VAWKUM Transactions on Computer Sciences
http://vfast.org/journals/index.php/VTCS@ 2026, ISSN(e): 2308-8168, ISSN(p): 2411-6335
Volume 14, Number 1, January-June 2026 pp:15-27

Enhanced Vertebral Segmentation and Cobb angle Calculation using Advanced Instance Segmentation Techniques

Inzamam Ul Haq [1*], **Izaz Ullah** [1], **Fazli Amin Khalil** [2], **Sakin Jan** [3], **Israr Ahmed Khan** [4]

[1]Department of Computer Science, Government Postgraduate College charsadda, KPK, Pakistan; [2]Department of Electronics, University of Peshawar, KPK, Pakistan; [3]Department of Computer Science, Government Abdul Ali Khan Degree College, charsadda, KPK, Pakistan; [4]Department of Electronics, Government Postgraduate College charsadda, KPK, Pakistan

Keywords: *Scoliosis, Cobb angle, Deep learning, Segmentation*

Journal Info:
Submitted:
July 20, 2025
Accepted:
February 11, 2026
Published:
February 22, 2026

ABSTRACT

Scoliosis, a spinal deformity affecting children and adolescents, is quantified using the Cobb angle. Traditional manual measurement by clinicians is time-consuming and subject to variability. This study introduces an automated method using the YOLACT++ instance segmentation model to detect vertebrae and calculate the Cobb angle on the SpineWeb dataset. By identifying the most tilted vertebrae using principal component analysis, our approach achieves a Symmetric Mean Absolute Percentage Error (SMAPE) of 8.11%, outperforming previous segmentation-based methods. This demonstrates improved accuracy and reliability, with potential for clinical decision support. The source code and other details are available at github link. https://github.com/inzamam-ulhaq-collab/Cobb-angle-measurement-code.git

***Correspondence author email address:** inzichd@gmail.com
DOI: 10.21015/vtcs.v14i1.2197

1 Introduction

An important part of the human body . is the spine, which serves several functions, including supporting weight and maintaining posture. The spine consists of 33 vertebrae, which are further divided into five regions. (c1-c7) are called cervical, (T1-T12) thoracic, (L1-L5) Lumbar, (S1-S5) sacrum and (Co1-Co4) coccyx. The first 24 vertebrae provide the spine with flexibility and are movable. A normal spine looks straight when viewed from the back. An abnormal curve of the spine from the center to the left or right, making an 'S' or 'C' shape when viewed from the back, is called scoliosis. Scoliosis affects 2-3% of the population in the United States of America, or 6-9 million people. Symptoms of scoliosis are: difference in arm height when standing straight, back pain, uneven hip height, and difference in shoulder height [1].

Doctors and professionals diagnose scoliosis using X-rays and CT scans to detect vertebral deviation from the midline and measure the degree of curvature on X-ray images [2, 3]. The Cobb angle is the preferred method for assessing the degree of deviation and the severity of the vertebrae to the left or right of the center [4]. Doctors manually measure the Cobb angle by experience and by scanning X-rays to identify the most tilted vertebrae above and below the apex vertebrae [5].

Apex vertebrae are the vertebrae that are farthest away from the center of normal vertebrae. The most tilted vertebrae are those that have the maximum angle above and below the apex vertebra. The professional draws a line on the upper end plate of the most tilted vertebrae and the lower end plate of the most tilted vertebrae. The region where the line intersects is the Cobb angle. There can be one Cobb angle in the 'C' shaped scoliosis and two angles in 'S' shaped scoliosis. The same method applies to the 2nd angle by considering the most tilted vertebrae in the first angle as the most tilted vertebrae in the 2nd angle as shown in figure 1. Scoliosis with an Angle below 10 degrees is considered normal, below 20 degrees is considered mild, and above 20 degrees is considered severe.

As doctors find the Cobb angle manually by drawing lines, which, if calculated by many professionals, can be subject to human errors [6], and it is time-consuming. To automate finding Cobb angle and remove inconsistencies, Machine learning and deep learning based methods can be used for accuracy and efficiency in finding the Cobb angle [7]. First, accurately detecting the vertebrae that segment the vertebrae, 2nd, correctly finding the corner points of the vertebrae, and finally, from these corner points, accurately measuring the Cobb angle.

Instance segmentation-based methods such as Mask R-CNN, Boostnet, ResNet, and YOLOv3 have been used to automate Cobb angle measurement; these methods have the potential to reduce human error in calculating Cobb angle. However, these methods have limited accuracy in detecting vertebrae, leading to suboptimal Cobb angle calculations. Therefore, this article applies an enhanced instance segmentation-based method to detect vertebrae, thereby improving the accuracy of Cobb angle calculation.

The rest of the paper continues as follows: Section 2 provides related work and limitations; Section 3 provides details on the YOLACT and YOLACT++ model architectures and their operation; and Section 4 provides an Experimental discussion of data processing, angle calculation, and metrics. Section 5 discusses the results. And finally, section 6 concludes our work.

2 Literature review

The traditional method for detecting spinal X-rays using computers typically involves analyzing the X-ray image's pixel data using common image processing techniques. These techniques involve threshold-based [8, 9] and edge-based segmentation [10, 11]. In reference [12] They use BoostNet, a framework that leverages the powerful feature detection capabilities of Convolutional Neural Networks (CNNs), employs regression techniques to identify key landmarks, and achieves a Mean Squared Error (MSE) of 0.0046. This regression-based method can be used for feature extraction from medical images, but it requires a high number of parameters and hence has a high computational cost. Consequently, high-resolution input images, which can reach dimensions of 2600×1600 pixels, must be reduced to a significantly smaller size (256×128 pixels) for both training and inference. This downscaling process restricts the model's effectiveness, as critical details are often lost during compression. Other Models, like the Active Contour Model [13], Customized Filters [14], and Charged-Particle Models [15], were employed to identify and locate specific vertebrae, enabling the calculation of the Cobb angle based on their orientations. These approaches rely on precise vertebrae segmentation and manual feature extraction, making them resource-intensive and prone to inaccuracies due to inconsistencies in X-ray image quality.

Machine learning techniques, including Support Vector Regression (SVR) [16] Random Forest Regression (RFR) [17] , and Convolutional Neural Networks (CNN) [18–20], have been widely applied to various biomedical tasks. However, their direct use in assessing Adolescent Idiopathic Scoliosis (AIS) faces several challenges: (1) the approach's reliability and ability to generalize can be undermined by outliers (such as human errors or imaging artifacts) in the training data [21], often necessitating extensive preprocessing, and (2) the interdependencies among multiple outputs (e.g., landmark coordinates) are not explicitly modeled, which is crucial for improving the accuracy of spinal landmark detection. Although [22] adapted SVR to account for output dependencies in identifying spinal landmarks, their method still relies on manual feature extraction, which struggles to handle

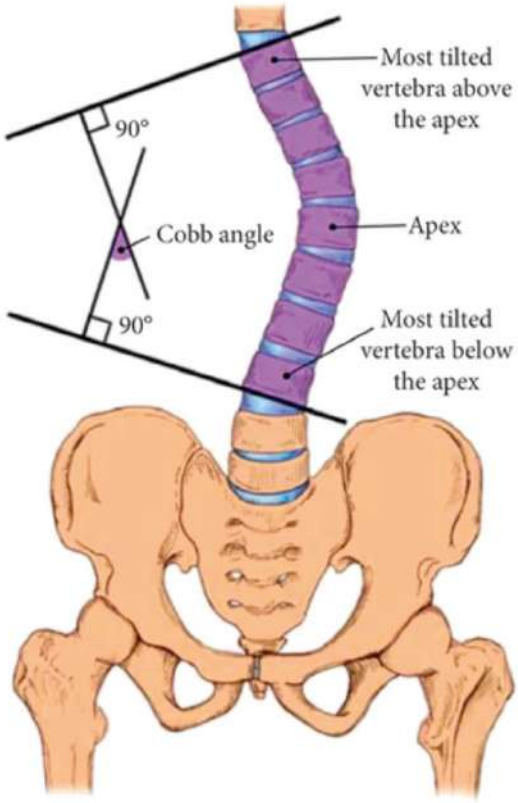

Figure 1. The Cobb angle by drawing on the most tilted vertebrae.

image anomalies effectively.

A segmentation model for vertebrae localization can resolve some issues. Horng et al. implemented a modified U-Net architecture on a dataset of 595 vertebral images, achieving a mean squared error (MSE) of 0.025 on the test set [23]. Another previously used segmentation model is Mask R-CNN. Pan et al. employed two separate Mask R-CNN models to independently identify the spine and vertebral bodies. The midpoints of the vertebral endplates were subsequently calculated based on the generated segmentation masks. This method achieved a Mean Absolute Difference (MAD) of 3.32 degrees, demonstrating the models' reliability [24]. [25] Introduces a new framework that uses a graph convolution network GCN to analyze 3D gait sequences. The authors demonstrate that modeling the spatiotemporal dynamics of skeletal motion provides a scalable and efficient alternative to the traditional, cost-intensive radiographic method. However, the process is designed for screening, not clinical diagnosis. While it effectively identifies scoliosis, it does not replace the Cobb angle measurement obtained from X-rays, which remains necessary for clinical treatment planning.

In general, the segmentation method provides acceptable results. Previous studies have shown that segmentation models often produce merged or distorted vertebral masks when processing low-quality input images. The [26] paper proposes an enhanced instance segmentation model that integrates a visual attention network into a ResNet50 backbone and utilizes a U-Net structure for precise pixel-level mask prediction. It introduces a "parameter link loss" to simplify hyperparameter tuning, significantly improving detection and segmentation performance. Although the paper claims efficiency, the inclusion of a U-Net-style up-sampling down-sampling path and a visual attention network adds architectural complexity that may impact real-time inference speed compared to "one-stage" models.

The above issues can be addressed using object detection methods. Khanal et al. introduced an innova-

tive technique that initially detects 17 vertebrae using a bounding-box object detector, then feeds the detected regions into a densely connected CNN to identify the four corners of each vertebra [27]. Faster R-CNN attains a Symmetric Mean Absolute Percentage Error (SMAPE) of 25.69% when evaluated on a dataset comprising 609 images [27]. Yi et al. first proposed a method to identify the center of each vertebra and then determine its four corners by leveraging learned corner offsets within a ResNet-based framework [28]. This method accounts for the interdependence between vertebrae and achieves an average Symmetric Mean Absolute Percentage Error (SMAPE) of 15.9% when tested on a dataset of 609 images [29]. In another paper, an instance segmentation-based method called YOLACT achieves an SMAPE of 10.76% on a dataset of 609 images .Despite existing methods, accurately and consistently measuring the Cobb angle remains a challenge, underscoring the need for further improvement. This study builds upon previous work by leveraging YOLACT++ for more refined instance segmentation, reducing SMAPE to 8.11% thus contributing toward more reliable scoliosis diagnosis. Table 1 presents different models from the literature review and compares their SMAPEs.

Table 1. Example LaTeX Table

Model Name	SMAPE
UNET	16.68%
S^2VR	37.08%
BoostNet	37.08%
FAST R-CNN	23.44%
ResNet	10.81%
YOLACT	10.76%

3 Methodology

YOLACT is a one-shot instance segmentation model that utilizes a ResNet backbone to generate a feature pyramid from the input image [30]. The extracted feature pyramid is then processed through two distinct pathways. The first pathway, known as Protonet, consists of four convolutional layers and performs segmentation. The second pathway, inspired by RetinaNet, determines the mask coefficients to generate image masks and bounding boxes. Ultimately, the algorithm combines the two outputs via a linear combination and applies a sigmoid activation function to produce the final mask. Figure 2 shows the YOLACT++ architecture for the proposed methodology for this study.

The Protonet is designed as a Fully Convolutional Network (FCN) and is connected to the deepest feature layer to generate high-quality and robust masks. The arrows in Figure 3 denote 3×3 convolutional layers, except the final layer, which is a 1×1 convolution. The increase in dimensions is attributable to an up-sampling process followed by a convolutional layer. The final layer contains k channels, with each channel corresponding to a prototype. These prototypes serve as candidate masks and retain spatial information. In this approach, the number of prototypes is set to 32 (k = 32) to maintain an optimal balance between speed and accuracy. Figure 3 presents some examples of these prototypes. Unlike traditional semantic segmentation, this method does not impose a direct loss on the prototypes. Instead, their supervision occurs after the mask assembly process.

The second pathway follows a structure similar to RetinaNet and is responsible for predicting mask coefficients. Typically, anchor-based object detectors have two prediction branches: one for class confidence scores and another for bounding box regression. YOLACT adds a third parallel branch dedicated to predicting k mask coefficients. Figure 4 shows it. To improve stability and reduce nonlinearity, a tanh activation function is applied to the mask coefficients.

In the final step, YOLACT combines the outputs from the Protonet and the mask coefficient branch through matrix multiplication and applies a sigmoid activation function to generate the final masks. This process follows Equation 1, where P represents a w × h × k matrix containing prototype masks, and C is an n × k matrix of mask coefficients:

$$M = \sigma(PC^T)M \tag{1}$$

After assembling the masks, the model is trained using three loss components: classification loss, box regression loss, and mask loss, with respective weights of 1, 1.5, and 6.125. The classification and box regression losses follow the same formulation as in the Single Shot

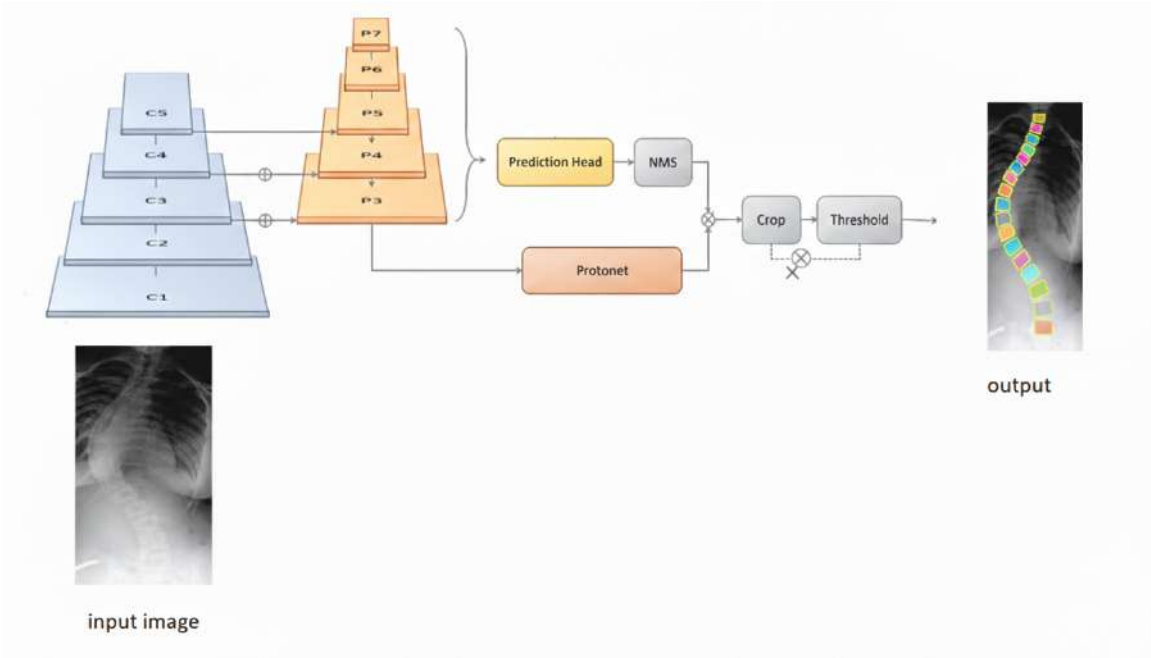

Figure 2. Architectural overview of the YOLACT++ model.

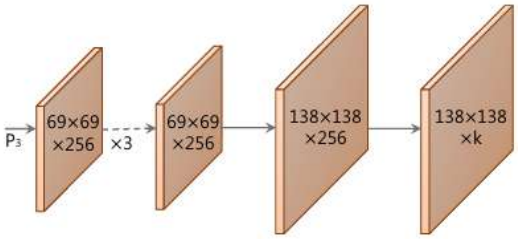

Figure 3. Visualization of prototype masks in the Protonet.

Multi-Box Detector (SSD). The mask loss is computed as pixel-wise binary cross-entropy between the predicted and ground truth masks.

YOLACT++ is an advanced version of the original YOLACT with real-time instance segmentation. YOLACT++ segments objects efficiently in images. It enhances mask prediction using a deformable convolution network, DCNv2, which adapts to the shape of objects and results in more segmentation outputs [31]. YOLACT++ uses ResNet-101 with a Feature Pyramid Network (FPN), replacing the 3x3 convolution layer in each ResNet block with a 3x3 deformable convolution layer for C3 to C5. As YOLACT is a single-shot method and it does not require a re-sampling strategy compared to the two-stage method, it improves the sampling strategy. The feature pyramid is processed through two paths. The first path, called protonet, creates a prototype of the input image using four convolution layers and is responsible for generating segmentation masks.

The second path calculates the mask coefficient based on Retina-Net, which produces bounding boxes and associated masks.YOLACT++introduces a Fast Mask Re-Scoring Network that adjusts the predicted masks based on their Intersection over Union (IoU) with the ground truth. The Fast Mask Re-scoring is structured as a fully convolutional Network FCN comprising six layers, each followed by a ReLU activation function, and has a final global pooling layer. This network takes the cropped mask prediction from YOLACT and computes the IoU score for each object. To enhance the accuracy of the mask score, the method multiplies the predicted IoU for each category by the classification confidence derived from the model classification head.

4 Experiment

The dataset comprises a total of 609 anterior-posterior spinal X-ray images available within the AASCE SpineWeb

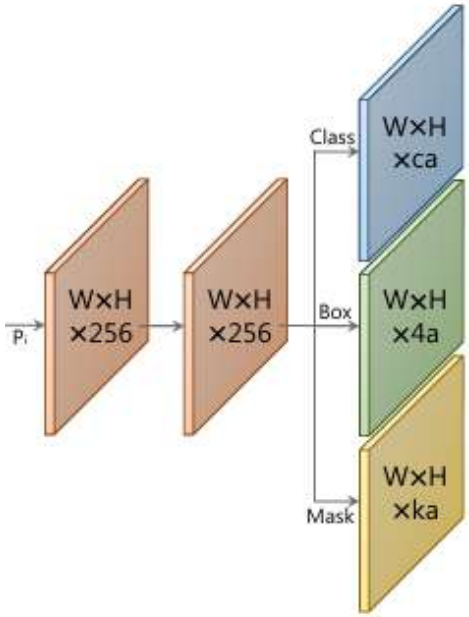

Figure 4. Parallel prediction branches for mask coefficients and bounding boxes.

dataset [32]. 80% of the images are allocated for training, while 20% are reserved for testing. All images have been preprocessed using Contrast-Limited Adaptive Histogram Equalization (CLAHE), as depicted in Figure 5. This technique is recognized as a robust preprocessing method in medical image processing for enhancing image contrast.

We resize all the images to 550x550. YOLACT++accepts input data in COCO Format. We annotate all images on LabelMe using polygons. The training has been performed on Google Colab GPU.YOLACT++ accurately detected vertebrae, producing precise, high-quality segmented masks (Figure 6).

5 Angle Calculation

Following the training phase, segmentation points are obtained and subsequently used to derive contour points using OpenCV [33]. From these contours, a minimum area rectangle (minAreaRect) is constructed, and the corner points of each vertebra are identified. Utilizing these corner points, the centroid of each vertebra is computed. A reference line is then drawn connecting the centroids of the first and last vertebrae. Based on the signed distance, vertebrae positioned to the left of this reference line are categorized as left vertebrae, while those to the right are classified as right vertebrae. The vertebra at the greatest distance from the reference line is designated the apex vertebra.

Subsequently, Principal Component Analysis (PCA), a dimensionality reduction algorithm, is employed to determine the most tilted vertebrae above and below the apex. PCA analyzes the spatial arrangement and orientation of the vertebrae based on their coordinates. The first principal component, PC1, corresponds to the primary orientation or axis of the vertebra. The angle formed between PC1 and the vertical axis quantifies the vertebra's tilt. The deviation of all vertebrae from the vertical axis, both above and below the apex, is calculated. The algorithm then identifies the vertebrae exhibiting the maximum deviation, which are considered the most tilted vertebrae above and below the apex, as shown in Figure 7.

Finally, the Cobb angle is computed by drawing lines along the upper endplates of the most tilted vertebrae above the apex and the lower endplates of the most tilted vertebrae below the apex. The intersection of these lines yields the Cobb angle, a critical metric for assessing spinal curvature. Calculating the slope (m1)

Figure 5. (a)Original image (b)after preprocessing image

Figure 6. (a) Input image (b) bbox detection (c) Segmentation

and y-intercept (c1) of the first line, x1, x2, and y1, y2 are the coordinates of the upper plates of the most tilted vertebrae above the apex.

$$m_1 = \frac{y_2 - y_1}{x_2 - x_1} \tag{2}$$

$$c_1 = y_1 - m_1 x_1 \tag{3}$$

Calculating the slope (m2) and y-intercept (c2) of the second line, x3, x4, and y3, y5 are the coordinates of the lower plates of the most tilted vertebrae below the apex.

$$m_2 = \frac{y_4 - y_3}{x_4 - x_3} \tag{4}$$

$$c_2 = y_3 - m_2 x_3 \tag{5}$$

$$\theta_{\text{Cobb}} = \tan^{-1} \left| \frac{m_2 - m_1}{1 + m_1 m_2} \right| \tag{6}$$

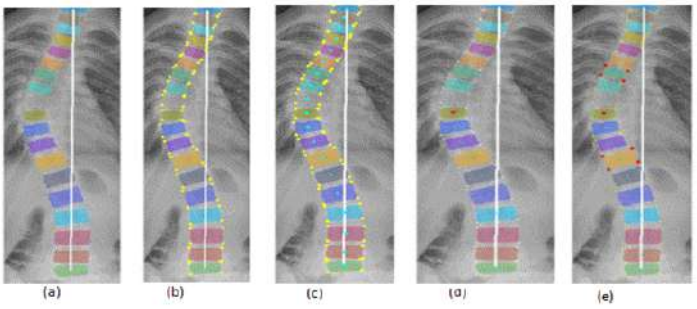

Figure 7. (a) Drawing lines from the first and last vertebrae. (b)Detecting corner points (c), getting center points (d), apex vertebrae (e), and the most tilted vertebrae above and below the apex

Subsequently, the algorithm determines whether a second Cobb angle is present. If a second angle is identified, the process is repeated by deciding the most distant vertebra from the reference line, which is then selected as the apex vertebra for the second angle. The most tilted vertebra below the apex in the first Cobb angle is designated as the most tilted vertebra above the

apex for the second angle. Similar to the first angle, Principal Component Analysis (PCA) is employed to identify the most tilted vertebra below the apex for the second angle.

Using the same methodology as previously described, the second Cobb angle is calculated in Figure 8. Lines are drawn along the upper endplates of the most tilted vertebra above the apex and the lower endplates of the most tilted vertebra below the apex. The intersection of these lines determines the second Cobb angle, thereby providing a comprehensive assessment of spinal curvature across multiple segments.

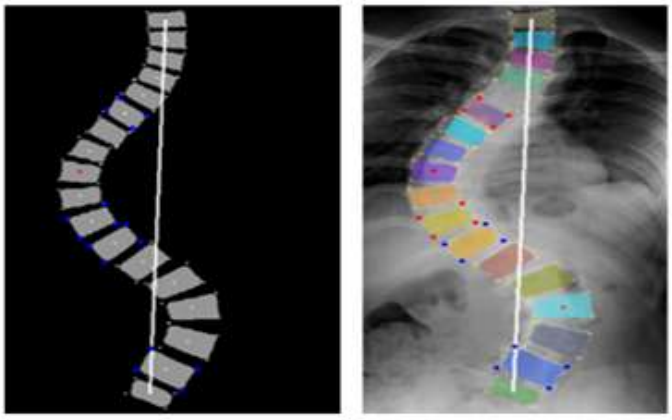

Figure 8. (a) Getting segmented vertebrae with a black background (b) red points first angle, blue points 2nd angle

6 Evaluation Formula

The Symmetric Mean Absolute Percentage Error (SMAPE) is a metric used to measure the accuracy of predictions, particularly in regression tasks. It is expressed as a percentage and is symmetric, meaning it treats over-predictions and under-predictions equally. The formula for SMAPE is:

$$\text{SMAPE} = \frac{1}{n}\sum_{i=1}^{n}\frac{|y_i - \hat{y}_i|}{(|y_i| + |\hat{y}_i|)/2}100\% \quad (7)$$

The variables utilized in the mathematical framework are defined as follows:

- y_i: The ground truth Cobb angle, representing the manual measurement performed by a clinical expert.
- $\hat{y}_i$: The predicted Cobb angle, representing the automated measurement generated by the Yolact++ model.
- n: The total number of observations (X-ray samples) evaluated from the SpineWeb dataset.

7 Results

The model of vertebrae segmentation was trained over 25,000 iterations showing powerful and consistent learning throughout the process.

7.1 Loss Profile:

Initially, the Training Loss curve shows fluctuation quite a bit as shown in Figure 9, followed by gradual convergence. The training became much smoother after initial phase. By iteration 20,000, the total loss has stabilized, that means the model had successfully reached a reliable local minimum.

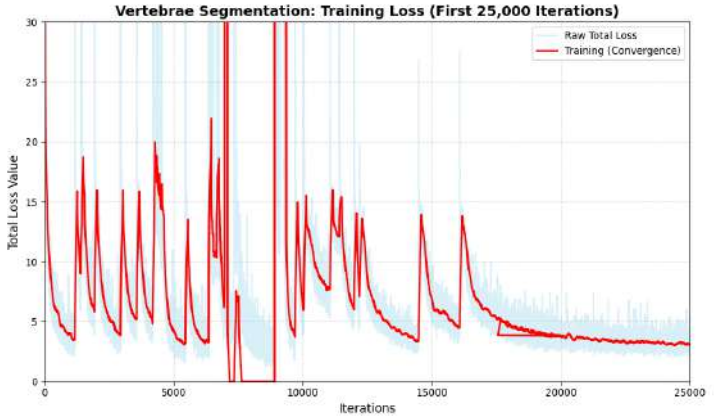

Figure 9. loss curve

7.2 Segmentation Accuracy:

The mAP (mean Average Precision) curves show this stability. Figure 10 show mAP@50, the model achieved good performance (approaching 98.67%) for both bounding box detection and mask segmentation. Even under a much stricter test of mAP@50-95, the model maintained an appealing accuracy of approximately 73–75%. It indicates that the model shows good precision in defining the exact edges and contours of each vertebra as illustrated in Figure 11.

To test how well the model performs in real-world scenario, we compared the Yolact++ automated Cobb angle

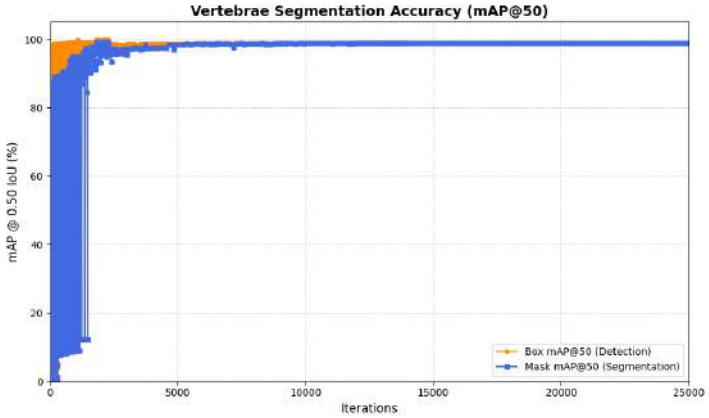

Figure 10. Accuracy curve @50

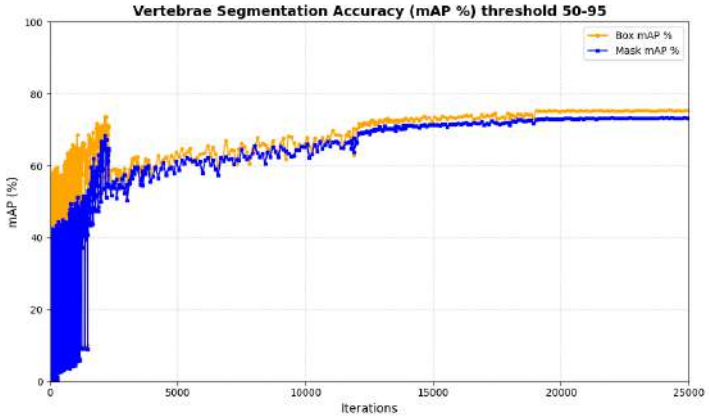

Figure 11. Accuracy curve@50-75

predictions with expert ground truth measurements.

7.3 Linear Correlation:

The scatter plot in Figure 12 shows an incredibly strong correlation, with an R2 value of 0.966. The data points closely follow the identity line (y = x), which proves that the model does not systematically over-estimating or under-estimating the scoliosis severities across a wide range of angles (from 0 degree to nearly 80 degree).

7.4 Agreement Analysis:

The Bland-Altman plot in Figure 13 gives the most human-centric validation. In our results an average bias of 1.10 degree is observed, which is considered insignificant in clinical setting. However, inter-observer variability among two human radiologists is usually between 3 to 5 degrees. Since most data points fall

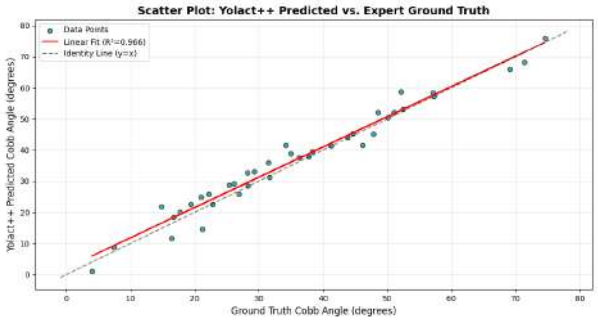

Figure 12. Scatter Plot

within the 1.96 which is a standard limits of agreement (-5.16 to 7.35 degree), it confirms that this automated system is effectively reliable as a second human opinion.

However, some outliers exist on which the model shows low performance, specifically in the assessment of larger Cobb angles. This shortcoming can be attributed to the fact that large spinal angles lead to considerable overlap of the vertebrae, thereby causing structural distortion, which may not follow the training patterns. Future work aims to improve the model architecture that could able to capture these outliers. It could be the use of advanced data augmentation methods, includes a diverse set of high-degree Scoliosis images or enhanced model architecture. Moreover, an ensemble method of these all methods may output better results.

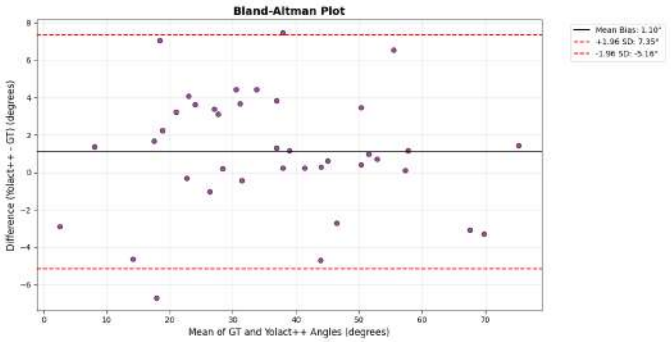

Figure 13. Bland-Altman Plot

In a comparative evaluation using the SMAPE metric, the YOLACT++ model achieved the lowest error rate of 8.11%, indicating the highest performance among the considered approaches. While YOLACT++ was trained and tested in this study, the results for other

Table 2. Detailed Comparison of Cobb angle Measurements: Expert vs. YOLACT++

Image ID	Upper Vert.	Lower Vert.	Expert Angle (°)	YOLACT++ (°)	Diff.
65	DV-5	DV-12	74.57	76.00	+1.43
	DV-11	LV-4	52.49	53.19	+0.70
146	DV-4	DV-10	57.13	58.30	+1.17
	DV-9	LV-3	50.06	50.46	+0.40
150	DV-6	LV-2	44.67	45.29	+0.60
185	DV-6	LV-1	51.02	52.01	+1.00
193	DV-6	DV-11	43.84	44.14	+0.30
214	DV-9	LV-3	26.14	29.26	+3.12
243	DV-5	DV-12	28.27	28.46	+0.20
279	DV-5	DV-12	69.08	66.00	-3.08
	LV-1	LV-5	71.39	68.11	-3.28
281	DV-5	LV-1	38.38	39.54	+1.20
296	DV-5	LV-1	28.26	32.69	+4.43
310	DV-6	DV-12	17.76	20.00	+2.24
347	DV-8	LV-1	22.18	25.81	+3.63
384	DV-6	LV-1	20.93	25.00	+4.07
404	DV-5	LV-2	34.99	38.82	+3.83
405	DV-5	DV-12	31.50	35.95	+4.45
408	DV-10	LV-3	41.23	41.45	+0.22
412	DV-5	DV-12	47.83	45.11	-2.72
428	DV-11	LV-4	14.85	21.89	+7.04
445	DV-6	DV-11	7.44	8.80	+1.36
452	DV-5	LV-2	25.34	28.74	+3.40
453	DV-5	LV-2	29.30	33.00	+3.70
490	DV-9	LV-4	31.63	31.21	-0.42
500	DV-5	LV-1	52.15	58.68	+6.53
516	DV-2	LV-1	48.51	52.00	+3.49
605	DV-5	LV-1	37.78	38.00	+0.22
102	DV-11	LV-4	26.88	25.84	-1.04
157	DV-8	DV-12	46.18	41.48	-4.70
160	DV-5	DV-10	21.26	14.56	-6.70
182	DV-6	LV-2	22.83	22.53	-0.30
189	DV-5	DV-10	4.00	1.10	-2.90
196	DV-5	DV-11	57.23	57.35	+0.12
200	DV-6	LV-1	36.27	37.59	+1.32
204	DV-9	LV-4	19.39	22.62	+3.23
210	DV-4	LV-1	34.17	41.63	+7.46
303	DV-10	LV-3	16.42	11.78	-4.64
320	DV-11	LV-2	16.68	18.36	+1.68

models—such as S2VR (37.08%), FAST R-CNN (25.69%), and BoostNet (23.44%)—were taken from existing literature. Additional comparative models include Multiview Extrapolation Net (18.95%), U-Net (16.48%), ResNet (10.81%), and YOLACT (10.76%). These findings highlight YOLACT++ as the most accurate method for the given task, according to the SMAPE metric. Also, previous approaches like ResNet use two models, whereas YOLACT++ uses only one and achieves a remarkable SMAPE of 8.11%. The findings indicates that the Yolact++ architecture is more than just a technically accurate tool in a laboratory setting but clinically "safe" to assist with medical diagnoses.

The high map values at 25,000 iterations suggest that the model has mastered the challenging task of identifying overlapping spinal structures in X-ray images. Furthermore, the high precision shown in the Bland-Altman analysis suggests that this tool could significantly reduce the time needed for scoliosis screening while maintaining a level of precision that matches expert performance. By automating the "monotonous" task of manual vertebral marking, we allow clinicians to focus more on treatment planning and patient interaction, ultimately humanizing the diagnostic process through AI-driven efficiency.

In the above Table 2, various images from the dataset's image field display the image number, followed by the upper and lower vertebrae the DV represent dorsal vertebrae which is vertebrae located in thoracic region of spine and are from DV1 upto DV12 and LV shows lumbar vertebrae which count from LV1 to LV5, next column show the angle determined by doctors or professionals. The second-to-last column shows the angle calculated by the YOLACT++ model, while the final column presents the difference between the YOLACT++ model's angle and those calculated by others professionals, a difference of 5 degree is clinically acceptable. The ± symbol represent angle are above or below than those calculated by professionals.

8 Conclusion and Future work

In this research study, the YOLACT++ instance segmentation model is adopted to calculate the Cobb angle measurement on a dataset comprised of 609 images. YOLACT++ is considered more accurate, stable, and computational efficient. An SMAPE of 8.11% was achieved for the Cobb angle estimation which demonstrates that YOLACT++ is precise for segmentation and Cobb angle measurement. Further improvements could be possible by targeting cases includes severe curvature and a dataset of diverse nature like high vertebral rotation and overlapping structures.

9 Acknowledgement

The authors would like to thank the original dataset providers for making the data publicly available, which enabled this research.

10 Dataset availability

The dataset analyzed in this research study is publicly available at the https://aasce19.github.io/

Author Contributions

Inzamam Ul Haq: Conceptualization, Methodology, Software **Izaz Ullah**: Data curation, Writing- Original draft preparation. **Fazli Amin Khalil**: Visualization, Investigation. **Sakin Jan**: Supervision.: **Israr Ahmed Khan**: Software, Validation.

Compliance with Ethical Standards

It is declare that all authors don't have any conflict of interest. It is also declare that this article does not contain any studies with human participants or animals performed by any of the authors. Furthermore, informed consent was obtained from all individual participants included in the study.

Funding Information

This research received no external funding.

References

[1] C. Du, J. Yu, J. Zhang, J. Jiang, H. Lai, W. Liu, *et al.*, "Relevant areas of functioning in patients with adolescent idiopathic scoliosis on the International Classification of Functioning, Disability and Health: The patients' perspective," *J. Rehabil. Med.*, vol. 48, no. 9, pp. 806–814, 2016.

[2] T. R. Kuklo, B. K. Potter, T. M. Schroeder, and M. F. O'Brien, "Comparison of manual and digital measurements in adolescent idiopathic scoliosis," *Spine*, vol. 31, no. 11, pp. 1240–1246, 2006.

[3] R. T. Morrissy, G. Goldsmith, E. Hall, D. Kehl, and G. Cowie, "Measurement of the Cobb angle on radiographs of patients who have scoliosis. Evaluation of intrinsic error," *JBJS*, vol. 72, no. 3, pp. 320–327, 1990.

[4] S. Langensiepen, O. Semler, R. Sobottke, *et al.*, "Measuring procedures to determine the Cobb angle in idiopathic scoliosis: a systematic review," *Eur. Spine J.*, vol. 22, no. 11, pp. 2360–2371, 2013.

[5] M. C. Tanure, A. P. Pinheiro, and A. S. Oliveira, "Reliability assessment of Cobb angle measurements using manual and digital methods," *The Spine J.*, vol. 10, no. 9, pp. 769–774, 2010.

[6] M. Gstoettner, K. Sekyra, N. Walochnik, *et al.*, "Inter-and intraobserver reliability assessment of the Cobb angle: manual versus digital measurement tools," *Eur. Spine J.*, vol. 16, no. 10, pp. 1587–1592, 2007.

[7] K. Chen, X. Zhai, K. Sun, *et al.*, "A narrative review of machine learning as promising revolution in clinical practice of scoliosis," *Ann. Transl. Med.*, vol. 9, no. 1, p. 67, 2021.

[8] F. Bao, D. Wang, H. Zhao, and B. Xu, "Application of adaptive threshold image segmentation algorithm in orthopedic CT imaging," *J. Med. Imaging Health Inf.*, vol. 9, no. 8, pp. 1736–1740, 2019.

[9] J. Mao, K. Wang, Y. Hu, W. Sheng, and Q. Feng, "GrabCut algorithm for dental X-ray images based on full threshold segmentation," *IET Image Process.*, vol. 12, no. 12, pp. 2330–2335, 2018.

[10] L. Ťepelea, L. Gavriluţ, and A. Gacsádi, "Edge Based CNN Image Segmentation Methods for Medical Imaging," *J. Comput. Sci. & Control Syst.*, vol. 3, no. 2, 2010.

[11] Y. Mohamed Ben Ali, "Edge-based segmentation using robust evolutionary algorithm applied to medical images," *J. Signal Process. Syst.*, vol. 54, no. 1, pp. 231–238, 2009.

[12] H. Wu, C. Bailey, P. Rasoulinejad, and S. Li, "Automatic landmark estimation for adolescent idiopathic scoliosis assessment using BoostNet," in *Proc. MICCAI*, 2017, pp. 127–135.

[13] G. K. Prabhu, "Automatic quantification of spinal curvature in scoliotic radiograph using image processing," *J. Med. Syst.*, vol. 36, no. 3, pp. 1943–1951, 2012.

[14] H. Anitha, A. Karunakar, and K. V. N. Dinesh, "Automatic extraction of vertebral endplates from scoliotic radiographs using customized filter," *Biomed. Eng. Lett.*, vol. 4, no. 2, pp. 158–165, 2014.

[15] T. A. Sardjono, M. H. F. Wilkinson, A. G. Veldhuizen, *et al.*, "Automatic Cobb angle determination from radiographic images," *Spine*, vol. 38, no. 20, pp. E1256–E1262, 2013.

[16] M. Sánchez-Fernández, M. de-Prado-Cumplido, J. Arenas-García, and F. Pérez-Cruz, "SVM multiregression for nonlinear channel estimation in multiple-input multiple-output systems," *IEEE Trans. Signal Process.*, vol. 52, no. 8, pp. 2298–2307, 2004.

[17] X. Zhen, Z. Wang, A. Islam, *et al.*, "Multi-scale deep networks and regression forests for direct bi-ventricular volume estimation," *Med. Image Anal.*, vol. 30, pp. 120–129, 2016.

[18] T. Kooi, G. Litjens, B. Van Ginneken, *et al.*, "Large scale deep learning for computer aided detection of mammographic lesions," *Med. Image Anal.*, vol. 35, pp. 303–312, 2017.

[19] P. F. Christ, M. E. A. Elshaer, F. Ettlinger, *et al.*, "Automatic liver and lesion segmentation in CT using cascaded fully convolutional neural networks and 3D conditional random fields," in *Proc. MICCAI*, 2016, pp. 415–423.

[20] S. Jan, F. A. Khalil, S. B. S. Abid, I. Ullah, I. Ullah, and I. U. Haq, "Optimization of Feature Selection using Firework Algorithm for Machine Learning Algorithm," *VAWKUM Trans. Comput. Sci.*, vol. 13, no. 2, pp. 250–262, 2025.

[21] E. Acuña and C. Rodriguez, "On detection of outliers and their effect in supervised classification," *Univ. Puerto Rico Mayaguez*, vol. 15, 2004.

[22] H. Sun, X. Zhen, C. Bailey, *et al.*, "Direct estimation of spinal Cobb angles by structured multi-output regression," in *Proc. IPMI*, 2017, pp. 529–540.

[23] M. Horng, C. Kuok, M. Fu, *et al.*, "Cobb angle measurement of spine from X-ray images using convolutional neural network," *Comput. Math. Methods Med.*, vol. 2019, Art. no. 6357171, 2019.

[24] Y. Pan, Q. Chen, T. Chen, *et al.*, "Evaluation of a computer-aided method for measuring the Cobb angle on chest X-rays," *Eur. Spine J.*, vol. 28, no. 12, pp. 3035–3043, 2019.

[25] Z. Peng, Z. Wang, M. Sun, *et al.*, "Graph convolutional networks for 3D skeleton-based scoliosis screening using gait sequences," *The Vis. Comput.*, pp. 1–13, 2025.

[26] Q. Liu, Z. Lu, R. Gao, X. Bu, and N. Hanajima, "SimpleMask: parameter link and efficient instance segmentation," *The Vis. Comput.*, vol. 41, no. 3, pp. 1573–1589, 2025.

[27] B. Khanal, L. Dahal, P. Adhikari, and B. Khanal, "Automatic Cobb angle detection using vertebra detector and vertebra corners regression," in *Proc. Spine Imaging Workshop*, 2019, pp. 81–87.

[28] J. Yi, P. Wu, Q. Huang, H. Qu, and D. N. Metaxas, "Vertebra-focused landmark detection for scoliosis assessment," in *Proc. IEEE ISBI*, 2020, pp. 736–740.

[29] C. Chen, K. Namdar, Y. Wu, *et al.*, "Automating Cobb angle measurement for adolescent idiopathic scoliosis using instance segmentation," in *Proc. IEEE EMBC*, 2024, pp. 1–5.

[30] D. Bolya, C. Zhou, F. Xiao, and Y. J. Lee, "YOLACT: Real-time instance segmentation," in *Proc. IEEE/CVF ICCV*, 2019, pp. 9157–9166.

[31] C. Zhou, *YOLACT++ Better Real-Time Instance Segmentation*, Davis, CA: Univ. California, 2020.

[32] J. Kim, T. Kim, T. Kim, *et al.*, "Morphology-aware interactive keypoint estimation," in *Proc. MICCAI*, 2022, pp. 675–685.

[33] V. Kukreja *et al.*, "Segmentation and Contour Detection for handwritten mathematical expressions using OpenCV," in *Proc. IEEE DASA*, 2022, pp. 305–310.