

Copyright Undertaking

This thesis is protected by copyright, with all rights reserved.

By reading and using the thesis, the reader understands and agrees to the following terms:

1. The reader will abide by the rules and legal ordinances governing copyright regarding the use of the thesis.
2. The reader will use the thesis for the purpose of research or private study only and not for distribution or further reproduction or any other purpose.
3. The reader agrees to indemnify and hold the University harmless from and against any loss, damage, cost, liability or expenses arising from copyright infringement or unauthorized usage.

IMPORTANT

If you have reasons to believe that any materials in this thesis are deemed not suitable to be distributed in this form, or a copyright owner having difficulty with the material being included in our database, please contact lbsys@polyu.edu.hk providing details. The Library will look into your claim and consider taking remedial action upon receipt of the written requests.

**DEVELOPMENT OF 3D SPATIAL
MEASUREMENT METHOD BASED ON
COMPUTED TOMOGRAPHY IMAGES OF
ADOLESCENT IDIOPATHIC SCOLIOSIS**

LI LENING

PhD

The Hong Kong Polytechnic University

2025

The Hong Kong Polytechnic University

Department of Biomedical Engineering

**Development of 3D Spatial Measurement Method
based on Computed Tomography Images of
Adolescent Idiopathic Scoliosis**

Li Lening

A thesis submitted in partial fulfilment of the
requirements of the Degree of Doctor of Philosophy

June 2024

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

_____(Signed)

Li Lening (Name of student)

ABSTRACT

This thesis presents a deep learning framework specifically designed to automatically measure the 3D spatial angles from computed tomography (CT) images for pre-surgical patients with adolescent idiopathic scoliosis (AIS). AIS is a common spinal disorder among adolescents, characterized as a complex three-dimensional deformity that includes lateral curvature and vertebral rotation. Severe cases require surgical treatment.

Traditional scoliosis assessment is based on 2D radiographic Cobb angle measurements. However, due to the three-dimensional complexity of the spinal deformity, 2D assessment may not capture the true spinal deformity, leading to inaccurate surgical planning and prognosis. In addition, segmentation of vertebrae can be a challenging task due to the morphological variations of the deformed vertebrae and the proximity of adjacent anatomical structures, which complicates identification and characterization.

The 3D spatial angle gives a more comprehensive view of spinal alignment by considering the curvature in three dimensions rather than just two. This method is useful for surgeons to make more accurate surgical planning and expected outcomes. This study firstly utilized a dataset of 116 scoliosis patients to perform spine segmentation using U-net and a new developed neural network nnformer++. In addition, a new spine curve fitting network called NURBS-Net was developed using non-uniform rational B-spline curves (NURBS). The 3D spatial angle was then calculated by recognizing the maximum angular deviation between vertebrae along the curve.

U-net and nnformer++ both have better performance in severe scoliosis spine segmentation compared to recent studies. The application of NURBS curves in spine curve fitting significantly outperforms traditional methods by providing finer control with fewer parameters, thereby minimizing the risk of overfitting and improving the reliability of the measurements. The 3D spatial angle predicted by the deep learning model correlated strongly with the traditional 2D Cobb annotated by the surgeon, with a Pearson correlation coefficient as high as 0.983.

In conclusion, this method not only validates the feasibility of accurate, automated 3D spatial angle measurements preoperatively in scoliosis patients, but also emphasizes its potential for medical imaging and surgical planning. By providing a detailed 3D view of the spinal deformity, this method is expected to significantly improve surgical accuracy and outcomes for patients with AIS.

PUBLICATIONS ARISING FROM THE THESIS

Journal Papers

Li, L., Zhang, T., Lin, F., Wong MS. Automated 3D Cobb Angle Measurement Using U-Net in CT Images of Preoperative Scoliosis Patients. J Digit Imaging. Inform. med. (2024). <https://doi.org/10.1007/s10278-024-01211-w>.

Li, L., Wong, MS. The application of machine learning methods for predicting the progression of adolescent idiopathic scoliosis: a systematic review. BioMed Eng OnLine 23, 80 (2024). <https://doi.org/10.1186/s12938-024-01272-6>

Conference papers

Lening Li, Man-Sang Wong. Automatic Segmentation of 3D CT Images for Patients with Scoliosis. In: the 7th Singapore Rehabilitation Conference and Asian Prosthetics and Orthotics Scientific Meeting (SRC-APOSM 2022), Singapore, 8 - 9 October 2022

Lening Li, Man-Sang Wong. Enhanced 3D Cobb Angle Estimation in Idiopathic Scoliosis Using U-Net Based CT Image Segmentation. In: International Research Society of Spinal Deformities (IRSSD) Scientific Meeting 2024. Hong Kong, 21-23 June 2024.

Under-reviewed journal paper

Li, L., Zhang, T., Lin, F., Wong MS. A Transformer-Based Neural Network for Vertebra Segmentation in Scoliosis Patients. Journal of Applied Clinical Medical Physics, under review.

Acknowledgements

First and foremost, I wish to express my deepest gratitude to Prof. Man-Sang Wong, my supervisor, for giving me the opportunity to pursue my PhD studies. Throughout the lengthy journey of my PhD, he has provided me with invaluable insights. "Keep balance" has been his consistent advice, which I have struggled to understand and continuously achieve. He is a selfless clinician who use his knowledge to help more patients, guiding my own career in medical science. He has taught me that the most important thing of scientific research is to benefit patients directly, as serving the patient's real needs is the true responsibility of a researcher.

I would like to thank Dr. Lin Fan from the Second People's Hospital of Shenzhen, from whose department all my research data collected. During the COVID-19, it was difficulty to complete the ethical approval for our trial data and gradually accumulate clinical patient data week by week. I will forever remember the weekends spent reviewing patient information in the PACS system together.

Special thanks to Dr. Zhang Teng and Mr. Qin Shuo, who have greatly assisted me with deep learning algorithms and coding. Without their help, I could not have achieved the results I have today.

I am also deeply grateful to my family: my parents, my husband, and my daughter. My daughter, in particular, came into my life just when I began my PhD studies. Her growth has paralleled my study journey. She is my first child, and this PhD is the second.

Lastly, I want to thank myself. As a part-time PhD student, my full-time job is venture capital within the healthcare industry. The busy work, coupled with the high needs of a young baby from mom, leaves very little time for my own research. I have only been able to follow Professor Wong's advice to "keep balance" and struggle for equilibrium between work, life, and study. There were so many times I thought about giving up, but

the drive to be an idol for my daughter has kept me going. Thanks to myself for completing this meaningful journey.

TABLE OF CONTENTS

CHAPTER 1 INTRODUCTION	1
1.1 Background.....	1
1.2 Study Objectives.....	4
CHAPTER 2 LITERATURE REVIEW	5
2.1 Anatomy of spine and vertebra.....	5
2.2 Adolescent Idiopathic Scoliosis	9
2.2.1 Prevalence.....	10
2.2.2 Symptoms and Progression	11
2.2.3 Assessment.....	13
2.2.4 Classification	16
2.3 Image modality of AIS	22
2.3.1 Radiology	22
2.3.2 EOS	25
2.3.3 Computed Tomography.....	26
2.3.4 MRI	28
2.3.5 Ultrasound	29
2.3.6 Summary	31
2.4 Treatment of AIS	31
2.4.1 Non-surgical treatments	32
2.4.2 Surgical treatments.....	35
2.5 Image analysis based on deep learning.....	36
2.5.1 Radiography	36
2.5.2 EOS.....	39
2.5.3 CT.....	42
2.5.4 MRI.....	46
2.5.5 Ultrasound	47
2.6 Summaries	48
CHAPTER 3 AUTOMATICALLY VERTEBRA SEGMENTATION	
BASED ON U-NET	50
3.1 Introduction	50
3.2 Methodology.....	51
3.3 Results	66
3.4 Discussion.....	67
3.5 Conclusion	68

CHAPTER 4 AUTOMATICALLY VERTEBRA SEGMENTATION	
BASED ON NNFORMER++	69
4.1 Introduction	69
4.2 Methodology.....	70
4.3 Results	82
4.4 Discussion.....	85
4.5 Conclusion	86
CHAPTER 5 3D SPATIAL ANGLE MEASUREMENT	88
5.1 Introduction	88
5.2 Methodology.....	89
5.2.1 Brief methodology of NURBS	89
5.2.2 Optimization of Spinal Curve Reconstruction: From K-means	
Clustering to NURBS-net Prediction	94
5.2.3 <i>Measurement of 3D spatial angle</i>	100
5.3 Results	102
5.4 Discussion.....	106
5.5 Conclusion	108
CHAPTER 6 CONCLUSION	109

LIST OF FIGURES

Figure 1 The articulated vertebral column in an adult – anterior, posterior and lateral views [43]	5
Figure 2. Essential features of typical cervical, thoracic and lumbar vertebrae[43]	7
Figure 3. A spinal motion segment[43]	8
Figure 4. Illustration of the six Risser stages of skeletal age[57].....	12
Figure 5. Forward bend test for scoliosis screening[62].	14
Figure 6. Scoliometer[67].	15
Figure 7. Cobb angle measurement[69]	16
Figure 8. The King classification[70].	17
Figure 9. The Lenke clasification.	19
Figure 10. Overview of the most common imaging modalities in AIS[99].	24
Figure 11. Estimated average radiation doses during dedicated periods in AIS management. Total average estimated cumulative dose 5572.74 mrem (55.72mSv)[113].	28
Figure 12. 3D printed Boston Brace[159]	35
Figure 13. Result from the deep learning method, the most pronounced curve was characterized by a Cobb angle of 88° for the principal curve, with corresponding upper and lower curves measured at 30° and 49°, respectively[164].	38
Figure 14. The predicted outcomes for two subjects from the testing set encompass significant data points such as UV (upper vertebrae), AV (apical vertebrae), and LV (lower vertebrae), along with CSVL and C7PL[166]	38
Figure 15. Activation maps for one patient samples, wherein regions of interest, specifically the iliac crests, humeral heads, and femoral heads, exhibit pronounced activation. In addition to these primary regions, the activation maps also reveal enhanced activation in the cranial bones and rib cage[174].	41
Figure 16. Bio-informed Mechanistic predictions (FFNNCR-BM) with ground truth from EOS at different age stages. Comparisons are made at 160 months (within the trained data range), 179 months, and 187 months (both outside the trained data range), using various 2D and 3D views, including detailed geometric reconstructions[176].	42
Figure 17. Example of the sagittal Cobb angle thoracic kyphosis (left) and lumbar lordosis (right) measurement from the method[41].....	43
Figure 18. The magnitude of differences of the automatic segmentations from the ground truth segmentations	44
Figure 19. Binary spine segmentation	46
Figure 20. The network of CNN.	52

Figure 21. The process of convolutional computation	54
Figure 22. Three active functions	55
Figure 23. the Max pooling and average pooling	57
Figure 24. Different connection method	60
Figure 25. The architecture of U-net	61
Figure 26. Segmentation result in Clinical Cases: Displayed are the segmentation outcomes from two distinctive cases within the study cohort. Figure (a) showcases a sample from the training set, whereas figure (b) presents a corresponding example from the validation set. Each image has been carefully trimmed to minimize the extensive non-relevant background, ensuring focused representation of the spinal region	67
Figure 27. Structure of nnFormer++	71
Figure 28. The best segmentation performance in scoliosis[214],	76
Figure 29. Segmentation result in in-house data: Displayed are the segmentation outcomes from two distinctive cases within the study cohort. Figure (a) showcases a sample from the training set, whereas figure (b) presents a corresponding example from the validation set. Each image has been carefully trimmed to minimize the extensive non-relevant background, ensuring focused representation of the spinal region	84
Figure 30 3D mesh of spine in testing dataset	85
Figure 31. (a) displays a cubic B-spline curve and its control polygon, while (b) shows the basis functions for the B-spline curve.	90
Figure 32. The illustration of knot	93
Figure 33. 3rd-order curve with 6 points	94
Figure 34the Spoke Kernel Filtering process. (A) At the core of this technique is the placement of the spoke kernel's center upon each pixel within the spinal segmentation area. (B) Diameters extend from this central point, and endpoints are scrutinized. Should both endpoints reside outside the vertebral region (indicated by red points), the intermediate pixels will be excised from the segmentation mask. Conversely, if any endpoint is within the spinal boundary, the connecting pixels will be maintained (shown by blue dashed lines).	97
Figure 35. outlines the deep learning network specifically architected for spinal curve fitting. It employs ResNet as its backbone, processing spine segmentation as input and yielding control points and knots of NURBS curves as outputs.	99
Figure 36. Illustration of the PRED-3D-CA calculated method	101
Figure 37. Post-processing and curve fitting. (a) and (b) are post-processing results of two patients from the training dataset.	103
Figure 38. (a) The X-ray, (b) CT and (c) NURBS-curve of one patient from the testing dataset.	104
Figure 39. The Bland–Altman plot assessing the difference of the Cobb angle measurements, i.e., PRED-3D-CA, MAP-2D-CA and XRAY-CA.	106

LIST OF TABLES

Table 1. A survey of AIS prevalence and female-to-male ratio.	11
Table 2. A general overview of the modalities and their clinical parameters available for imaging of the chest in AIS [99]	24
Table 3. Bracing studies using different brace types[151].	34
Table 4. Summary of the data used in this study	64
Table 5. Label information in VerSe dataset	78
Table 6. Label information in VerSe dataset	78
Table 7. Result compared with other method in Verse 19	83
Table 8. Result compared in U-net, nnformer and nnformer++.	84
Table 9. Difference between XRAY-CA, PRED-3D-CA and MAP-2D-CA. .	105

CHAPTER 1 INTRODUCTION

1.1 Background

Adolescent Idiopathic Scoliosis (AIS) is defined as a three-dimensional deformity of the spine with unknown etiology, and it accounts for 90% of idiopathic scoliosis cases in children[1]. According to the international consensus, curvature in the coronal plane larger than 10° is defined as AIS. The overall prevalence of AIS in the whole world is 0.47%-0.52% [1].. Gender is an important factor in scoliosis. The incidence is higher in girls than boys, with a female-to-male ratio of 1.5:1[2] to 3:1[3] indicative of this discrepancy. AIS is not merely a musculoskeletal disease, but it can also result in back pain[4], shoulder imbalance[5], and even compromise cardiopulmonary function[6, 7] in severe deformity. Additionally, recent studies have increasingly focused on mental health[8, 9], due to the potential for AIS to induce psychological conditions such as depression and anxiety.

The management of AIS mainly involves two approaches: non-surgical treatment, typically facilitated through orthotic bracing, and surgical intervention. These approaches largely hinge on the Cobb angle, as measured from the coronal radiograph[10]. However, the influence of curve rotation or positioning is comparatively minimal[11]. As a result, the evaluation and management of AIS have seen an increasing focus on a range of parameters derived from three-dimensional (3D) models. These parameters are critical in providing a more comprehensive understanding of the condition. For instance, the efficacy of treatment, particularly post-bracing, benefits from the consideration of parameters such as thoracic and lumbar Cobb angles, in addition to thoracic kyphosis (T1-T12) and lumbar lordosis (L1-S1) [12]. Furthermore, the Plane of Maximum Curvature (PMC) has been recognized as a particularly significant parameter in numerous aspects of AIS assessment[13, 14], classification[15], and treatment[16].

With the accelerating progression of artificial intelligence technology, the utilization of deep learning in medical image processing has become increasingly widespread. The attainment of precise and satisfactory results in deep learning, particularly when employing the latest transformer methodologies, frequently necessitates the use of extensive datasets. However, securing a substantial volume of efficacious data for algorithmic training presents considerable challenges in specific medical scenarios, such as in cases involving AIS patients. Given the concerns related to radiation dosage, both 2D radiographs and 3D CTs scans are generally obtained exclusively from patients in the pre-surgical stage. However, surgical instances constitute only an insignificant proportion of AIS cases. Consequently, the restricted availability of large-scale datasets severely constrains the training of segmentation models.

Traditional deep learning methods, such as UNet, generally segment images into patches, treating each patch as an independent instance for learning. However, this strategy results in the loss of spatial relationships among patches, hindering the Convolutional Neural Network (CNN) model's capacity to effectively capture dependencies between distant pixels. In contrast to CNN models, the Graph Neural Network (GNN) model [17-19] exhibits the ability to learn relationships between distant entities while substantially maintaining topographic information in graph data. In recent advancements, GNN has demonstrated its value in addressing the issue of sample insufficiency. The GNN extracts image features through a feature extraction network and subsequently represents them as nodes within a graph convolutional network. The transfer of information between nodes takes place through graph convolution operations, culminating in image segmentation predicated on the feature information resident in the nodes.

This research[20] initially utilized GNN in small sample images and proposed the construction of a comprehensive graph network. Each node feature is associated with the corresponding class tag, after which the label information is propagated by updating the node feature through the attention mechanism of the graph network. This work laid

the foundation for the study of small-sample learning based on GNN. The research published in 2018[21] introduced a Transductive Propagation Network (TPN) method and applied transductive settings to this model. The labels in transductive graphs are propagated from the support set to the query set through the Laplacian matrix, subsequently measuring similarity. Another methodology established[22] a novel Edge-labeling Graph Neural Network (EGNN), constructed based on the similarity and dissimilarity between samples. It can dynamically update node and edge features in complex interactions. A series of attention mechanisms have been proposed to address overfitting or over-smoothing problems[23]. However, few studies have focused on how to reduce the depth and number of features in a neural network. Given the complex topological structures present in medical images, a large number of features and depth could lead to suboptimal segmentation results and increased time consumption.

In severe instances, surgical intervention becomes imperative to address the spinal curvature[24]. The relevance of CA is extensively established, demonstrating its critical role in determining fusion strategies, predicting postoperative curve corrections[25], and evaluating spontaneous curve correction in specific patient cohorts[26]. Additionally, CA is instrumental in identifying factors that may contribute to postoperative complications[27] and in improving the precision of predictive models for surgical outcomes[6, 7]. It significantly influences spinal height correction, with substantial height gains observed following surgery [28-30], and impacts important surgical considerations such as the duration of the procedure, blood loss, and the necessity for blood transfusions[31-33]. Furthermore, CA is predictive of the 'adding-on' phenomenon after surgery[34] and the success rates of selective thoracic fusions[35]. The magnitude of the preoperative CA is also associated with surgical challenges and risk assessments[36, 37], highlighting its extensive impact on scoliosis treatment and postoperative improvements, including enhancements in patient self-image[38].

Given the spine's intricate anatomy and the three-dimensional nature of the deformity, measuring CA using computed tomography (CT) scans can provide more detailed

information[11, 39]. While many studies and public databases have implemented vertebral segmentation using CT, advancements in CA measurement are seldom discussed. Previous studies on CT-based CA calculations have mainly focused on mild to moderate scoliosis cases[40-42]. Although CT imaging is typically acquired for presurgical assessment in severe cases to minimize unnecessary radiation exposure, the evaluation of spinal curves in this patient group remains underexplored.

1.2 Study Objectives

- 1) To develop a spine segmentation framework based on U-net and conduct segmentation training.
- 2) To develop a transformer-based spine segmentation network and validation of public and in-house datasets.
- 3) To engineer a rapid and resilient method on 3D spatial angle estimation.

CHAPTER 2 LITERATURE REVIEW

2.1 Anatomy of spine and vertebra

The spinal column holds a principal position within the entire skeletal system and acts as a support and protective mechanism for the internal organs. It is the primary structure for bearing weight, movement, absorbing shocks, and balancing the body.

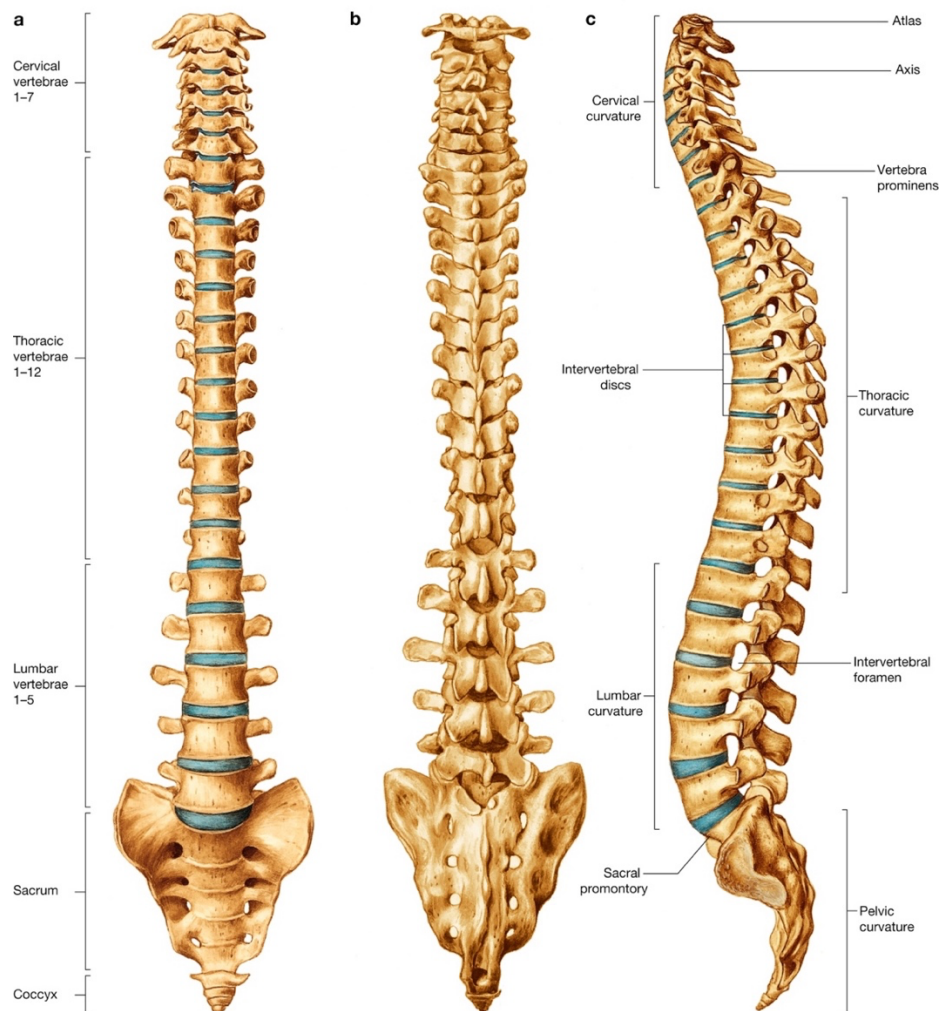


Figure 1 The articulated vertebral column in an adult – anterior, posterior and lateral views [43]

The spine consists of 33 vertebrae, forming a flexible pillar. In adults, there are five sacral vertebrae fused into the sacrum and four coccygeal vertebrae forming the coccyx, leaving only 24 vertebrae that can move: 7 in the neck (cervical), 12 in the chest (thoracic), and 5 in the lower back (lumbar). Except for the unique shapes of the first two cervical vertebrae, the remaining vertebrae share common features, including the vertebral body, vertebral appendages, pedicles, laminae, superior and inferior articular processes, transverse processes, and spinous processes.

The size of the vertebrae increases gradually from top to bottom, with cervical vertebrae being small, and lumbar vertebrae being large and sturdy. The spinal column has four physiological curves: the cervical and lumbar vertebrae curve forward, and the thoracic and sacral vertebrae curve backward, with the junctions between the curves being most susceptible to fractures. From an anatomical function perspective, the spinal column can be divided into several basic parts:

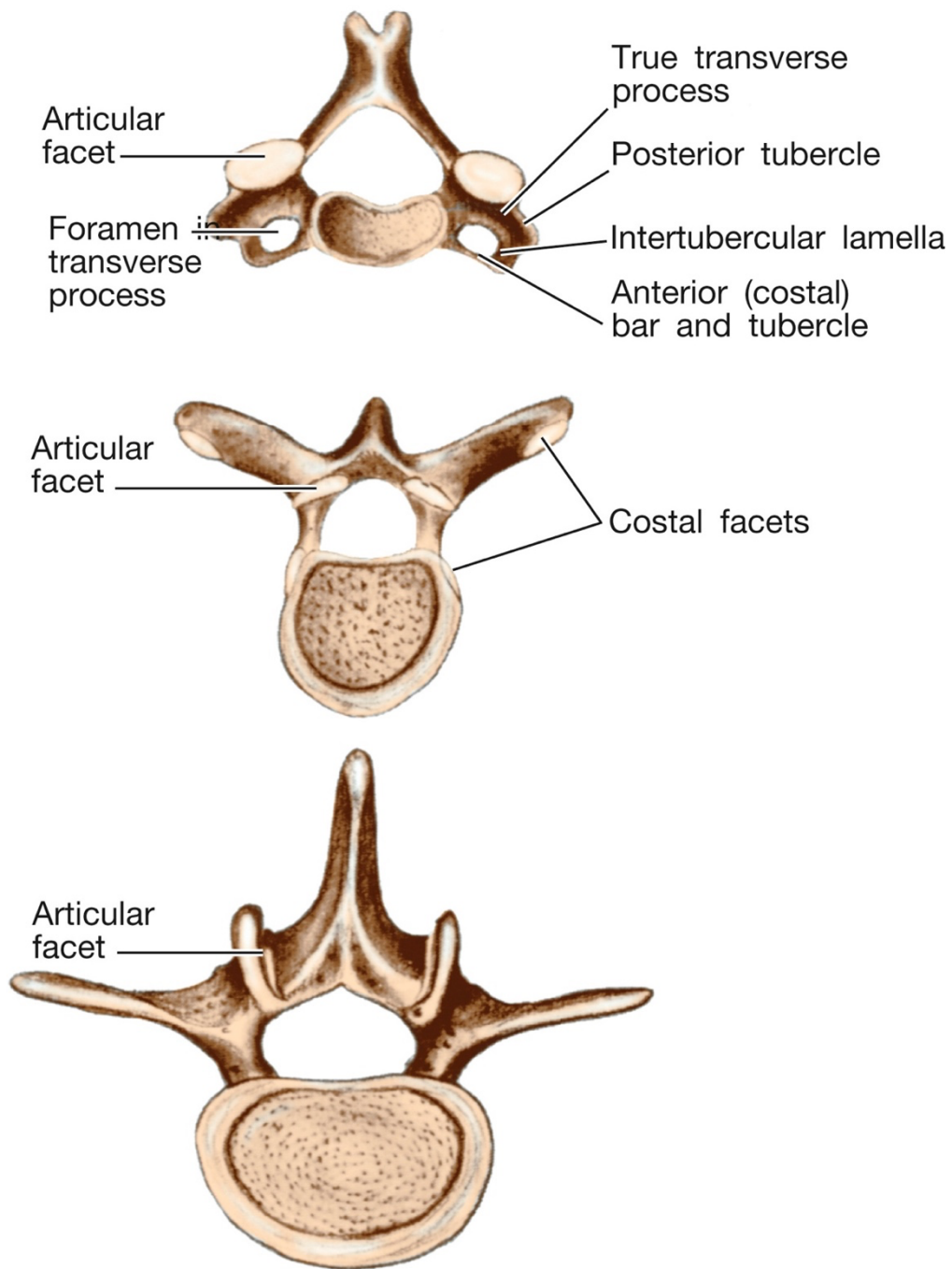


Figure 2. The structure of typical cervical, thoracic and lumbar vertebrae[43]

Vertebral Bodies and Intervertebral Discs: Shaped like short columns, the vertebral bodies mainly consist of spongy bone, with intervertebral discs between them. This section bears weight and acts like a pillar for the trunk. Intervertebral discs also function as shock absorbers.

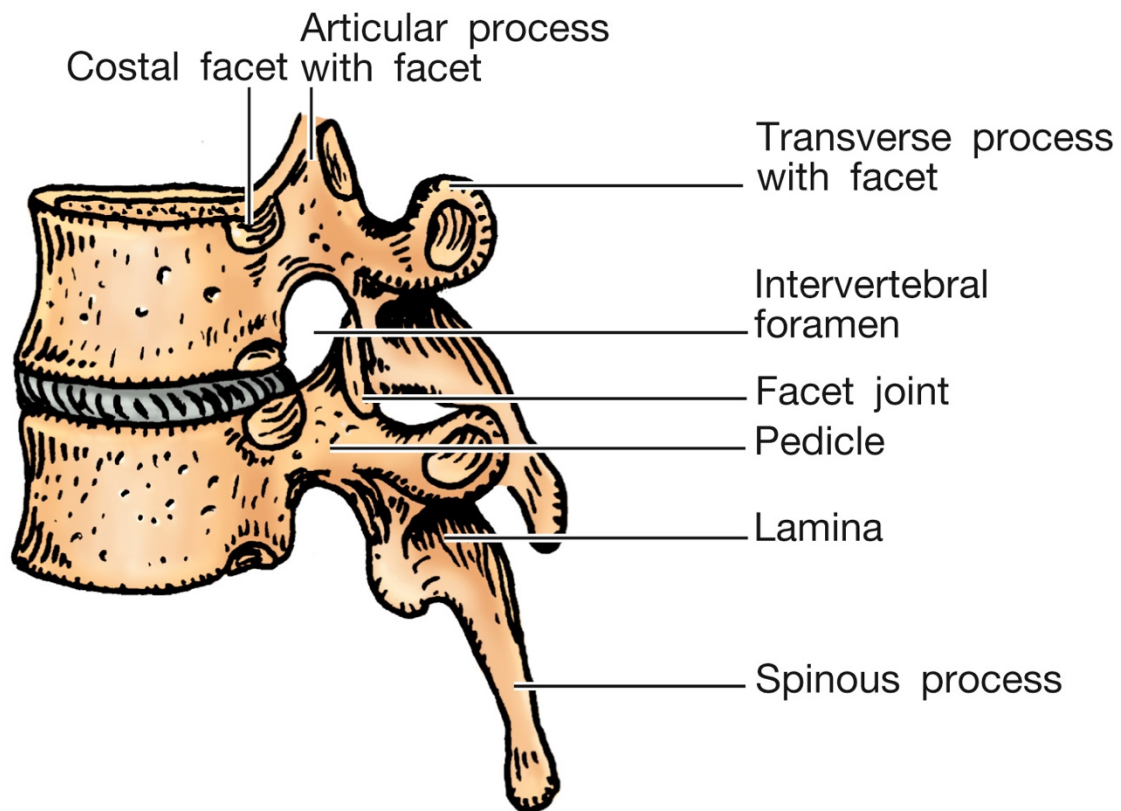


Figure 3. A spinal motion segment[43]

Neural Arch and Facet Joints: The neural arch is a curved structure formed by the fusion of the pedicles and laminae. It primarily protects the spinal cord. Facet joints increase the spine's stability and prevent dislocation.

Spinous Processes and Transverse Processes: The spinous processes prevent the spine from overextending, while the transverse processes provide attachment points for muscles that stabilize the spine.

Ligaments: There are three main ligaments that run the full length of the spine, especially the anterior longitudinal ligament, which is extremely strong and plays a vital role in preventing overextension of the spine.

Vertebral Canal: This canal is created by the vertebral body and arch merging together and houses the spinal cord. The canal's dimensions change along the spine, being broader in the cervical area and more constricted in the thoracic region.

Intervertebral Foramen: Intervertebral Foramen: Spine nerves could go through the holes between adjacent vertebrae.

Spinal Cord and Nerves: The spinal cord begins at the base of the skull and ends at the first lumbar vertebra, with a total of 31 pairs of spinal nerves. Fractures or dislocations below the second lumbar vertebra often result in injury to the cauda equina nerves, while above this level, spinal nerve injury may occur.

Additionally, the spinal cord features two prominent regions of expansion: the cervical expansion (spanning from C3 to C7) and the lumbar expansion (stretching from T10 to L1), with the latter serving as a crucial hub for controlling bladder functions autonomously.

Scoliosis Description: Scoliosis is identified by its distinct spinal anomalies, which include a lateral bend and a rotational displacement of the vertebrae. The etiological factors contributing to scoliosis are multifaceted, and can be systematically categorized into five primary classifications: congenital abnormalities, neuromuscular conditions, syndromic causes, idiopathic origins, and secondary-induced spinal curvature

2.2 Adolescent Idiopathic Scoliosis

Scoliosis is defined as a 3D spinal deformity, evidenced by both a lateral deviation and a rotational misalignment within the vertebrae. The underlying causes that contribute to scoliosis are complex, and they can be organized into three main categories: congenital, syndromic, and idiopathic types. Congenital scoliosis is attributed to malformation in the vertebrae, leading to the observed deformity. Syndromic scoliosis relates to disorders affecting the neuromuscular, skeletal, or connective tissue systems; this includes, but is not limited to, conditions such as neurofibromatosis. Idiopathic scoliosis, on the other hand, has no identifiable etiology. Notably, adolescent idiopathic scoliosis stands as the predominant form of spinal deformity encountered in clinical practice.

2.2.1 Prevalence

Adolescent Idiopathic Scoliosis (AIS) is defined as a three-dimensional deformity of the spine with unknown etiology, and it accounts for 90% of idiopathic scoliosis cases in children[1]. The prevalence of AIS (Adolescent Idiopathic Scoliosis) exhibits certain variations across different regions, but a consistent pattern is that the incidence is higher among females than males, and the majority of cases are found to be mild at the time of detection[16, 44-51]. Furthermore, the prevalence of AIS is correlated with BMI (Body Mass Index), with a lower BMI being associated with a higher prevalence rate[52, 53]. AIS has now become a serious disease threatening adolescents, a trend that is connected to the increasing pressures of academic studies.

Study	Country	Children	female-to-male ratio	Age	Prevalence combined (%)	Prevalence girls(%)
Fu[44]	China	79122	0.8291	10-16	2.4	3.12
Hu[45]	CHina	10731	0.9447	11-16	2	2.65
Hurriyet[47]	Turky	16347	1.0367	10-15	2.3	3.1
Penha[46]	Brazil	2562	1.4283	10-14	1.5	2.2
Kamtsiuris[48]	Germany	17641	0.9623	0-17	5.2	11–13 years:8.3 14–17 years:13.5
Daruwalla[50]	Singapore	110744	1.1896	6–7 11–12 16–17	1	6–7 years: 0.15 11–12 years: 1.67

						16–17 years: 3.12
Suh[49]	Korea	1,134, 890	0.9415	10-14	3.26	4.65
Thomas[51]	America	1782	2.083	10-18	0.5225	0.7177

Table 1. A survey of AIS prevalence and female-to-male ratio

2.2.2 Symptoms and Progression

Scoliosis that gets worse can lead to noticeable changes in the shape and look of the body[5]. These changes can include shifts in the chest wall, ribs that stick out, shoulders that are at different heights, and a twisted torso. People with AIS often have shoulders that don't line up, a waistline that appears uneven (with one hip sticking out more than the other), or a rib that is more noticeable. These changes are often first seen by family members. Complaints of back pain are sometimes the first symptom[4]. In addition to the impact on physical posture, the undeniable effect on cardiopulmonary function is evident in patients with severe AIS[6, 7]. Furthermore, due to the influences on appearance and pain, an increasing number of studies are now discovering the psychological issues, such as depression and anxiety, caused by this disease[8, 9].

Research into the disorder's evolution during growth phases has been documented[54, 55]. Such investigations point out that factors like chronological age, sex, curvature size (as measured by the Cobb angle), the status of the first menstrual cycle in girls, and the Risser sign are vital in forecasting how the curve will develop while growing. The Risser sign points to the visual manifestation of the pelvis's iliac apophysis and serves as a tool for assessing skeletal maturity[56]. Six stages of Risser, ranging from naught to five, signify the transition of the apophysis from the front to the back of the iliac spine, culminating in its union with the iliac bone[57]. There exists a documented relationship between the progression of untreated AIS, the Risser sign, and the extent

of the curve. With regard to curvature spanning 20° - 29° in an undeveloped youth exhibiting a Risser sign of either 0 or 1, progression was observed in 68% of cases. For angles of less than 19° in a fully grown young person with a Risser sign of 2 or higher, the progression rate was merely 1.6%. When dealing with minor curves of under 19° in an undeveloped youth (Risser sign 0 or 1), and more pronounced curves (20° - 29°) in a developed child (Risser sign of 2 or above), the rate of progression was nearly identical, at approximately 22% and 23% correspondingly[58].

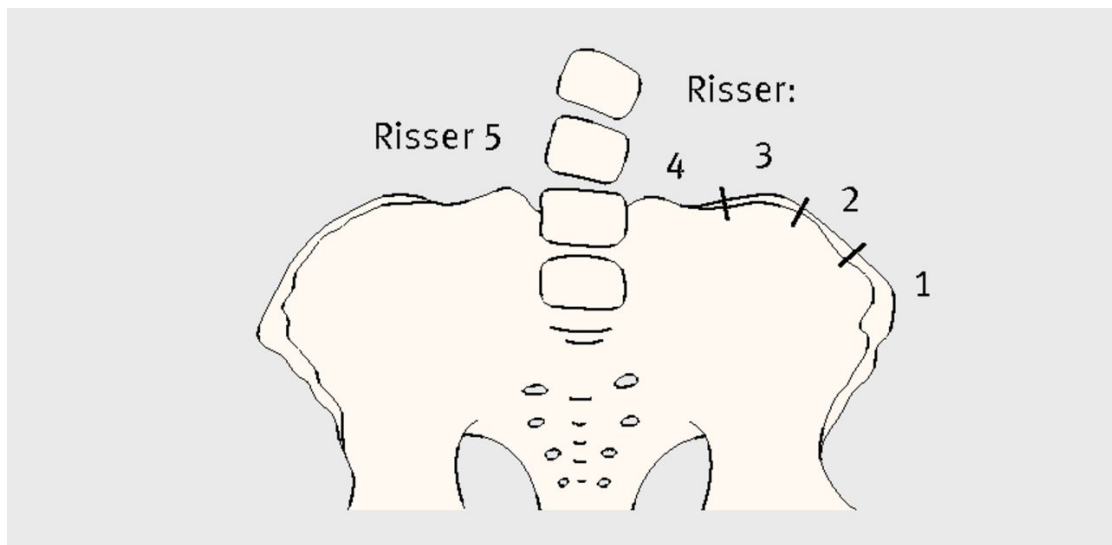


Figure 4. Illustration of the six Risser stages of skeletal age[57]

One terrible aspect is that spinal curvature does not cease with the conclusion of adolescence. An examination of scoliosis's progression was conducted through a prospective series of cases, comprising 133 subjects[59]. They were tracked for an average duration of 40.5 years (with a range of 31-53 years), and it was found that 68% of the deformities in adolescent idiopathic scoliosis continued to develop beyond the stage of complete bone growth. Curves in the thoracic area that were greater than 50° developed at an average pace of 1° per year, while curves in the thoracolumbar section increased by 0.5° each year, and those in the lumbar region expanded by 0.24° annually. Curvatures in the thoracic region below 30° showed no change.

Back pain is more prevalent in adult spinal scoliosis. Previous studies [60, 61] indicate that the pain may be due to spinal imbalance (such as coronal or sagittal imbalance), facet joint pathology, muscle fatigue, or central or foraminal stenosis. Asymmetrical loading of the intervertebral discs and facet joints leads to dysfunction in one or more spinal segments, followed by segmental instability, whether in the sagittal plane (i.e., spinal slippage), or the more common coronal plane shift (i.e., lateral slippage), or three-dimensional rotational subluxation. Osteophytes form at the vertebral endplates and facet joints, and hypertrophy of the ligamentum flavum and joint capsules leads to central and/or lateral recess stenosis and subsequent nerve root pathology. In concavities, foraminal stenosis is common and may also result in nerve root pain. Increases in curve length are associated with more severe pain, as is the degree of degenerative changes at the curve apex.

2.2.3 Assessment

The forward bend test is a widely accepted procedure to identify AIS, and it is frequently used in conjunction with instruments like a scoliometer or the technique known as Moiré topography. To carry out this test, the patient stands straight, and the examiner checks from behind to detect any noticeable curvature in the spine (Figure [62]). The patient then bends forward to make the spine level with the ground, allowing the arms to hang, palms touching, and knees pointing ahead. The examiner looks for any bump in the thoracic or lumbar region that may indicate scoliosis.

This test is highly sensitive, detecting 92% to 100% of thoracic scoliosis cases in patients with a Cobb angle over 20 degrees[63]. However, it's less effective for spotting lumbar scoliosis. If a curve is suspected, a scoliometer is often used, and it helps decide whether or not to order an X-ray for a definitive diagnosis. The scoliometer measures the angle of trunk rotation, a value that can help estimate the exact Cobb angle, which is measured using radiography.

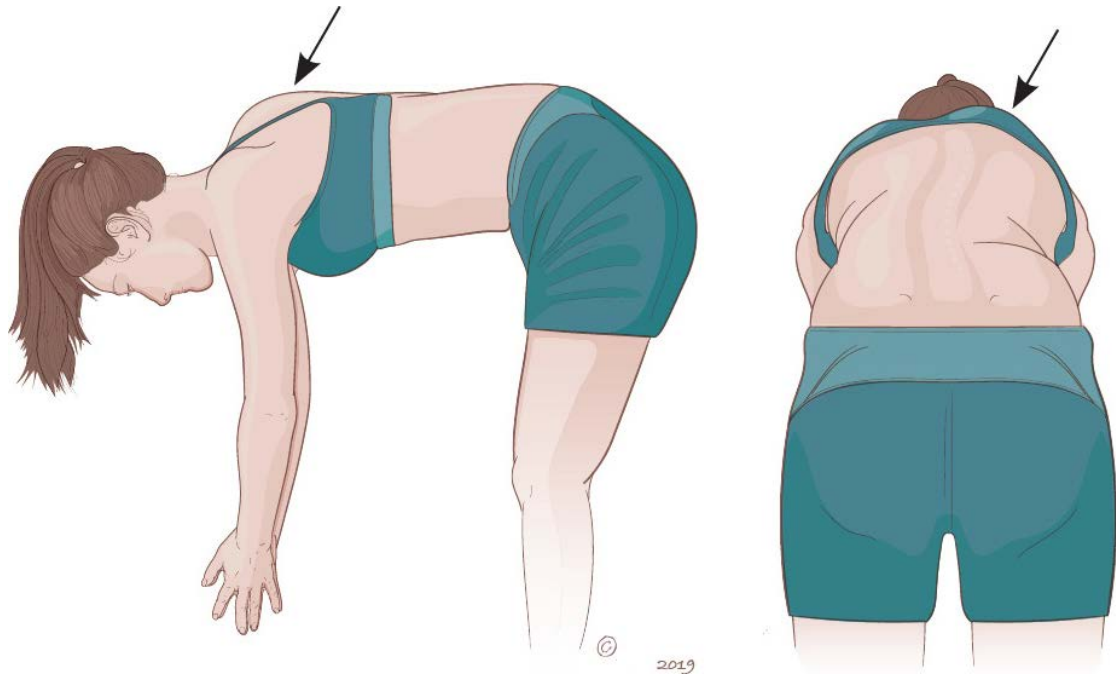


Figure 5. Forward bend test for scoliosis screening[62].

The specific points at which an X-ray is ordered vary, usually between 5 to 7 degrees of trunk rotation as measured by a scoliometer[64, 65]. This translation between scoliometer rotation and Cobb angle isn't precise. For example, in normal-weight patients, a 7-degree angle of trunk rotation corresponds roughly to a Cobb angle of 20 degrees. The scoliometer is about 83% sensitive and 87% specific when using a cutoff of 7 degrees of trunk rotation for a Cobb angle greater than 10 degrees. When a 5-degree threshold is used, sensitivity increases to 100%, but specificity falls to 47%[66]. This lower cutoff should be used in patients with a body mass index in the 85th percentile or higher[65].

X-ray test should be ordered for any adolescent with a noticeable curve observed during a physical exam or if there is clear unevenness in the chest or lower back, especially in those with a family history of scoliosis. This can also be vital for tracking the development of scoliosis in patients who have already been diagnosed.

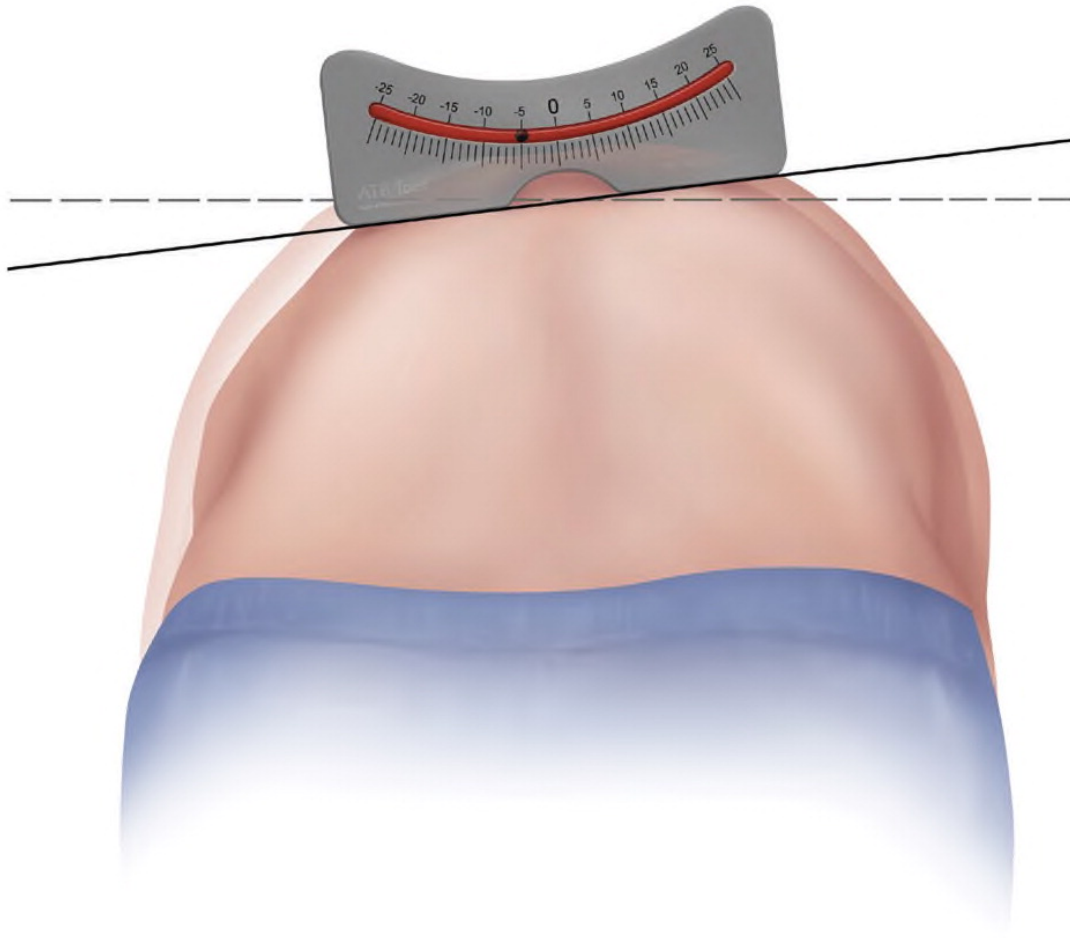


Figure 6. Scoliosimeter[67]

The Cobb angle measurement is a widely accepted method for diagnosing and assessing the severity of AIS[10]. This technique involves measuring the side-to-side curvature of the spine using X-ray images taken from the front or back[68]. Specifically, the measurement focuses on the two most slanted vertebrae at the top and bottom of the curve. By evaluating the Cobb angle, healthcare professionals can determine the degree of spinal curvature. It's essential to measure this angle accurately, as treatment decisions are based on specific established values[68]. Additionally, the Cobb angle helps in tracking how scoliosis changes over time. A change of more than 5° in repeated measurements is a sign that the curve is progressing[58].

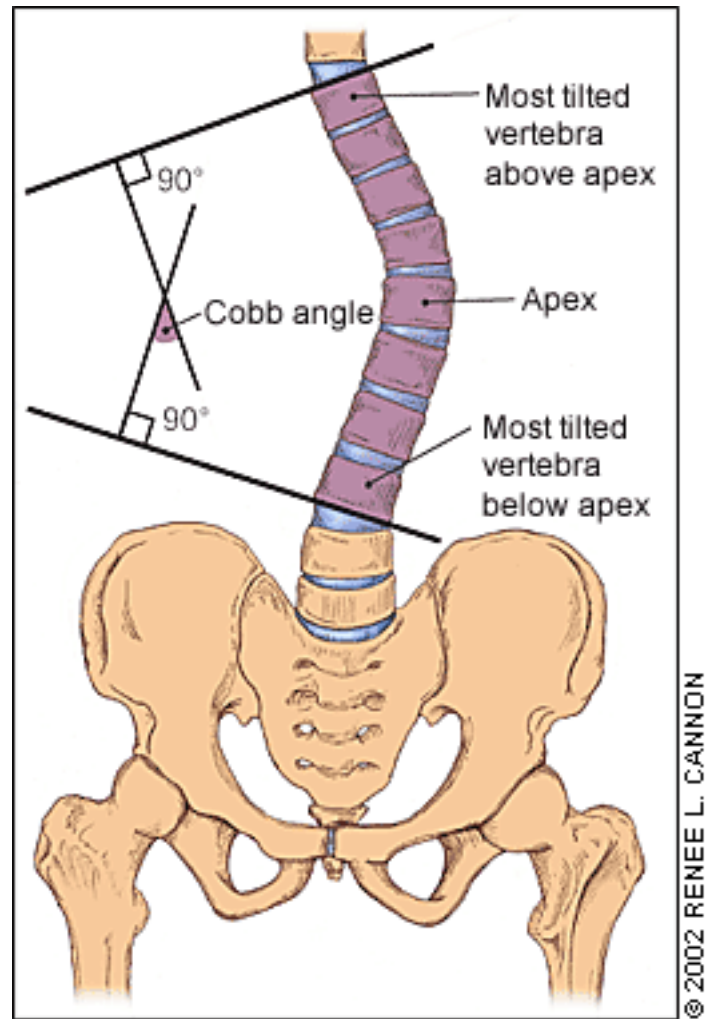


Figure 7. Cobb angle measurement[69]

2.2.4 Classification

In 1983, Howard King and colleagues analyzed the data of 405 patients with scoliosis, classifying AIS into five types based on characteristics such as the position of the apical vertebra, the severity of the lateral curvature, and the flexibility of the spine, with the thoracic curve being the primary consideration[70].

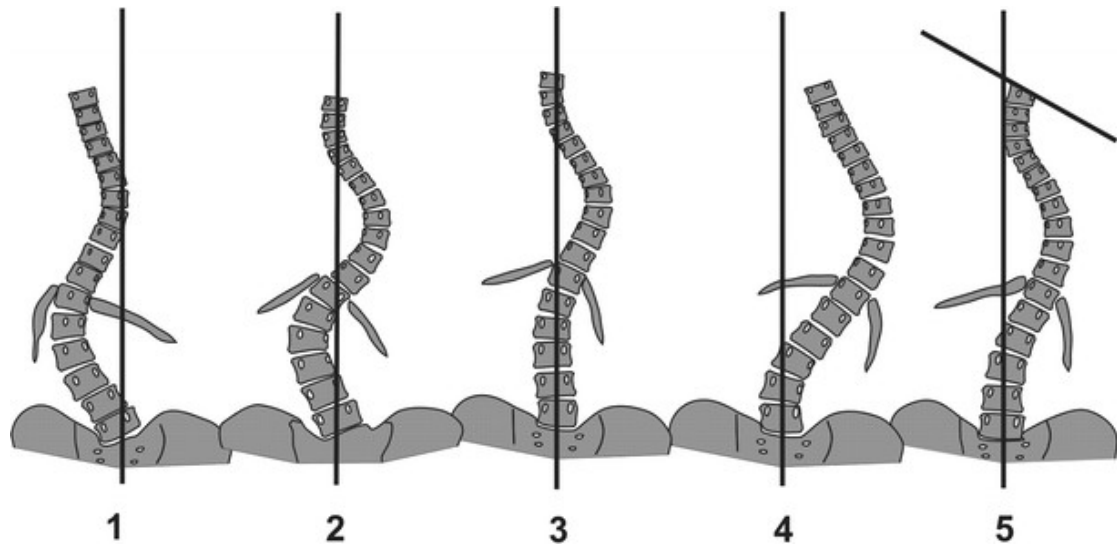


Figure 8. The King classification[70]

Type I: Both the lumbar and thoracic curves extend beyond the center sacral vertical line (CSVL), and the lumbar curve's Cobb angle is larger, with its flexibility being worse than the thoracic curve (if the thoracic curve is greater than the lumbar curve in the standing position but the lateral curvature is more flexible towards the thoracic curve, it is also classified as Type I).

Type II: Both the thoracic and lumbar curves extend beyond the CSVL, and the thoracic curve's Cobb angle is larger, with its flexibility being worse.

Type III: A single thoracic curve, with its compensatory lumbar curve not extending beyond the CSVL.

Type IV: This category features an extended thoracic curve where the center sacral vertical line bisects L5, and L4 is inclined towards the lengthy thoracic curvature.

Type V: Structural double thoracic curve, with T1 tilting towards the concave side of the upper thoracic curve or the convex side of the lower thoracic curve.

The proposal of this classification has milestone significance in the development history of spinal orthopedic surgery. However, with the growing preference for

segmental instrumentation methods supplanting the Harrington rods, this categorization methodology was found lacking in providing precise and dependable directives for the selection of suitable fusion levels[71-73]. Furthermore, numerous studies have indicated the existence of minimal consistency between different observers (inter-observer) and the same observer over time (intra-observer) with regard to this categorization system[74, 75].

In 2001, Lenke[76] addressed the deficiencies found in King's classification, proposing a two-dimensional classification method that includes types of scoliosis (I to VI), lumbar curve modifiers (A, B, C), and sagittal thoracic curve modifiers (-, N, +). In recent years, the Lenke classification has become the standard method for classifying idiopathic scoliosis internationally. Based on the position of the primary lateral curve and the structural characteristics of the secondary lateral curve, the Lenke classification can be specifically divided into 6 types.

















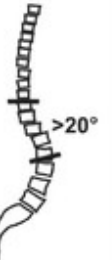

Curve type (1-6)						
Lumbar deviation (A-C)	Type 1 Single thoracic	Type 2 Double thoracic	Type 3 Double major	Type 4 Triple curve	Type 5 Thor.-lumb. or lumbar	Type 6 Thor.-lumb. or lumbar
A minimal	 1A	 2A	 3A	 4A		
B moderate	 1B	 2B	 3B	 4B		
C severe	 1C	 2C	 3C	 4C	 5C	 6C
Sagittal plane	 Normal	 Cerv.-thor. kyphosis	 Thor.-lumb. kyphosis	 Cerv.-thor+ thor.-lumb. kyphosis		

Figure 9. The Lenke clasification

Type 1: Main thoracic curve; the thoracic curve is the primary curve, with proximal thoracic and thoracolumbar/lumbar being non-structural secondary curves.

Type 2: This type consists of a double thoracic curvature where the primary curve is in the thoracic region. The proximal thoracic curve is structurally secondary, whereas the thoracolumbar/lumbar curve is not structural.

Type 3: Known as a double major curve, this classification features structural thoracic and thoracolumbar/lumbar curves. The proximal thoracic curve is non-structural. The thoracic curve is considered the principal lateral curve if its Cobb angle is equal to or exceeds that of the thoracolumbar/lumbar, or if the difference between them does not surpass 5 degrees.

Type 4: Characterized by a triple major curve, this type includes structural curves in the proximal thoracic, thoracic, and thoracolumbar/lumbar regions. Both thoracic and thoracolumbar/lumbar curves may function as the primary lateral curves.

Type 5: This type is defined by a thoracolumbar or lumbar curve, where the thoracolumbar/lumbar region forms the structural primary lateral curve, and both the proximal thoracic and thoracic curves are non-structural.

Type 6: Thoracolumbar/lumbar and thoracic curve; thoracolumbar/thoracic is the primary lateral curve with at least a 5° greater angle than thoracic. Thoracic is a structural secondary curve, and proximal thoracic is a non-structural curve.

Second Step: Based on the relationship between the center sacral vertical line (CSVL) and the position of the lumbar curve, the lumbar curve is further corrected into three types, A, B, and C.

Type A: CSVL passes between the pedicles of lumbar vertebrae below the stable vertebra. If there is doubt whether CSVL passes between the bilateral pedicles, it is classified as Type B, and the deformity must coincide with the apex vertebra at or above the T11/T12 intervertebral space.

Type B: CSVL is located from the lateral edge of the concave side pedicle to the lumbar vertebrae or the edge of the intervertebral disc. If there is doubt whether CSVL touches the edge of the vertebrae or disc, it is classified as Type B. This type does not include thoracolumbar/lumbar curvature.

Type C: CSVL is located outside the lumbar vertebrae or the edge of the intervertebral disc. The main lateral deformity may be located in the thoracic, lumbar, and/or thoracolumbar regions. If there is doubt about whether CSVL touches the edge of the vertebrae or disc, it is classified as Type B. Type C may include all deformities with the primary thoracic curvature and necessarily includes all thoracolumbar/lumbar curvatures.

In the treatment of AIS, although the proportion requiring surgery is relatively low, comprehensive and accurate pre-operative planning is essential for optimal surgical outcomes. The standard surgical treatment for AIS includes vertebral fusion and internal fixation. It is crucial to choose an appropriate number of vertebrae to minimize surgical complications while maximizing deformity correction. Opting for the least number of vertebrae for fusion, while achieving adequate spinal correction, minimizes the loss of flexibility, especially in the lumbar region. The two AIS classification standards mentioned above are based on two-dimensional parameters for three-dimensional deformity classification. An extensive body of literature has documented its limitations in surgical planning [77-80]. Associations exist between spinal deformations across three anatomical planes among subgroups of AIS patients, due to the influence of three-dimensional spinal alignment on posture balance and biomechanical functions.[81], some research has tried to use three-dimensional parameters for AIS classification to improve prognostic predictions for each subtype.

Aiming to overcome the limitations of existing 2D approaches, the 3D classification of the spine has been introduced. Numerous three-dimensional parameters have been established to describe the mechanics and geometric features of spinal curves[81-87]. Duong et al.[88] and Stokes[89] proposed various three-dimensional classifications

using spinal 3D reconstruction models and the metrics derived from torsional components and axial rotation. Implementing these methods in clinical practice has been challenging due to their complexity, the need for advanced computational requirements, and the difficulty of visually assessing parameters. These methods have been restricted to scientific research in the AIS classification[90]. Duong and colleagues utilized fuzzy clustering techniques on all Lenke-type AIS patients to delineate 11 distinct subtypes[91]. Sangole and their team analyzed clinical variables from the primary thoracic AIS curve spinal deformity, pinpointing three unique axial rotation subgroups within thoracic AIS[6]. Kadoury and associates applied dimensionality reduction and clustering techniques to study Lenke 1-type AIS patients.[92]. Their results identified four subgroups within Lenke 1-type AIS with different sagittal curve features: (1) normal kyphosis/excessive lordosis, (2) mild kyphosis/normal lordosis, (3) mild kyphosis/excessive lordosis, and (4) excessive kyphosis. Saba et al. [93]categorized five sagittal plane types: normal, lower thoracolumbar kyphosis without proximal kyphosis, lower thoracolumbar kyphosis accompanied by proximal kyphosis, a flat sagittal plane lacking proximal kyphosis but featuring a high inflection point, and a flat sagittal plane exhibiting proximal kyphosis with a low inflection point.

Although various 3D classification methodologies have been developed in the field, none have yet succeeded in supplanting the Lenke classification system as the gold standard. In this domain, there continues to be a pronounced need for more refined and accessible classification techniques.

2.3 Image modality of AIS

2.3.1 Radiology

As a progressive condition, AIS disease management demands ongoing clinical monitoring and periodic review of treatment, particularly in the adolescent growth phase. Patients at this developmental stage need regular imaging, often in the form of

annual or semiannual full spine radiography[94], to track the progression of the deformity. The imaging is still needed even after interventions such as bracing or surgical procedures. Radiography stands as the most readily available imaging technique, maintaining a pivotal role in AIS evaluation. It involves taking two-dimensional pictures from two angles: anteroposterior (front to back) and lateral (side to side). This approach allows doctors to assess vital diagnostic factors related to spinal curvature, including the type of curve, coronal Cobb angle, rotation at the apex of the vertebrae, or balance within the coronal plane [95]. Additionally, it can display the severity of rib prominence.

However, traditional X-rays have limitations, especially in evaluating the rotation of individual vertebrae and certain aspects of the pelvis. A significant drawback to this method is the harmful radiation it emits. Repeated exposure to this radiation carries risks of long-term health issues[96]. There is substantial evidence showing that excessive radiation during growing years can greatly increase the risk of cancer later in life, especially in children[97]. Past research reveals that depending on the treatment approach – whether surgical, bracing, or observational – AIS patients may undergo, on average, 12.2, 5.7, or 3.5 plain radiographs a year. This translates to annual radiation doses of 1400, 700, and 400 mrem respectively[98].

Modality	Radiation Exposure	Assessment of Rib Deformity	Costs	Principle Clinical Role
Radiography	Low	Planar evaluation of rib deformity	Low	Evaluation of scoliosis severity and progression during follow-up
EOS	Ultra-Low	Deformity evaluation of the bone structures. Possibility for 3D reconstruction	Low	Evaluation of scoliosis severity and progression during follow-up
CT	High	High accuracy of rib cage and	High	Sometimes used for visualizing complex

		internal organ structures. Possibility for 3D reconstructions and automatic segmentation		osseous abnormalities, for preoperative planning or as input for spinal navigation surgery.
3D ultrasound	None	Visualization of spinal posterior elements. Only first dorsal part of rib deformity included	Low	Mostly used in research settings for screening for scoliosis and evaluation of scoliosis severity.
MRI	None	Mainly used to assess soft tissue rather than bone structures	High	Evaluation for abnormalities of the spinal cord

Table 2. A general overview of the modalities and their clinical parameters available for imaging of the chest in AIS [99]



Figure 10. Overview of the most common imaging modalities in AIS[99]

2.3.2 EOS

The EOS imaging system, originating from the ultra-sensitive X-ray detection technology awarded the 1992 Nobel Prize in Physics, is a type of X-ray device that operates with low radiation doses [100]. The EOS imaging system's average radiation dose ranges from 0.11~0.30 mGy, reducing the radiation dosage experienced by patients by 80%~90% compared to traditional X-ray images [101].

In 2012, Somoskeöy et al.[102] utilized the three-dimensional EOS imaging system to measure three-dimensional vertebral vector parameters of the coronal and sagittal surfaces of the scoliotic spine. They compared these measurements with the Cobb angle measured by the two-dimensional EOS imaging system. They suggested that the three-dimensional vertebral vector parameters from the EOS imaging system could more effectively reflect the three-dimensional characteristics of spinal deformities and could guide the three-dimensional classification of these deformities.

Currently, some researchers are attempting to use the EOS imaging system for the three-dimensional classification of spinal deformities. They compared this method with the traditional two-dimensional classification and proposed that the three-dimensional classification could guide treatment more accurately when the Cobb angle exceeds 55°[103].

Chung et al.[104] used a spinal model to simulate 32 different degrees of curvature and measured the Cobb angle using computerized X-ray radiography, digital X-ray radiography, and EOS imaging technology. Their statistical analysis found the reliability, consistency, and compatibility of the three methods satisfactory, with the EOS imaging technology's measurements demonstrating strong predictive power. They suggested that the EOS imaging system's measurements play a crucial role in the treatment evaluation and progress monitoring of spinal deformities.

The EOS imaging system not only has a low radiation dosage but also provides full-length, standing spine images that display the morphological changes of each vertebra. This helps avoid multi-plane deformity changes that may occur in the non-weight-bearing supine position, reduces measurement errors caused by the supine position, and enables three-dimensional reconstruction of the skeletal system, including the pelvis.

Some researchers have performed three-dimensional reconstruction of the full spine in the standing position using the EOS imaging system. They compared the measurements of thoracolumbar vertebral rotation degrees and pelvic parameters among different observers and concluded that the EOS imaging system exhibits excellent reproducibility and reliability[105]. Nevertheless, the study encompassed only patients presenting with mild to moderate severity, evidenced by an average Cobb angle of 18.2°. The system may not be suitable for the three-dimensional reconstruction of severe scoliosis due to the difficulty of identifying anatomical landmarks.

Somoskeőy et al.[106] found that even for experienced professionals, it took an average of 20~30 minutes to complete the reconstruction of full-length standing spine images using the EOS imaging system. The process was even more time-consuming for patients with severe scoliosis. Moreover, compared with CT and MRI, the EOS imaging system does not provide soft tissue information. Besides, the equipment is expensive, hard to acquire, and inconvenient for population screening.

2.3.3 Computed Tomography

Owing to its ultra-high resolution and the ability to provide authentic three-dimensional skeletal deformity information, Computed Tomography (CT) has become an indispensable imaging modality in AIS surgery, encompassing the entire process from preoperative planning[107] and intraoperative navigation [108] to postoperative assessment[109, 110]. Moreover, to simulate spinal growth and the effects of brace treatment[111], finite element analysis must utilize CT data as the foundation for modeling. Additionally, due to the universality of CT devices, the current cost of chest

or lumbar CT in China is only 400 RMB, and it is generally covered by medical insurance, posing a relatively minor financial burden for patients.

However, it is undeniable that CT has unavoidable shortcomings. The first is the positioning; commonly used radiography and EOS are performed in the standing posture, whereas CT is conducted in the supine or prone posture [112]. This difference in positioning leads to certain flaws in the measurement of spinal parameters. The second is radiation; AIS patients require X-ray photography for assessment and treatment guidance. While this is a necessary component, it raises the risk of cancer in pediatric patients due to radiation exposure. Children's body tissues are more sensitive to radiation effects compared to adult tissues, and since children have a longer expected lifespan, they are more likely to manifest adverse effects from radiation. Furthermore, during the insertion of pedicle screws and the process of segmental fusion, surgeons may employ different fluoroscopic techniques to determine the implantation position of the implant, which lead to more radiation dose. According to statistics from sources such as [113, 114], the estimated total radiation dose for AIS patients over a 2-year period is below the annual safety limit of 50 millisieverts proposed by the Nuclear Regulatory Commission and the Occupational Safety and Health Administration (OSHA).

Hence, the enduring issue in the AIS field is how to reduce radiation exposure while maintaining resolution.

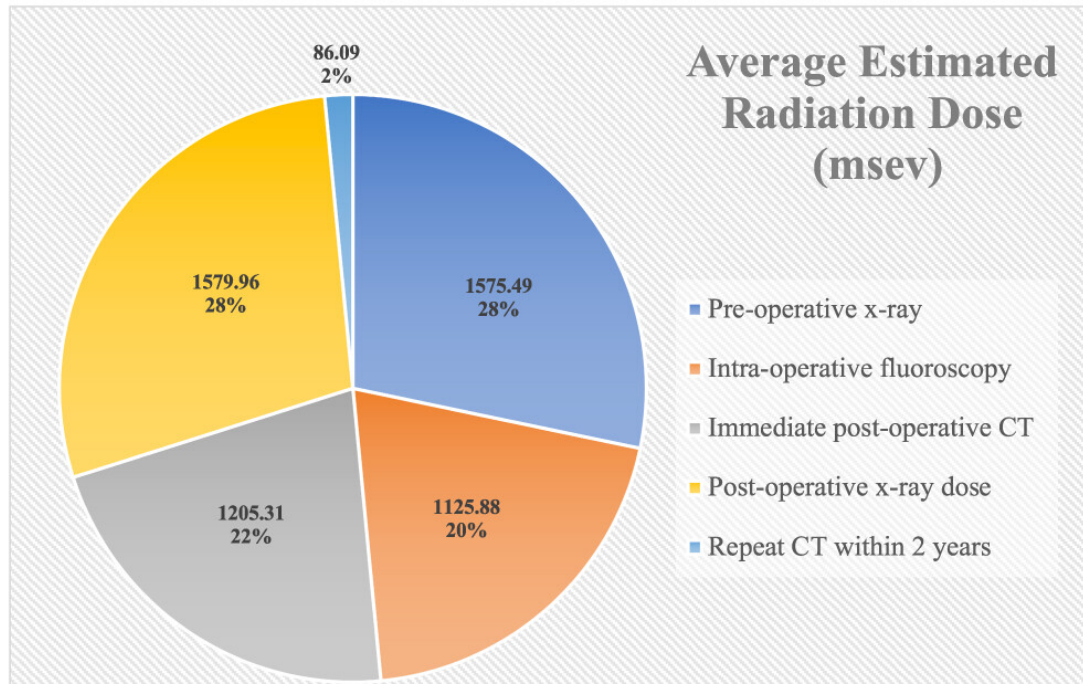


Figure 11. Estimated average radiation doses during dedicated periods in AIS management. Total average estimated cumulative dose 5572.74 mrem (55.72mSv)[113].

2.3.4 MRI

Magnetic Resonance Imaging (MRI), a non-ionizing imaging technique, has been employed to be a strong means in the research of AIS pathologies and morphological disparities.

MRI have been harnessed to probe the disparities in neurological characteristics between AIS sufferers and healthy controls. Investigations by Shi et al.[115] discerned white matter attenuation in designated cerebral areas, while Wang et al. [116] pinpointed anomalies in cortical thickness. Besides, Lee et al. [117] detected asymmetry within the brainstem. Studies focusing on the cerebellum have revealed insights such as tonsillar ectopia in 7.3% of AIS cases by Cheng et al. [118] and a 48% incidence of cerebellar tonsillar descent in upright MRI by Lee et al. [119]. Subsequent studies emphasized cerebellar volume increase [120] and protracted SEP within AIS with cerebellar ectopia [121]. Supportive research for modifications in cerebral white

matter connectivity among AIS patients has also been noted [122], along with imbalances and structural discrepancies in the AIS patients' vestibular system [123].

Considering the radiological characteristics of MRI, research has been directed towards back muscles in AIS patients. A study employing T1 MRI examined texture characteristics within muscle to differentiate between flexible and rigid scoliotic curves [124]. Similarly, another study using intravoxel incoherent motion (IVIM) MR perfusion imaging found marked asymmetry in muscle perfusion in AIS patients post-Roman chair exercise [125]. Further analysis of MRI data revealed greater asymmetry in the deep paraspinal muscles at the apex of the spinal curvature in AIS compared to healthy controls [126].

MRI has also been instrumental in assessing alterations in the vertebral body (VB) and intervertebral disc (IVD) in AIS [127]. Insights into vertebral wedging and its contribution to lateral deformations were gained, along with observations on the modest ratio of change in the lateral curve angle in comparison to VB. In painful AIS patients, MRI has proven to detect underlying pathologies more substantially (35.5%) compared to other imaging modalities [128].

MRI has also been employed in AIS morphological analysis, showing similar conus medullaris locations between AIS patients and healthy controls [129]. Additionally, a 21% incidence rate of breast asymmetry has been found in AIS patients, irrespective of chest wall deformity [130].

2.3.5 Ultrasound

Ultrasound, as the advantages of non-radiative, cost-effective, portable, and real-time imaging characteristics, has been applied in lots of research areas. Since the 1980s, ultrasound technology has been adopted for assessing spinal curvature, the coronal Cobb angle, and vertebral rotation[131, 132]. Subsequent research has confirmed the ability of three-dimensional ultrasound to guide orthopedic treatment for AIS by

examining the morphology of the posterior structures of the vertebrae [133, 134]. Additionally, its application has evolved to include the precise tracking and measurement of spinal curvature angles using ultrasound techniques[135]. One study [136] demonstrated specificity and sensitivity for detecting curve progression at 0.91 and 0.83 respectively, indicating the effectiveness of this technique. Moreover, the correlation coefficient between ultrasound and radiographic methods in a different study [137] was between 0.78 and 0.84, which is within the clinically acceptable error margin of 5°.

Building on these foundations, in 2016, Wang et al. [138] reported measurements of the rotational degrees of the apex vertebra in 16 patients. They utilized supine three-dimensional ultrasound imaging and compared it with the Aaro-Dahlborn method employed in MRI. A statistical analysis revealed high intra- and inter-group reliability with no significant difference between the two methods, exhibiting high consistency. This evidences the efficacy of employing three-dimensional ultrasound for measuring vertebral rotation in AIS patients, though it also highlights that the magnitude of the Cobb angle may affect the accuracy of this method.

The body of research confirming the reliability and efficacy of three-dimensional ultrasound for assessing spinal deformities is growing. The method has proven effective for measuring spinous angles, transverse angles, and laminar angles[139, 140]. Moreover, three-dimensional ultrasound can be used to guide orthopedic treatment by assessing the spinal curvature elasticity of AIS patients in the prone position[141], and displaying the coronal curvature change patterns during spinal flexion[142].

Building on these advances, Zheng et al. [143] introduced the Scolioscan system, which is composed of an ultrasonic scanner with a built-in linear probe, a frame, an electromagnetic tracker, and specialized software. This system operates by scanning a patient's back from the bottom to the top, while the tracker continuously monitors the probe's position. Subsequently, the acquired B-mode images, along with their spatial

data, are transformed into a three-dimensional view via the Volume Projection Imaging (VPI) technique[144]. This technique has already been commercialized.

However, these ultrasound measurement techniques still have some challenges. The requirement for complex imaging equipment and software somewhat limits their clinical application. Ultrasound imaging is also prone to interference from internal metallic implants, body shape, and the degree of spinal deformity. Additionally, ultrasound can only scan from the back in an order from C7 to S1 or from L5 to C7. This heavy reliance on the posterior structures of the vertebrae could overlook the deformity features of the vertebrae. Finally, while ultrasound measurements include both supine and standing positions, the standing position, although it can reduce measurement errors, is less stable than the supine position. This necessitates a careful consideration of trade-offs when selecting between them.

2.3.6 Summary

Traditional imaging techniques for measuring spinal deformity, despite the radiation exposure, stand as the prevalent and standardized clinical methodologies. Emerging computer and AI-assisted approaches promise to augment the precision and efficiency of these measurements but remain nascent. Technologies such as the EOS imaging system, three-dimensional ultrasound, and MRI mitigate radiation risks, facilitating three-dimensional scoliotic deformity evaluation. However, their clinical application is limited by some challenges, including equipment accessibility, complex imaging systems, and the necessity for specialized technical operation. The future trajectory of measurement tool development should emphasize radiation safety, precision, reliability, technological simplicity, robust operability, cost-effectiveness, and adaptability for widespread screening and evaluation. The integration and synergistic application of different measurement methods, AI, and novel imaging technologies are likely to be instrumental in fulfilling these goals.

2.4 Treatment of AIS

The therapeutic strategy for AIS is designed to correct, arrest, or minimize further distortion of the spine, and to counteract or minimize the effects of the condition. Individuals with a AIS diagnosis and curvature measurements below 7 degrees are generally placed under clinical observation, with check-ups every six months. If the curvature extends beyond 7 degrees, the use of radiography is usually advised. For those with a Risser grade from 0 to 2 and curves less than 25 degrees, observation is scheduled every three to six months. This continues until there is an increase of more than 5 degrees between two consecutive follow-up or the curvature exceeds 25 degrees. Treatments are needed in this condition. Curvatures within the range of 25 to 40 degrees often lead to the usage of a brace, whereas curves exceeding 45 or 50 degrees typically require surgical intervention. During the phase when patients are nearing skeletal maturity (marked as Risser 3 or above), those with curves less than 25 degrees are monitored every six to nine months. After reaching skeletal maturity, patients should continue to be observed for at least one year (this corresponds to Risser 4 in girls and Risser 5 in boys). Following this period, patients with curves under 30 degrees can often be discharged from ongoing care, as these curves are unlikely to progress further. On the other hand, those with curves greater than 50 degrees may continue to change into adulthood, thus ongoing observation remains essential[145, 146].

2.4.1 Non-surgical treatments

Bracing therapy is the most prevalent non-surgical treatment method. Within continuous follow-up care, bracing is typically recommended for patients who are skeletally immature, with curvature degrees ranging between 25 to 40 degrees, or when the curvature progression exceeds 5 degrees. The primary objective of bracing is to prevent further development of the curvature rather than to correct it. Bracing is unsuitable for patients who are skeletally mature or those presenting a Cobb angle exceeding 50 degrees. Patients undergoing brace treatment should be monitored approximately every six months. Bracing continues until the end of the growth spurt (roughly 2 or 4 years after the onset of menstruation in girls, and approximately 5 years

in boys). When the curvature does not surpass 5 degrees, the bracing is considered successful[146].

The Milwaukee brace[147], created in the 1940s, is a type of back brace designed to help fix problems in the middle and lower parts of the spine. It has a base that fits around the hips made of a strong plastic. The brace extends upwards with a front piece and two back pieces that connect to a neck ring at the top. This neck ring can have a special mold for the front of the neck and cushions for the back of the head, or a simpler, more comfortable neck ring. It also has special pads that can be adjusted to help correct the spine's shape.

The Boston brace[148], developed in the 1970s at Harvard University, is a type of back brace designed like the TLSO brace. It's made from a tough plastic material for the hip area and has a softer plastic layer inside for comfort. This brace is shaped to fit the curve of the lower back. It can be used for most kinds of scoliosis (a condition where the spine curves sideways), but if the curve is in the upper back and is very pronounced, it's recommended to use the Milwaukee brace.

The Dynamic Spine-Cor brace[149], developed between 1992 and 1993, uses a unique Adjustment Motion based on the type of spine curve. This specific movement is designed to work well, with the brace set up following guidelines from the Spine-Support Helper software. For the best results and to help with muscle and nerve coordination, it's crucial that the brace supports and increases this adjustment movement over time.

The Providence brace[150] came into use after discovering that precise, strong forces from a stiff mold could effectively straighten curves in the spine. This brace works well for treating both single and double (bifurcated) spine curvatures. Originally, the mold aimed to help doctors understand how flexible a patient's spine was while lying down. The brace measured forces directly and from different angles to gently guide the spine

towards or even beyond its central position. To do this, a cast of the patient's torso is made on the mold, targeting the spine with these corrective forces.

Type of Brace	Number of patients	average initial curve	Conclusion
Milwaukee	1020	30-35	Immediate bracing for greater than 25° and Risser 0 prevents curve progression
Boston	51	36-45	61% did not progress greater than 5° from onset of brace use to discontinuation
Dynamic Spine-Cor brace	249	24-40	Spine-Cor brace able to prevent progression of curve in 60% of patients
Providence	102	27	61% to 79% success rate; the Providence brace is effective in curves less than 35°

Table 3. Bracing studies using different brace types[151]

The most challenge in orthotic therapy is compliance, particularly due to the stipulated duration of wear[152, 153]. However, in summers or for patients located in subtropical and tropical climates, the humid weather is not conducive to wearing orthoses. Moreover, long-term wear often leads to numerous skin-related issues. Furthermore, wearing brace can make adolescents feel different from their peers, thus intensifying their reluctance towards its usage. All these factors contribute to a reduced wearing duration and suboptimal orthotic outcomes.

To address this issue, previous studies have adopted various strategies to make these inherently "anti-humanity" products more acceptable and monitorable. Some researchers have incorporated temperature[154] or pressure[155] sensors into the brace to track the duration of wear. Besides, engaging patients during therapy, which includes collaborating on individualized designs and emotional support strategies, is essential to enhance brace comfort and improve compliance[156]. Others have employed 3D printing techniques[157, 158] to customize orthoses for patients. However, based on clinical feedback, 3D printing is more costly (compared to low-temperature thermoplastics) due to material limitations. Furthermore, the fatigue resistance of these materials is relatively low. As brace requires daily donning and doffing, current 3D printing materials tend to fracture under frequent flexing.



Figure 12. 3D printed Boston Brace[159]

2.4.2 Surgical treatments

Surgical intervention serves as a commendable treatment approach for severe AIS patients, addressing various issues stemming from spinal deformities. In the current domain, empowered by a myriad of techniques, surgeons ardently work towards minimizing post-operative complications and enhancing efficacy. However, such complications are still inevitable. Given the relatively younger age of AIS patients at the time of surgery, it becomes imperative to judiciously evaluate the potential surgical complications against the persistent impact posed by the ailment itself.

2.5 Image analysis based on deep learning

2.5.1 Radiography

X-ray imaging constitutes the most accessible image modality for AIS diagnosis and follow-up, lots of Artificial intelligence (AI) methodologies have been applied on this imaging technique. The topic of automated Cobb angle measurement and vertebra detection are the most popular aspects based on anterior-posterior x-ray images.

Convolutional neural network (CNNs)[160] were utilized to gauge the Cobb angle from anterior-posterior radiographs, enhance the precision of vertebrae detection within localized X-ray images through the use of CNNs, and to discern specific vertebrae (apex, superior, and inferior end) for the measurement of the Cobb Angle (CA). The accuracy rate was 93.6%, substantiated by an intraclass correlation coefficient exceeding 0.95. In a related study[161], CNNs was formulated specifically for the detection of vertebrae within the spine, the measurement of the Cobb angle, and the classification of curvature severity within X-ray images, achieving an accuracy rate of 0.9 (90%). Unlike other research, the study referenced in[162] utilized CNNs not only to measure the Cobb angle but also expanded its analysis to evaluate different curves in the spine, specifically the proximal thoracic (PT), main thoracic (MT), and thoracolumbar/lumbar (TL/L) curves. However, the method showed limitations in analyzing data from patients with severe conditions, mainly because it relied heavily on the precise vertices of the vertebral boundaries for its calculations. The previously published deep learning models for Cobb angle measurement have rarely applied to severe AIS patient data[163]. In contrast, the research cited in[164] introduced an innovative method using deep learning that enables a more efficient and automatic measurement of the Cobb angle across a arrange of AIS severity, from mild to severe cases.

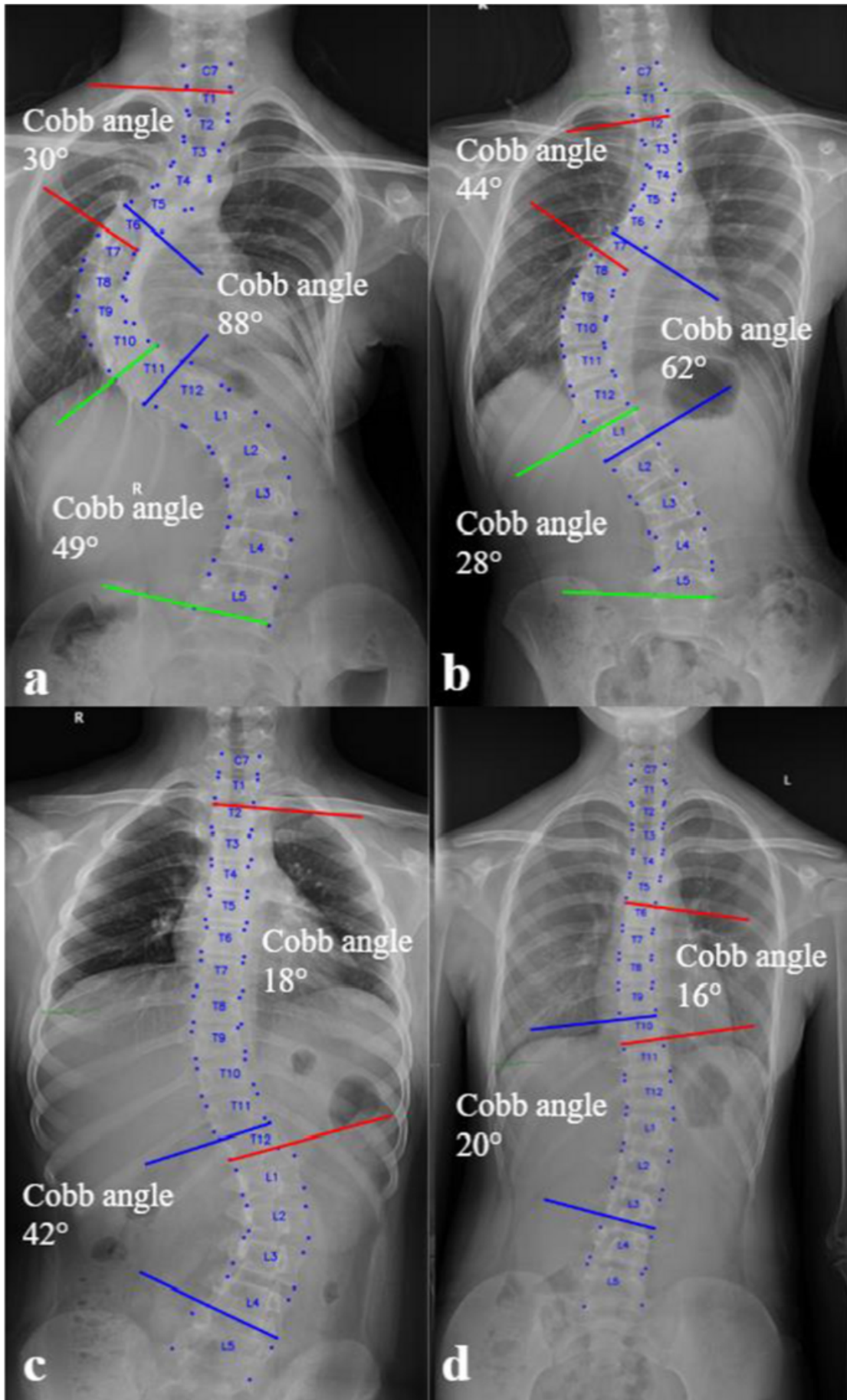


Figure 13. Result from the deep learning method, the most pronounced curve was characterized by a Cobb angle of 88° for the principal curve, with corresponding upper and lower curves measured at 30° and 49° , respectively[164].

Convolutional Neural Networks (CNNs) may exhibit susceptibility to variations in image textures, a phenomenon documented in[165]. This particular sensitivity may present limitations when applied to the diagnosis of complex spinal structure disorders such as scoliosis, where a consistent analysis of the global spinal structure is essential. In an effort to overcome these limitations, a novel methodology was established, outlined in[166]. This approach introduces a specialized deep-learning architecture known as MSE-Net, designed for the automatic evaluation and diagnosis of scoliosis through the comprehensive analysis of full spinal X-ray images. The MSE-Net methodology incorporates functionalities for precise anchor point detection, landmark identification, and a series of post-processing techniques to enhance the diagnostic accuracy.

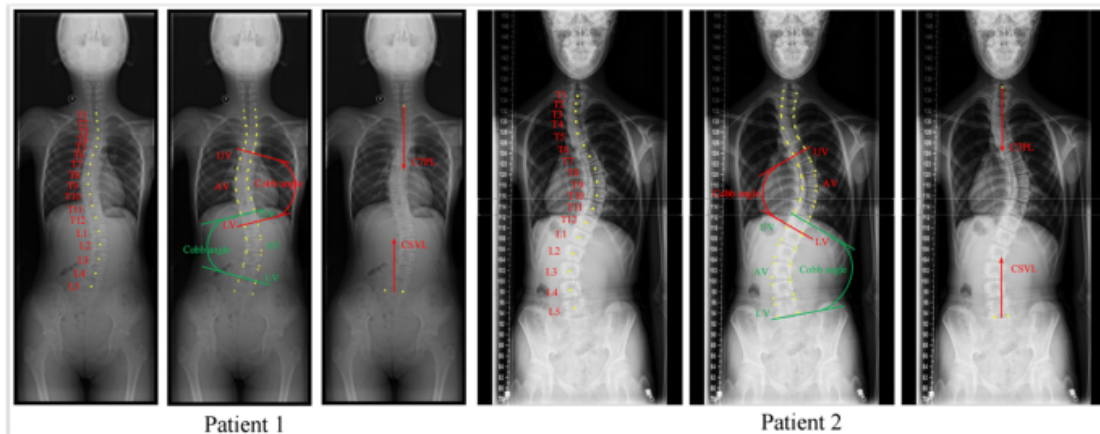


Figure 14. The predicted outcomes for two subjects from the testing set encompass significant data points such as UV (upper vertebrae), AV (apical vertebrae), and LV (lower vertebrae), along with CSVL and C7PL[166]

Additionally, deep learning could also be used in prediction in AIS. This methodology [167] has culminated in the creation of a predictive model for assessing the risk of scoliosis progression in patients with AIS, utilizing a deep convolutional neural

network (DCNN). The integration of transfer learning, through the employment of a pretrained DCNN, has proven pivotal in augmenting the prediction accuracy of this model, achieving a statistical accuracy rate of 69%.

Another study developed a comprehensive machine learning model to identify AIS curves at risk of progression by analyzing a large patient cohort[168]. The model, which includes patient demographics, vertebral morphology, and bone quality, demonstrated high accuracy, achieving an accuracy of 83.2%, sensitivity of 80.9%, specificity of 83.6%, and an AUC of 0.84. A different research effort focused on predicting AIS progression through a two-stage transfer learning framework, showing notable accuracy with an accuracy of 82.1%, sensitivity of 80.8%, and specificity of 88.6%. The most significant morphological indicator being the apical vertebral rotation of the major curve[169].

Besides, research[170] introduced a machine learning approach utilizing 3-D spine models from biplanar X-ray images, aiming for precise progression prediction through a novel embedding technique that optimizes class separation between progressing and non-progressing scoliosis. The technique demonstrated encouraging outcomes, achieving an 81% success rate in classifying progressive versus non-progressive patients and accurately predicting differences in main curve angulation to within 2.1° .

2.5.2 EOS

EOS can obtain X-ray images of the whole spine through the low radiation dose, and the supporting software can generate three-dimensional models of the whole spine. These advantages lead to the gradual increase of EOS-based image analysis, and AI-based research hotspots mainly focus on feature extraction and disease classification.

A stacked auto-encoder analysis technique[171] was used for the classification of spinal deformities from tailored 3D models of the spinal geometry. This method crystallized the 3D spinal models' complexity while preserving traditional attributes, with a

balanced dataset from EOS and conventional radiography, which is a good for scoliosis surgical planning.

Research[172] involved the employment of Convolutional Neural Networks (CNNs) to orchestrate an automated quantification of axial vertebral rotation (AVR) in AIS patients, capitalizing on the Stokes method in alignment with EOS images. The aggregate accuracy of these automated assessments culminated at 81%, juxtaposed with manual techniques. However, this methodology was singularly applied to cases with milder conditions, demarcated by a Cobb angle not exceeding 45°.

In Study[173], CNNs were harnessed to conduct an examination of 1830 patient cases via EOS imaging or radiographic techniques to deduce the Risser stage. With an attained accuracy of 78% in juxtaposition with evaluations conducted by medical professionals. Likewise, study[174] introduces a machine-learning technique for Risser sign skeletal maturity evaluation employing EOS radiographs, wherein 24 image features are harvested from designated regions of interest via a ResNet101-type CNN pre-trained from the ImageNet database. A support vector machine (SVM) algorithm was invoked for the ultimate classification, achieving accuracies of 84%, 78%, and 80% for iliac crests, humeral heads, and femoral heads, respectively.

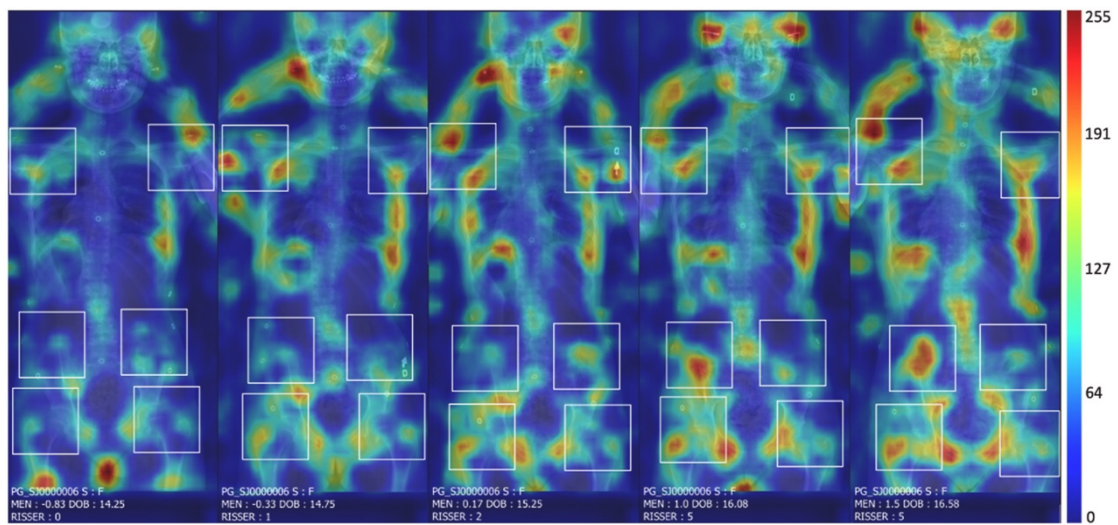


Figure 15. Activation maps for one patient samples, wherein regions of interest, specifically the iliac crests, humeral heads, and femoral heads, exhibit pronounced activation. In addition to these primary regions, the activation maps also reveal enhanced activation in the cranial bones and rib cage[174].

In [175], an innovative model, designated as Spine TK, was used to autonomously generate the Cobb angle from anterior-posterior EOS radiographs. The methodology achieved a mean absolute angle measurement discrepancy of 1.16° . Finally, the research[176] introduces a bio-informed mechanistic deep learning model designed to predict the spinal morphology. By integrating clinical data obtained from EOS imaging with mechanistic attributes like the stress distribution on the growing surface of the vertebrae, extracted from a surrogate spinal model and coinciding with the bone growth model, a comprehensive analysis is formulated.

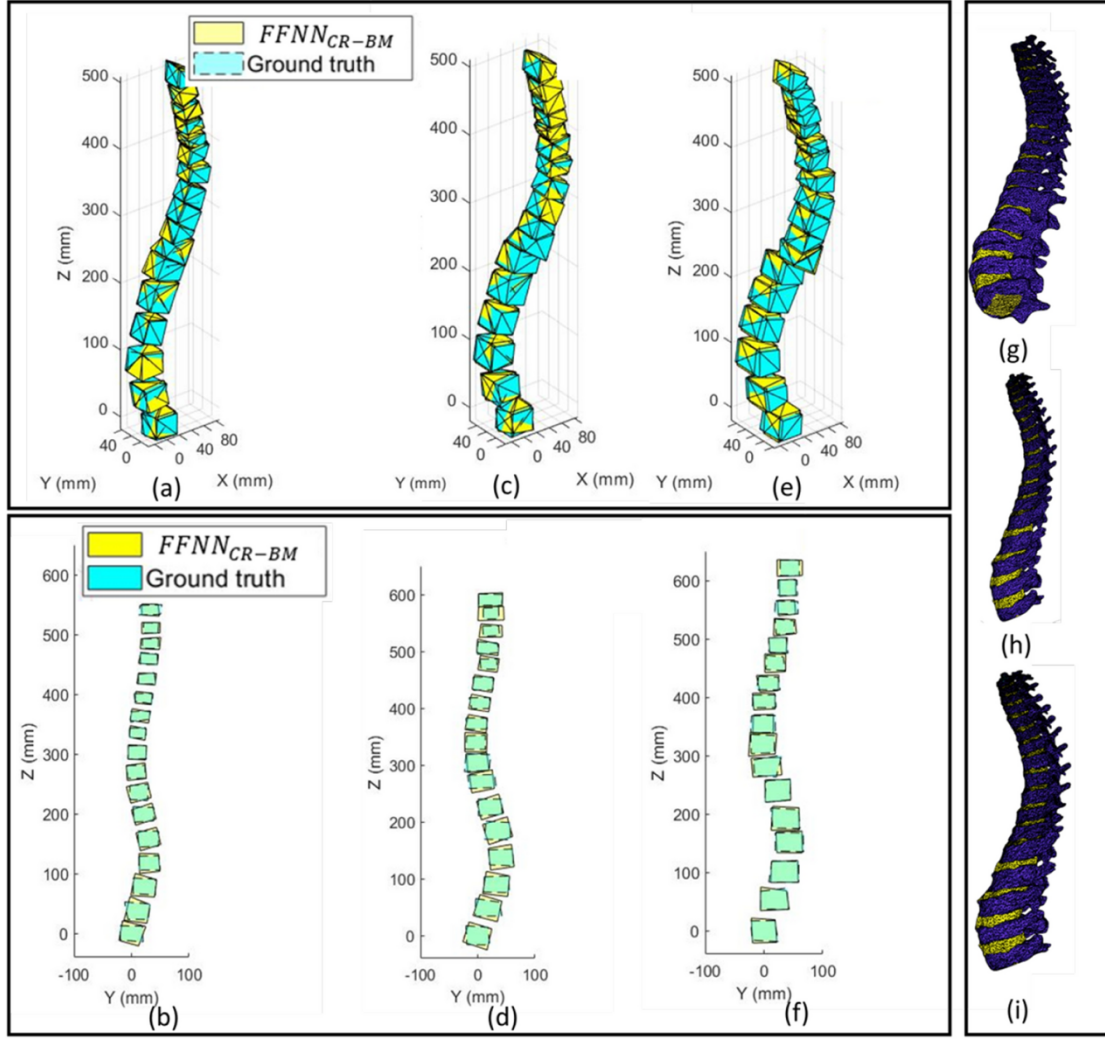


Figure 16. Bio-informed Mechanistic predictions (FFNNCR-BM) with ground truth from EOS at different age stages. Comparisons are made at 160 months (within the trained data range), 179 months, and 187 months (both outside the trained data range), using various 2D and 3D views, including detailed geometric reconstructions[176].

2.5.3 CT

Computed Tomography (CT) imaging, due to the ability in the three-dimensional architecture of the skeletal system, has become an important imaging modality in the assessment of AIS, particularly in surgical interventions. Both preoperative and postoperative CT evaluations are necessary components of the therapeutic assessment. The AI-driven image analysis based on CT scans is mainly focus on the segmentation

of the spinal structure. The objective is to develop an expeditious, authentic, and universally applicable method of fully automatic segmentation. This may subsequently serve as the foundation for more intricate analyses, such as morphological measurement or finite element analysis.

This study[41] employed a three-step approach using U-net for vertebral body (VB) and intervertebral disk (IVD) segmentation. The first generated a coarse probability map of spine location; the second detected VB centers; the third segmented the vertebral region for each VB center. The synthesized segmented areas resulted in a multi-label spine mask, with a DCS of 0.90 in a 160 non-scoliosis CT dataset.

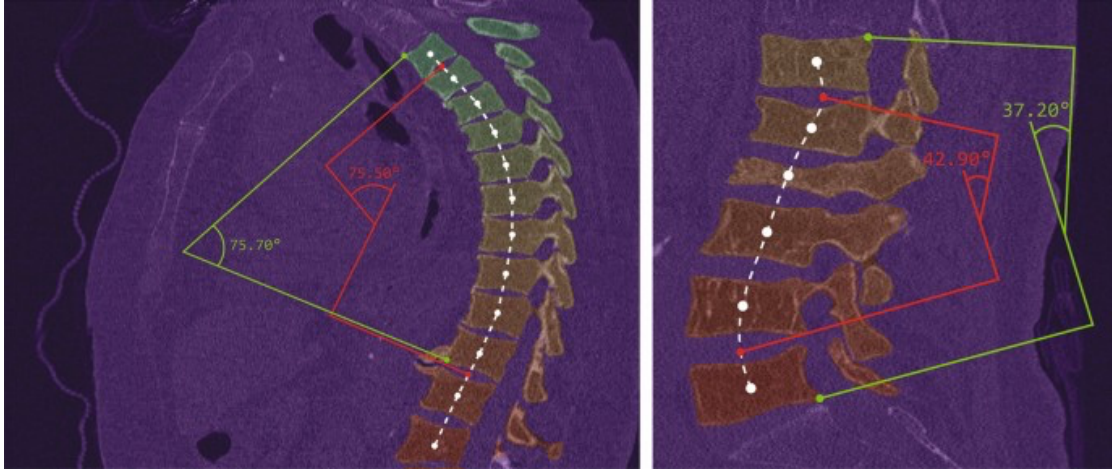


Figure 17. Example of the sagittal Cobb angle thoracic kyphosis (left) and lumbar lordosis (right) measurement from the method[41].

Nikolas et al. [177] implemented a convolutional neural network (CNN) architecture to segment the thoracic and lumbar vertebrae. Using an iterative instance-by-instance method across CT and MRI modalities, they recorded segmentation accuracies of $94.9 \pm 2.1\%$ for CT datasets and $94.4 \pm 3.3\%$ for MR datasets.

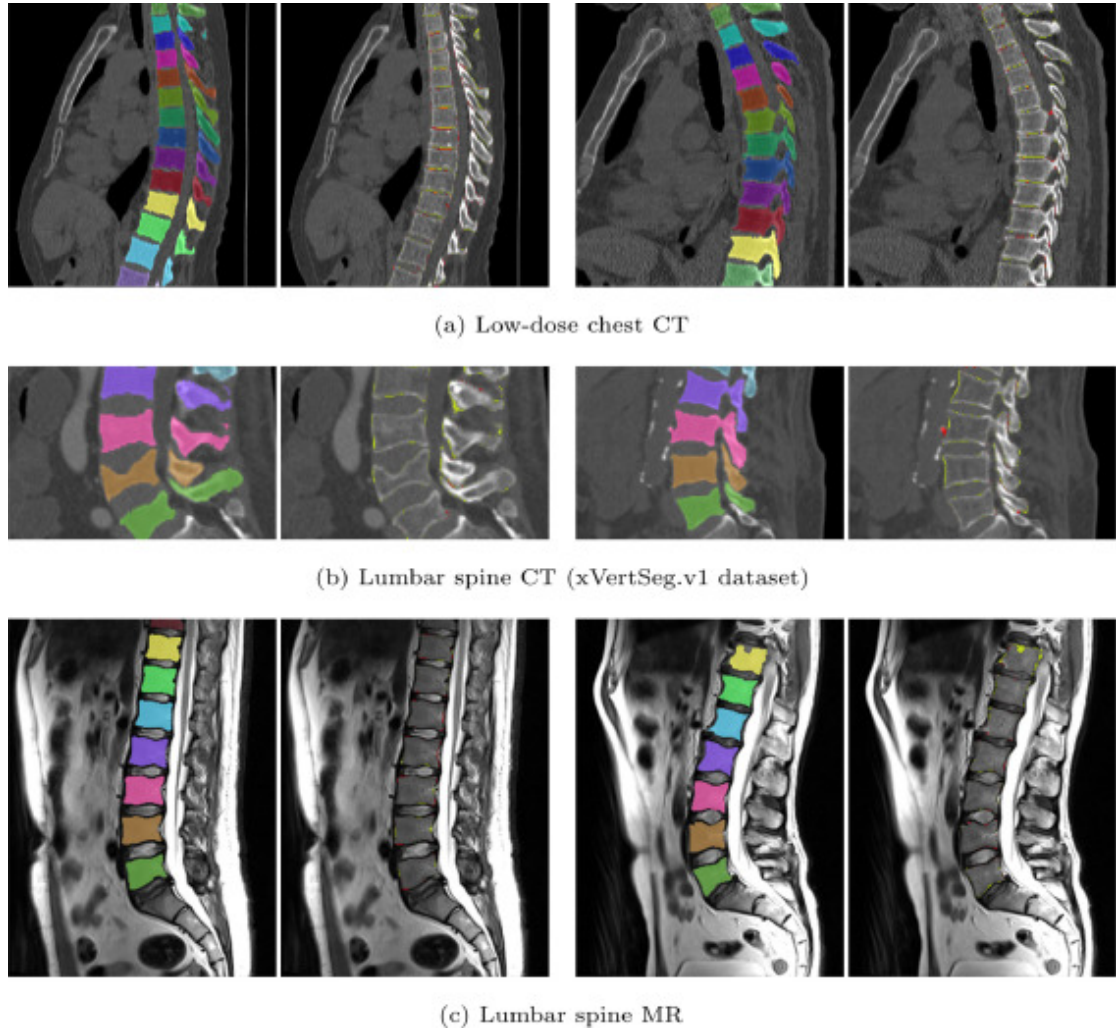


Figure 18. The magnitude of differences of the automatic segmentations from the ground truth segmentations

A patch-based deep belief networks (PaDBN) model[178] applied an automatic feature selection process, combining unsupervised feature reduction and supervised fine-tuning. The model utilized CT image-derived regions of interest (ROIs) and a contrastive divergence algorithm for weight optimization, achieving an accuracy of 93.3%, sensitivity of 91.1%, and specificity of 93.4%.

An innovative method referred to as "FU-Net"[179] combined traditional region-based level set with deep learning to accurately predict vertebral bone shapes, addressing the complexities of fractured cases. The method was evaluated on two distinct CT datasets,

achieving Dice scores of $96.4 \pm 0.8\%$ and $92.8 \pm 1.9\%$ in the CSI 2014 dataset, and $95.2 \pm 1.9\%$ and $95.4 \pm 2.1\%$ in the CSI 2016 datasets.

A 2D U-Net model[180] was evaluated on 41 cervical vertebrae CT scans. This study comprised 24 scans from healthy controls for training and 17 from diseased subjects for validation. The method achieved mean Dice similarity coefficient (DSC) values ranging from 88.67% to 96.23%, although the performance was found to be suboptimal with greater slice thickness.

In a similar research[181], a multi-faceted approach was employed for cervical vertebra segmentation. This technique integrated PointNet++ on a much larger dataset, utilized an adaptive threshold filter to isolate cervical vertebra tissue from CT images, and applied PointNet++ to segment individual vertebrae, followed by a convergence segmentation technique to differentiate adjacent vertebrae edges, thereby enhancing the overall segmentation accuracy.

A combination of CNN, K-means Clustering, and k-NN was deployed in[182] to offer the advantage of eliminating the need for individual vertebra-level annotations during training. This approach was validated on 12 CT scans, representing varied scoliosis severity, with multi-class DSCs recorded at $49.79 \pm 24.90\%$, $77.76 \pm 15.05\%$, and $83.48 \pm 12.56\%$ for severe, moderate, and mild scoliosis patients, respectively.

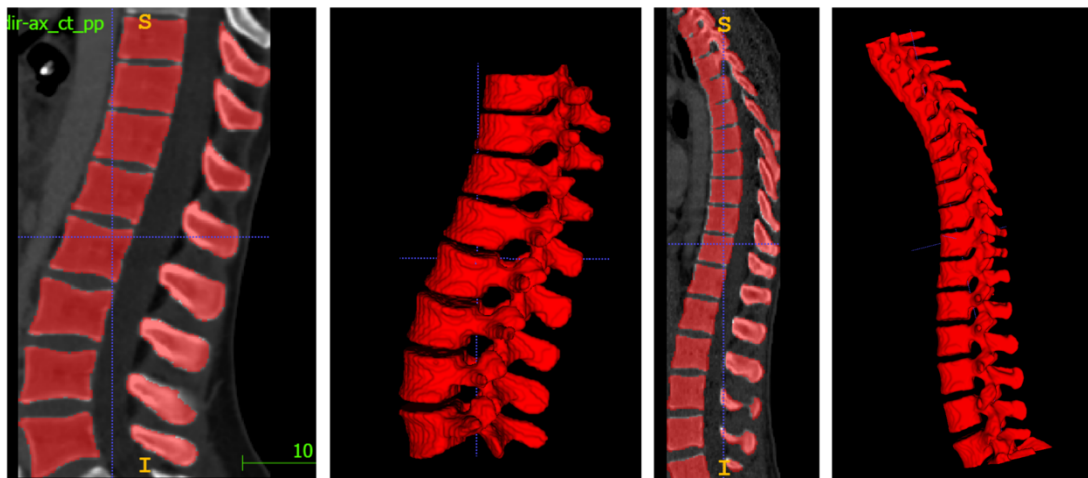


Figure 19. Binary spine segmentation

A recent study[183] introduced a novel patch-based deep learning methodology. This method leverages a stacked sparse autoencoder (SSAE) to extract discriminative features from unlabeled data. Validated across three publicly available datasets, 2D image slices were utilized, and overlapped patches were extracted for model training. The proposed model captured high-level feature representations of pixel intensity, and a sigmoid layer was used to effectively differentiate between vertebrae and nonvertebrae patches.

In summary, AI-based vertebra segmentation using CT has achieved a series of advancements at present. However, many existing methods are trained on non-AIS disease data or healthy individuals, with limited AIS patient data. The performance are not good within moderate to severe spinal scoliosis images. Among AIS patients, the complex topological structures of severe vertebra deformities and fusion still arise challenge in segmentation.

2.5.4 MRI

MRI is also playing an increasingly important role in AIS disease management due to its non-radiation characteristics. Although its advantage lies in soft tissue imaging, bone tissue segmentation based on it has also been involved in more and more studies.

A approach [184] achieved optimal vertebra segmentation in polynomial time through task-specific CNNs, graph cut formulation, and star-convexity constraints. This method reached a Dice coefficient of $93.8 \pm 2.6\%$. In contrast, an unsupervised deep learning pipeline [185] combined rule-based methods with voting mechanisms for vertebral segmentation. A method employing transfer learning [186] accomplished an accuracy of 98.6% for lumbar spine localization. A sophisticated architecture was introduced in study[187] that seamlessly integrated manifold regularization techniques, demonstrating robust performance in lumbar spinal indices estimation.

A fully convolutional network was developed[188] that localized and segmented intervertebral discs from multi-modality 3D MR images. Similarly, machine learning algorithms [189] for the segmentation of lumbar spinal canal areas and intervertebral discs, with results comparable to the ground truth from human experts. Novel networks such as DMML-Net [190] and CNNs [191] were employed for target organ salience and segmentation of lumbar neural foraminal stenosis. Similar deep learning methodologies [192] yielded accuracies of 86.2% for central stenosis detection and 85.2% for disc herniation detection.

Expanding beyond these applications, AI-based methodologies have extended to spinal ailment treatment. Research into distinguishing between tuberculous and pyogenic spondylitis [193] achieved accuracy nearly consistent with radiology specialists. Artificial neural network models were utilized to predict cervical spondylotic myelopathy (CSM) diagnosis and severity [194]. An image-analysis pipeline developed for spinal cord and lesion segmentation [195] achieved a Dice coefficient of 0.93. Studies have also demonstrated the potential of CNNs in differentiating between spinal tumors such as schwannoma and meningioma [196].

Furthermore, the feasibility of using supervised machine learning for predicting radiographic progression in axial spondyloarthritis (axSpA) has been investigated [197]. High-performing models, including GLM and SVM, achieved an average ROC AUC exceeding 0.78 and maintained a balanced accuracy greater than 65%. Another study [198] utilized machine learning to cluster axSpA patients, revealing distinct clinical characteristics and further emphasizing the burgeoning potential of AI methodologies in the field of spinal imaging and diagnosis.

2.5.5 Ultrasound

Compared to other imaging modalities, studies focusing on AI analysis based on ultrasound are less prevalent. However, considering the long-term follow-up characteristics required for AIS management, the adoption of non-radiative ultrasound

technology is likely to increase. Consequently, corresponding image analysis techniques will also become more extensive.

Huang, Z. et al. [199] proposed an innovative RSN-U-net approach for efficient ultrasound spinal image segmentation, showcasing resilience against speckle and consistent occlusion disturbances. The study utilized the Total Variance (TV) loss technique to enhance the neural network's training, improving its resistance to noise, resulting in consistent and superior segmentation outcomes.

The Advanced Segmentation Network (ASN) [200] effectively delineates vertebral features in ultrasonographic images amidst interference. By blending a modified U-Net design with efficient convolution and feature-recognizing gates, and linking encoder-decoder units via multi-dimensional pathways, ASN has shown superior segmentation efficacy against comparable systems like U-Net.

Researchers have introduced a novel AI-integrated model for spinal ultrasound image segmentation, termed UGBNet [201]. This model amalgamates spatial and channel-specific insights through a specialized global guidance module, ensuring comprehensive capture of feature interdependencies and scale nuances. Its performance, benchmarked against established segmentation frameworks like UNet, indicated a significant improvement with a Dice score metric of 74.2%.

2.6. Summaries

This chapter has reviewed the recent advancements in the application of AI in spine imaging]. Improvements in modern computer power, along with the availability of various imaging techniques, all contributes to the disease managment. As is the case with many medical disciplines, the clinical applicability of AI in AIS research remains nascent. With a growing public cognizance and focus on this disease, it is anticipated that an increased number of patients will be identified at earlier stages of the disease in the future. Consequently, how AI can augment disease screening, longitudinal follow-

up, and enhance treatment efficacy constitutes a direction that necessitates more nuanced and in-depth research in the future.

CHAPTER 3 AUTOMATICALLY VERTEBRA SEGMENTATION BASED ON U-NET

3.1 Introduction

Ct-based spinal segmentation and 3D reconstruction play an important role in the diagnosis, follow-up and treatment of scoliosis. The previous segmentation methods were mostly carried out manually or semi-automatically (like level set[202] and atlas-based[203]), but they were inefficient and less robust. With the development of deep learning technology, there are more and more applications in vertebral segmentation.

Deep learning models are available for different image modalities such as MRI, ultrasound, OCT and CT. U-net, an optimized model of CNN, is widely used in medical image segmentation because of its simple model, fewer parameters and better effect.

The U-Net framework identified within deep learning applications through its distinctive "encoder-decoder configuration," drawing upon comprehensive convolutional networks. The U-Net [41] [179] [180] architecture has emerged as a prominent tool for vertebral segmentation, offering high precision and versatility across diverse scenarios, including traumatic injuries and varying degrees of scoliosis. Other innovative contributions include patch-based deep belief networks[178], enhancements through PointNet++ integration[181], novel method by converting the 3D labeling problem into a 2D task[204], and hybrid approaches combining CNNs with clustering techniques[182]. Nevertheless, challenges persist, particularly in AIS datasets. The complex three-dimensional spinal distortion and apparent vertebral abnormalities in AIS cases pose significant obstacles. Achieving automatic spine segmentation in such a complex topology remains a daunting task. In this chapter, U-net model will be used to achieve the segmentation of the spine.

3.2 Methodology

CNN

Within the realm of image analysis using deep learning, Convolutional Neural Networks (CNNs) [205] have emerged as a leading approach. These networks are characterized by their multi-layered, stacked architecture. Each layer in a CNN can be regarded as a computational module, and by cascading these modules, a complete convolutional neural network is formed. This approach transforms different tasks into mathematical problems, meaning that the CNN has the ability in finding the appropriate mapping function from input to output. Specifically for medical image segmentation tasks, the objective of a CNN is to learn the nonlinear relationships between medical images and their segmentation outcomes, to identify different features within medical images. CNNs are capable of automatically extracting features from large datasets and are characterized by local connectivity and shared weights. Local connectivity refers to each hidden unit mapping to a subregion of the input image, known as the receptive field. Weight sharing means that one set of weights is shared across a subregion. In contrast, dense networks require assigning a unique weight to each unit, each connected to every neuron in the subsequent layer, which can be summarized as forming a fully connected network. Due to the operation of linear activation functions in hidden layers, this results in a high number of parameters and computational costs. Therefore, compared to dense networks, CNNs reduce the demand for memory and parameters, enhancing the network efficiency.

A convolutional neural network typically consists of the following five layers: the input layer, the convolutional (CONV) layer, the Rectified Linear Unit (ReLU) layer, the pooling layer, and the fully connected (FC) layer.

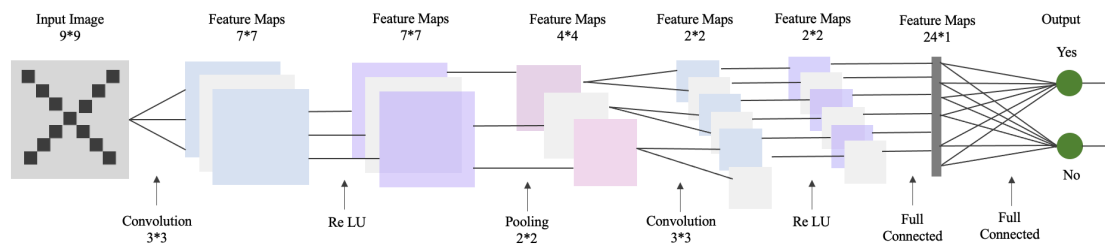


Figure 20. The network of CNN

Input Layer

The primary task of this layer is to preprocess the raw image data, which includes:

Mean Subtraction: This involves centering each dimension of the input data around zero. The purpose is to bring the center of the samples back to the origin of the coordinate system.

Normalization: This process scales the magnitude of the data to the same range. It reduces the interference caused by differences in the range of values across dimensions. For example, if there are two features, A ranging from 0 to 10 and B from 0 to 1,000, couldn't use these features directly. A good practice is to normalize them, scaling both A and B to a range from 0 to 1.

PCA/Whitening: PCA is used for dimensionality reduction; whitening normalizes the amplitude of the data along each feature axis.

Convolutional Layer

This crucial layer forms the core of the CNN, giving the "Convolutional Neural Network" its name. Its main purpose is to extract features from the input data, achieved by the convolutional kernels. The convolutional layer may include one or several kernels. Essentially, a kernel is a scanner of a specific window size that continually scans the input data for feature extraction. Typically, lower-level convolutional layers learn basic local features, such as textures and colors of the image; higher-level layers

can capture more abstract information, such as the positional information of the image. In summary, lower-level feature maps contain more detailed information, while higher-level feature maps incorporate more global information. Figure shows the process of convolutional computation.

The parameters representing the output image, subscript 'filter' for the convolutional kernel parameters, 'S' for stride, and 'P' for padding around the image edges.

Stride controls how the filter convolves around the input content. The filter moves over the input content by shifting one unit at a time, where the distance moved is the stride. Stride settings typically ensure that the output dimensions are integers, not fractions.

In the early layers of the network, aims to retain as much information from the original input as possible, thus extracting low-level features. For instance, if the same convolutional layer want to apply but also maintain the output dimensions as 32x32x3, zero padding of size 2 can apply. Zero padding adds zeros around the boundaries of the input content. Using two zeros for padding would result in a 36x36x3 input matrix.

$$O = \frac{W-K+2P}{S} + 1 \quad (3.1)$$

where O is the output size, W is the input size, K is the filter size, P is the padding, and S is the stride.

Input Volume						Filter			Output Volume							
0	0	0	0	0	0		1	0	1	5	1	2	-1			
0	1	2	-1	0	0		0	0	2	4	-1	4	-1			
0	2	1	-1	1	0		1	0	1	0	3	2	0			
0	1	0	0	0	0	$S = 1$							-2	3	2	0
0	1	-1	1	1	0											
0	0	0	0	0	0											

Figure 21. The process of convolutional computation

In the convolutional layer, each neuron is connected to a data window with fixed weights, focusing on a single feature. Neurons act as filters in image processing, such as the Sobel filter used for edge detection. Each filter in the convolutional layer focuses on a specific image feature, such as vertical edges, horizontal edges, colors, textures, etc. Collectively, these neurons function as a comprehensive feature extractor for the entire image.

Activation Layer

Since the features extracted by CNNs are linear, they cannot address nonlinear problems. Activation functions enable nonlinear transformations, thereby enhancing the model's expressive capacity. Early CNNs predominantly used sigmoid-shaped activation functions (e.g., Sigmoid and Tanh), which map input values to a specific range; Sigmoid maps to (0,1) and Tanh maps to (-1,1).

$$f(x) = \frac{1}{1+e^{-x}} \quad (3.2)$$

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (3.2)$$

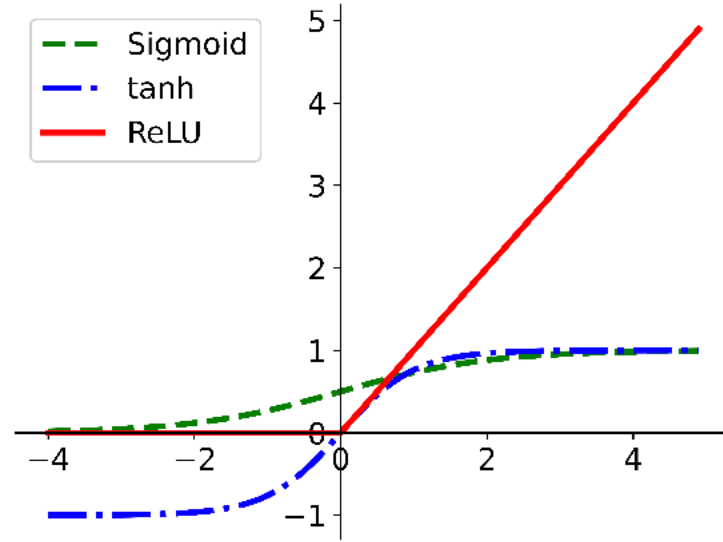


Figure 22. Three active functions

The images of the above three activation functions are shown in the figure. The issue with sigmoid-shaped activation functions in deeper networks is that their derivatives become very small, hindering the network's ability to update and converge the weights in the first layer.

The Rectified Linear Unit (ReLU)[206] activation function sets all negative values to zero and is mathematically defined as:

$$f(x) = \max(0, x) \quad (3.3)$$

where x is the input. ReLU operates with computational simplicity and efficiency, enabling rapid convergence. Unlike Sigmoid and other functions that involve exponential calculations, derivatives, and division during activation and error gradient backpropagation, ReLU has no saturation region and therefore does not suffer from the vanishing gradient problem. When Sigmoid approaches saturation, its transformation slows and its derivative tends toward zero, causing the gradient to vanish and affecting the training of deep networks. ReLU induces sparsity in the network by setting some

neuron outputs to zero, reducing interdependency among parameters and avoiding the overfitting problem. Due to these reasons, ReLU is widely used in CNNs.

Pooling Layer

The pooling layer is a method for extracting core features from the input data, typically the feature maps generated by the convolutional layer following convolution operations. Positioned between successive convolutional layers, it not only compresses the original data, reducing its size, but also decreases the computational parameters of the network model, effectively lowering the risk of overfitting. The main characteristics of the pooling layer include:

Feature Dimension Reduction: Pooling operations reduce the size of the feature maps. While an image contains abundant information with numerous features, some are useless for specific image tasks. Pooling can remove such redundant information and extract the most important features, thereby reducing computational complexity.

Feature Extraction: Pooling operations extract the most significant features from local areas. These operations help in extracting features like texture, shape, and edges from the image, providing useful information for subsequent classification and segmentation tasks.

Translation Invariance: Pooling provides translation invariance, meaning minor shifts or translations in input features do not significantly impact the output of the pooling operation, enhancing the model's generalization ability.

Control Overfitting: Pooling can also control overfitting to some extent. By reducing the size of the feature maps, pooling decreases the number of parameters in the model and reduces its complexity, helping to prevent the model from overfitting the training data.

The most commonly used pooling methods are max pooling and average pooling. Max pooling preserves the most significant features by selecting the maximum value from the local area, while average pooling captures the overall characteristics by averaging the values within the local area. Max pooling is more frequently used in practical algorithms.

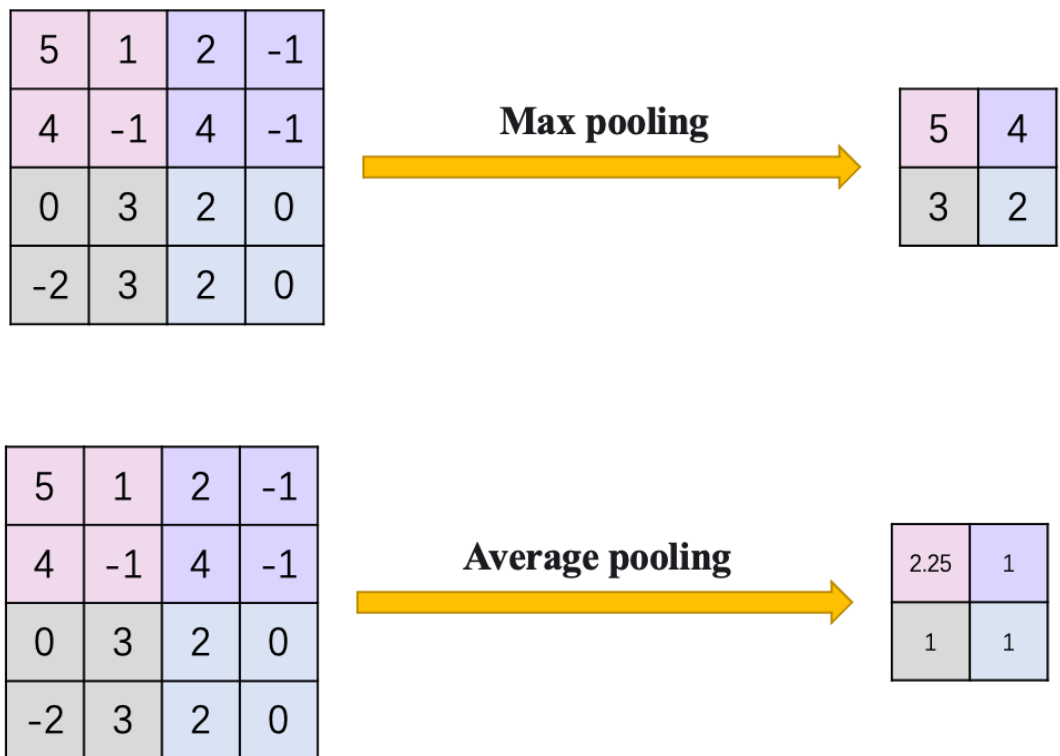


Figure 23. the Max pooling and average pooling

Fully Connected Layer

In CNNs, convolution layers extract local features, the fully connected (FC) layer reassembles these local features through a weight matrix into a complete representation of the image. Since it utilizes all local features, it is termed "fully connected." The primary role of the fully connected layer is to transform the outputs from the convolution and pooling layers into final classification or regression outcomes. It introduces a fully connected operation at the network's last layer, converting the feature maps extracted from previous layers into class probabilities or numeric predictions.

Every neuron in an FC layer is connected to all neurons in the preceding layer, implementing what is known as a fully connected configuration. This is typically achieved using an Affine layer. After several convolutional and pooling layers in a CNN structure, one or more fully connected layers follow. Each neuron in an FC layer is fully connected with all neurons from the previous layer, synthesizing the features extracted before. Due to its fully connected nature, this layer usually has the most parameters. To enhance CNN performance, the activation function for each neuron in the fully connected layer is usually ReLU.

The FC layer has the following characteristics:

Feature Integration: In a CNN, earlier layers, including convolution and pooling layers, are responsible for extracting various features from the input data. The FC layer integrates these features, combining them into higher-level representations to enable the network to make more complex decisions and classifications.

Classification Decisions: The fully connected layer is typically located at the top of the neural network, where it transmits the integrated features to an activation function that then generates scores or probabilities for each category. For classification tasks, a softmax function is often used to convert these scores into class probabilities, thus determining which category the input data belongs to.

Parameter Learning: The fully connected layer contains a large number of learnable parameters, which are adjusted during training through backpropagation and gradient descent. This adjustment allows the network to adapt to training data and make accurate predictions. By learning appropriate weights and biases, the network can perform specific tasks and generalize.

Nonlinear Modeling: The fully connected layer typically includes activation functions, such as ReLU, to introduce nonlinearity. This is a crucial factor in the powerful

representational capability of neural networks, allowing them to learn complex relationships in data.

FCN

Jonathan Long introduced the Fully Convolutional Network (FCN) in 2015[207], marking a foundational advancement in the use of deep neural networks for semantic segmentation. This model is distinctive because it replaces the FC layers of a CNN with transposed convolution layers (also known as deconvolution layers), enabling pixel-wise predictions for images of any size.

FCN typically consist of an encoder and a decoder. The encoder is composed of several convolutional layers, often adapted from classical CNN architectures (like VGG, ResNet) omitting their fully connected layers to efficiently extract abstract features and produce feature maps. The decoder then upsamples these compressed feature maps to restore them to their original output dimensions, achieving results for semantic segmentation. In this step, low-dimensional feature maps are mapped to a higher-dimensional original space. This process is facilitated by a special type of convolutional layer known as the deconvolution layer, also referred to as a transposed convolutional layer. Different from traditional CNNs, as FCN can input images of any size and use transposed convolutions to restore feature dimensions. FCN are widely used because they: 1) employ 1x1 convolution layers, allowing the network to adapt to inputs of any size; 2) use transposed convolution layers to enlarge feature map dimensions, thereby containing sufficient spatial information for finer results; and 3) utilize skip connections to merge feature maps from different layers, smoothing information transfer.

U-net

U-net[208] is a typical model of FCN specifically designed for medical image application. It has been widely used in medical image computing. Its encoder shares a similar structure with the FCN-32s model, undergoing four pooling operations

involving five different feature dimensions. U-net's architecture is completely symmetrical. Utilizing five transposed convolution layers, it gradually expands the compressed features. Unlike traditional FCN, which combine shallow and deep features by “addition”, U-net uses “concatenation”.

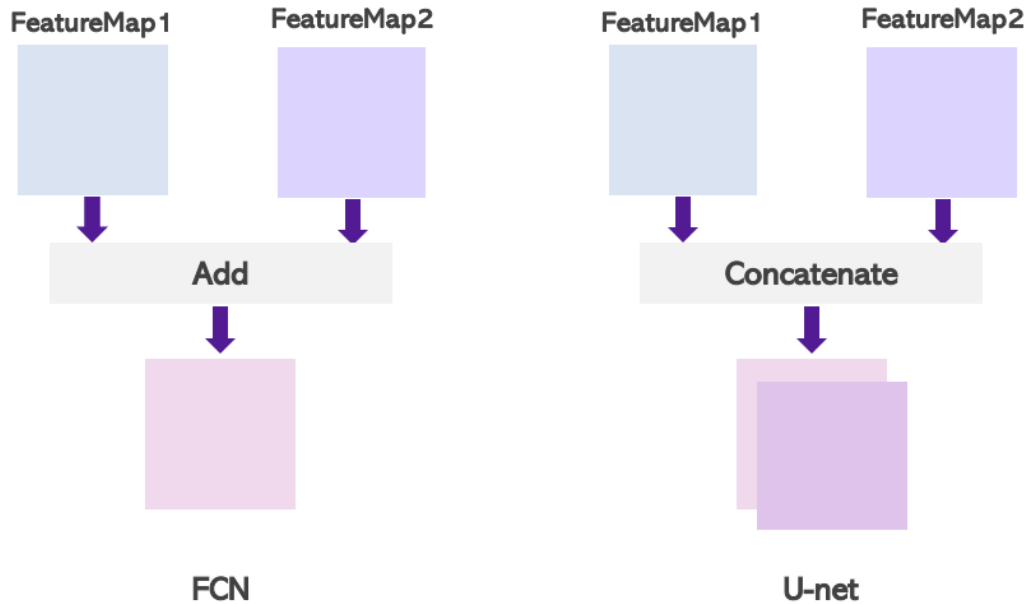


Figure 24. Different connection method

Shallow feature maps tend to represent basic feature elements like points, lines, and edge contours, containing more spatial information. In contrast, deep feature maps are more inclined to convey semantic information of the image, containing less spatial but more semantic features. The U-shaped architecture of U-net allows for a fuller integration of shallow and deep features, enhancing data reusability and maximizing the extraction of boundaries in segmentation targets. Thus, U-net demonstrates excellent segmentation performance on small-sample grayscale datasets and remains a classic and effective model in medical image segmentation.

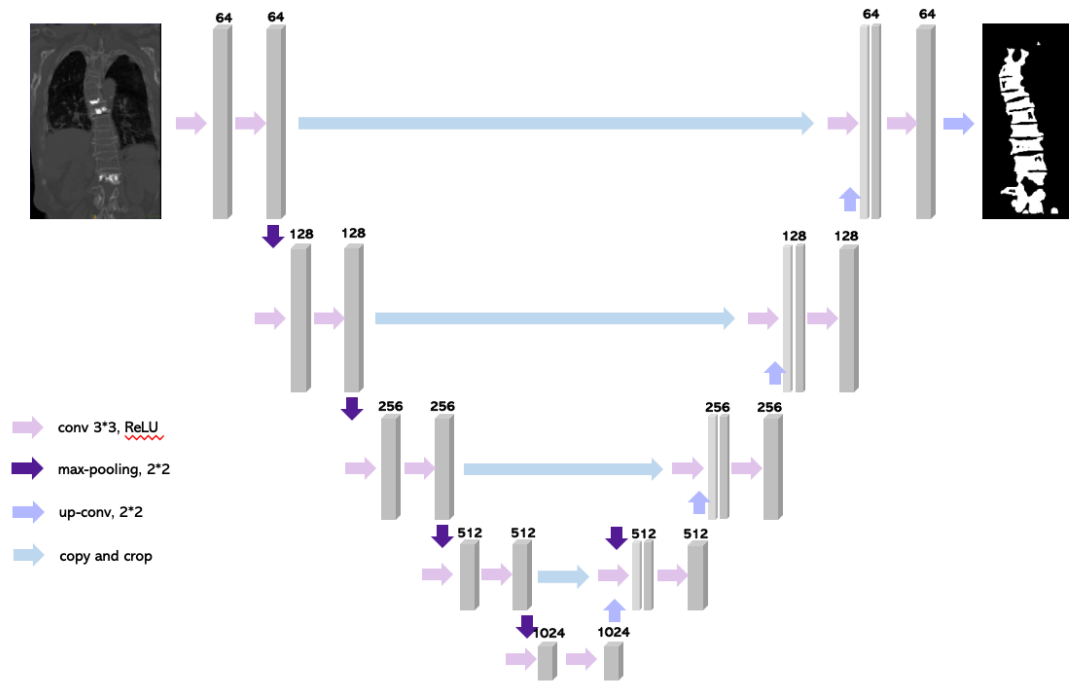


Figure 25. The architecture of U-net

The U-net architecture, resembling the letter 'U', is named for its distinctive shape and comprises an encoder (feature extraction side) and a decoder (upsampling side), collectively referred to as the Encoder-Decoder structure.

U-net Encoder - Contracting Path

Feature Extraction and Resolution Reduction: The encoder consists of four blocks, each containing two 3x3 convolution operations with a ReLU activation and a downsampling step. The use of 3x3 convolution kernels aims to minimize network complexity while maintaining segmentation accuracy.

Initial Operation: U-net begins by performing a mirror edge operation on images sized 388x388 pixels. This operation extends the image by adding a mirrored border, which enhances the network's ability to handle edge information effectively. After this operation, the image size increases to 572x572 pixels.

Convolutions and Activation: The encoder repeats the sequence of two 3x3 convolutions followed by ReLU activation. The ReLU function helps accelerate convergence and prevents gradient vanishing.

Downsampling: A 2x2 max pooling operation reduces image resolution while preserving crucial information. However, some features may be lost during this pooling process. After four such downsampling steps, the feature map dimension increases from an initial 64 to 512, with the final feature map size reduced to 32x32 pixels.

U-net Decoder - Expansive Path

Image Detail Restoration and Precise Localization: The decoder also consists of four blocks, each including two 3x3 convolutions followed by ReLU activation and an upsampling step.

Upsampling: This step is designed to decode and restore the downsampled abstract features back to the original image size. Each upsampling operation doubles the size of the feature maps while halving the number of channels, culminating in an output feature map of 388x388 pixels.

Skip Connections: Skip connections are incorporated between the contracting and expansive paths. These connections involve cropping feature maps from the contracting path to the same size as those in the expansive path and then concatenating them. This process helps restore information lost during downsampling.

U-net has widely used in medical image processing due to following five key features:

- It employs extensive data augmentation techniques such as random rotations, flips, and scaling to solve the limited number of medical images, enhancing model generalizability and robustness.

- As a lightweight network with fewer parameters, U-net reduces the overfitting and allows for faster training.
- It can efficiently utilize information from different modalities, accommodating multi-channel and multi-modal inputs like MRI images, which may include T1, T2, and FLAIR channels.
- Additionally, U-net is particularly suited to medical imaging due to its ability to process complex structures with simple semantic information. Medical images, such as CT scans of the spine, brain MRIs, or OCT images of the retina, are less dense in information compared to everyday photos but require precise targeting.
- U-net's skip connections combine low-resolution data (for identifying object categories) with high-resolution data (for precise segmentation), making it highly effective for these specialized imaging tasks.

Subjects

In this analysis, participants were chosen adhering to the specified entry criteria: identification of scoliosis; age bracket as defined: idiopathic scoliosis in adolescents: 10-16 years, congenital scoliosis: 0-10 years, and degenerative scoliosis: 55-85 years; a coronal Cobb angle ranging from 10° to 85°; the patient both included CT and posteroanterior (PA) radiographs encompassing the full spinal. Criteria for non-inclusion entailed any historical spinal surgical procedures or concurrent disorders that could affect the spinal contour. Table 1 introduced the data utilized for the training and testing processes. The whole data consisted of 106 individuals: 89 were in the training cohort, while 27 were in the testing cohort. A retrospective assessment of patient files from The First Affiliated Hospital of Shenzhen University was performed, targeting those who received spine fusion surgery or AIS patient who had more than 1 year of imaging follow-up record. during the period of 2016 to 2023. Ethical consent for this

inquiry was appropriately procured from the ethical oversight committee of the authors' institution, as substantiated by the authorization code 20220920004.

Property	Database	
	Training	Testing
Patients		
No. of patients	89	27
No. of patients with spinal fusion	36	27
Gender, male/female	30/59	4/23
Age, mean \pm SD (years)	32.4 \pm 24.5	17.3 \pm 5.8
# with idiopathic scoliosis	54	27
# with degenerative scoliosis	35	-

Table 4. Summary of the data used in this study

Image Acquisition

All the participants in the study underwent pre-operative imaging in the prone position using a CT scanner (Siemens) configured with the following parameters: 120 kV, 212 mAS, a slice interval of 5 mm, a pitch of 0.8, and a slice thickness of 1.5 mm.

Image Pre-processing

Data preprocessing is conducted to optimize the computational efficiency and accuracy of the analysis. The dataset is cropped exclusively within the non-zero regions to reduce computational load. Resampling is applied due to the presence of varying spacings within the dataset; it is automatically normalized to the median spacing of all the dataset

values. For the original data, third-order spline interpolation is utilized; for the mask, nearest-neighbor interpolation is employed.

The UNet Cascade employs a specific resampling strategy when the median size is more than four times the maximum manageable size under the memory constraints (with a batch size of 2). In such cases, a cascading strategy is used to downsample the data (by multiples of two) until the requirements are satisfied. If the data resolution differs across the three axes, the axis with the highest resolution is downsampled first to equalize the resolutions, followed by a simultaneous downsampling of all three axes until they meet the specified criteria. The pixel values within the mask across the dataset are statistically analyzed for their Hounsfield Unit (HU) ranges. The HU values are clipped to the [0.05, 99.5] percentile range and subsequently normalized using the z-score method.

Spine Segmentation

Manual Segmentation

An expert in radiology meticulously outlined the spinal anatomy using the ITK-SNAP tool, a resource made available by the University of Pennsylvania and University of North Carolina at Chapel Hill. This detailed work was later verified by a second radiology specialist for accuracy. Considering the limited number of scans and the differences in their coverage, the entire spinal column was marked as one unit, rather than identifying each spine bone on its own. To make sure the space between the images was consistent, this study adjusted the size of all the original CT scans and the outlines to a fixed measurement of 0.90 by 0.35 by 0.35 millimeters. After resizing, randomly separated these images into two groups: one with 80 images for the main training part of our study and another with 10 images to check how well our methods worked.

Processing image

In this analysis, participants were chosen adhering to the specified entry criteria: identification of scoliosis; age bracket as defined: idiopathic scoliosis in adolescents:

Training the U-Net Architecture

A training protocol utilizing the nnU-Net framework, accessible via its GitHub repository, was employed for the development of a 3D U-net model tailored for medical image analysis. The network processed image patches of dimensions $112 \times 160 \times 128$ pixels, with each training batch comprising two such patches. This setup involved the model undergoing a series of five downsampling stages, a methodology designed to incrementally reduce the spatial resolution while simultaneously expanding the representational capacity of the network. Each downsampling stage effectively doubled the number of convolutional channels, beginning from an initial count of 32, thereby enhancing the model's ability to capture complex features in the data. The activation function selected for this network was the ReLU, known for its efficacy in maintaining a healthy gradient flow during training phases. Optimization of the network was systematically carried out by minimizing key performance metrics, namely the dice coefficient and cross-entropy loss, employing a method of stochastic gradient descent with a learning rate set at 0.01. The comprehensive training phase was rigorously conducted over a span of 1,000 epochs, aiming to refine the model's accuracy and reliability in predicting medical imaging outcomes.

3.3 Results

The congruence between the ground truth segmentation delineated by radiologists and the spinal masks generated automatically was quantified using dice coefficients across both training and testing datasets. Within the training dataset, the model demonstrated exemplary performance metrics: the dice coefficient was recorded at 0.940 with a standard deviation of 0.002, the F1-score at 0.945 with a variability of 0.026, precision at 0.920 with a deviation of 0.047, and recall at 0.973 with a standard deviation of 0.014. In contrast, the testing dataset exhibited slightly reduced metrics with a dice coefficient

of 0.935 plus or minus 0.062, an F1-score of 0.877 with a standard deviation of 0.062, precision at 0.818 with a larger variability of 0.099, and recall at 0.954 with a standard deviation of 0.045.

Visual assessments further corroborated the quantitative findings, confirming that the segmentation produced by the deep learning model closely aligned with the ground truth, as exemplified in the figure below. It is noteworthy that the automated method not only replicated but also enhanced the delineation of vertebral bodies, capturing both the superior and inferior extents more completely than the manual annotations, thus offering a refined accuracy that surpassed traditional manual segmentation methods.

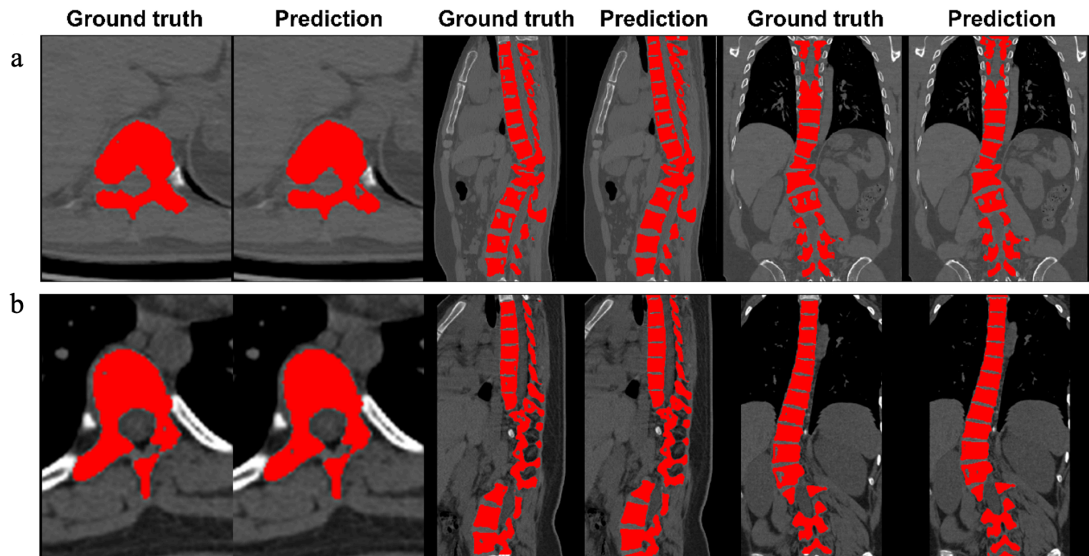


Figure 26. Segmentation result in Clinical Cases: Displayed are the segmentation outcomes from two distinctive cases within the study cohort. Figure (a) showcases a sample from the training set, whereas figure (b) presents a corresponding example from the validation set. Each image has been carefully trimmed to minimize the extensive non-relevant background, ensuring focused representation of the spinal region

3.4 Discussion

The effectiveness of this approach was validated through its application on preoperative CT images from 27 patients scheduled for spinal fusion surgery. The deep learning

algorithms used in this study demonstrated substantial accuracy in segmenting the spine, achieving dice coefficients of 0.940 in the training dataset and 0.935 in the testing dataset. These results significantly exceed those of previous methodologies, as cited in the literature[41, 177, 183], and show particular improvement in handling cases of moderate to severe spinal deformities [182]. This advancement underscores the potential of this automated method to enhance diagnostic accuracy and support surgical planning in orthopedic practice.

3.5 Conclusion

U-net is a classical model in medical image segmentation, renowned for its modest data requirements and rapid training capabilities. It has found widespread application in the field of medical imaging segmentation, particularly where labeled data is scarce. Many innovative algorithms based on the U-net architecture have also achieved favorable results. Considering that U-net is not an innovative model, as researchers in previous MICCAI vertebral segmentation studies have also employed U-net for vertebral segmentation, so did not validate it with a public dataset in this chapter. Instead, this chapter mainly focused on a model tailored for scoliosis using an in-house dataset. The architecture of the network did not change but the parameters were adjusted to meet a better segmentation result for the in-house dataset. However, U-net performs well in detail features, but is weak in the control of global information, so more promising advancements in deep learning are needed, especially with the rise of transformer models. The next chapter will explore vertebral segmentation using an improved transformer network.

CHAPTER 4 AUTOMATICALLY VERTEBRA SEGMENTATION BASED ON NNFORMER++

4.1 Introduction

Based on architectures utilizing self-attention mechanisms, especially the Transformer, it has become the model of choice for natural language processing (NLP). Due to the Transformer's ability to effectively capture and utilize long-term dependencies between pixels or voxels, there has recently been a significant emergence of models and networks that combine CNNs and Transformers for medical image processing. Most results suggest that embedding Transformer-like structures in appropriate positions within a CNN can effectively enhance the network's performance.

The evolution of vertebra segmentation techniques has shifted from traditional CNN architectures to more advanced transformer-based models, driven by the need for better handling of global contextual information and complex anatomical variability. For example, the LumVertCancNet[209] utilizes a hybrid approach, blending Swin Transformers with CNNs to enhance lumbar vertebra segmentation, achieving an impressive Dice similarity coefficient (DSC) of 96.29%, illustrating its superior accuracy in capturing intricate anatomical features. Similarly, the RUnT[210] model combines Residual U-Net with transformers, achieving notable DSCs of 88.4% on CTSpine1K and 81.5% on VerSe 20, which underscores its capability to improve vertebral edge feature detection through enhanced global and local context processing.

The superiority of transformers over CNNs in vertebra segmentation is increasingly evident. For instance, EG-Trans3DUNet[211] employs a sophisticated edge detection combined with global information processing, reaching a Dice score of 86.82% on the challenging VerSe'20 dataset. This performance is indicative of the model's ability to effectively integrate detailed spatial relationships across the spine, a task at which

CNNs are less efficient due to their inherent limitations in processing long-range dependencies.

Transformers have redefined the landscape of medical imaging by addressing the limitations of CNN architectures, particularly in applications requiring the integration of extensive spatial contexts, such as in spine imaging. Models like Spine-Transformers[212] showcase this progression by adeptly managing arbitrary fields of view in spine CT scans, demonstrating the transformers' advantage in dynamically adapting to varied imaging scenarios.

In an effort to better combine the advantages of CNNs and transformers, Zhou et al. [213] proposed a new imaging segmentation network, Not-aNother transFORMER (nnFormer). Compared to nnUnet, which were used in the previous chapter, nnFormer still achieved slight improvements in multi-organ segmentation tasks based on public data. Therefore, both a public dataset and an in-house dataset were used to evaluate the algorithm's performance.

4.2 Methodology

Brief methodology of nnFormer++

nnFormer[213] is a sophisticated 3D transformer-based architecture tailored for volumetric medical image segmentation. Its design is a strategic response to the inadequacies of traditional CNNs in capturing long-range dependencies within medical images. By innovatively combining convolutional operations with transformer technology, nnFormer transcends the spatial limitations of CNNs, offering enhanced segmentation accuracy through a deeper contextual understanding of medical images.

nnFormer++ enhances the nnFormer model by integrating interleaved conv-transformer blocks with a nested U-Net framework. This combination allows the model to dynamically capture long-range relationships and fully leverage multi-scale semantic features in medical images. The model operates on 3D CT images by first splitting them

into multiple overlapping 3D image patches. These patches are then processed sequentially through the encoder, bottleneck, and decoder components of nnFormer++, transforming the feature maps into the final vertebral segmentation mask.

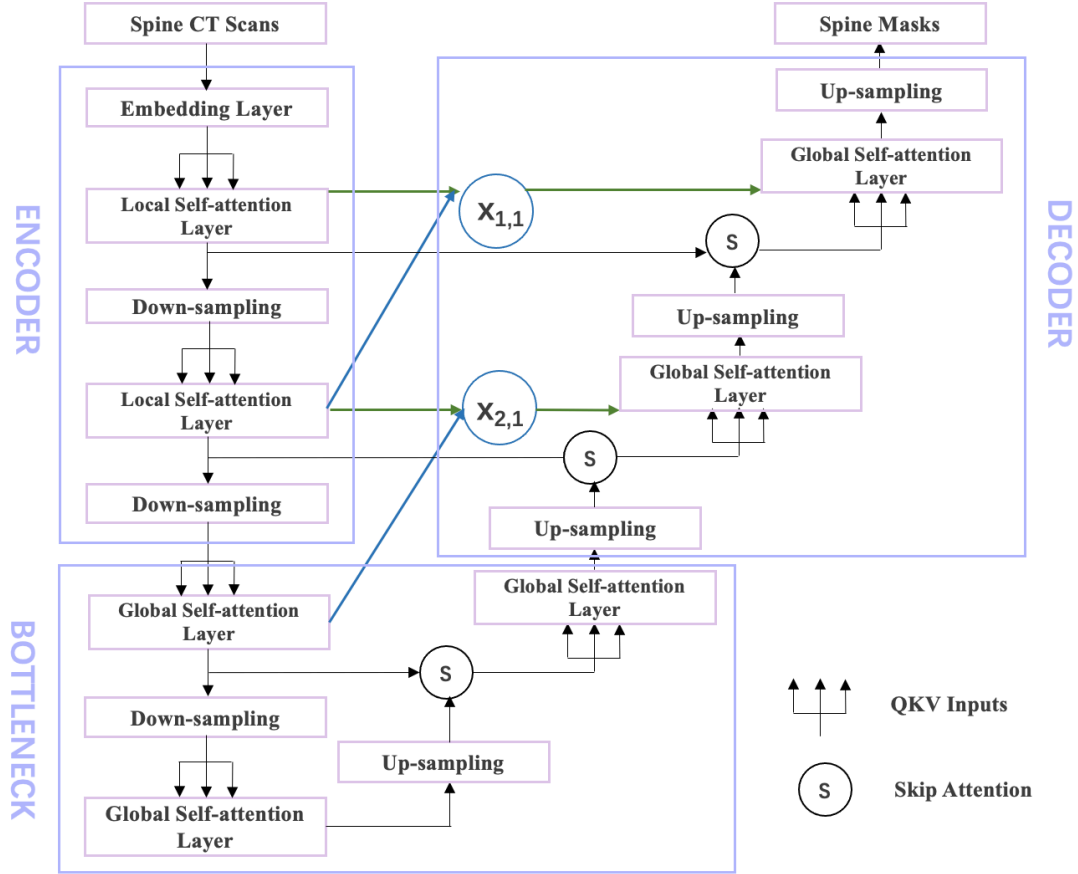


Figure 27. Structure of nnFormer++

Architecture and Components

Figure 26 illustrates a modified architecture of the nnFormer++ network tailored for processing spine CT scans, structured into Encoder, Bottleneck, and Decoder segments. In this architecture, input images first pass through an embedding layer in the Encoder, followed by a sequence of local self-attention layers interspersed with downsampling steps to condense the feature information. This process continues into the Bottleneck, which utilizes global self-attention layers and additional downsampling to intensify feature extraction at the deepest network level. Subsequently, the Decoder reconstructs

the spatial resolution through up-sampling and global self-attention layers to produce detailed spine masks.

Distinctively, compared to nnformer, nnformer++ redesigned skip connections (illustrated in blue) that integrate convolutional layers to bridge the semantic gap between deep and shallow network features, ensuring a more effective feature translation across layers. Moreover, denser skip connections (shown in green) enhance the inter-layer information flow, improving the overall segmentation performance. This sophisticated framework leverages the strengths of self-attention mechanisms to handle complex spatial hierarchies in medical images, optimizing the accuracy and efficiency of spine segmentation tasks.

Encoder:

The encoder begins by transforming the input image into a format that the model can efficiently process. This is done using a convolutional layer that prepares the image by highlighting initial features like edges and textures. After this initial preparation, the encoder uses a mix of downsampling and transformer blocks to refine these features. Downsampling reduces the image size and helps the model to focus on important parts, while transformer blocks analyze these parts to understand broader patterns and relationships within the image.

The encoder starts by converting the input image into a set of initial features. This is done through a convolutional layer:

$$X_{embedded} = Conv(X) \quad (4.1)$$

Where X is the input image and $X_{embedded}$ represents the feature-rich version of the input. As the encoder processes these features, it alternates between reducing the image size (downsampling) and applying transformer blocks for detailed analysis:

$$X_{downsampled} = Downsample(X_{prev}) \quad (4.2)$$

Here, X_{prev} is the output from the previous layer, and $X_{downsampled}$ is the reduced version to focus processing on essential features.

The local volume-based multi-head self-attention(LV-MSA) is designed to capture fine-grained local spatial relationships within small sub-regions of the input volume. This mechanism focuses on understanding the intricate structures within a localized area of the medical image, which is crucial for identifying detailed anatomical features. The transformer blocks use a local self-attention mechanism to analyze small regions within the image:

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (4.3)$$

Where Q , K , and V are the query, key, and value components derived from the features, and d_k is a scaling factor that helps stabilize the training. This formula helps normalize the influence of different features, focusing the model's attention based on the relevance of local features.

Multi-head Mechanism

Each 'head' in the multi-head attention looks at different representation subspaces at different positions. This diversification allows the model to capture various aspects of the local information:

$$MultiHead(Q, K, V) = Concat(head_1, head_2, \dots, head_h W^o) \quad (4.4)$$

where each $head_i = Attention(QW_i^Q, KW_i^K, VW_i^V)$ and W^o , W_i^Q , W_i^K , W_i^V are parameter matrices that are learned during training.

Bottleneck:

At the core of nnFormer is the bottleneck, which uses a powerful mechanism called Global Volume-based Multi-head Self-attention. This part of the model looks at the

entire image at once to gather global information that helps distinguish between different regions of the image, such as different types of tissues in a medical scan.

The bottleneck captures the global context using Global Volume-based Multi-head Self-attention(GV-MSA). Unlike LV-MSA, GV-MSA processes the entire input volume to capture high-level, global contextual relationships. This global perspective is vital for understanding the overall structure and spatial relationships between different anatomical parts across the entire image. Just like LV-MSA but applied to the whole volume, GV-MSA also uses the softmax-scaled dot-product attention, allowing the model to integrate information across the entire field of view:

$$GV - MSA(Q, K, V) = softmax(\frac{QK^T}{\sqrt{d_k}} + B)V \quad (4.5)$$

Here, B represents a bias or a positional encoding that adds more context to the attention mechanism, helping to maintain spatial relationships in the global context.

$$GV - MSA(X) = Attention(X, X, X) \quad (4.6)$$

This function processes the entire volume of data to ensure that global relationships within the image are well-understood, improving segmentation accuracy.

Decoder:

The decoder reverses the process of the encoder. It starts with the condensed information from the bottleneck and gradually adds back details using upsampling. At the same time, it reconnects information from earlier in the encoder through a novel method called skip attention. This approach helps the decoder focus on important details by highlighting features that are crucial for accurate segmentation.

In the decoder, the process of downsampling is reversed by upsampling, gradually restoring the image's original dimensions while integrating detailed features:

$$X_{upsampled} = Upsample(X_{prev}) \quad (4.7)$$

Additionally, the skip attention mechanism selectively reintegrates features from the encoder using an attention-based method. Skip Attention is an innovative approach used in the skip connections of nnFormer, enhancing the traditional U-Net architecture by replacing simple concatenation or summation with a more selective attention-driven approach. Skip Attention uses the attention mechanism to selectively integrate features from the encoder directly into the decoder, based on their relevance to the segmentation task:

$$Skip - Attention(Q, K, V) = softmax(\frac{QK^T}{\sqrt{d_k}})V \quad (4.8)$$

In this setting, Q is derived from the decoder features aiming to query the corresponding K and V from the encoder features, ensuring that only the most relevant features are passed through the skip connections.

Interleaved Convolution and Transformer Blocks: This setup allows nnFormer to efficiently handle detailed local information while also understanding the broader context, leading to more accurate segmentation.

Local and Global Self-Attention: By focusing on small and large areas of the image, nnFormer can accurately recognize and outline various structures within the image, from small anomalies to larger organs.

Skip Attention Mechanism: Unlike traditional methods that simply stitch together features from different layers, skip attention uses a more selective approach, focusing only on relevant features, which improves the accuracy of the final image segmentation.

Subjects

In this chapter, we validate the effectiveness of our algorithm using two datasets: a public dataset and an in-house dataset. The public dataset employed is the VerSe

dataset[214], a large-scale, multi-device, multi-center CT image dataset for spinal segmentation, encompassing 374 scans from 355 patients. This dataset aggregates data from the MICCAI 2019 and 2020 VerSe challenges, known as VerSe19 and VerSe20, respectively. Specifically, VerSe19 includes 160 scans from 141 patients, while VerSe20 involves 319 scans from 300 patients. However, both datasets include data from 86 patients in common, so the total number of unique patients in VerSe is not simply the sum of the two datasets.

The VerSe dataset divides these 374 scans into 141 for training, 120 for validation, and 113 for testing, with all scans and labels publicly available. Among the 26 spinal annotation categories, besides the typical C1-C7, T1-T12, and L1-L5 categories totaling 24, special attention is given to the rare T13 and L6 vertebrae.

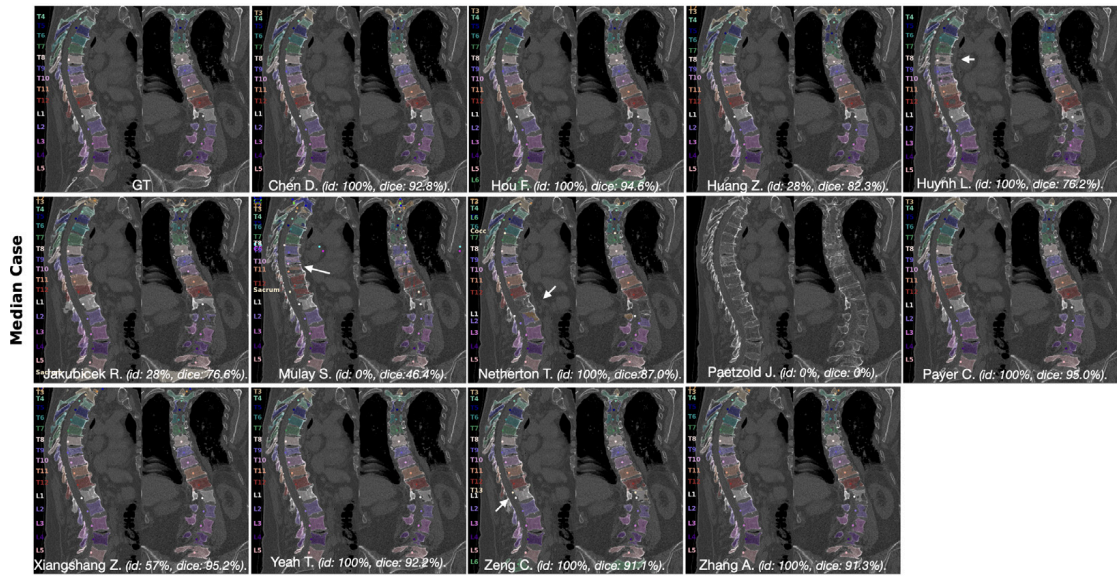


Figure 28. The best segmentation performance in scoliosis[214],

The dataset includes some CT data of scoliosis, categorized under the 'Median Case'. However, the number of patients with scoliosis in the dataset is limited, with the following label information:

Label	Anatomical structure	Frequency of occurrence	Proportion of occurrence	Max Volume (cm ³)	Mini Volume (cm ³)	Mean Volume (cm ³)
1	C1	76	20.32%	25	7.82	14.24
2	C2	76	20.32%	28.23	11.73	18.38
3	C3	80	21.39%	20.66	8.26	12.98
4	C4	81	21.66%	19.69	7.57	12.51
5	C5	90	24.06%	19.61	8.01	13.08
6	C6	98	26.2%	22.2	8.59	15
7	C7	129	34.49%	26.97	10.19	17.34
8	T1	180	48.13%	35.99	12.15	22.64
9	T2	198	52.94%	36.18	13.23	22.61
10	T3	182	48.66%	35.57	7.37	21.72
11	T4	174	46.52%	39.59	12.75	22.34
12	T5	165	44.12%	39.86	13.71	24.04
13	T6	160	42.78%	43.19	15.04	26.11
14	T7	166	44.39%	46.05	17.03	29.78
15	T8	176	47.06%	49.81	17.72	32.2
16	T9	215	57.49%	59.08	19.04	34.85
17	T10	249268	66.58%	61.32	13.79	38.72
18	T11	272	71.66%	66.97	22.54	41.98
19	T12	298	72.73%	74.25	23.19	46.4

20	T13	6	1.6%	66.24	31.41	57.06
21	L1	297	79.68%	82.81	28.54	51.46
22	L2	295	79.41%	87.89	30.93	56.54
23	L3	292	78.88%	99.59	27.32	62.5
24	L4	276	78.07%	108.07	37.8	64.66
25	L5	50	73.8%	106.12	38	63.76
26	L6	297	13.37%	101.18	47.52	70.55
27	sacrum	0	0%	0	0	0
28	coccyx	0	0%	0	0	0

Table 5. Label information in VerSe dataset

In the MICCAI challenge, the performance of the top five algorithms is presented in the following table.

	VERSe	PUBLIC		HIDDEN	
		All	Top-5	All	Top-5
<i>id.rate</i>	2019	61.4±44.5	83.3±30.7	61.6±43.6	82.4±31.6
	2020	72.5±39.3	93.9±21.0	68.6±42.1	94.4±17.5
Dice	2019	71.2±33.7	82.5±25.9	71.3±32.6	78.9±28.4
	2020	75.2±28.5	89.3±17.9	71.1±32.2	88.8±16.7

Table 6. Label information in VerSe dataset

The in-house dataset employed consists of the same spinal scoliosis patient data as discussed in the previous chapter. Compared to the public dataset, the in-house dataset includes scoliosis of varying severities. Detailed information about the data has already been introduced in the previous chapter and will not be reiterated here.

Spine Segmentation

Manual Segmentation

Training the nnFormer++ Architecture

Patch Dimensions

Patch Embedding:

- The input images are divided into smaller patches, and each patch is projected into a higher-dimensional space.
- Initial projection involves Conv3d layers with kernel sizes of (3, 3, 3) and strides of (2, 2, 2), followed by another Conv3d layer with kernel size (3, 3, 3) and stride (1, 1, 1).
- This results in patches that are of reduced spatial dimensions but increased feature depth.

Window Sizes

Swin Transformer Blocks:

- The Swin Transformer introduces the concept of window-based multi-head self-attention (W-MSA) to limit the scope of self-attention to non-overlapping local windows.
- The window size is crucial as it defines the region over which self-attention is calculated. In the training process:
- Each ‘SwinTransformerBlock’ processes features within a window size that balances the computational efficiency and the ability to capture local context.

Layer Details

‘LayerNorm’ and GELU Activation:

- ‘LayerNorm’ is used to normalize the inputs across the features, improving stability and performance.
- GELU (Gaussian Error Linear Unit) is used as the activation function, offering a smooth approximation which helps in the optimization process.

Patch Merging:

- After each stage of the Swin Transformer blocks, ‘PatchMerging’ layers are applied to reduce the spatial resolution and increase the feature dimension.
- This merging process involves ‘Conv3d’ layers with a kernel size of (3, 3, 3) and stride of (2, 2, 2), followed by ‘LayerNorm’.

Encoder-Decoder Structure

Encoder:

- Consists of multiple ‘BasicLayer’ modules that include a series of ‘SwinTransformerBlock’ layers.
- Each ‘BasicLayer’ downsamples the input feature map using ‘PatchMerging’.

Decoder:

Mirrors the encoder structure but includes ‘PatchExpanding’ layers to upsample the feature maps back to the original input resolution.

‘BasicLayer_up’ modules contain ‘SwinTransformerBlock’ layers designed for upsampling.

Upsampling

Patch Expanding:

- 'ConvTranspose3d' layers are used in 'Patch_Expanding' to upsample the feature maps. The kernel sizes and strides are designed to match the dimensions back to the original input size.
- 'LayerNorm' is applied after each upsampling step to ensure normalized feature distributions.

Final Projection

Final Patch Expanding:

The final layers involve 'ConvTranspose3d' layers to project the upsampled feature maps into the desired output dimensions, ensuring that the segmentation maps are correctly sized and aligned with the original input images.

Training Dynamics

Epoch Details:

- Training is conducted over many epochs, each involving forward and backward passes, weight updates, and performance evaluations.
- Each epoch logs training loss, validation loss, and Dice coefficients to track progress and model performance.

Learning Rate Scheduling:

The learning rate starts at 0.01 and gradually decreases as training progresses. This helps in fine-tuning the model weights and achieving better convergence.

4.3 Results

The Table 7 presents a comparative analysis of vertebra segmentation performance on the Verse19 dataset, focusing on two key metrics: the Dice Similarity Coefficient (DSC) and the 95th percentile Hausdorff Distance (HD95). These metrics are used to evaluate the accuracy and robustness of various segmentation methods.

The DSC measures the overlap between the predicted segmentation and the ground truth, with higher values indicating better performance. In this comparison, our method achieved a DSC of 0.94, which is comparable to the top-performing methods such as Zhang et al. [215] with a DSC of 0.942 and Zhou et al. [213] with a DSC of 0.934. This indicates that our method has a high degree of overlap between the predicted segmentation and the actual anatomical structures.

The HD95 metric assesses the distance between the predicted segmentation boundaries and the ground truth, where lower values denote better boundary precision. Our method achieved an HD95 of 1, which is the lowest among the compared methods. This signifies that our method provides the most precise boundary delineation, outperforming other methods like Zhou et al. [213], which had an HD95 of 1.36, and Zhang et al. [215], which recorded an HD95 of 6.24.

These results highlight the effectiveness of our segmentation approach. The high DSC indicates that our method captures the overall shape and volume of the vertebrae accurately, while the low HD95 demonstrates superior boundary precision. Such precise segmentation is crucial in clinical settings, particularly for planning surgical interventions and evaluating treatment outcomes. By combining high overlap accuracy with precise boundary delineation, our method ensures reliable and clinically relevant segmentation results.

Ref. authors	Evaluated on public test data	
	DSC	HD95

Payer et al.[216]	0.910	6.35
Tao et al.[217]	0.911	6.34
Sekuboyina et al.[214]	0.930	6.39
Zhang et al.[215]	0.942	6.24
Zhou et al.[213]	0.934	1.36
Our method	0.94	1

Table 7. Result compared with other method in Verse 19

Table 8 presents a comparative analysis of the segmentation performance of three models: U-Net, nnFormer, and nnFormer++ at epoch 330. The evaluation metrics used for comparison are the Average Global Foreground Dice and the 95th percentile Hausdorff Distance (HD95). The Average Global Foreground Dice measures the overlap between the predicted segmentation and the ground truth, with higher values indicating better performance. In this comparison, nnFormer++ achieved the highest Dice score of 0.946, indicating superior segmentation accuracy compared to U-Net (0.935) and nnFormer (0.912). This suggests that nnFormer++ is more effective in capturing and delineating the vertebral structures accurately.

The HD95 metric evaluates the boundary accuracy by measuring the 95th percentile of the Hausdorff distance between the predicted segmentation boundaries and the ground truth. A lower HD95 value signifies better boundary precision. In this study, nnFormer++ obtained an HD95 of 9.87, which is lower than that of U-Net (10.15), but slightly higher than nnFormer (9.44). This indicates that while nnFormer has a slight edge in boundary precision, nnFormer++ still provides competitive performance, balancing both overall segmentation accuracy and boundary delineation.

Figures 29 and 30 further illustrate the segmentation outcomes and validate the performance of the models. Figure 29 shows segmentation results for two cases from the study cohort, with (a) representing a training set sample and (b) a validation set sample. These images demonstrate that the predictions closely align with the ground

truth annotations, highlighting the model's capability to accurately segment vertebral structures. Figure 30 displays the 3D mesh reconstructions of the spine from the testing dataset, showcasing the model's ability to generate precise and clinically relevant 3D representations of the spine. These visualizations confirm the effectiveness of nnFormer++ in producing accurate segmentation outcomes, which are essential for pre-surgical planning and assessment. Overall, the nnFormer++ model demonstrates significant advancements in vertebral segmentation, offering improved accuracy and practical utility in clinical applications.

		U-net	nnformer	nnformer++
		(Epoch350)	(Epoch350)	(Epoch350)
Average	Global	0.935	0.912	0.946
Foreground Dice				
HD95		10.15	9.44	9.87

Table 8. Result compared in U-net, nnformer and nnformer++

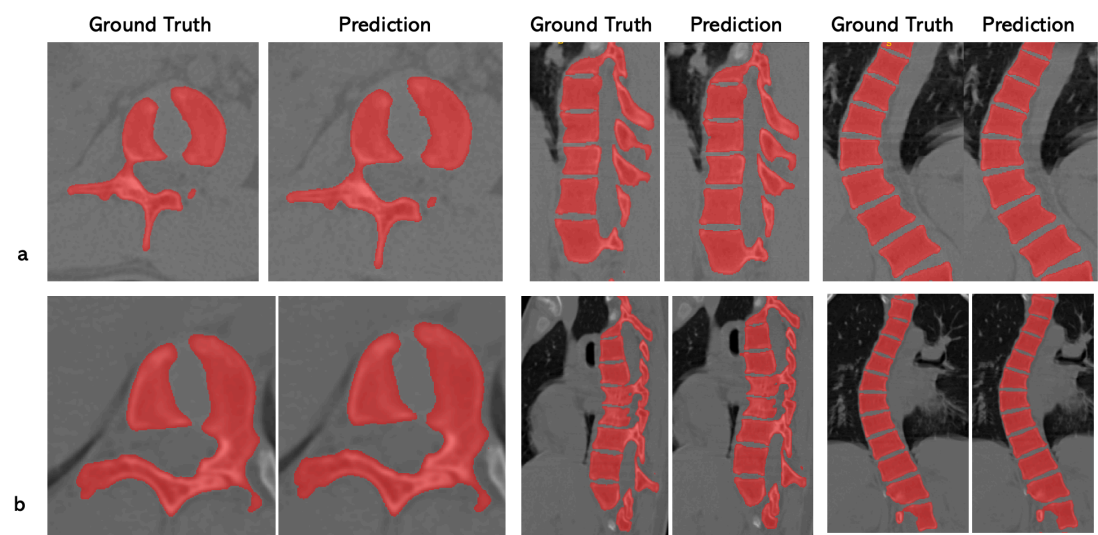


Figure 29. Segmentation result in in-house data: Displayed are the segmentation outcomes from two distinctive cases within the study cohort. Figure (a) showcases a sample from the training set, whereas figure (b) presents a corresponding example from

the validation set. Each image has been carefully trimmed to minimize the extensive non-relevant background, ensuring focused representation of the spinal region

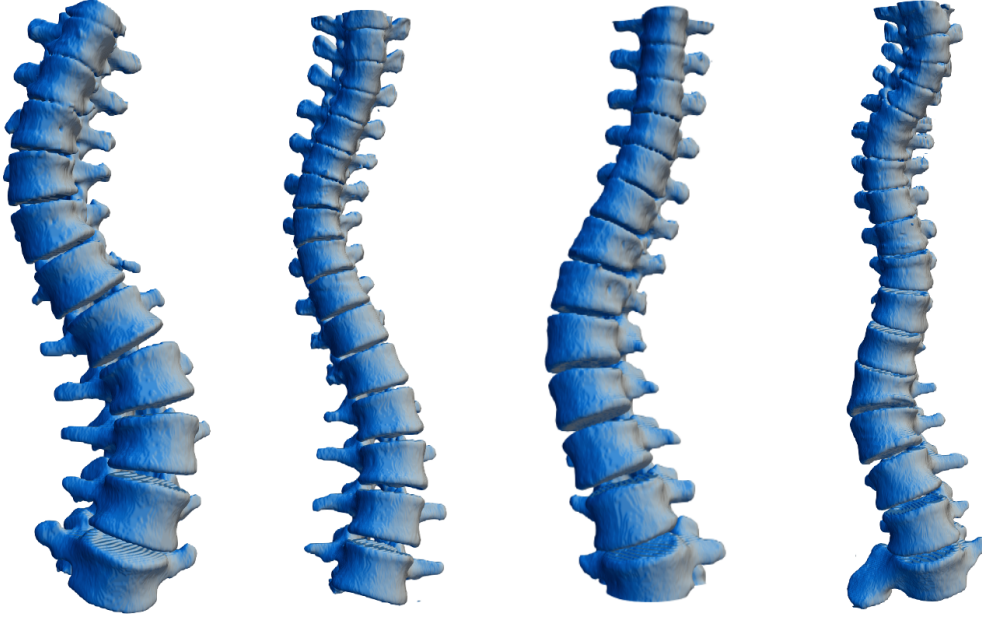


Figure 30. 3D mesh of spine in testing dataset

4.4 Discussion

Compared with nnformer and Unet, NNform ++ has a slight advantage in the segmentation performance based on in-house dataset. Table 8 shows that nnFormer++ achieves the highest Average Global Foreground Dice score of 0.946, outperforming U-Net (0.935) and nnFormer (0.912). Additionally, nnFormer++ has a lower HD95 value of 9.87 compared to U-Net's 10.15, although it is slightly higher than nnFormer's 9.44.

Furthermore, Table 7 highlights the effectiveness of our method in comparison to other existing methods on the Verse19 dataset. Our method achieves a Dice score of 0.94 and an HD95 of 1. These results are superior to other methods, such as those by Sekuboyina et al. (0.930, 6.39) and Zhang et al. (0.942, 6.24), indicating that our model provides both high segmentation accuracy and precise boundary delineation.

Figures 29 and 30 further illustrate the robustness of nnFormer++ in practical applications. Figure 29 displays segmentation outcomes from both training and validation sets, showcasing the close alignment between predictions and ground truth annotations. Accurate segmentation is particularly important for scoliosis patients, as it directly impacts the measurement of the Cobb angle, a key metric in diagnosing and planning treatment for scoliosis[215]. The 3D mesh reconstructions in Figure 30 provide a comprehensive view of the spine, confirming the model's ability to generate precise and clinically relevant 3D representations. These 3D models facilitate detailed examination and pre-surgical planning, ensuring better patient outcomes.

The advancements brought by transformer-based methods in medical image segmentation are evident from the performance of nnFormer++. By leveraging the strengths of both UNet++ and transformer architectures, nnFormer++ captures long-range dependencies and global contextual information, leading to more accurate and reliable segmentation results. Previous studies, such as those by Zhang et al. and Zhou et al., have demonstrated the efficacy of combining CNNs with transformers for enhanced segmentation performance[211, 218]. Moreover, transformer-based models like the Swin Transformer have shown superior performance in various segmentation tasks by effectively capturing hierarchical features[209, 213]. The superior results of nnFormer++ align with these findings, validating the model's robustness and applicability in clinical settings.

4.5 Conclusion

In conclusion, the nnFormer++ model demonstrates better performance in vertebral segmentation for severe scoliosis patients, achieving higher accuracy and better boundary delineation compared to the studies before both in public dataset and in-house dataset. The enhanced segmentation performance of nnFormer++ is crucial for accurate 3D spatial angle measurement, facilitating improved diagnosis and treatment planning for scoliosis. The integration of transformer-based methodologies with CNN

architectures in nnFormer++ leverages the advantages of both approaches, capturing detailed anatomical features and ensuring precise segmentation outcomes.

CHAPTER 5 3D SPATIAL ANGLE MEASUREMENT

5.1 Introduction

In the preceding chapter, the segmentation of vertebral bodies was accomplished. On this basis, it is now necessary to undertake the fitting of conical curves. This step is crucial for further evaluating spinal scoliosis using three-dimensional curves.

The evolution of curve fitting techniques has progressed from polynomial fitting to Bézier curves, and then to B-splines, culminating in Non-Uniform Rational B-Splines (NURBS). Bézier curves are limited in local editing capabilities, and with increasing order, managing the curve's shape becomes more complex. Furthermore, their order also limits the control's degrees of freedom. In contrast, B-splines provide enhanced control over the curves. B-Spline separates the control points from the order, removing their direct relationship and adding a knot vector. B-splines, unlike Bézier curves with only a single degree of freedom via control points, offer up to three degrees of freedom. Dr. Ken V. Versprille first extensively formulated Non-Uniform Rational B-Splines (NURBS) in 1975 in his doctoral thesis, "Computer-Aided Design Applications of the Rational B-Spline Approximation Form." Presently, NURBS curves are a prevalent method of curve representation in computer graphics. In contrast to B-splines, NURBS incorporates a weight factor w into its formula. NURBS are frequently utilized mathematical models in computer graphics to generate and depict curves and surfaces. They are characterized by the following primary features:

1. Non-Uniformity (Highlighted Change): NURBS curves allow for non-uniform distribution of control points, facilitating detailed control in specific sections of the curve.
2. Rationality (Highlighted Change): Based on rational Bézier curves, NURBS curves integrate a weight parameter, enhancing their flexibility to represent diverse curve shapes.

3. B-Spline Basis (Highlighted Change): NURBS employs spline functions as its base, offering robust local control - moving or adding control points impacts only the curve's specific areas.
4. Variable Order (Highlighted Change): The order of NURBS curves can be adjusted, typically utilizing third or fifth-order curves for smoother curve continuity.

NURBS curves' principal idea revolves around using control points and weights to establish a control polygon, which then creates smooth curves via spline basis functions. This technique merges the strengths of Bézier and spline curves, providing powerful representational capabilities.

5.2 Methodology

5.2.1 Brief methodology of NURBS

Among parametric polynomial curves, the Bézier curve was the first to be defined using polynomial basis functions, specifically the Bernstein basis functions. The Bézier curve, influenced by its control vertices b_j , can be defined as:

$$C(u) = \sum_{j=0}^n b_j B_{j,n}(t), 0 \leq t \leq 1 \quad (5.1)$$

where $b_0 = a_0$ and $b_j = b_{j-1} + a_j$ ($j = 1, 2, \dots, n$) for $j = 1, 2, \dots, n$. Here, the b_j are known as the control vertices, also referred to as Bézier points. The a_j (for $j = 1, 2, \dots, n$) are the edge vectors of the control polygon. This representation facilitates interactive design, as slight adjustments to the control vertices can easily alter the curve's shape.

To meet the significant demand for curve representation and design, researchers expanded on the Bézier technique by introducing B-splines, which better capture complex shapes. The B-spline curve equation is generally represented as:

$$C(u) = \sum_{i=0}^n N_{i,p}(u)P_i \quad (5.2)$$

Here P_i , for i ranging from 0 to n , are the control points of the B-spline curve. The $N_{i,p}$ functions are the basis functions of the B-spline curve, where the second subscript p indicates the degree of these basis functions. The B-spline basis functions are defined over a non-periodic and non-uniform canonical knot vector U , which can be represented as:

$$U = \left\{ \underbrace{0, 0, \dots, 0}_{k+1}, u_{k+1}, u_{k+2}, \dots, u_n, \underbrace{1, 1, \dots, 1}_{k+1} \right\} \quad (5.3)$$

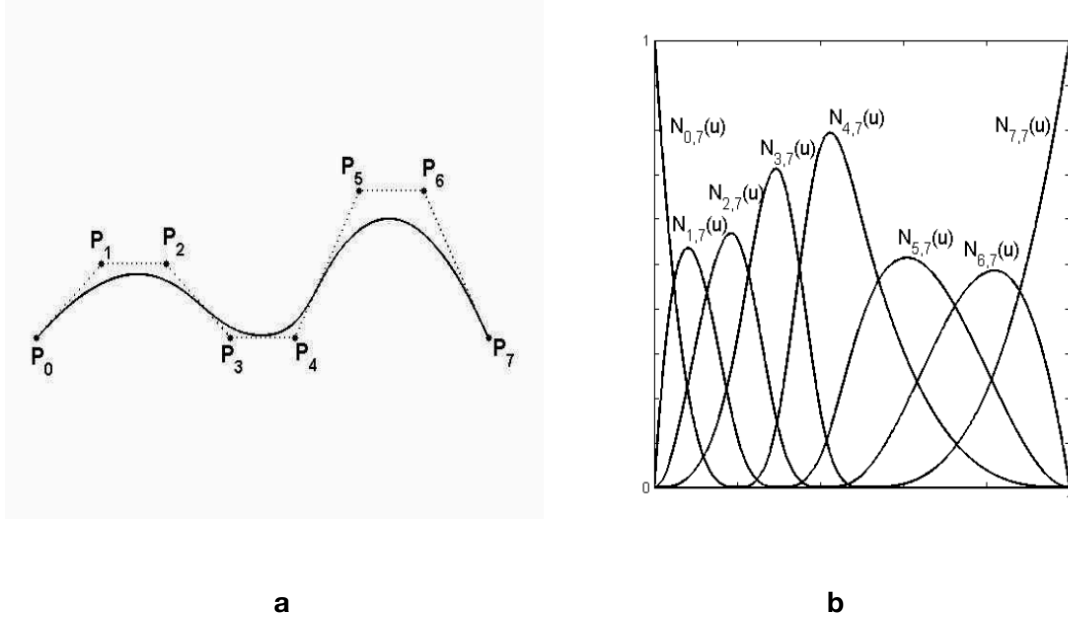


Figure 31. (a) displays a cubic B-spline curve and its control polygon, while (b) shows the basis functions for the B-spline curve.

Following the Bézier and B-spline methods, the next significant development was Non-Uniform Rational B-Splines (NURBS), which were needed for the representation of quadratic curves and surfaces. NURBS allows for the incorporation of both straight lines and free-form curves in a unified representation. It introduces a flexible and freely adjustable weight factor, adding a new dimension of control. NURBS curves consist of the core elements: knot vectors, weights, and control points, forming the backbone of their structure.

For representation in homogeneous coordinates P_i , it is expressed as:

$$P_i = \begin{bmatrix} x_i \\ y_i \\ z_i \\ 1 \end{bmatrix} \quad (5.4)$$

Where w_n , represents the weight of the control point. When control points are defined in homogeneous coordinates P_i^w , they maintain their influence on the curve shape but allow for the introduction of rationality through weight factors. This process doesn't alter the point position due to the properties of homogeneous coordinates. Each control point is then associated with a weight value w_n , and its coordinates are multiplied by this weight to give a new representation in homogeneous coordinates:

$$P_i^w = \begin{bmatrix} w_i x_i \\ w_i y_i \\ w_i z_i \\ w_i \end{bmatrix} \quad (5.5)$$

This new representation does not change the location of P_i since multiplying by a nonzero scalar in homogeneous coordinates results in the same point. Substituting this into the B-spline equation, then get $C^w(u)$ in homogeneous coordinates as follows:

$$C^w(u) = \sum_{i=0}^n N_{i,p}(u) P_i^w = \sum_{i=0}^n N_{i,p}(u) \begin{bmatrix} w_i x_i \\ w_i y_i \\ w_i z_i \\ w_i \end{bmatrix} = \begin{bmatrix} \sum_{i=0}^n N_{i,p}(u) (w_i x_i) \\ \sum_{i=0}^n N_{i,p}(u) (w_i y_i) \\ \sum_{i=0}^n N_{i,p}(u) (w_i z_i) \\ \sum_{i=0}^n N_{i,p}(u) w_i \end{bmatrix} \quad (5.6)$$

$C^w(u)$ is the representation of the B-spline curve in homogeneous coordinates. By dividing $C^w(u)$ by its fourth coordinate (the weight), the curve is transformed back into Cartesian coordinates. This leads to the final formula for the NURBS curve:

$$C(u) = \frac{1}{\sum_{i=0}^n N_{i,p}(u) w_i} \sum_{i=0}^n N_{i,p}(u) w_i P_i \quad (5.7)$$

The key components for defining a NURBS curve are:

- Control points, $P_0, P_1, P_2, \dots, P_n$, where $n+1$ is the number of control points.
- Weights, $w_0, w_1, w_2, \dots, w_n$, corresponding to each control point.
- The knot vector $u_0, u_1, u_2, \dots, u_m$, where $m+1$ is the number of knots.

The control points and weights define the shape of the NURBS curve.

Degree

The degree of the curve is an essential factor in defining its smoothness and flexibility. A higher degree allows for smoother transitions within the curve, but it also requires more control points. In NURBS, the degree affects the local control over the curve. The degree, in conjunction with the knot vector, defines the continuity of the curve at the knots. A higher degree results in greater continuity and smoothness.

Knots

Knots in a NURBS curve affect how the basis functions define the curve's geometry. They influence the curve's partitioning and, consequently, its shape.

The distribution of knots can create varying effects on the curve's behavior. A uniform knot vector results in a uniform influence across the entire curve, whereas a non-uniform knot vector allows for non-uniform changes in the curve's shape, offering localized control. This distinction is crucial for applications where precise control is needed, as it allows for adjustments that are specific to a region of the curve. Knot vectors can start and end with a multiplicity equal to the degree plus one, to ensure that the curve passes through the first and last control points.

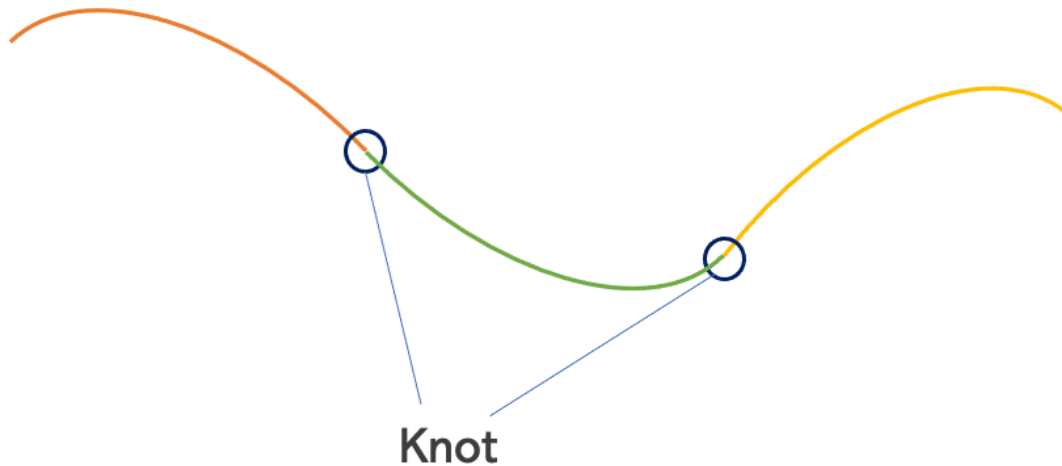


Figure 32. The illustration of knot

Control Points

Control points, positioned outside the curve, determine the shape of the curve and share the same concept as the control points in Bézier curves. However, unlike Bézier curves, the order of the B-spline polynomial can be independent of the number of control points.

Control Point Weights

The influence of control points on the curve or surface is determined by the weights assigned to them. A higher weight results in the curve or surface closely following the control point. Essentially, this alters the coefficients preceding the control point.

Editing points, all positioned above the curve, have the same number as control points. Editing points are calculated by averaging the knot values. For example, for a 3rd-order curve with knot vector $(0, 0, 0, 1, 2, 3, 3, 3)$, the editing point vector would be $(0, 1/3, 1, 2, 8/3, 3)$, potentially overlapping with the knots.

Both editing points and control points can modify the curve's shape. However, moving an editing point affects the entire curve's shape, while moving a control point only alters a portion of the curve's shape.

Relationship between Control Points, Knots, and Order: In B-spline curves, the number of control points equals the number of internal knots (excluding the two endpoints) plus the order plus one. The relationship between n , m , and p (degree) satisfies $m=n+p+1$.

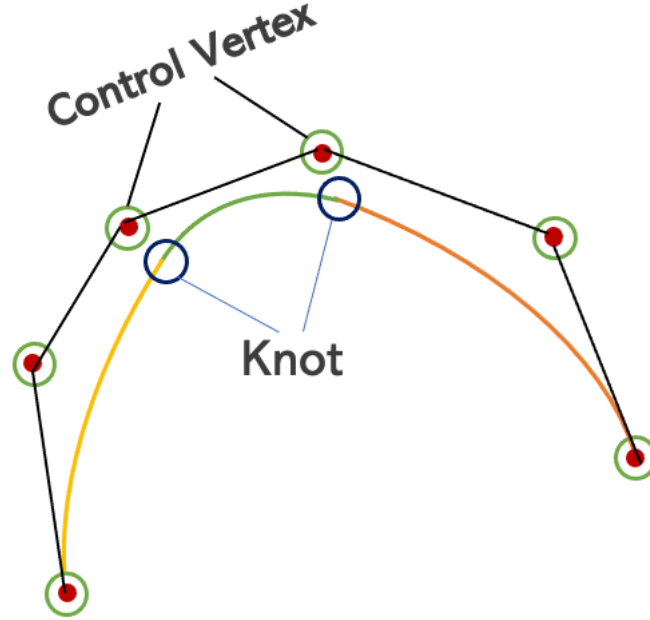


Figure 33. 3rd-order curve with 6 points

5.2.2 Optimization of Spinal Curve Reconstruction: From K-means Clustering to NURBS-net Prediction

Post-processing of segmentation result

The deep learning algorithm initially delineated the entire vertebral regions, encapsulating the vertebral body, transverse processes, and spinous processes. Owing to the pronounced structural variances present in scoliosis patients, the transverse and spinous processes might skew the accuracy of spinal curve estimation. To mitigate this issue, spoke kernel filtering was used to refine the segmentation.

Spoke kernel filtering can be classified as a method within the domain of image processing and computer vision. This method is used to enhance the accuracy of

segmentations by selectively filtering out specific structures based on geometric considerations. Here's a breakdown of the classification:

1. **Image Segmentation:** At its core, spoke kernel filtering is applied to segment images more precisely by identifying and removing unwanted elements that might otherwise skew the analysis. This places it firmly within image segmentation methods, which are fundamental to many medical imaging applications where accurate delineation of anatomical structures is crucial.

2. **Morphological Operations:** The technique could also be viewed as a form of advanced morphological operation. Morphological operations in image processing involve the manipulation of structures in an image, usually applied to binary images. This includes operations like dilation, erosion, opening, and closing, which help in removing noise, filling in gaps, and connecting disjoint elements. Spoke kernel filtering, while not a standard morphological operation, shares a similar goal of refining the image based on structural characteristics.

3. **Geometric Transformations:** The use of algorithms like the midpoint circle algorithm and the midpoint line algorithm suggests that spoke kernel filtering incorporates geometric transformations to create and manipulate shapes (circles and lines, respectively) within the image data. These transformations are crucial for accurately positioning the filtering kernel in relation to the anatomical structures of interest.

4. **Filtering Techniques:** Given that spoke kernel filtering involves the traversal of a kernel over an image to apply a specific filtering criterion (removing pixels that do not align with the spinal structure), it also falls under the category of filtering techniques. These are used widely in image processing to enhance image quality or extract meaningful information from an image based on predefined criteria.

5. Computational Geometry: The description also suggests an element of computational geometry, which deals with algorithms and techniques to solve geometric problems in computing. The systematic exclusion of certain pixels based on their geometric alignment with the spine involves computational geometric problems in computing. The systematic exclusion of certain pixels based on their geometric alignment with the spine involves computational geometric principles.

The spoke kernel filtering technique is adept at excluding both transverse and spinous processes and is instrumental in isolating adjacent vertebral regions that may appear conjoined due to severe scoliosis deformities. As depicted in Figure 1, the primary objective of the spoke kernel is the targeted removal of constricted regions akin to "bottlenecks." This method allows for a more precise separation of neighboring vertebrae as well as transverse and spinous processes.

Kernel Design: The spoke kernel is designed as a circular mask, created using the midpoint circle algorithm. This involves plotting points that form a circle based on a defined midpoint.

Line Generation: From this circular mask, line segments (spokes) are generated that radiate outward from the center to the circumference. These line segments are created using the midpoint line algorithm, which connects diametrically opposite points on the circle.

Segmentation Refinement: During the filtering process, the kernel is systematically moved along the spine within the segmentation mask of each vertebral column. As it moves, it checks whether the ends of each line segment align with expected spinal structures.

Exclusion of Non-spinal Elements: If the extremities of a line segment do not correspond to the spinal structure (i.e., they fall on the transverse or spinous processes), the pixels along this line segment are flagged for exclusion from the segmentation mask.

Cleanup: After a thorough scan, the identified non-spinal coordinates (pixels) are discarded, leading to a more accurate representation of the spinal anatomy, particularly by isolating adjacent vertebrae that may appear fused due to severe scoliosis deformities.

The primary purpose of the spoke kernel filtering is to facilitate a more precise separation of the vertebral bodies by removing regions that do not contribute to the accurate depiction of the spinal curve. This method is particularly crucial in complex cases where conventional segmentation might fail due to the presence of structural abnormalities like those seen in scoliosis.

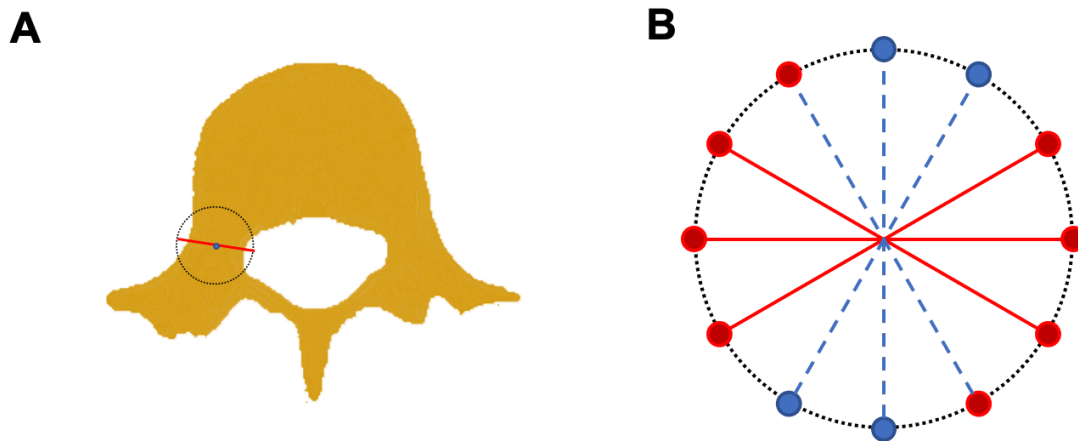


Figure 34 the Spoke Kernel Filtering process. (A) At the core of this technique is the placement of the spoke kernel's center upon each pixel within the spinal segmentation area. (B) Diameters extend from this central point, and endpoints are scrutinized. Should both endpoints reside outside the vertebral region (indicated by red points), the intermediate pixels will be excised from the segmentation mask. Conversely, if any endpoint is within the spinal boundary, the connecting pixels will be maintained (shown by blue dashed lines).

This filtration technique can be conceptualized as a specialized variant of the morphological opening operation, selectively excising narrow isthmuses that link expansive regions. It is important to note that while the boundary voxels might be

slightly reduced through spoke kernel filtering, the integrity of the vertebral shape is preserved, and contours are refined. The kernel's radius was initially set to 8mm, subject to manual adjustments to address variations in image scale. Post filtration, this approach ensures the elimination of diminutive, disconnected regions, employing a size threshold of 500 for removal.

NURBS Fitting for Ground Truth Enhancement

After the application of Spoke Kernel Filtering, the segmentation of vertebral domains became distinctly demarcated. To process each vertebral body, an enhanced K-means clustering algorithm was utilized to congregate voxels into clusters. The number of clusters was judiciously determined as 1/1000 of the total voxel count, allowing for a precise and manageable cluster size. Subsequently, a three-dimensional nonuniform rational B-spline (NURBS) fitting process was engaged. Here, cluster centroids provided the necessary data points for the application of the least squares fitting method. For this fitting process, cubic B-spline basis functions with a degree of three were utilized. This selection ensures a balance between curve flexibility and control. A provisional assignment of 8 control points was established to define the curve, creating an initial structure for the ground truth utilized in the subsequent deep learning framework.

Refined Approach to Deep Learning-based Curve Fitting

Post-segmentation, the NURBS-net was employed to infer the NURBS control points and knot vectors from the spinal segmentation results obtained before. The NURBS-net, structured upon the robust framework of ResNet (refer to Figure 32 for architectural details), underwent an extensive training regime utilizing the same dataset as vertebra segmentation. During this phase, the spine segmentation and the “Ground-truth curve” were subjected to random affine transformations, a strategic move to diversify the training data and thereby enhance the robustness of the model.

One of the most important note in the curve fitting process is the loss function, meticulously crafted to compare the ground truth control points and knot vectors with those predicted by the model. The loss function is composed as follows:

$$loss = \frac{\|ctrlpts_{GT} - ctrlpts_{pred}\|_{L2}^2}{\|ctrlpts_{GT}\|_{L2}^2} + \frac{\|knots_{GT} - knots_{pred}\|_{L2}^2}{\|knots_{GT}\|_{L2}^2} \quad (5.8)$$

Here, 'ctrlpts' and 'knots' represent the control points and knot vectors, while 'ctrlpts' and 'knots' denote the ground truth and predicted values, respectively. This equation ensures that the deviations in both the control points and knot vectors are normed by the ground truth's magnitude, emphasizing proportional accuracy.

This evaluation strategy, characterized by a learning rate of 0.001, paved the way for a nuanced understanding over a span of 1000 epochs. To assess the precision between the ground truth and the model's predicted curves, the symmetric mean of minimum distances was employed, providing a symmetric assessment of spatial deviations between the curves.

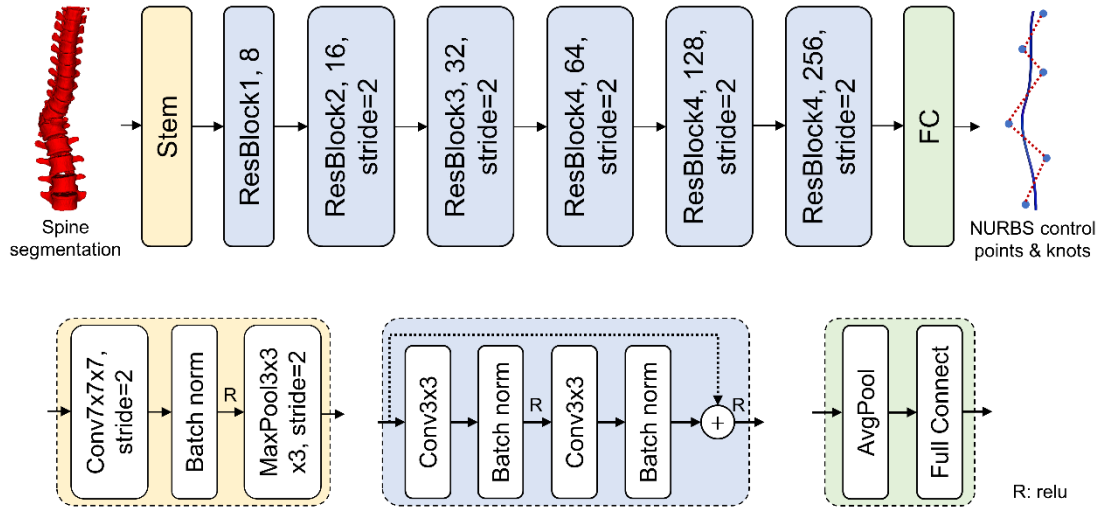


Figure 35. outlines the deep learning network specifically architected for spinal curve fitting. It employs ResNet as its backbone, processing spine segmentation as input and yielding control points and knots of NURBS curves as outputs.

5.2.3 Measurement of 3D spatial angle

Utilizing the predicted NURBS control points and knot vectors, synthesized deep learning-based curves, henceforth termed "NURBS-curves," comprising 100 discrete points accompanied by their respective tangent vectors for each patient's spinal curvature representation. To ensure accuracy, the pairwise angles between these tangent vectors were meticulously calculated over a central range—from the 5th to the 95th points—excluding the extremities of the curves. This consideration stems from the observation that the terminal points of the curves might be influenced by incomplete vertebral structures often encountered at the top and bottom ends in CT images, leading to potential distortions in angle measurement.

To ascertain the most significant 3D angle within this range, a systematic traversal along the NURBS-curve was conducted. For every pair of points, denoted by p and q , the angle between their corresponding tangent vectors T_p and T_q was computed using the arccosine of their dot product, which mathematically is described by the formula:

$$Angle(p, q) = \arccos(T_p \cdot T_q). \quad (5.9)$$

A 2D "angle matrix" was subsequently constructed, where the pixel intensity at the (i, j) position corresponded to the angle between the i-th and j-th curve points. To enhance the stability of the angle measurements, this matrix was then processed using a Gaussian smoothing filter. In this section, two types of 3D spatial angle were obtained:

The predicted 3D Cobb angle (PRED-3D-Cobb angle): which is the maximum value in the smoothed angle map derived from the NURBS curve.

The 2D mapping Cobb angle (MAP-2D-Cobb angle): which is the maximal angle formed by the tangent vectors along the projected 2D spinal curve.

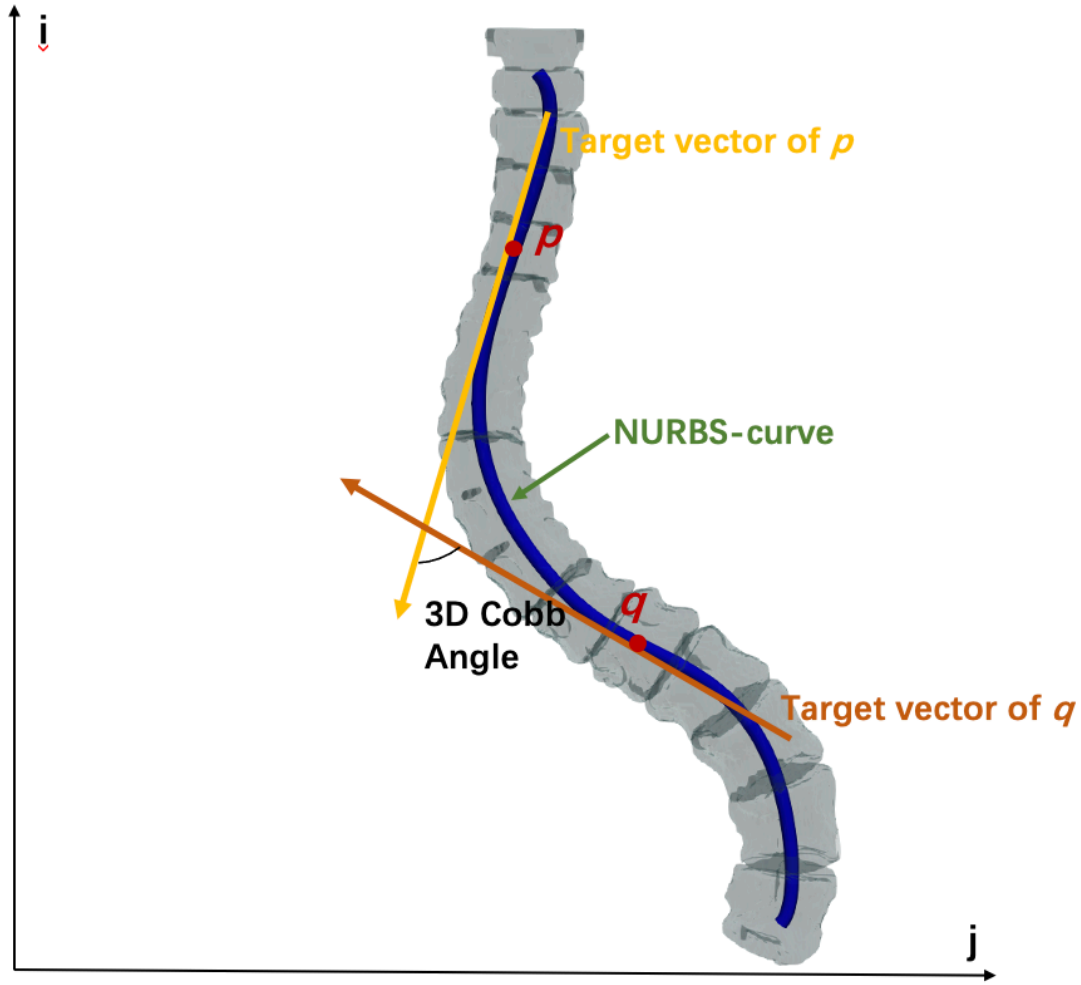


Figure 36. Illustration of the PRED-3D-CA calculated method

In-Depth Statistical Analysis

The statistical analysis was conducted with precision using SPSS software (version 21, IBM, Chicago, IL, USA), where a threshold of 0.05 were used for determining statistical significance. The established benchmark for spinal curvature, known as the traditional 2D Cobb angle and annotated by surgeons from 2D radiographic images, is herein referred to as “XRAY-CA”.

The Pearson correlation coefficient, a measure of linear correlation, was implemented to delve into the relationships between the deep learning-based angle measurements

(PRED-3D-CA and MAP-2D-CA) and the expert annotations classified under XRAY-CA. This statistical method provided insights into how well the automated measures aligned with the conventional manual annotations performed by surgeons.

Further, to gauge the precision and accuracy of the deep learning model, the mean absolute error (MAE) and standard deviation (SD) were calculated. The MAE served as a clear indicator of the average magnitude of errors between the XRAY-CA and our model's outputs without considering their direction. The SD provided a snapshot of the variance in these differences, informing us of the consistency of our model.

To complement these metrics, Bland-Altman plots were generated, offering a visual representation of the agreement between the XRAY-CA and the automated measures. These plots are particularly valuable as they visually articulate not just the average difference but also the range of agreement across the spectrum of measurements, allowing us to detect any patterns or biases.

Through these comprehensive statistical evaluations, this section sought to establish a robust validation of the deep learning-based measurements, ensuring that they serve as reliable, objective, and reproducible methods in clinical settings for assessing spinal deformities.

5.3 Results

Efficacy of Spoke Kernel Filtering and Subsequent Curve Reconstruction

The spoke kernel filtering technique proved to be highly effective in delineating vertebral bodies from surrounding anatomical structures, as visualized in Figure 3. For the majority of the patient cohort (88%), a spoke kernel with an 8 mm radius yielded optimal segmentation results. A subset of patients—7 with smaller spinal structures—required a finer 6 mm radius for adequate separation, whereas 4 patients with more expansive anatomical features necessitated larger radii of 10 or 12 mm. This adaptive

approach, involving interactive determination of the kernel radius based on individual patient anatomy, ensured precision in the segmentation process.

Upon successful segmentation, NURBS-net was employed to reconstruct the NURBS-curves, which were then accurately positioned at the spinal region's center (shown in Figure 32) before. Comparative analysis between these reconstructed curves and the ground-truth data revealed an average discrepancy of 2.81 ± 1.09 mm within the training set and a slightly greater divergence of 3.65 ± 1.45 mm within the testing set.

Visualization and Comparison of Curve Fitting Results

Figures 33 provide visual summaries of the segmentation and post-processing outcomes and highlights the meticulous post-processing and fitting of NURBS-curves, contrasting morphology-based curves with the deep learning-derived reconstructions for clarity.

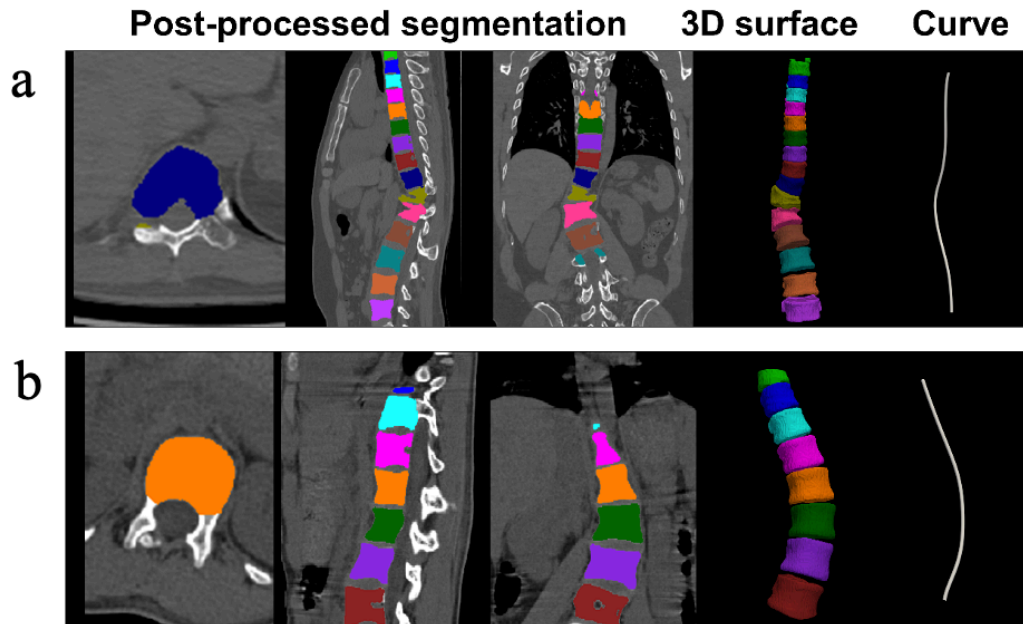


Figure 37. Post-processing and curve fitting. (a) and (b) are post-processing results of two patients from the training dataset.

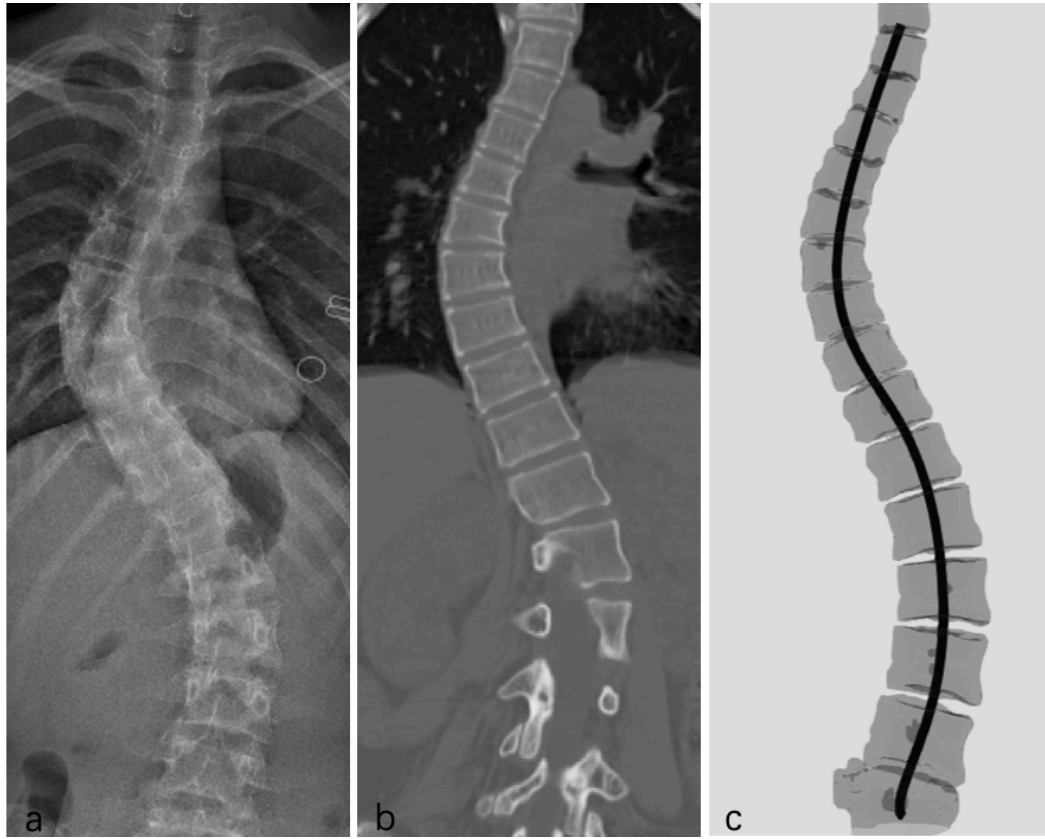


Figure 38. (a) The X-ray, (b) CT and (c) NURBS-curve of one patient from the testing dataset.

Figure 35 presents a comprehensive visual set comprising an X-ray, CT, and the reconstructed NURBS-curve from a testing dataset patient, offering a holistic view of the diverse imaging modalities and the computational reconstruction process.

Table 7 offers a comparative perspective, juxtaposing the traditional XRAY-CA measurements against the deep learning-derived PRED-3D-CA and MAP-2D-CA. Notably, the mean Cobb angle measured radiographically was $56.8^{\circ} \pm 12.8^{\circ}$, reflecting significant variability among patients. The computational techniques, however, demonstrated remarkable concordance with the radiographic measurements: PRED-3D-CA at $55.7^{\circ} \pm 13.8^{\circ}$ and MAP-2D-CA at $56.7^{\circ} \pm 13.8^{\circ}$, underpinning the reliability of the deep learning methods.

The correlation analysis reinforced these findings, with the Pearson Correlation Coefficient reaching 0.983 ($p<0.05$) for PRED-3D-CA and 0.934 ($p<0.05$) for MAP-2D-CA when compared to XRAY-CA, signifying a high degree of alignment.

Insights from Error Metrics and Agreement Analysis

The Mean Absolute Error (MAE) further substantiated the models' accuracy, with PRED-3D-CA showing a lower MAE of 2.4 ± 2.6 degrees, signifying a tighter congruence with XRAY-CA, as opposed to MAP-2D-CA's MAE of 4.1 ± 4.9 degrees.

Figure 35 delineates the Bland-Altman plot, which elucidates the agreement between XRAY-CA and computational predictions. The plot for XRAY-CA versus PRED-3D-CA showed a mean difference of 1.1 degrees, with the limits of agreement tightly bound between -4.0 and 6.2 degrees, denoting a substantial agreement. Conversely, XRAY-CA versus MAP-2D-CA revealed a mean difference of 0.2 degrees, albeit with a wider agreement range from -9.5 to 9.8 degrees, suggesting greater variability in MAP-2D-CA measurements.

Data Category	XRAY-CA	PRED-3D-CA	MAP-2D-CA
Mean Cobb (mean \pm SD)	56.8 $^{\circ}$ \pm 12.8 $^{\circ}$	55.7 $^{\circ}$ \pm 13.8 $^{\circ}$	56.7 $^{\circ}$ \pm 13.8 $^{\circ}$
Pearson Correlation coefficient	-	0.983 ($p<0.05$)	0.934 ($p<0.05$)
MAE \pm SD	-	2.4 \pm 2.6	4.1 \pm 4.9

SD standard deviation, *MAE* Mean Absolute Error

Table 9. Difference between XRAY-CA, PRED-3D-CA and MAP-2D-CA

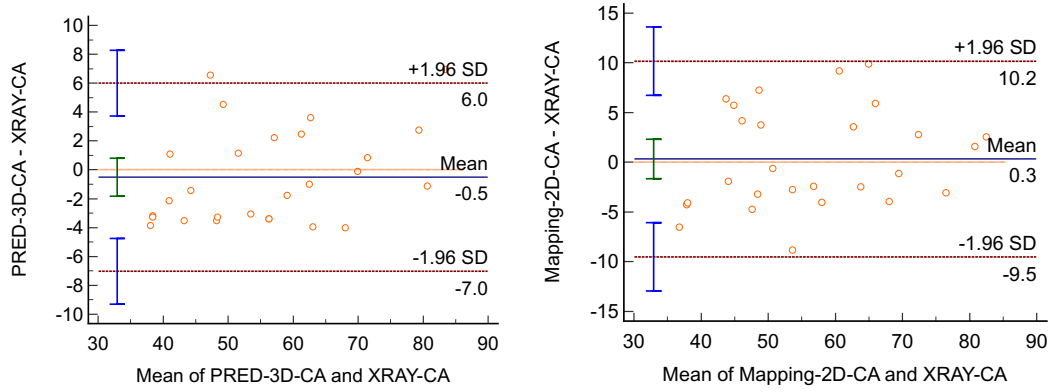


Figure 39. The Bland–Altman plot assessing the difference of the Cobb angle measurements, i.e., PRED-3D-CA, MAP-2D-CA and XRAY-CA.

5.4 Discussion

This chapter presents an innovative automated methodology for evaluating the 3D spatial angle in patients with idiopathic scoliosis, leveraging CT imaging data. Based on the vertebra segmentation result, a subsequent deep learning model, NURBS-net were used for precise spinal curve fitting. Two variants of the 3D spatial angle were derived from the NURBS-curves, and the practicality of this method was validated through its application to preoperative CT scans of 27 patients scheduled for spinal fusion surgery.

As a critical preparatory step, Spoke Kernel Filtering was introduced to refine segmentation outcomes for subsequent curve fitting, meticulously isolating the vertebral body from contiguous osseous structures. This filtering technique discerns between vertebral bodies and proximate anatomical features, such as the lamina and spinous process, by verifying the alignment of distant pixels with the vertebral segmentation. This ensures that the fitted NURBS curve closely mirrors the vertebral body's contour, minimizing the impact of adjacent spinal sections.

NURBS curve fitting, informed by K-means clustering on the segmentation mask, simplifies the representation of spinal trajectories, generating smooth curves that absorb

minor segmentation deviations. This integration of morphological analysis with NURBS curve fitting constitutes a comprehensive protocol for spinal curve assessment, enabling efficient and streamlined calculations.

Historically, 3D spatial angle estimations fall into two principal categories: “endplate-based” and “curve-based” methods. Endplate-based approaches calculate the maximum angle across vertebral endplates from 3D segmentations, whereas curve-based techniques derive a spinal curve from vertebral segmentations. Our proposed method is categorized within the latter but differentiates itself by utilizing deep learning to determine control points and knots. This reduces the predicted points, mitigating the risk of overfitting that could occur with curve-based methods reliant on vertebral centroids.

Despite its innovative contributions, this study is subject to several limitations. The validation was constrained to a modest cohort of internal samples rather than a broad, external dataset, predominantly due to the challenges in obtaining comprehensive CT and X-ray data for AIS patients. Inconsistencies in imaging perspectives between CT and X-ray led to a narrowed sample size, introducing potential biases. Moreover, the comparison of our 3D spatial angle results was limited to 2D X-ray-derived Cobb angles, acknowledging the inherent differences in patient positioning and imaging modalities between CT scans and 2D X-rays. While a MAP-2D-CA was derived to approximate the 2D X-ray measurement, it inherently remains a 3D assessment, warranting further validation against other 3D spatial angle methodologies.

Lastly, the study did not differentiate among various scoliosis types or curve locations and omitted a subgroup analysis that could offer nuanced insights into the condition's heterogeneity. Future investigations could address these limitations by including larger, more diverse datasets and by comparing results across different scoliosis classifications.

5.5 Conclusion

In the current study, this chapter introduced a computational technique to assess the three-dimensional Cobb angle from CT images for patients afflicted with severe scoliosis. Our results demonstrate that this innovative method can quantify the 3D spatial angle with a degree of accuracy that shows significant consistency when compared to the traditional X-Ray measurements, particularly in patients who are candidates for surgical intervention. The computational model not only aligns with the traditional two-dimensional assessments but provide a more comprehensive three-dimensional spinal analysis.

Future work will focus not only on expanding the in-house dataset to improve neural network performance but also on providing comprehensive spatial information, including the precise location and affected vertebral level of the maximum curvature, rather than just a single 3D spatial angle value, to enhance the clinical utility of the model.

CHAPTER 6 CONCLUSION

This thesis delves into the complex application of advanced deep learning techniques for assessing AIS by measuring 3D spatial angles from CT images. The framework was developed to improve the accuracy of preoperative assessment and potentially improve surgical outcomes by visualizing the spinal deformity in greater detail, beyond the capabilities of traditional 2D measurements.

Firstly, a classical neural network called U-net was used to do the vertebra segmentation in in-house dataset. Then developed nnformer++ with the combination of two most advanced networks in medical image segmentation - U-Net and Transformer - to initially segment the spinal structure based on both in-house dataset and public dataset. Subsequent spine curve fitting using a customized deep learning model, NURBS-net, generates a NURBS curve of the spine. Based on this curve, 3D spatial angles are accurately calculated at the maximum value in a smooth angle map, aiming to automate and refine what has traditionally been a labor-intensive and error-prone process.

The efficacy of these models is demonstrated by their ability to effectively segment the spine, capturing vertebrae that are often difficult to accurately delineate. The introduction of spoke kernel filtering further enhances the differentiation and focusing of vertebral regions, improving the overall accuracy of the spine curve analysis. The Cobb angle measurements derived from the deep learning approach correlated strongly with the measurements labeled by the surgeon on the 2D radiographs, achieving a high Pearson's correlation coefficient, highlighting the accuracy and reliability of the model.

Despite these advances, this study encountered some significant limitations. Foremost among these is the reliance on a limited internal dataset for validation rather than a larger and more diverse external dataset. This limitation stems primarily from the difficulty of obtaining comprehensive, publicly available datasets that include CT data and radiographs from patients with AIS. In addition, the study did not differentiate between various types of scoliosis and did not perform subgroup analyses, which may

be more revealing of the specific applicability and validity of 3D measurements in different scoliosis.

In future work, it will be essential not only to expand the in-house dataset to further enhance the performance of the neural network but also to provide more comprehensive information for clinical treatment. The current study yields only a single 3D spatial angle value without indicating the precise location of the maximum curvature or identifying the specific vertebral level affected. However, such information would be more intuitive and valuable for clinical decision-making. Moving forward, the goal is to provide not only the maximum three-dimensional curvature but also its spatial location to assist clinicians in better understanding the condition.

In summary, by combining deep learning models with traditional medical imaging techniques, this thesis demonstrates the potential to significantly improve the accuracy of scoliosis assessment. The transition from traditional 2D assessment to more comprehensive 3D analysis, facilitated by this research, promises better patient outcomes in the management of scoliosis. Further research is essential, which will require not only technological innovation but also collaboration in data sharing and method development.

Reference

1. Konieczny, M.R., H. Senyurt, and R. Krauspe, *Epidemiology of adolescent idiopathic scoliosis*. Journal of children's orthopaedics, 2013. **7**(1): p. 3-9.
2. Kamtsiuris, P., et al., *Prevalence of somatic diseases in German children and adolescents. Results of the German Health Interview and Examination Survey for Children and Adolescents (KiGGS)*. Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz, 2007. **50**(5-6): p. 686-700.
3. Daruwalla, J.S., *IDIOPATHIC SCOLIOSIS (PREVALENCE AND ETHNIC DISTRIBUTION IN SINGAPORE SCHOOLCHILDREN)*. Journal of pediatric orthopaedics, 1985. **5**(5): p. 621.
4. Balagun, F. and F. Pellisn, *Adolescent idiopathic scoliosis and back pain*. Scoliosis, 2016. **11**(1): p. 27-27.
5. Yagi, M., M. Takemitsu, and M. Machida, *Clavicle chest cage angle difference (CCAD): a novel predictor of postoperative shoulder imbalance in patients with adolescent idiopathic scoliosis*. Spine (Philadelphia, Pa. 1976), 2013. **38**(12): p. E705-E712.
6. Agabegi, S.S., et al., *Natural History of Adolescent Idiopathic Scoliosis in Skeletally Mature Patients: A Critical Review*. Journal of the American Academy of Orthopaedic Surgeons, 2015. **23**(12): p. 714-723.
7. Liu, L., et al., *Prevalence of cardiac dysfunction and abnormalities in patients with adolescent idiopathic scoliosis requiring surgery*. Orthopedics (Thorofare, N.J.), 2010. **33**(12): p. 882-882.
8. Mitsiaki, I., et al., *Adolescent Idiopathic Scoliosis and Mental Health Disorders: A Narrative Review of the Literature*. Children (Basel), 2022. **9**(5): p. 597.
9. Wang, H., et al., *Mental health of patients with adolescent idiopathic scoliosis and their parents in China: a cross-sectional survey*. BMC psychiatry, 2019. **19**(1): p. 147-147.
10. Jada, A., et al., *Evaluation and management of adolescent idiopathic scoliosis: a review*. Neurosurgical focus, 2017. **43**(4): p. E2-E2.
11. Pasha, S., et al., *Characterizing the differences between the 2D and 3D measurements of spine in adolescent idiopathic scoliosis*. European spine journal, 2016. **25**(10): p. 3137-3145.
12. Pasha, S., *3D spinal and rib cage predictors of brace effectiveness in adolescent idiopathic scoliosis*. BMC musculoskeletal disorders, 2019. **20**(1): p. 384-384.
13. Wu, H.D., et al., *Assessment of the plane of maximum curvature for patients with adolescent idiopathic scoliosis via computed tomography*. 2020.
14. Vo, Q.N., L.H. Le, and E. Lou, *A semi-automatic 3D ultrasound reconstruction method to assess the true severity of adolescent idiopathic scoliosis*. Medical & Biological Engineering & Computing, 2019. **57**(10): p. 2115-2128.
15. Shen, J., et al., *Geometric Torsion in Adolescent Idiopathic Scoliosis: A New Method to Differentiate Between Lenke 1 Subtypes*. Spine, 2017. **42**(9).
16. Cobetto, N., et al., *3D correction of AIS in braces designed using CAD/CAM and FEM: a randomized controlled trial*. Scoliosis and Spinal Disorders, 2017. **12**(1): p. 24.
17. Duvenaud, D., et al., *Convolutional Networks on Graphs for Learning Molecular Fingerprints*. 2015.

18. Li, G., et al., *DeepGCNs: Making GCNs Go as Deep as CNNs*. IEEE transactions on pattern analysis and machine intelligence, 2021. **PP**: p. 1-1.
19. Atwood, J. and D. Towsley, *Diffusion-Convolutional Neural Networks*. 2015.
20. Garcia, V. and J. Bruna, *Few-Shot Learning with Graph Neural Networks*. 2017.
21. Liu, Y., et al., *Learning to Propagate Labels: Transductive Propagation Network for Few-shot Learning*. 2018.
22. Kim, J., et al., *Edge-labeling Graph Neural Network for Few-shot Learning*. 2019.
23. Cheng, H., et al., *Attentive Graph Neural Networks for Few-Shot Learning*. 2020.
24. Bachmann, K.R., et al., *The Lumbosacral Takeoff Angle Can Be Used to Predict the Postoperative Lumbar Cobb Angle Following Selective Thoracic Fusion in Patients with Adolescent Idiopathic Scoliosis*. J Bone Joint Surg Am, 2020. **102**(2): p. 143-150.
25. Chen, I.H., et al., *Simultaneous Hypercorrection of Lowest Instrumented Vertebral Tilt and Main Thoracic Curve is Associated With Progression of Residual Lumbar Curve in Adolescent Idiopathic Scoliosis*. Spine (Phila Pa 1976), 2022. **47**(19): p. 1362-1371.
26. Wan, S.H., et al., *Patient and surgical predictors of 3D correction in posterior spinal fusion: a systematic review*. Eur Spine J, 2023. **32**(6): p. 1927-1946.
27. Bai, J., et al., *Proximal junctional kyphosis in Lenke 5C adolescent idiopathic scoliosis after selective posterior thoracolumbar/lumbar fusion: risk factors and predictive index*. J Orthop Surg Res, 2024. **19**(1): p. 24.
28. Shi, B., et al., *Accurate prediction of height loss in adolescent idiopathic scoliosis: Cobb angle alone is insufficient*. European Spine Journal, 2016. **25**(10): p. 3341-3346.
29. van Popta, D., J. Stephenson, and R. Verma, *Change in spinal height following correction of adolescent idiopathic scoliosis*. The Spine Journal, 2016. **16**(2): p. 199-203.
30. Hwang, S.W., et al., *A multicenter analysis of factors associated with change in height after adolescent idiopathic scoliosis deformity surgery in 447 patients*. Journal of Neurosurgery: Spine, 2013. **18**(3): p. 298-302.
31. Kwan, M.K., et al., *Perioperative Outcome of Single Stage Posterior Spinal Fusion for Severe Adolescent Idiopathic Scoliosis (AIS) (Cobb Angle $\geq 90^\circ$): The Role of a Dual Attending Surgeon Strategy*. Spine, 2019. **44**(6): p. E348-E356.
32. Mehta, N., et al., *Predictors of Operative Duration in Posterior Spinal Fusion for Adolescent Idiopathic Scoliosis: A Retrospective Cohort Study*. International Journal of Spine Surgery, 2022. **16**(3): p. 559-566.
33. Soliman, H.A.G., et al., *Predicting lowest hemoglobin level and risk of blood transfusion in spinal fusion surgery for adolescent idiopathic scoliosis*. European Spine Journal, 2019. **28**(6): p. 1342-1348.
34. Yang, C., et al., *Adding-on Phenomenon After Surgery in Lenke Type 1, 2 Adolescent Idiopathic Scoliosis: Is it Predictable?* Spine, 2016. **41**(8): p. 698-704.
35. Schulz, J., et al., *Optimal Radiographical Criteria After Selective Thoracic Fusion for Patients With Adolescent Idiopathic Scoliosis With a C Lumbar Modifier: Does Adherence to Current Guidelines Predict Success?* Spine, 2014. **39**(23): p. E1368-E1373.
36. Nugent, M., et al., *Influence of curve magnitude and other variables on operative time, blood loss and transfusion requirements in adolescent idiopathic scoliosis*. Irish Journal of Medical Science (1971 -), 2016. **185**(2): p. 513-520.
37. Yu, X., et al., *Prediction of Massive Blood Loss in Scoliosis Surgery From Preoperative*

- Variables*. Spine, 2013. **38**(4).
38. Mimura, T., et al., *Can surgery for adolescent idiopathic scoliosis of less than 50 degrees of main thoracic curve achieve good results?* Journal of Orthopaedic Science, 2018. **23**(1): p. 14-19.
 39. Jin, C., et al., *A Review of the Methods on Cobb Angle Measurements for Spinal Curvature*. Sensors (Basel), 2022. **22**(9).
 40. Huo, X., et al., *An Integrative Framework for 3D Cobb Angle Measurement on CT Images*. Computers in biology and medicine, 2017. **82**: p. 111-118.
 41. Alukaev, D., et al., *A deep learning framework for vertebral morphometry and Cobb angle measurement with external validation*. European spine journal, 2022. **31**(8): p. 2115-2124.
 42. Wang, C., et al., *Deep learning model for measuring the sagittal Cobb angle on cervical spine computed tomography*. BMC Med Imaging, 2023. **23**(1): p. 196.
 43. Mahadevan, V., *Anatomy of the vertebral column*. Surgery (Oxford), 2018. **36**(7): p. 327-332.
 44. Zheng, Y., Y. Dang, X. Wu, Y. Yang, J. Reinhardt, C. He D., and M. Wong., *Epidemiological Study of Adolescent Idiopathic Scoliosis in Eastern China*. Journal of Rehabilitation Medicine., May 2017. **vol. 49**(no. 6): p. 512-9.
 45. Hu, M., et al., *Prevalence and determinants of adolescent idiopathic scoliosis from school screening in Huangpu district, Shanghai, China*. Am J Transl Res, 2022. **14**(6): p. 4132-4138.
 46. Penha, P.J., et al., *Prevalence of Adolescent Idiopathic Scoliosis in the State of São Paulo, Brazil*. Spine, 2018. **43**(24).
 47. Yilmaz, H., et al., *Prevalence of adolescent idiopathic scoliosis in Turkey: an epidemiological study*. The Spine Journal, 2020. **20**(6): p. 947-955.
 48. Kamtsiuris, P., et al., *Prävalenz von somatischen Erkrankungen bei Kindern und Jugendlichen in Deutschland*. Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz, 2007. **50**(5): p. 686-700.
 49. Suh, S.-W., et al., *Idiopathic scoliosis in Korean schoolchildren: a prospective screening study of over 1 million children*. European Spine Journal, 2011. **20**(7): p. 1087-1094.
 50. Daruwalla, J.S., et al., *Idiopathic scoliosis. Prevalence and ethnic distribution in Singapore schoolchildren*. The Journal of Bone & Joint Surgery British Volume, 1985. **67-B**(2): p. 182-184.
 51. Thomas, J.J., et al., *Trends in Incidence of Adolescent Idiopathic Scoliosis: A Modern US Population-based Study*. J Pediatr Orthop, 2021. **41**(6): p. 327-332.
 52. Tarrant, R.C., et al., *Prevalence and impact of low body mass index on outcomes in patients with adolescent idiopathic scoliosis: a systematic review*. European Journal of Clinical Nutrition, 2018. **72**(11): p. 1463-1484.
 53. Bondar, K., et al., *The Demographics and Epidemiology of Infantile, Juvenile, and Adolescent Idiopathic Scoliosis in a Southern California Integrated Health Care System*. Spine, 2021. **46**(21).
 54. Goldberg, C.J., et al., *Adolescent idiopathic scoliosis: natural history and prognosis*. Stud Health Technol Inform, 2002. **91**: p. 59-63.
 55. Weinstein, S.L., et al., *Adolescent idiopathic scoliosis*. Lancet, 2008. **371**(9623): p. 1527-37.
 56. Risser, J.C., *The Iliac apophysis; an invaluable sign in the management of scoliosis*. Clin

- Orthop, 1958. **11**: p. 111-9.
57. Sanders, J.O., *Maturity indicators in spinal deformity*. J Bone Joint Surg Am, 2007. **89 Suppl 1**: p. 14-20.
 58. Lonstein, J.E. and J.M. Carlson, *The prediction of curve progression in untreated idiopathic scoliosis during growth*. J Bone Joint Surg Am, 1984. **66**(7): p. 1061-71.
 59. Weinstein, S.L. and I.V. Ponseti, *Curve progression in idiopathic scoliosis*. J Bone Joint Surg Am, 1983. **65**(4): p. 447-55.
 60. Schwab, F., et al., *Adult scoliosis: a health assessment analysis by SF-36*. Spine (Phila Pa 1976), 2003. **28**(6): p. 602-6.
 61. Kostuik, J.P. and J. Bentivoglio, *The incidence of low back pain in adult scoliosis*. Acta Orthop Belg, 1981. **47**(4-5): p. 548-59.
 62. Reamy, B.V. and J.B. Slakey, *Adolescent idiopathic scoliosis: review and current concepts*. Am Fam Physician, 2001. **64**(1): p. 111-6.
 63. Côté, P., et al., *A study of the diagnostic accuracy and reliability of the Scoliometer and Adam's forward bend test*. Spine (Phila Pa 1976), 1998. **23**(7): p. 796-802; discussion 803.
 64. Bunnell, W.P., *Outcome of spinal screening*. Spine (Phila Pa 1976), 1993. **18**(12): p. 1572-80.
 65. Margalit, A., et al., *Body Mass Hides the Curve: Thoracic Scoliometer Readings Vary by Body Mass Index Value*. J Pediatr Orthop, 2017. **37**(4): p. e255-e260.
 66. Ashworth, M.A., et al., *Scoliosis Screening An Approach to Cost/Benefit Analysis*. Spine, 1988. **13**(10).
 67. Horne, J.P., R. Flannery, and S. Usman, *Adolescent idiopathic scoliosis: diagnosis and management*. Am Fam Physician, 2014. **89**(3): p. 193-8.
 68. Langensiepen, S., et al., *Measuring procedures to determine the Cobb angle in idiopathic scoliosis: a systematic review*. Eur Spine J, 2013. **22**(11): p. 2360-71.
 69. Greiner, K.A., *Adolescent idiopathic scoliosis: radiologic decision-making*. Am Fam Physician, 2002. **65**(9): p. 1817-22.
 70. King, H.A., et al., *The selection of fusion levels in thoracic idiopathic scoliosis*. JBJS, 1983. **65**(9).
 71. Richards, B.S., *Lumbar curve response in type II idiopathic scoliosis after posterior instrumentation of the thoracic curve*. Spine (Phila Pa 1976), 1992. **17**(8 Suppl): p. S282-6.
 72. Roye, D.P., Jr., et al., *Results of spinal instrumentation of adolescent idiopathic scoliosis by King type*. Spine (Phila Pa 1976), 1992. **17**(8 Suppl): p. S270-3.
 73. Knapp, D.R., Jr., et al., *Choosing Fusion Levels in Progressive Thoracic Idiopathic Scoliosis*. Spine, 1992. **17**(10).
 74. Cummings, R.J., et al., *Interobserver Reliability and Intraobserver Reproducibility of the System of King et al. for the Classification of Adolescent Idiopathic Scoliosis**. JBJS, 1998. **80**(8).
 75. Behensky, H., et al., *MULTI SURGEON ASSESSMENT OF CORONAL PATTERN CLASSIFICATION SYSTEMS OF ADOLESCENT IDIOPATHIC SCOLIOSIS*. Orthopaedic Proceedings, 2004. **86-B**(SUPP_II): p. 113-113.
 76. Lenke, L.G., et al., *Multisurgeon Assessment of Surgical Decision-Making in Adolescent Idiopathic Scoliosis: Curve Classification, Operative Approach, and Fusion Levels*. Spine,

2001. **26**(21).
77. Pasha, S. and J. Flynn, *Data-driven Classification of the 3D Spinal Curve in Adolescent Idiopathic Scoliosis with an Applications in Surgical Outcome Prediction*. Scientific Reports, 2018. **8**(1): p. 16296.
 78. Abelin-Genevois, K., et al., *Sagittal classification in adolescent idiopathic scoliosis: original description and therapeutic implications*. European Spine Journal, 2018. **27**(9): p. 2192-2202.
 79. Pasha, S., J.M. Flynn, and W.N. Sankar, *Outcomes of selective thoracic fusion for Lenke 1 adolescent idiopathic scoliosis: predictors of success from the sagittal plane*. European Spine Journal, 2018. **27**(9): p. 2223-2232.
 80. Cahill, P.J., et al., *Youth and Experience: The Effect of Surgeon Experience on Outcomes in Cerebral Palsy Scoliosis Surgery*. Spine Deformity, 2018. **6**(1): p. 54-59.
 81. Pasha, S., et al., *Three-Dimensional Spinopelvic Relative Alignment in Adolescent Idiopathic Scoliosis*. Spine, 2014. **39**(7).
 82. Birchall, D., et al., *Demonstration of vertebral and disc mechanical torsion in adolescent idiopathic scoliosis using three-dimensional MR imaging*. European Spine Journal, 2005. **14**(2): p. 123-129.
 83. Kadoury, S., J. Shen, and S. Parent, *Global geometric torsion estimation in adolescent idiopathic scoliosis*. Medical & Biological Engineering & Computing, 2014. **52**(4): p. 309-319.
 84. Shen, J., et al., *Geometric Torsion in Adolescent Idiopathic Scoliosis: A Surgical Outcomes Study of Lenke Type 1 Patients*. Spine, 2016. **41**(24).
 85. Atmaca, H., et al., *Axial plane analysis of Lenke 1A adolescent idiopathic scoliosis as an aid to identify curve characteristics*. The Spine Journal, 2014. **14**(10): p. 2425-2433.
 86. Courvoisier, A., et al., *Transverse plane 3D analysis of mild scoliosis*. European Spine Journal, 2013. **22**(11): p. 2427-2432.
 87. Shen, J., et al., *Towards a new 3D classification for adolescent idiopathic scoliosis*. Spine Deformity, 2020. **8**(3): p. 387-396.
 88. Duong, L., F. Cheriet, and H. Labelle, *Three-Dimensional Classification of Spinal Deformities Using Fuzzy Clustering*. Spine, 2006. **31**(8).
 89. Stokes, I.A., A.P. Sangole, and C.E. Aubin, *Classification of scoliosis deformity three-dimensional spinal shape by cluster analysis*. Spine (Phila Pa 1976), 2009. **34**(6): p. 584-90.
 90. Donzelli, S., et al., *State of the art of current 3-D scoliosis classifications: a systematic review from a clinical perspective*. Journal of NeuroEngineering and Rehabilitation, 2015. **12**(1): p. 91.
 91. Duong, L., et al., *Three-dimensional Subclassification of Lenke Type 1 Scoliotic Curves*. Clinical Spine Surgery, 2009. **22**(2).
 92. Kadoury, S. and H. Labelle, *Classification of three-dimensional thoracic deformities in adolescent idiopathic scoliosis from a multivariate analysis*. European Spine Journal, 2012. **21**(1): p. 40-49.
 93. Pasha, S., et al., *A hierarchical classification of adolescent idiopathic scoliosis: Identifying the distinguishing features in 3D spinal deformities*. PLOS ONE, 2019. **14**(3): p. e0213406.
 94. Girdler, S., et al., *Emerging Techniques in Diagnostic Imaging for Idiopathic Scoliosis in Children and Adolescents: A Review of the Literature*. World Neurosurgery, 2020. **136**: p.

- 128-135.
95. Ng, S.Y. and J. Bettany-Saltikov, *Imaging in the Diagnosis and Monitoring of Children with Idiopathic Scoliosis*. Open Orthop J, 2017. **11**: p. 1500-1520.
 96. Karpel, I., et al., *A Survey of Methods and Technologies Used for Diagnosis of Scoliosis*. Sensors (Basel), 2021. **21**(24).
 97. Luan, F.-J., et al., *Cancer and mortality risks of patients with scoliosis from radiation exposure: a systematic review and meta-analysis*. European Spine Journal, 2020. **29**(12): p. 3123-3134.
 98. Presciutti, S.M., T. Karukanda, and M. Lee, *Management decisions for adolescent idiopathic scoliosis significantly affect patient radiation exposure*. The Spine Journal, 2014. **14**(9): p. 1984-1990.
 99. San Román Gaitero, A., et al., *Imaging Methods to Quantify the Chest and Trunk Deformation in Adolescent Idiopathic Scoliosis: A Literature Review*. Healthcare (Basel), 2023. **11**(10).
 100. Dubousset, J., et al., *[A new 2D and 3D imaging approach to musculoskeletal physiology and pathology with low-dose radiation and the standing position: the EOS system]*. Bull Acad Natl Med, 2005. **189**(2): p. 287-97; discussion 297-300.
 101. Deschênes, S., et al., *Diagnostic Imaging of Spinal Deformities Reducing Patients Radiation Dose With a New Slot-Scanning X-ray Imager*. Spine, 2010. **35**: p. 989-94.
 102. Somoskeöy, S., et al., *Clinical validation of coronal and sagittal spinal curve measurements based on three-dimensional vertebra vector parameters*. Spine J, 2012. **12**(10): p. 960-8.
 103. Post, M., et al., *New sagittal classification of AIS: validation by 3D characterization*. Eur Spine J, 2019. **28**(3): p. 551-558.
 104. Chung, N., et al., *Spinal phantom comparability study of Cobb angle measurement of scoliosis using digital radiographic imaging*. J Orthop Translat, 2018. **15**: p. 81-90.
 105. Rehm, J., et al., *3D-modeling of the spine using EOS imaging system: Inter-reader reproducibility and reliability*. PLoS One, 2017. **12**(2): p. e0171258.
 106. Somoskeöy, S., et al., *Accuracy and reliability of coronal and sagittal spinal curvature data based on patient-specific three-dimensional models created by the EOS 2D/3D imaging system*. Spine J, 2012. **12**(11): p. 1052-9.
 107. Miller, D.J., et al., *Posterior Correction Techniques for Adolescent Idiopathic Scoliosis*. JAAOS - Journal of the American Academy of Orthopaedic Surgeons, 2020. **28**(9).
 108. Chan, A., et al., *Does image guidance decrease pedicle screw-related complications in surgical treatment of adolescent idiopathic scoliosis: a systematic review update and meta-analysis*. European Spine Journal, 2020. **29**(4): p. 694-716.
 109. Chan, C.Y.W. and M.K. Kwan, *Safety of Pedicle Screws in Adolescent Idiopathic Scoliosis Surgery*. Asian Spine J, 2017. **11**(6): p. 998-1007.
 110. Kwan, M.K., et al., *Accuracy and Safety of Pedicle Screw Placement in Adolescent Idiopathic Scoliosis Patients: A Review of 2020 Screws Using Computed Tomography Assessment*. Spine, 2017. **42**(5).
 111. Wei, W., et al., *Finite element analysis in brace treatment on adolescent idiopathic scoliosis*. Medical & Biological Engineering & Computing, 2022. **60**(4): p. 907-920.
 112. Chiyanka, C. and W.C.W. Chu, *Weight-bearing Imaging in Adolescent Idiopathic Scoliosis*. Semin Musculoskelet Radiol, 2019. **23**(06): p. 584-593.

113. Duke, A., et al., *Radiation in Adolescent Idiopathic Scoliosis Management: Estimated Cumulative Pre-Operative, Intra-Operative, and Post-Operative Exposure*. Orthopedic Research and Reviews, 2022. **14**: p. 487-493.
114. McArthur, N., D.P. Conlan, and J.R. Crawford, *Radiation exposure during scoliosis surgery: a prospective study*. Spine J, 2015. **15**(3 Suppl): p. S33-s36.
115. Shi, L., et al., *Volume-based morphometry of brain MR images in adolescent idiopathic scoliosis and healthy control subjects*. American journal of neuroradiology, 2009. **30**(7): p. 1302-1307.
116. Wang, D., et al., *Abnormal cerebral cortical thinning pattern in adolescent girls with idiopathic scoliosis*. Neuroimage, 2012. **59**(2): p. 935-942.
117. Geissele, A.E., et al., *Magnetic resonance imaging of the brain stem in adolescent idiopathic scoliosis*. Spine, 1991. **16**(7): p. 761-763.
118. Cheng, J.C., et al., *Correlation between curve severity, somatosensory evoked potentials, and magnetic resonance imaging in adolescent idiopathic scoliosis*. Spine, 1999. **24**(16): p. 1679.
119. Lee, R.K., et al., *Effect of upright position on tonsillar level in adolescent idiopathic scoliosis*. European radiology, 2015. **25**: p. 2397-2402.
120. Shi, L., et al., *Volumetric changes in cerebellar regions in adolescent idiopathic scoliosis compared with healthy controls*. The Spine Journal, 2013. **13**(12): p. 1904-1911.
121. Chau, W.W., et al., *Anatomical origin of abnormal somatosensory-evoked potential (SEP) in adolescent idiopathic scoliosis with different curve severity and correlation with cerebellar tonsillar level determined by MRI*. Spine, 2016. **41**(10): p. E598-E604.
122. Noriega-Gonzalez, D.C., et al. *Cerebral White Matter Connectivity in Adolescent Idiopathic Scoliosis: A Diffusion Magnetic Resonance Imaging Study*. Children, 2022. **9**, DOI: 10.3390/children9071023.
123. Zeng, W., et al. *Shape analysis of vestibular systems in adolescent idiopathic scoliosis using geodesic spectra*. in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010: 13th International Conference, Beijing, China, September 20–24, 2010, Proceedings, Part III* 13. 2010. Springer.
124. Chevrefils, C., et al., *To distinguish flexible and rigid lumbar curve from MRI texture analysis in adolescent idiopathic scoliosis: A feasibility study*. Journal of Magnetic Resonance Imaging, 2018. **48**(1): p. 178-187.
125. Federau, C., et al., *Demonstration of asymmetric muscle perfusion of the back after exercise in patients with adolescent idiopathic scoliosis using intravoxel incoherent motion (IVIM) MRI*. NMR in Biomedicine, 2020. **33**(3): p. e4194.
126. Duncombe, P., et al., *Quantifying Muscle Size Asymmetry in Adolescent Idiopathic Scoliosis Using Three-dimensional Magnetic Resonance Imaging*. Spine, 2023.
127. Labrom, F.R., et al., *Sequential MRI reveals vertebral body wedging significantly contributes to coronal plane deformity progression in adolescent idiopathic scoliosis during growth*. Spine Deformity, 2020. **8**(5): p. 901-910.
128. Ramírez, N., et al., *Prevalence and clinical relevance of underlying pathological conditions in painful adolescent idiopathic scoliosis: a MRI-based study*. Spine Deformity, 2020. **8**(4): p. 663-668.
129. Hesarikia, H., et al., *Magnetic resonance imaging investigations of position of conus*

- medullaris in adolescent idiopathic scoliosis as a peripheral neuropathy*. Int J Clin Exp Med, 2015. **8**(4): p. 5918-24.
130. Eder, M., et al., *Three-dimensional evaluation of breast contour and volume changes following subpectoral augmentation mammoplasty over 6 months*. Journal of Plastic, Reconstructive & Aesthetic Surgery, 2011. **64**(9): p. 1152-1160.
 131. Letts, M., et al., *Computerized ultrasonic digitization in the measurement of spinal curvature*. Spine (Phila Pa 1976), 1988. **13**(10): p. 1106-10.
 132. Suzuki, S., et al., *Ultrasound measurement of vertebral rotation in idiopathic scoliosis*. J Bone Joint Surg Br, 1989. **71**(2): p. 252-5.
 133. Darrieutort-Laffite, C., et al., *Ultrasonography of the lumbar spine: sonoanatomy and practical applications*. Joint Bone Spine, 2014. **81**(2): p. 130-6.
 134. Li, M., et al., *Could clinical ultrasound improve the fitting of spinal orthosis for the patients with AIS?* European spine journal : official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society, 2012. **21**: p. 1926-35.
 135. Ungi, T., et al., *Spinal curvature measurement by tracked ultrasound snapshots*. Ultrasound Med Biol, 2014. **40**(2): p. 447-54.
 136. Young, M., et al., *Reliability and accuracy of ultrasound measurements with and without the aid of previous radiographs in adolescent idiopathic scoliosis (AIS)*. Eur Spine J, 2015. **24**(7): p. 1427-33.
 137. Zheng, R., et al., *Intra- and Inter-rater Reliability of Coronal Curvature Measurement for Adolescent Idiopathic Scoliosis Using Ultrasonic Imaging Method-A Pilot Study*. Spine Deform, 2015. **3**(2): p. 151-158.
 138. Wang, Q., et al., *Validity Study of Vertebral Rotation Measurement Using 3-D Ultrasound in Adolescent Idiopathic Scoliosis*. Ultrasound Med Biol, 2016. **42**(7): p. 1473-81.
 139. Lee, T.T., et al., *A Novel Method to Measure the Sagittal Curvature in Spinal Deformities: The Reliability and Feasibility of 3-D Ultrasound Imaging*. Ultrasound Med Biol, 2019. **45**(10): p. 2725-2735.
 140. Brink, R.C., et al., *A reliability and validity study for different coronal angles using ultrasound imaging in adolescent idiopathic scoliosis*. Spine J, 2018. **18**(6): p. 979-985.
 141. He, C., et al., *An effective assessment method of spinal flexibility to predict the initial in-orthosis correction on the patients with adolescent idiopathic scoliosis (AIS)*. PLoS One, 2017. **12**(12): p. e0190141.
 142. Jiang, W.W., et al., *Patterns of coronal curve changes in forward bending posture: a 3D ultrasound study of adolescent idiopathic scoliosis patients*. Eur Spine J, 2018. **27**(9): p. 2139-2147.
 143. Zhou, G.Q. and Y.P. Zheng. *Assessment of scoliosis using 3-D ultrasound volume projection imaging with automatic spine curvature detection*. in *2015 IEEE International Ultrasonics Symposium (IUS)*. 2015.
 144. Cheung, C.W., et al., *Ultrasound Volume Projection Imaging for Assessment of Scoliosis*. IEEE Trans Med Imaging, 2015. **34**(8): p. 1760-8.
 145. Weiss, H.R. and D. Goodall, *The treatment of adolescent idiopathic scoliosis (AIS) according to present evidence. A systematic review*. Eur J Phys Rehabil Med, 2008. **44**(2): p. 177-93.

146. Burton, M.S., *Diagnosis and treatment of adolescent idiopathic scoliosis*. *Pediatr Ann*, 2013. **42**(11): p. 224-8.
147. Lonstein, J.E. and R.B. Winter, *The Milwaukee brace for the treatment of adolescent idiopathic scoliosis. A review of one thousand and twenty patients*. *JBJS*, 1994. **76**(8).
148. O'Neill, P.J., et al., *Decreased Orthotic Effectiveness in Overweight Patients with Adolescent Idiopathic Scoliosis*. *JBJS*, 2005. **87**(5).
149. Coillard, C., et al., *Effectiveness of the SpineCor Brace Based on the New Standardized Criteria Proposed by the Scoliosis Research Society for Adolescent Idiopathic Scoliosis*. *Journal of Pediatric Orthopaedics*, 2007. **27**(4).
150. Roland d' Amato, C., S. Griggs, and B. McCoy, *Nighttime Bracing With the Providence Brace in Adolescent Girls With Idiopathic Scoliosis*. *Spine*, 2001. **26**(18).
151. Schiller, J.R., N.A. Thakur, and C.P. Ebersson, *Brace Management in Adolescent Idiopathic Scoliosis*. *Clinical Orthopaedics and Related Research®*, 2010. **468**(3): p. 670-678.
152. Landauer, F., C. Wimmer, and H. Behensky, *Estimating the final outcome of brace treatment for idiopathic thoracic scoliosis at 6-month follow-up*. *Pediatr Rehabil*, 2003. **6**(3-4): p. 201-7.
153. Koroivessis, P., et al., *Comparative multifactorial analysis of the effects of idiopathic adolescent scoliosis and Scheuermann kyphosis on the self-perceived health status of adolescents treated with brace*. *Eur Spine J*, 2007. **16**(4): p. 537-46.
154. Donzelli, S., et al., *Adolescents with idiopathic scoliosis and their parents have a positive attitude towards the Thermobrace monitor: results from a survey*. *Scoliosis and Spinal Disorders*, 2017. **12**(1): p. 12.
155. Fuss, F.K., et al., *Pressure Sensor System for Customized Scoliosis Braces*. *Sensors (Basel)*, 2021. **21**(4).
156. Cheung, M.-C., et al. *Adolescents' Experience during Brace Treatment for Scoliosis: A Qualitative Study*. *International Journal of Environmental Research and Public Health*, 2022. **19**, DOI: 10.3390/ijerph191710585.
157. Zhang, Y., et al., *3D-printed brace in the treatment of adolescent idiopathic scoliosis: A study protocol of a prospective randomised controlled trial*. *BMJ Open*, 2020. **10**: p. e038373.
158. Storm, F.A., et al. *Additive Manufacturing of Spinal Braces: Evaluation of Production Process and Postural Stability in Patients with Scoliosis*. *Materials*, 2022. **15**, DOI: 10.3390/ma15186221.
159. Wynne, J.H. and L.R. Houle *Short-Term Outcomes of the Boston Brace 3D Program Based on SRS and SOSORT Criteria: A Retrospective Study*. *Children*, 2022. **9**, DOI: 10.3390/children9060842.
160. Caesarendra, W., et al., *Automated Cobb Angle Measurement for Adolescent Idiopathic Scoliosis Using Convolutional Neural Network*. *Diagnostics*, 2022. **12**(2): p. 396.
161. Caesarendra, W., et al. *AutoSpine-Net: Spine Detection Using Convolutional Neural Networks for Cobb Angle Classification in Adolescent Idiopathic Scoliosis*. in *Proceedings of the 2nd International Conference on Electronics, Biomedical Engineering, and Health Informatics*. 2022. Singapore: Springer Nature Singapore.
162. Huang, X., et al., *The Comparison of Convolutional Neural Networks and the Manual Measurement of Cobb Angle in Adolescent Idiopathic Scoliosis*. *Global Spine Journal*,

- 2022: p. 21925682221098672.
163. Yan, S., et al. *Automated Cobb Angle Measurement in Adolescent Idiopathic Scoliosis: Validation of a Previously-Published Deep Learning Method*. in *2022 IEEE 10th International Conference on Healthcare Informatics (ICHI)*. 2022.
 164. Sun, Y., et al., *Comparison of manual versus automated measurement of Cobb angle in idiopathic scoliosis based on a deep learning keypoint detection technology*. *European Spine Journal*, 2022. **31**(8): p. 1969-1978.
 165. Zhao, S., et al., *Automatic Spine Curvature Estimation by a Top-Down Approach*, in *Computational Methods and Clinical Applications for Spine Imaging: 6th International Workshop and Challenge, CSI 2019, Shenzhen, China, October 17, 2019, Proceedings*. 2019, Springer-Verlag: Shenzhen, China. p. 75–80.
 166. Wu, C., et al., *A multi-stage ensemble network system to diagnose adolescent idiopathic scoliosis*. *European Radiology*, 2022. **32**(9): p. 5880-5889.
 167. Yahara, Y., et al., *A deep convolutional neural network to predict the curve progression of adolescent idiopathic scoliosis: a pilot study*. *BMC Musculoskeletal Disorders*, 2022. **23**(1): p. 610.
 168. Wang, H., et al., *An intelligent composite model incorporating global / regional X-rays and clinical parameters to predict progressive adolescent idiopathic scoliosis curvatures and facilitate population screening*. *eBioMedicine*, 2023. **95**.
 169. Wang, H., et al., *Application of deep learning upon spinal radiographs to predict progression in adolescent idiopathic scoliosis at first clinic visit*. *EClinicalMedicine*, 2021. **42**: p. 101220.
 170. Kadoury, S., et al., *3-D Morphology Prediction of Progressive Spinal Deformities From Probabilistic Modeling of Discriminant Manifolds*. *IEEE Trans Med Imaging*, 2017. **36**(5): p. 1194-1204.
 171. Thong, W., et al., *Three-dimensional morphology study of surgical adolescent idiopathic scoliosis patient from encoded geometric models*. *European Spine Journal*, 2016. **25**(10): p. 3104-3113.
 172. Logithasan, V., et al., *Using machine learning to automatically measure axial vertebral rotation on radiographs in adolescents with idiopathic scoliosis*. *Medical Engineering & Physics*, 2022. **107**: p. 103848.
 173. Kaddioui, H., et al., *Convolutional Neural Networks for Automatic Risser Stage Assessment*. *Radiology: Artificial Intelligence*, 2020. **2**(3): p. e180063.
 174. Magnide, E., et al., *Automatic bone maturity grading from EOS radiographs in Adolescent Idiopathic Scoliosis*. *Computers in Biology and Medicine*, 2021. **136**: p. 104681.
 175. Suri, A., et al., *Conquering the Cobb Angle: A Deep Learning Algorithm for Automated, Hardware-Invariant Measurement of Cobb Angle on Radiographs in Patients with Scoliosis*. *Radiology: Artificial Intelligence*, 2023. **5**(4): p. e220158.
 176. Tajdari, M., et al., *Next-generation prognosis framework for pediatric spinal deformities using bio-informed deep learning networks*. *Engineering with Computers*, 2022. **38**(5): p. 4061-4084.
 177. Lessmann, N., et al., *Iterative fully convolutional neural networks for automatic vertebra segmentation and identification*. *Medical image analysis*, 2019. **53**: p. 142-155.
 178. Furqan Qadri, S., et al., *Automatic Deep Feature Learning via Patch-Based Deep Belief*

- Network for Vertebrae Segmentation in CT Images*. Applied sciences, 2018. **9**(1): p. 69.
179. Rehman, F., et al., *A Region-Based Deep Level Set Formulation for Vertebral Bone Segmentation of Osteoporotic Fractures*. Journal of digital imaging, 2020. **33**(1): p. 191-203.
 180. Bae, H.-J., et al., *Fully automated 3D segmentation and separation of multiple cervical vertebrae in CT images using a 2D convolutional neural network*. Computer methods and programs in biomedicine, 2020. **184**: p. 105119-105119.
 181. Zhang, L. and H. Wang, *A novel segmentation method for cervical vertebrae based on PointNet++ and converge segmentation*. Computer methods and programs in biomedicine, 2021. **200**: p. 105798-105798.
 182. Altini, N., et al., *Segmentation and Identification of Vertebrae in CT Scans Using CNN, k-Means Clustering and k-NN*. Informatics (Basel), 2021. **8**(2): p. 40.
 183. Qadri, S.F., et al., *CT-Based Automatic Spine Segmentation Using Patch-Based Deep Learning*. International Journal of Intelligent Systems, 2023. **2023**: p. 2345835.
 184. Rak, M., et al., *Combining convolutional neural networks and star convex cuts for fast whole spine vertebra segmentation in MRI*. Computer Methods and Programs in Biomedicine, 2019. **177**: p. 47-56.
 185. Kuang, X., et al. *MRI-SegFlow: a novel unsupervised deep learning pipeline enabling accurate vertebral segmentation of MRI images*. in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. 2020.
 186. Zhou, Y., et al., *Automatic Lumbar MRI Detection and Identification Based on Deep Learning*. Journal of Digital Imaging, 2019. **32**(3): p. 513-520.
 187. Pang, S., et al., *Direct automated quantitative measurement of spine by cascade amplifier regression network with manifold regularization*. Medical Image Analysis, 2019. **55**: p. 103-115.
 188. Li, X., et al., *3D multi-scale FCN with random modality voxel dropout learning for Intervertebral Disc Localization and Segmentation from Multi-modality MR Images*. Medical Image Analysis, 2018. **45**: p. 41-54.
 189. Gaonkar, B., et al., *Quantitative Analysis of Spinal Canal Areas in the Lumbar Spine: An Imaging Informatics and Machine Learning Study*. American Journal of Neuroradiology, 2019. **40**(9): p. 1586.
 190. Han, Z., et al., *Automated Pathogenesis-Based Diagnosis of Lumbar Neural Foraminal Stenosis via Deep Multiscale Multitask Learning*. Neuroinformatics, 2018. **16**(3): p. 325-337.
 191. Lu, J.-T., et al., *Deep Spine: Automated Lumbar Vertebral Segmentation, Disc-Level Designation, and Spinal Stenosis Grading using Deep Learning*, in *Proceedings of the 3rd Machine Learning for Healthcare Conference*, D.-V. Finale, et al., Editors. 2018, PMLR: Proceedings of Machine Learning Research. p. 403--419.
 192. Kai-Uwe Lewandrowski, I., et al., *Feasibility of Deep Learning Algorithms for Reporting in Routine Spine Magnetic Resonance Imaging*. International Journal of Spine Surgery, 2020. **14**(s3): p. S86.
 193. Kim, K., et al., *Performance of the deep convolutional neural network based magnetic resonance image scoring algorithm for differentiating between tuberculous and pyogenic spondylitis*. Scientific Reports, 2018. **8**(1): p. 13124.

194. Hopkins, B.S., et al., *Machine Learning for the Prediction of Cervical Spondylotic Myelopathy: A Post Hoc Pilot Study of 28 Participants*. World Neurosurgery, 2019. **127**: p. e436-e442.
195. McCoy, D.B., et al., *Convolutional Neural Network–Based Automated Segmentation of the Spinal Cord and Contusion Injury: Deep Learning Biomarker Correlates of Motor Impairment in Acute Spinal Cord Injury*. American Journal of Neuroradiology, 2019. **40**(4): p. 737.
196. Maki, S., et al., *A Deep Convolutional Neural Network With Performance Comparable to Radiologists for Differentiating Between Spinal Schwannoma and Meningioma*. Spine, 2020. **45**(10).
197. Joo, Y.B., et al., *Machine learning–based prediction of radiographic progression in patients with axial spondyloarthritis*. Clinical Rheumatology, 2020. **39**(4): p. 983-991.
198. Joo, Y.B., et al., *Novel classification of axial spondyloarthritis to predict radiographic progression using machine learning*. Clin Exp Rheumatol, 2021. **39**(3): p. 508-518.
199. Huang, Z., et al. *Bone Feature Segmentation in Ultrasound Spine Image with Robustness to Speckle and Regular Occlusion Noise*. in *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. 2020.
200. Banerjee, S., et al. *Light-Convolution Dense Selection U-Net (LDS U-Net) for Ultrasound Lateral Bony Feature Segmentation*. Applied Sciences, 2021. **11**, DOI: 10.3390/app112110180.
201. Jiang, W., F. Mei, and Q. Xie, *Novel automated spinal ultrasound segmentation approach for scoliosis visualization*. Frontiers in Physiology, 2022. **13**.
202. Aslan, M.S., et al., *Probabilistic shape-based segmentation method using level sets*. IET computer vision, 2014. **8**(3): p. 182-194.
203. Fu, Y., et al., *Automatic and hierarchical segmentation of the human skeleton in CT images*. Physics in medicine & biology, 2017. **62**(7): p. 2812-2833.
204. Dabiri, S., et al., *Multi-view parallel vertebra segmentation and identification on computed tomography (CT) images*. Informatics in Medicine Unlocked, 2022. **34**: p. 101091.
205. LeCun, Y., et al., *Backpropagation applied to handwritten zip code recognition*. Neural computation, 1989. **1**(4): p. 541-551.
206. Nair, V. and G.E. Hinton. *Rectified linear units improve restricted boltzmann machines*. in *Proceedings of the 27th international conference on machine learning (ICML-10)*. 2010.
207. Long, J., E. Shelhamer, and T. Darrell. *Fully convolutional networks for semantic segmentation*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015.
208. Ronneberger, O., P. Fischer, and T. Brox. *U-Net: Convolutional Networks for Biomedical Image Segmentation*. in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. 2015. Cham: Springer International Publishing.
209. Zhang, Y., et al., *LumVertCancNet: A novel 3D lumbar vertebral body cancellous bone location and segmentation method based on hybrid Swin-transformer*. Computers in Biology and Medicine, 2024. **171**: p. 108237.
210. Xu, H., et al., *RUnT: A network combining Residual U-net and Transformer for vertebral edge feature fusion constrained spine CT image segmentation*. IEEE Access, 2023.
211. You, X., et al. *EG-Trans3DUNet: a single-staged transformer-based model for accurate*

- vertebrae segmentation from spinal CT images. in *2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI)*. 2022. IEEE.
212. Tao, R. and G. Zheng. *Spine-transformers: Vertebra detection and localization in arbitrary field-of-view spine ct with transformers*. in *Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part III* 24. 2021. Springer.
 213. Zhou, H.-Y., et al., *nnformer: Volumetric medical image segmentation via a 3d transformer*. IEEE Transactions on Image Processing, 2023.
 214. Sekuboyina, A., et al., *VerSe: A Vertebrae labelling and segmentation benchmark for multi-detector CT images*. Med Image Anal, 2021. **73**: p. 102166.
 215. Zhang, Y., et al., *A Spine Segmentation Method under an Arbitrary Field of View Based on 3D Swin Transformer*. International Journal of Intelligent Systems, 2023. **2023**.
 216. Payer, C., et al. *Coarse to Fine Vertebrae Localization and Segmentation with SpatialConfiguration-Net and U-Net*. in *VISIGRAPP (5: VISAPP)*. 2020.
 217. Tao, R., W. Liu, and G. Zheng, *Spine-transformers: Vertebra labeling and segmentation in arbitrary field-of-view spine CTs via 3D transformers*. Medical Image Analysis, 2022. **75**: p. 102258.
 218. Zhang, Y., et al., *Automatic Segmentation of Lumbar Vertebra Anatomical Region Based on Hybrid Swin-Transformer Network*. 2023. 1–5.

Appendix

伦理审查批件

批件号	20220920004		
项目名称	脊柱侧弯中 2D 到 3D 的图像处理技术研究		
审查日期	2022.09.28	主审委员	邵超鹏 郑施施
项目类型	<input type="checkbox"/> 药品 <input type="checkbox"/> 器械 <input type="checkbox"/> 诊断试剂 <input type="checkbox"/> 医疗新技术 <input checked="" type="checkbox"/> 其他 <u>临床科研</u>		
审查类别	<input checked="" type="checkbox"/> 初始审查 <input type="checkbox"/> 跟踪审查 <input type="checkbox"/> 复审		
审查方式	<input type="checkbox"/> 会议审查 <input checked="" type="checkbox"/> 快速审查 <input type="checkbox"/> 紧急审查		
研究科室	放射科	主要研究者	林帆
申办方	无		
审查文件	初始审查申请清单文件		
审查决定	同意		
跟踪审查频率	12 个月	批件有效期	12 个月
声明	<div>1. 本伦理委员会的职责、人员组成和工作程序遵循 ICH-GCP、中国 GCP，符合赫尔辛基宣言的原则，并遵守中国相关法律法规的规定。</div> <div>2. 研究应遵循 ICH-GCP 和国家法律法规的要求，按照本伦理委员会批准的方案开展临床研究，保护受试者的健康与权益。</div> <div>3. 研究过程中，若对临床研究方案、知情同意书、招募材料等有任何修改，请递交修正案审查申请，获得批准后方可执行。</div> <div>4. 本伦理委员会对研究项目进行跟踪审查，自批件签发之日起，请研究者在规定的跟踪审查截止日期或批件失效前 1 个月提交研究进展报告。</div> <div>5. 获得本伦理批件后，研究者需登录医学研究备案系统 https://www.medicalresearch.org.cn/login，项目注册成功后，方可按要求开展临床科研。</div> <div>6. SAE/SUSAR 报告：本中心发生的 SAE 必须在研究者获知之后 24 小时以内报告；中国其他中心发生的 SAE，至少每 3 个月一次及时报告；SUSAR 至少每 6 个月一次报告伦理委员会。方案偏离报告应在发生后 1 个月内报告伦理委员会。</div> <div>7. 暂停/终止研究报告、分中心小结表、结题报告等，须及时提交伦理委员会。</div>		
主任委员/ 副主任委员签名		签发日期	2022.9.29

深圳市第二人民医院临床科研伦理委员会（盖章）