. 2022;35(2):213–25.
15. Aubert B, Vazquez C, Cresson T, Parent S, de Guise JA. Toward automated 3D spine reconstruction from biplanar radiographs using CNN for statistical spine model fitting. IEEE Trans Med Imaging. 2019;38(12):2796–806.
16. Cho BH, Kaji D, Cheung ZB, Ye IB, Tang R, Ahn A, Carrillo O, Schwartz JT, Valliani AA, Oermann E. Automated measurement of lumbar lordosis on radiographs using machine learning and computer vision. Glob Spine J. 2020;10(5):611–8.
17. Weng C-H, Wang C-L, Huang Y-J, Yeh Y-C, Fu C-J, Yeh C-Y, Tsai T-T. Artificial intelligence for automatic measurement of sagittal vertical axis using ResUNet framework. J Clin Med. 2019;8(11):1826.
18. Jiang T, Lu P, Zhang L, Ma N, Han R, Lyu C, Li Y, Chen K. Rtmpose: real-time multi-person pose estimation based on mmpose. 2023.
19. Lu P, Jiang T, Li Y, Li X, Chen K, Yang W. RTMO: towards high-performance one-stage real-time multi-person pose estimation. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2024; pp. 1491–1500.
20. Wang W, Wang Z, Liu Z, Zhu Z, Zhu F, Sun X, Lam TP, Cheng JC, Qiu Y. Are there gender differences in sagittal spinal pelvic inclination before and after the adolescent pubertal growth spurt? Eur Spine J. 2015;24:1168–74.
21. Zhang Y-P, Qian B-P, Qiu Y, Qu Z, Mao S-H, Jiang J, Zhu Z. Sagittal vertical axias, spinosacral angle, spinopelvic angle, and T1 pelvic angle: which parameters may effectively predict the quality of life in ankylosing spondylitis patients with thoracolumbar kyphosis? Clin Spine Surg. 2017;30(7):E871–6.
22. Protopsaltis TS, Lafage R, Smith JS, Passias PG, Shaffrey CI, Kim HJ, Mundis GM, Ames CP, Burton DC, Bess SJS. The lumbar pelvic angle, the lumbar component of the T1 pelvic angle, correlates with HRQOL, PI-LL mismatch, and it predicts global alignment. Spine. 2018;43(10):681–7.
23. Karam M, Ghanem I, Vergari C, Khalil N, Saadé M, Chaaya C, Rteil A, Ayoub E, Saad E, Kharrat K. Global malalignment in adolescent idiopathic scoliosis: the axial deformity is the main driver. Eur Spine J. 2022;31(9):2326–38.
24. Pesenti S, Prost S, Pomero V, Authier G, Severyns M, Roscigni L, Boulay C, Blondel B, Jouve J-L. Early dynamic changes within the spine following posterior fusion using hybrid instrumentation in adolescents with idiopathic scoliosis: a gait analysis study. Arch Orthop Trauma Surg. 2022;142(12):3613–21.
25. Wang L, Liu X. Cervical sagittal alignment in adolescent idiopathic scoliosis patients (Lenke type 1–6). J Orthop Sci. 2017;22(2):254–9.
26. Aebi M. The adult scoliosis. Eur Spine J. 2005;14:925–48.
27. Chen H-C, Lin C-J, Wu C-H, Wang C-K, Sun Y-N. Automatic Insall-Salvati ratio measurement on lateral knee x-ray images using model-guided landmark localization. Phys Med Biol. 2010;55(22):6785.
28. Galbusera F, Niemeyer F, Wilke H-J, Bassani T, Casaroli G, Anania C, Costa F, Brayda-Bruno M, Sconfienza L. Fully automated radiological analysis of spinal disorders and deformities: a deep learning approach. Eur Spine J. 2019;28:951–60.
29. Kyrölä KK, Salme J, Tuija J, Tero I, Eero K, Arja H. Intra-and interrater reliability of sagittal spinopelvic parameters on full-spine radiographs in adults with symptomatic spinal disorders. Neurospine. 2018;15(2):175.
30. Nguyen TP, Chae D-S, Park S-J, Kang K-Y, Yoon J. Deep learning system for Meyerding classification and segmental motion measurement in diagnosis of lumbar spondylolisthesis. Biomed Signal Process Control. 2021;65:102371.
31. Chae DS, Nguyen TP, Park S-J, Kang K-Y, Won C, Yoon J. Decentralized convolutional neural network for evaluating spinal deformity with spinopelvic parameters. Comput Methods Programs Biomed. 2020;197:105699.
32. 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(1):9083916.
33. Liu J, Yuan C, Sun X, Sun L, Dong H, Peng Y. The measurement of Cobb angle based on spine X-ray images using multi-scale convolutional neural network. Phys Eng Sci Med. 2021;44:809–21.

Publisher's Note