

B-spine: Learning B-spline Curve Representation for Robust and Interpretable Spinal Curvature Estimation

Hao Wang¹, Qiang Song², Ruofeng Yin^{2*}, Rui Ma^{1,3*}

¹ School of Artificial Intelligence, Jilin University

² China-Japan Union Hospital, Jilin University

³ Engineering Research Center of Knowledge-Driven Human-Machine Intelligence, MOE, China
{whao22, songqiang20}@mails.jlu.edu.cn, {yrf.wind, ruim}@jlu.edu.cn

Abstract

Spinal curvature estimation is important to the diagnosis and treatment of the scoliosis. Existing methods face several issues such as the need of expensive annotations on the vertebral landmarks and being sensitive to the image quality. It is challenging to achieve robust estimation and obtain interpretable results, especially for low-quality images which are blurry and hazy. In this paper, we propose B-Spine, a novel deep learning pipeline to learn B-spline curve representation of the spine and estimate the Cobb angles for spinal curvature estimation from low-quality X-ray images. Given a low-quality input, a novel SegRefine network which employs the unpaired image-to-image translation is proposed to generate a high quality spine mask from the initial segmentation result. Next, a novel mask-based B-spline prediction model is proposed to predict the B-spline curve for the spine centerline. Finally, the Cobb angles are estimated by a hybrid approach which combines the curve slope analysis and a curve-based regression model. We conduct quantitative and qualitative comparisons with the representative and SOTA learning-based methods on the public AASCE2019 dataset and our new proposed JLU-CJUH dataset which contains more challenging low-quality images. The superior performance on both datasets shows our method can achieve both robustness and interpretability for spinal curvature estimation.

1 Introduction

Scoliosis is a condition of spinal deformity which can lead to a side-to-side spine curve. Serious scoliosis can affect the breathing and movement of the patients and may even cause the disability if they are not treated in time. A quantitative measurement of the spinal curvature is important to diagnose the scoliosis and provide the guidance for proper treatment. Conventionally, the spinal curvature estimation is operated as the Cobb angle measurement in the clinical practice. Given an AP view X-ray spine image of a person in standing position, the Cobb angles, main thoracic (MT), proximal thoracic (PT) and thoracolumbar/lumbar (TL), can be measured by identifying the most tilted vertebra and computing the angles between the lines drawn from the endplates of involved vertebra. However, manual measurement of Cobb angles requires skills from a professional clinician,

*Corresponding authors.

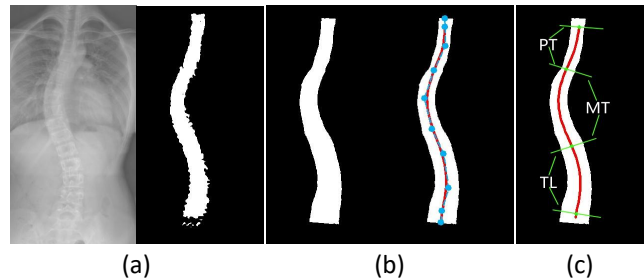


Figure 1: Given a low-quality X-ray image, our method can robustly learn a B-spline curve representation for the spine centerline and estimate the Cobb angles in an interpretable manner. (a) Input X-ray image and the initial spine mask obtained by an existing segmentation network; (b) Refined mask and the learned B-spline (note the blue points are the predicted control points); (c) Estimated location of endplates for computing the Cobb angles (MT, PT, TL).

while the process is time-consuming and may lead to subjective errors due to high inter- and intra-observer variability (Gstoettner et al. 2007).

In recent years, computer-aided spinal curvature or Cobb angle estimation has been widely studied. Early methods such as (Zhang et al. 2010; Samuvel et al. 2012; Shaw et al. 2012; Prabhu et al. 2012; Akbar et al. 2013; Wu et al. 2014) employ traditional image processing techniques to identify the upper and lower end vertebra for computing the Cobb angles. These methods mainly have limited accuracy and may be affected by subjective errors. With the development of machine learning and deep learning, data-driven approaches have achieved promising results by learning Cobb angle estimation based on annotated datasets such as AASCE2019 (Li and Wang 2019). Existing learning-based methods mainly predict the Cobb angles based on certain intermediate representations of the spine including spine segmentation mask, vertebral landmarks and the centerline of the spine. Despite relatively accurate estimation can be obtained by these learning-based methods, they are still limited in the *robustness* or may lack the *interpretability* of the results.

For example, methods like (Wang, Huang, and Wang 2020; Lin et al. 2021) which directly regress the Cobb angles from the segmentation mask cannot explain where the

scoliosis occurs. In contrast, the landmarks-based methods (Wu et al. 2017; Lin et al. 2020) aim to reconstruct the shape of each vertebra which can lead to well interpretable results. However, predicting accurate landmarks is challenging since the performance of the model highly relies on the expensive landmarks annotation on the training data and the results are also sensitive to the image quality. Moreover, for low-quality images which look blurry and hazy as Figure 1 (a), it is difficult to annotate the ground truth landmarks of each vertebra even by human. Hence, the landmarks-based methods may not be suitable or robust for low-quality images. On the other hand, it is possible to attain both the robustness and interpretability by estimating the centerline of the spine and then predicting the Cobb angles based on the curvature analysis. Yet, existing centerline-based methods such as (Tu et al. 2019; Dubost et al. 2020) usually represent the centerline as a fitted polynomial curve or just a collection of connected points, and such curve representations are sensitive to the noise in the input images. A more proper curve representation which is smooth, stable and easy to learn is needed to improve the robustness of the centerline-based methods.

In this paper, we propose B-Spine, a novel deep learning pipeline to learn B-spline curve representation of the spine centerline and estimate the Cobb angles from low-quality X-ray spine images. Our goal is to achieve both robustness and interpretability for the spinal curvature estimation by representing the spine with the powerful and well-suited B-spline (Unser 2002). Our pipeline consists of three modules: Spine Mask Segmentation, B-Spline Centerline Prediction and Cobb Angle Estimation. First, we propose a novel SegRefine network which employs the unpaired image-to-image translation, i.e., CycleGAN (Zhu et al. 2017), to refine the initial segmentation mask obtained from the low-quality X-ray image. Next, we learn a B-spline curve representation for the centerline from the segmentation mask via a novel mask-based B-spline prediction model, while the control points and knots for constructing the B-spline curve are jointly predicted under supervision from the GT parameters and a new point-based *resample loss*. Finally, we propose a hybrid approach which combines the curve slope analysis and a curve-based regression model to obtain final Cobb angles.

For evaluation, we perform experiments using 609 images in the public AASCE2019 dataset and compare the results with representative and SOTA methods which are based on segmentation mask, landmarks and centerline. In addition, we collect and annotate a more challenging new dataset, named JLU-CJUH dataset, which contains 584 low-quality X-ray spine images for evaluation. Extensive quantitative and qualitative results show our B-Spine can generally outperform the compared methods. We also conduct ablation studies to verify our main modules. The code is available at <https://github.com/JLU-ICL/B-Spine>.

In summary, our contribution are as follows:

- We propose B-Spine, a novel deep learning pipeline to learn B-spline curve representation of the spine centerline and estimate the Cobb angles from low-quality X-ray images. To our best knowledge, we are the first to represent the spine using the B-spline curve and our novel multi-stage framework can predict the B-spline and Cobb

angles in a robust and interpretable manner.

- We propose a novel SegRefine network that employs the unpaired image-to-image translation to generate a high-quality spine mask from the initial segmentation result. With the refined mask, our B-spline prediction is less affected by the quality of the input image.
- We conduct quantitative and qualitative comparisons with the representative and SOTA learning-based methods on the public AASCE2019 dataset and our new proposed JLU-CJUH dataset. From the results, our B-Spine framework achieves superior performance especially for challenging low-quality images.

2 Related Work

Computer-aided Cobb angle estimation has been widely studied in the medical imaging and computer vision community. In this section, we mainly focus on reviewing recent learning-based methods and briefly reviewing the B-spline representation and its application in medical imaging.

Estimation based on the spine segmentation mask.

Thanks to the rapid development of segmentation techniques (Ronneberger, Fischer, and Brox 2015; Huang et al. 2020), the spine segmentation mask can be extracted in a relatively accurate and efficient manner. Based on the learned segmentation mask, (Wang, Huang, and Wang 2020) trains an angle regression network to predict the Cobb angles. Instead of training the segmentation and angle regression networks separately, Seg4Reg+ (Lin et al. 2021) jointly optimizes these two networks and achieves the SOTA performance for Cobb angle estimation. As these methods mainly perform the prediction by direct regression on the segmentation mask, their results are not well interpretable, i.e., it is unknown which image region the angles are corresponding to. In our B-Spine, the segmentation mask is used to predict the B-spline centerline which can be processed to obtain interpretable angle estimation. Furthermore, to improve the noisy segmentation mask obtained from the low-quality X-ray image, we take the inspiration from (Gu et al. 2021) and employ the unpaired image-to-image translation model CycleGAN (Zhu et al. 2017) to generate refined mask with most of noise removed. With the refined mask, we can perform more robust mask-based B-spline prediction.

Estimation based on vertebral landmarks. Besides the spine segmentation mask, the shape of each vertebra can provide more geometric and structural information of the spine. Normally, the four corners of each vertebra are annotated as the landmarks which can be predicted from an input X-ray image and be used to estimate the Cobb angles. The landmarks prediction is usually treated as a regression task. (Wu et al. 2017) proposes the BoostNet which integrates statistical outlier removal methods into CNN-based feature extraction to obtain robust image features for landmarks regression. The data from (Wu et al. 2017) which contain landmarks annotations on 17 vertebra composed of the thoracic and lumbar spine later become the base of the AASCE2019 dataset (Li and Wang 2019) and promote the advancement of landmarks-based spinal curvature estimation (Chen et al. 2019; Lin et al. 2020; Kim et al. 2020;

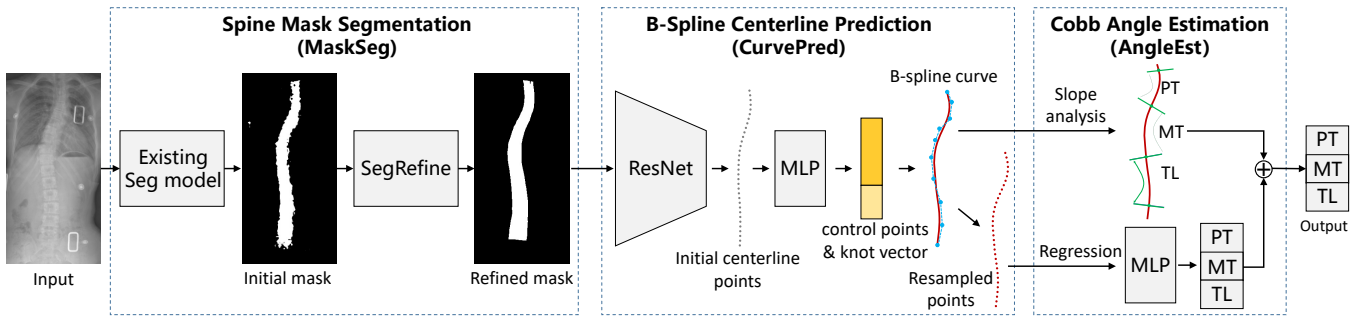


Figure 2: Overview of our B-Spine pipeline for estimating Cobb angles from a low-quality X-ray image.

Zhang et al. 2021). In addition to predicting the landmarks from the single AP view X-ray images, (Wu et al. 2018; Wang et al. 2019) utilize multi-view X-ray images for more robust Cobb angle estimation. Due to the annotation difficulty and estimation error, it is hard to obtain accurate landmarks for low-quality images. In contrast, we leverage the segmentation mask which is easier to obtain, and conduct mask-based B-spline centerline prediction.

Estimation based on the spine centerline. Since the spine itself is in a curve shape, the spinal curvature estimation can be directly performed on the spine centerline curve extracted based on input X-ray image, segmentation mask or the extracted vertebra shape. In general, the centerline can be represented in different curve representations. For example, (Alharbi et al. 2020) directly connects the center points of the vertebra estimated from the input X-ray image using CNN. In (Tu et al. 2019), the centerline curve is obtained by fitting a 6-degree polynomial curve to the center points extracted based on the spine segmentation mask. To improve the robustness of center points extraction, (Dubost et al. 2020) proposes CasNet which employs cascaded segmentation networks to obtain the segmentation mask of the centerline, from which the dense center points are extracted, connected and smoothed to form the centerline curve. For above methods, the Cobb angles can be obtained by analyzing slopes of the centerline curve. On the other hand, (Huo et al. 2021) jointly optimizes the centerline extraction and the centerline-based Cobb angle regression. Comparing to above methods, we adopt the B-spline curve for more robust centerline extraction and design a specific network to learn B-spline parameters from the segmentation mask. To further improve the robustness and interpretability of our pipeline, we use a hybrid scheme to combine results from the curve slope analysis and curve-based Cobb angle regression.

B-spline representation and its application. B-spline is a powerful curve or surface representation which is widely used in computer graphics and computer-aided design (Gordon and Riesenfeld 1974; Patrikalakis and Maekawa 2002). In the medical imaging domain, B-spline representation has also been investigated for image interpolation (Lehmann, Gonner, and Spitzer 1999), segmentation (Brigger, Hoeg, and Unser 2000) and registration (Klein, Staring, and Pluim 2007) etc. By using B-spline to represent the curves, the results are usually smoother and more robust (Williams et al.

2021). Meanwhile, to our best knowledge, there is no prior work that uses the B-spline to represent the spine centerline. Recently, some works apply deep learning to predict B-spline representation from point sequence (Laube, Franz, and Umlauf 2018) and segmentation boundary (Barrowclough et al. 2021). We take the similar idea of using deep learning to predict B-spline parameters, while our network is designed for the mask-based centerline approximation.

3 Method

In this section, we first briefly review the preliminaries on B-spline curve including the mathematical definition and construction algorithm. Then, we introduce details about the three modules of our B-Spine framework. Figure 2 shows an overview of our multi-stage pipeline.

3.1 Preliminaries on B-Spline Curve

Spline curve is a smooth curve widely used for curve fitting or interpolation. It is defined by using piecewise polynomials to approximate or pass through a series of given points. B-spline or basis spline curve is a specific type of spline curve which has good properties such as local support, flexibility and high approximation accuracy. Given $n + 1$ control points $\{P_0, P_1, \dots, P_n\}$ and a knot vector $\mathbf{U} = \{u_0, u_1, \dots, u_m\}$, a B-spline curve of degree p is defined as:

$$C(u) = \sum_{i=0}^n B_{i,p}(u)P_i, \tag{1}$$

where $B_{i,p}(u)$ are B-spline basis functions of degree p . The most common way to define the B-spline basis functions is using the Cox-de Boor recursion formula:

$$B_{i,0}(u) = \begin{cases} 1, & \text{if } u_i \leq u < u_{i+1}, \\ 0, & \text{otherwise,} \end{cases} \tag{2}$$

$$B_{i,p}(u) = \frac{u-u_i}{u_{i+p}-u_i} B_{i,p-1}(u) + \frac{u_{i+1}-u}{u_{i+1}-u_{i+1}} B_{i+1,p-1}(u). \tag{3}$$

Based on the definition of B-spline, the following relationship between the degree p , the number of control points $n+1$ and the length of knot vector $m+1$ needs to be satisfied: $m = n + p + 1$. In this paper, we adopt a special case of B-spline, i.e., clamped B-spline, for which the curve passes through the first and last control points and is also tangent to the first and last legs of its control polyline. To enable the clamped effect, it is required the first $p + 1$ and the last $p + 1$ knots must be identical. Please refer to (Patrikalakis and Maekawa 2002) for more introduction about B-spline.

3.2 Spine Mask Segmentation

The first module of our B-Spine framework is Spine Mask Segmentation or MaskSeg in short. Given an input X-ray image, we first segment the spine mask with a conventional semantic segmentation network. Here, we employ SSformer (Shi, Xu, and Gao 2022) which is a lightweight transformer-based method and enhance it with an additional structural similarity index (SSIM) loss (Wang et al. 2004) for better segmentation quality. Other semantic segmentation methods such as U-Net (Ronneberger, Fischer, and Brox 2015) can also be used in this step to obtain an initial segmentation mask. However, due to the low quality of X-ray images, there may be significant noise in the initial mask (see Figure 1 (a)). Also, according to our experiments, it is difficult to reduce the noise by simply fine-tuning the network or using other segmentation models.

Inspired by the application of generative adversarial networks (GANs) for denoising low quality medical images (Gu et al. 2021), we propose a SegRefine network which refines the initial noisy mask based on the unpaired image-to-image translation technique. Specifically, we train a CycleGAN model (Zhu et al. 2017) with a set of noisy masks predicted by the conventional segmentation network and a set of GT masks, while the images of these two sets are unpaired which can improve the generalization ability of the SegRefine network. After training, the SegRefine network can translate/generate a refined mask from the initial noisy mask, or vice versa. Comparing to the supervised learning based method, our unsupervised SegRefine model will mainly focus on converting images in one set into the *style* of images in the other set. Thus, it is more robust and more generalizable to the more diversified masks obtained from low quality images.

3.3 B-Spline Centerline Prediction

The second module of our pipeline is B-spline Centerline Prediction or CurvePred, which predicts B-spline parameters for representing the spine centerline from the segmentation mask. In our implementation, we choose the clamped B-spline of degree $p = 3$ and the number of control points and knots are 10 and 14, respectively. Following the constraints for the clamped B-spline, the first 4 knots are set to be 0 and the last 4 knots are set to be 1, i.e., the knot vector become to $\{0, 0, 0, 0, u_4, u_5, \dots, u_9, 1, 1, 1, 1\}$. Hence, the B-spline parameters to predict include positions of 10 control points and values of 6 knots. We formulate the B-spline prediction as a regression problem and use GT values as supervision to compute losses for these parameters. Besides, we also propose a new *resample loss* which is computed based on the centerline points resampled from the predicted and GT B-spline to enforce explicit point supervision on the B-spline curve. In Section 4.1, we provide details for how to prepare the GT centerline points. Overall, the CurvePred module contains three steps: 1) predict initial centerline points from the refined segmentation mask; 2) predict B-spline parameters; 3) resample centerline points from the predicted B-spline and use them to compute the resample loss.

Initial centerline points prediction. Directly predicting the B-spline parameters from the segmentation mask may

lead to unstable results. Therefore, we first predict a set of initial centerline points from the mask and then use these points as input for predicting the B-spline parameters. To predict the centerline points, we modify the last layer of ResNet101 (He et al. 2016) so that the coordinates of 34 centerline points are predicted under the supervision of GT centerline points. The loss for predicting initial centerline points is defined as:

$$L_{init} = \frac{1}{34} \sum_{i=1}^{34} \|C_i^{init} - C_i^{GT}\|_2^2, \quad (4)$$

where C_i^{init} and C_i^{GT} denote i -th predicted or GT centerline points and $\|\cdot\|_2$ is the L2 norm.

B-spline parameters prediction. From initial centerline points, we use a 4-layer MLP to predict the B-spline parameters, i.e., the control points and knots, and define the loss:

$$L_{paras} = \frac{1}{10} \sum_{i=0}^9 \|P_i^{pred} - P_i^{GT}\|_2^2 + \frac{1}{6} \sum_{i=4}^9 \|U_i^{pred} - U_i^{GT}\|_2^2, \quad (5)$$

where P_i^{pred} , P_i^{GT} , U_i^{pred} , U_i^{GT} denote the i -th predicted or GT control points or knots, respectively.

Resample loss. After predicting the control points and knots, we construct the B-spline curve $C(u)$ using the Cox-de Boor algorithm. As the B-spline parameters are predicted from the initial centerline points, the accuracy of the curve may be affected by the quality of the segmentation mask. To improve the robustness of curve prediction and boost the performance, we propose a new resample loss to introduce more constraints on the centerline points computed from the predicted curve. Specifically, we use the same set of parameters $\{u_i\}$ for generating the GT centerline points and obtain a set of resampled centerline points $\{C(u_i)\}$. Then, the resample loss is defined as:

$$L_{resample} = \frac{1}{34} \sum_{i=1}^{34} \|C(u_i) - C_i^{GT}\|_2^2, \quad (6)$$

where u_i is the i -th curve parameter obtained by uniform sampling from $[0, 1]$, following the same procedure for GT.

In summary, the total loss for CurvePred module is:

$$L_{CurvePred} = \lambda_1 \cdot L_{init} + \lambda_2 \cdot L_{paras} + \lambda_3 \cdot L_{resample}. \quad (7)$$

In our current implementation, we set $\lambda_1 = 1$, $\lambda_2 = 0.1$, $\lambda_3 = 0.1$ and jointly train the ResNet101 and the MLP for initial centerline points and B-spline parameter prediction.

3.4 Cobb Angle Estimation

The last module of our pipeline is Cobb Angle Estimation or AngleEst. From the predicted B-spline centerline, the Cobb angles can be directly estimated by performing curve slope analysis. On the other hand, as a multi-stage pipeline, the errors from previous modules may accumulate and lead to inaccurate B-spline curve. To alleviate this issue, we further learn a regression model to predict the Cobb angles from the learned B-spline curve. Then, we combine the results from both schemes and obtain the final Cobb angle estimation.

Cobb angles from curve slope analysis. Similar to previous methods (Dubost et al. 2020), we can estimate the Cobb

angles by finding the maximum and minimum slopes of the spine curve at sampled points. Given a predicted B-spline curve, we first sample 17 points, denoted as $\{S_i\}$, with equal arc length, while these points correspond to 12 thoracic vertebra and 5 lumbar vertebra. Next, for each point S_i , we sample another point S'_i that is sufficiently close to it so that we can compute the slope at S_i as:

$$\text{slope}_i = \frac{S'_{i,y} - S_{i,y}}{S'_{i,x} - S_{i,x}}, \quad (8)$$

where $S_{i,x}$, $S'_{i,x}$, $S_{i,y}$, $S'_{i,y}$ are the x and y coordinates for S_i and S'_i , respectively. Here, we assume the spine spans along the y direction and all S'_i are sampled at the same neighboring side of S_i . Then, we find the maximum and minimum slopes w.r.t the 17 sampled points and compute the MT angle as:

$$\text{Cobb}_{MT}^s = \frac{180}{\pi} \left| \arctan \frac{\text{slope}_{max} - \text{slope}_{min}}{1 + \text{slope}_{max} \cdot \text{slope}_{min}} \right| \quad (9)$$

Similarly, we can compute the PT or TL angle by finding the maximum or minimum slopes above or below the points used for computing MT. The results of the slope-based Cobb angles are denoted as $\text{Cobb}^s = \{\text{Cobb}_{MT}^s, \text{Cobb}_{PT}^s, \text{Cobb}_{TL}^s\}$.

Cobb angle regression. To regress the Cobb angles from the B-spline centerline, we reuse the 34 resampled points $\{C(u_i)\}$ as input and train a 4-layer MLP to predict the angles under the supervision of GT values, while MSE loss is used for each angle. The results of the predicted Cobb angles are denoted as $\text{Cobb}^r = \{\text{Cobb}_{MT}^r, \text{Cobb}_{PT}^r, \text{Cobb}_{TL}^r\}$.

Through our statistical analysis, Cobb^s and Cobb^r have different error distributions. When one method produces an over-estimated result w.r.t. to GT angle, the other one may produce an under-estimated result. Finally, we linearly combine the slope-based and regression-based Cobb angles and obtain the final results:

$$\text{Cobb} = \alpha \cdot \text{Cobb}^s + (1 - \alpha) \cdot \text{Cobb}^r, \quad (10)$$

where α can be selected based on the performance of each angle using the validation set of different datasets. From our experiments, we set $\alpha = 0.4$ for MT, and $\alpha = 0.5$ for PT and TL on AASCE2019 dataset, and $\alpha = 0.5$ for MT, and $\alpha = 0.8$ for PT and TL on JLU-CJUH dataset. More details about how α is selected are provided in Section 4.4.

4 Results

In this section, we first provide the details of two datasets and evaluation metrics we used for experiments. Then, we perform quantitative and qualitative comparisons with representative and SOTA methods to show the superiority of our method. Ablation studies on both datasets are also conducted to verify the effectiveness of proposed modules.

4.1 Datasets and Evaluation Metrics

AASCE2019 dataset. The MICCAI 2019 Challenge on Accurate Automated Spinal Curvature Estimation provides the dataset which is formed based on data from (Wu et al. 2017;

Wang et al. 2019). The dataset consists of 609 spinal AP X-ray images with 481 images used for training and 128 images for testing. Each image is annotated with the corners of each vertebra as landmarks (17×4 in total) and three Cobb angles (MT, PT and TL) by professional clinicians. To utilize this dataset in our framework, we first obtain the spine segmentation mask by connecting the left and right boarders of each vertebra to get the left and right mask contours, while the top and bottom boarders of the first and last vertebra become the top and bottom contours. Then, we extract the center of top and bottom boarders of each vertebra and get 34 initial centerline points. Then the GT parameters for the B-spline, including 10 control points, 14 knots (with 6 unknown knots), can be estimated by fitting the 3-degree clamped B-spline to the initial centerline points. In addition, we sample 34 GT centerline points from the B-spline curve by uniform sampling the curve parameter $u \in [0, 1]$.

JLU-CJUH dataset. To evaluate the performance of our method in more challenging scenarios, we build a new dataset which contains 584 spinal AP X-ray images collected from Jilin University China Japan Union Hospital. We split the dataset by 7:1:2, i.e., 408, 58 and 118 images for training, validation and testing, respectively. These images are normally in a lower quality than AASCE2019, e.g., blurrier and hazier, and it may be difficult to identify individual vertebra even by human. Thus, it is almost impossible to annotate the vertebral landmarks and train the landmarks-based methods on this dataset. On the other hand, it is easier to annotate the segmentation mask since only the rough shape needs to be labelled and the professionals have sufficient knowledge to identify the spine even from blurry images. Therefore, we design a labelling tool based on LabelMe (Russell et al. 2008) for manual annotation. Two professional local clinicians are trained to use our tool for annotating each image in JLU-CJUH dataset with the spine segmentation mask and three GT Cobb angles.

Evaluation metrics. Following previous works (Lin et al. 2020; Dubost et al. 2020), we evaluate the accuracy of the results using the mean absolute error (MAE) for each Cobb angle, in degrees. In addition, the symmetric mean absolute percent error (SMAPE) for all Cobb angles are computed:

$$\text{SMAPE} = \frac{1}{N} \sum_{i=1}^N \frac{\sum_{\theta \in \Theta} |A_{i,\theta} - A'_{i,\theta}|}{\sum_{\theta \in \Theta} (A_{i,\theta} + A'_{i,\theta})} \times 100\%, \quad (11)$$

Here, $\Theta = \{MT, PT, TL\}$ is the set of Cobb angles; $A_{i,\theta}$ and $A'_{i,\theta}$ are the ground truth and predicted Cobb angles; N is the number of images in the evaluation set.

4.2 Implementation Details

Our framework contains several networks which are trained in different manners. In MaskSeg, input X-ray images are resized to 512×256 following (Lin et al. 2020) and the SS-former is trained using the GT masks as supervision to produce the initial segmentation masks. Then, we train the CycleGAN using the initial segmentation masks and the GT masks as two sets of unpaired data, and the learned generator for the initial-to-refine mask translation is used as the SegRefine network. In CurvePred, we train the ResNet101

based image encoder and the MLP jointly, while the GT centerline points and GT B-spline parameters are used as supervision. In AngleEst, the MLP for regressing the Cobb angles is trained using the resampled B-spline centerline points as input and the GT angles as supervision. Note all networks are trained using images in the currently evaluated dataset. Our code is implemented in PyTorch and runs on a PC with a 3.6GHz CPU and one NVIDIA RTX 3060 GPU. In total, it takes 20h to train the whole pipeline for AASCE2019 or JLU-CJUH. For testing, it takes about 0.071s to estimate Cobb angles from an X-ray image.

4.3 Comparisons

Quantitative comparisons. We perform quantitative evaluations on both AASCE2019 and JLU-CJUH dataset to compare the performance of our B-Spine framework with existing representative and SOTA methods. For AASCE2019, the compared methods include: Seg4Reg (Lin et al. 2020), a representative landmarks-based method which wins the MICCAI 2019 AASCE Challenge; Seg4Reg+ (Lin et al. 2021), a mask-based method which achieves the SOTA performance on AASCE2019; CasNet (Dubost et al. 2020), a representative centerline-based method that predicts the dense centerline points based on cascaded segmentation networks; JointNet (Huo et al. 2021), a joint centerline prediction and Cobb angle estimation method which achieves the SOTA performance on AASCE2019 among the centerline-based works. We compare the MAE and SMAPE values reported in their paper. For JLU-CJUH dataset, we only compare with Seg4Reg and CasNet as only the code for Seg4Reg is available and the CasNet is relatively straightforward to implement. For Seg4Reg+ and JointNet, we do not try to implement their pipelines as they are quite involved and our implementation may not well reflect the capability of their methods. To produce the results, we use the released code of Seg4Reg and CasNet implemented by ourselves to train the models with images in JLU-CJUH dataset.

From Table 1, our method achieves the best performance except the SMAPE on AASCE2019. One possible reason is the SMAPE metric itself prefers the predicted angles which are larger than GT, i.e., since the predicted values also appear in the denominator (see Equation 11), SMAPE would produce smaller values when the predicted angle is larger, even if the relative absolute error from a smaller predicted angle is the same to the error of the larger prediction.

Qualitative comparisons. To quantitatively evaluate our method, in Figure 3, we show the visualization of our results and compare them with the results from the centerline-based method CasNet. The reason for choosing CasNet for comparison is other methods such as Seg4Reg and Seg4Reg+ are mainly regression-based and non-interpretable and no qualitative results can be visualized. From the visualized results, our method is able to perform robust and accurate prediction for low-quality images with noisy or partially broken masks.

4.4 Ablation Studies

Effectiveness of the SegRefine network. The SegRefine network improves the quality of the segmentation mask

Method	MT	PT	TL	SMAPE
Seg4Reg	4.71	5.79	5.65	10.85
Seg4Reg+	3.88	4.62	4.99	8.47
CasNet	5.12	5.41	5.43	11.42
JointNet	4.26	4.87	5.18	10.25
Ours	3.73	4.13	4.81	9.53
Seg4Reg	4.89	5.87	6.34	18.87
CasNet	4.73	7.79	7.28	22.22
Ours	4.17	4.59	4.20	15.53

Table 1: Quantitative comparisons on AASCE2019 (upper part) and JLU-CJUH dataset (lower part). MT, PT, TL stand for the mean absolute error (MAE) of the main thoracic (MT), proximal thoracic (PT) and thoracolumbar/lumbar (TL) angles, respectively.

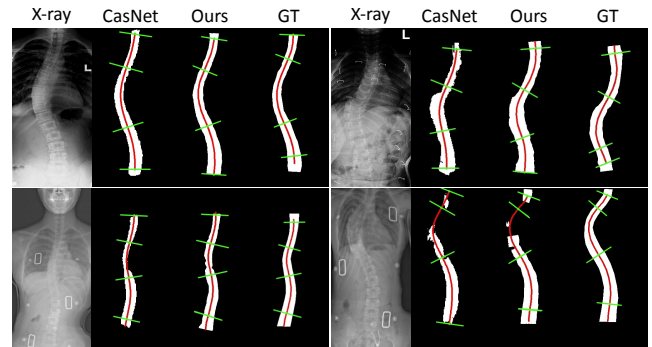


Figure 3: Qualitative comparisons on AASCE2019 (top) and JLU-CJUH (bottom) images. It can be observed our centerlines (red curve) and location of endplates (green segments) for estimating the Cobb angles are closer to the GT.

so that the following CurvePred module receives more refined input. To verify its effectiveness, we compare with the pipeline trained with the SegRefine network disabled, i.e., the initial segmentation mask from the SSformer is used for predicting the B-spline. From the comparison between the 2nd and last row of Table 2 and 3, it shows the SegRefine network can indeed lead to more accurate Cobb angle estimation. Furthermore, we quantitatively compare the segmentation masks obtained with different methods. We evaluate the intersection over union (IoU or Jaccard index), Dice coefficient (Dice), pixel-wise accuracy (pixel-AC) metrics of the output segmentation masks for the AASCE2019 dataset. The results reported from Seg4Reg+ (Lin et al. 2021) and our method without the SegRefine network are compared in Table 4. It can be seen our SegRefine network can indeed improve the accuracy of the segmentation mask and generate the results outperforming other baselines.

Effectiveness of the B-spline curve representation. In our CurvePred module, we use B-spline curve to fit the initial centerline points. To verify the benefits of B-spline over other curve representations, we conduct experiments with other representative curve representations for point fitting or interpolation. Specifically, we perform curve fitting with a 6-degree polynomial curve which is used in (Tu et al. 2019),

Method	Avg	MT	PT	TL	SMAPE
w/o SegRefine	4.64	4.00	4.64	5.28	10.43
6-Poly	5.36	3.88	5.91	6.30	11.02
CS	4.65	4.38	4.26	5.31	9.88
CHS	4.35	4.03	4.14	4.89	9.69
L_{paras} only	5.45	4.43	5.57	6.34	12.87
w/o $L_{resample}$	4.75	4.13	4.87	5.26	10.53
$Cobb^s$ only	4.73	4.37	4.66	5.16	10.89
$Cobb^r$ only	4.95	4.17	5.13	5.54	10.86
Ours	4.22	3.73	4.13	4.81	9.53

Table 2: Quantitative ablation studies on AASCE2019 dataset. Avg is the average value for the MAE of MT, PT and TL angles. 6-Poly stands for 6-degree polynomial curve, CS and CHS are the acronym for cubic spline and cubic Hermite spline curve, respectively.

Method	Avg	MT	PT	TL	SMAPE
w/o SegRefine	4.39	4.37	4.55	4.24	15.86
6-Poly	5.11	4.36	6.29	4.68	16.23
CS	4.44	4.40	4.74	4.17	15.57
CHS	4.49	4.40	4.82	4.25	15.55
L_{paras} only	5.00	4.20	4.81	5.98	16.77
w/o $L_{resample}$	4.54	4.33	4.72	4.58	15.82
$Cobb^s$ only	5.15	4.36	6.35	4.74	16.85
$Cobb^r$ only	5.74	4.73	6.53	5.95	16.96
Ours	4.32	4.17	4.59	4.20	15.53

Table 3: Quantitative ablation studies on JLU-CJUH dataset. The acronyms have the same meaning as in Table 2.

and point interpolation with cubic spline curve and cubic Hermite spline, to the initial centerline points. The curves obtained by above methods are passed to the AngleEst module to get the final results, while the quantitative results are shown in 3rd to 5th rows of Table 2 and 3. It can be observed our B-spline curve achieves the best performance and the cubic Hermite spline curve is the second best, and the curve fitted by the 6-degree polynomial produces the largest error. These results verify the B-spline curve is a more suitable spine representation which can support more robust angle estimation.

Effectiveness of the loss design. To learn the B-spline, we optimize the losses for B-spline parameters and the centerline points (Equation 7) in the CurvePred module. We conduct experiments to evaluate the effectiveness for these terms. First, we train the B-spline prediction model by only using the L_{paras} loss and the results are shown in 6th row in Table 2 and 3. It can be seen that directly predicting the B-spline parameters (i.e., only using L_{paras} loss) leads to inferior angle prediction accuracy. In addition to the supervision on B-spline parameters, the centerline points based losses L_{init} and $L_{resample}$ enforce the point-wise supervision for more stable B-spline learning. To verify their effectiveness, we train a new CurvePred module with L_{init} added to L_{paras} , i.e., without $L_{resample}$, and the results are shown in the 7th row of Table 2 and 3. The performance increase

Method	IoU	Dice	pixel-AC
MDC (Wei et al. 2018)	75.91	86.31	96.73
Seg4Reg+ (Lin et al. 2021)	77.86	87.55	95.49
w/o SegRefine	91.27	95.43	98.23
Ours	93.19	96.40	99.24

Table 4: Quantitative evaluation of various segmentation models on AASCE2019 dataset. Note in w/o SegRefine and ours, the re-trained SSformer is used for generating the initial segmentation mask. Values of MDC and Seg4Reg+ are reported from (Lin et al. 2021).

α	validation set			test set		
	MT	PT	TL	MT	PT	TL
0.2	4.30	5.68	5.38	4.37	4.99	4.88
0.3	4.13	5.39	4.91	4.24	4.85	4.51
0.4	4.07	5.15	4.63	4.17	4.72	4.38
0.5	4.11	4.72	4.34	4.20	4.59	4.20
0.6	4.22	4.92	4.41	4.36	4.78	4.48

Table 5: Analysis of α for combining the slope and regression based Cobb angles on JLU-CJUH dataset.

comparing to using L_{paras} only (6th row) indicates explicit point supervision is helpful for B-spline learning. Moreover, by adding the $L_{resample}$ loss, the performance is further increased in our full model. Such results verify the effectiveness of the resample loss for providing more constraints for learning the B-spline in a more robust manner.

Effectiveness of the hybrid scheme for Cobb angle estimation. In the AngleEst module, we combine the curve-slope based angle $Cobb^s$ and the regression-based angle $Cobb^r$ to obtain the final angle. To evaluate the effectiveness of such hybrid Cobb angle estimation scheme, we compare the results obtained by only using $Cobb^s$ or $Cobb^r$ in the 7th and 8th rows of Table 2 and 3. It can be found the errors of $Cobb^s$ are relatively smaller than $Cobb^r$, indicating the B-spline is well learned and the angles obtained by curvature analysis are fairly consistent with the human annotation. It is also worthy noting that our final angles obtained by combining $Cobb^s$ and $Cobb^r$ produces the smallest error. This is due to $Cobb^s$ and $Cobb^r$ may outperform each other in different cases. By combining them together, we can benefit from the advantages of both angle estimation manners.

Analysis of parameter α . The main parameter in our pipeline is the weight α for combining the Cobb angles obtained by the curve-slope analysis and curve-based regression. In our implementation, α is set based on the performance on the validation set. Also, since each type of Cobb angle (i.e., MT, PT, TL) has different result distribution, α is independently set for each angle type. Table 5 shows the experiment record when we selected α based on the validation set as well as the performance on the test dataset.

5 Conclusion

In this paper, we propose B-Spine, a novel deep learning pipeline for robust and interpretable spinal curvature estimation from low-quality X-ray images. To our best knowl-

edge, we are the first to represent the spine with a smooth and flexible B-spline curve. To improve the performance, we propose novel solutions for each module, including the Seg-Refine network to generate refined mask from the initial segmentation, the multi-step B-spline learning with the resample loss for explicit point supervision and the hybrid Cobb angle estimation by combining the curve analysis and data-driven approaches. Since the three modules of our pipeline can be trained independently, it is convenient to improve each module by using latest networks or new architectures. Meanwhile, as the later module still depends on the result of former one, the multi-stage pipeline may meet the issue for error accumulation, especially for severe low-quality input. How to improve the robustness of the pipeline while still keeping the interpretability of different modules is an interesting future work.

Benefited from our B-Spine, the workload of the clinicians can be alleviated as the spinal curvature can be estimated in an automatic and interpretable manner. Also, thanks to our robust pipeline, B-Spine can be applied to low-quality images obtained from the outdated devices and help to improve the medical service for underdeveloped regions. We believe our B-spline spine representation, the pipeline for annotating and learning the B-spline spine, as well as our JLU-CJUH dataset can inspire more follow-ups for more accurate, more robust and more interpretable Cobb angle estimation.

Acknowledgements

This work is supported in part by National Natural Science Foundation of China (No. 62202199) and Science and Technology Development Plan of Jilin Province (No. 20230101071JC).

References

- Akbar, M.; Terran, J.; Ames, C. P.; Lafage, V.; and Schwab, F. 2013. Use of Surgimap Spine in sagittal plane analysis, osteotomy planning, and correction calculation. *Neurosurgery Clinics*, 24(2): 163–172.
- Alharbi, R. H.; Alshaye, M. B.; Alkanhal, M. M.; Alharbi, N. M.; Alzahrani, M. A.; and Alrehaili, O. A. 2020. Deep learning based algorithm for automatic scoliosis angle measurement. In *2020 3rd International Conference on Computer Applications & Information Security (ICCAIS)*, 1–5. IEEE.
- Barrowclough, O. J.; Muntingh, G.; Nainamalai, V.; and Stangeby, I. 2021. Binary segmentation of medical images using implicit spline representations and deep learning. *Computer Aided Geometric Design*, 85: 101972.
- Brigger, P.; Hoeg, J.; and Unser, M. 2000. B-spline snakes: a flexible tool for parametric contour detection. *IEEE TIP*, 9(9): 1484–1496.
- Chen, B.; Xu, Q.; Wang, L.; Leung, S.; Chung, J.; and Li, S. 2019. An automated and accurate spine curve analysis system. *Ieee Access*, 7: 124596–124605.
- Dubost, F.; Collery, B.; Renaudier, A.; Roc, A.; Posocco, N.; Niessen, W.; and de Bruijne, M. 2020. Automated estimation of the spinal curvature via spine centerline extraction with ensembles of cascaded neural networks. In *International Workshop and Challenge on Computational Methods and Clinical Applications for Spine Imaging*, 88–94. Springer.
- Gordon, W. J.; and Riesenfeld, R. F. 1974. B-spline curves and surfaces. In *Computer aided geometric design*, 95–126. Elsevier.
- Gstoettner, M.; Sekyra, K.; Walochnik, N.; Winter, P.; Wachter, R.; and Bach, C. M. 2007. Inter- and intraobserver reliability assessment of the Cobb angle: manual versus digital measurement tools. *European Spine Journal*, 16(10): 1587–1592.
- Gu, J.; Yang, T. S.; Ye, J. C.; and Yang, D. H. 2021. CycleGAN denoising of extreme low-dose cardiac CT using wavelet-assisted noise disentanglement. *Medical Image Analysis*, 74: 102209.
- He, K.; Zhang, X.; Ren, S.; and Sun, J. 2016. Deep residual learning for image recognition. In *CVPR*, 770–778.
- Huang, H.; Lin, L.; Tong, R.; Hu, H.; Zhang, Q.; Iwamoto, Y.; Han, X.; Chen, Y.-W.; and Wu, J. 2020. Unet 3+: A full-scale connected unet for medical image segmentation. In *ICASSP*, 1055–1059. IEEE.
- Huo, L.; Cai, B.; Liang, P.; Sun, Z.; Xiong, C.; Niu, C.; Song, B.; and Cheng, E. 2021. Joint Spinal Centerline Extraction and Curvature Estimation with Row-Wise Classification and Curve Graph Network. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 377–386. Springer.
- Kim, K. C.; Yun, H. S.; Kim, S.; and Seo, J. K. 2020. Automation of spine curve assessment in frontal radiographs using deep learning of vertebral-tilt vector. *IEEE Access*, 8: 84618–84630.
- Klein, S.; Staring, M.; and Pluim, J. P. 2007. Evaluation of optimization methods for nonrigid medical image registration using mutual information and B-splines. *IEEE TIP*, 16(12): 2879–2890.
- Laube, P.; Franz, M. O.; and Umlauf, G. 2018. Deep learning parametrization for B-spline curve approximation. In *International Conference on 3D Vision (3DV)*, 691–699. IEEE.
- Lehmann, T. M.; Gonner, C.; and Spitzer, K. 1999. Survey: Interpolation methods in medical image processing. *IEEE transactions on medical imaging*, 18(11): 1049–1075.
- Li, S.; and Wang, L. 2019. AASCE: Accurate Automated Spinal Curvature Estimation, MICCAI Challenge.
- Lin, Y.; Liu, L.; Ma, K.; and Zheng, Y. 2021. Seg4Reg+: Consistency Learning Between Spine Segmentation and Cobb Angle Regression. In *MICCAI*, 490–499. Springer.
- Lin, Y.; Zhou, H.-Y.; Ma, K.; Yang, X.; and Zheng, Y. 2020. Seg4Reg networks for automated spinal curvature estimation. In *International Workshop and Challenge on Computational Methods and Clinical Applications for Spine Imaging*, 69–74. Springer.
- Patrikalakis, N. M.; and Maekawa, T. 2002. *Shape interrogation for computer aided design and manufacturing*, volume 15. Springer.

- Prabhu, G.; et al. 2012. Automatic quantification of spinal curvature in scoliotic radiograph using image processing. *Journal of medical systems*, 36(3): 1943–1951.
- Ronneberger, O.; Fischer, P.; and Brox, T. 2015. U-net: Convolutional networks for biomedical image segmentation. In *MICCAI*, 234–241. Springer.
- Russell, B. C.; Torralba, A.; Murphy, K. P.; and Freeman, W. T. 2008. LabelMe: a database and web-based tool for image annotation. *IJCV*, 77(1): 157–173.
- Samuvel, B.; Thomas, V.; M.G., M.; and J., R. K. 2012. A Mask Based Segmentation Algorithm for Automatic Measurement of Cobb Angle from Scoliosis X-Ray Image. In *2012 International Conference on Advances in Computing and Communications*, 110–113.
- Shaw, M.; Adam, C. J.; Izatt, M. T.; Licina, P.; and Askin, G. N. 2012. Use of the iPhone for Cobb angle measurement in scoliosis. *European Spine Journal*, 21(6): 1062–1068.
- Shi, W.; Xu, J.; and Gao, P. 2022. Ssformer: A lightweight transformer for semantic segmentation. In *2022 IEEE 24th International Workshop on Multimedia Signal Processing (MMSP)*, 1–5. IEEE.
- Tu, Y.; Wang, N.; Tong, F.; and Chen, H. 2019. Automatic measurement algorithm of scoliosis Cobb angle based on deep learning. In *Journal of Physics: Conference Series*, volume 1187, 042100. IOP Publishing.
- Unser, M. A. 2002. Splines: a perfect fit for medical imaging. In *Medical Imaging 2002: Image Processing*, volume 4684, 225–236. SPIE.
- Wang, L.; Xu, Q.; Leung, S.; Chung, J.; Chen, B.; and Li, S. 2019. Accurate automated Cobb angles estimation using multi-view extrapolation net. *Medical Image Analysis*, 58: 101542.
- Wang, S.; Huang, S.; and Wang, L. 2020. Spinal Curve Guide Network (SCG-Net) for Accurate Automated Spinal Curvature Estimation. In *Computational Methods and Clinical Applications for Spine Imaging*, 107–112.
- Wang, Z.; Bovik, A. C.; Sheikh, H. R.; and Simoncelli, E. P. 2004. Image quality assessment: from error measurement to structural similarity. *IEEE TIP*, 13: 600–613.
- Wei, Y.; Xiao, H.; Shi, H.; Jie, Z.; Feng, J.; and Huang, T. S. 2018. Revisiting Dilated Convolution: A Simple Approach for Weakly- and Semi-Supervised Semantic Segmentation. In *CVPR*.
- Williams, H.; Pedrosa, J.; Cattani, L.; Housmans, S.; Vercauteren, T.; Deprest, J.; and D’hooge, J. 2021. Interactive segmentation via deep learning and b-spline explicit active surfaces. In *MICCAI*, 315–325. Springer.
- Wu, H.; Bailey, C.; Rasoulinejad, P.; and Li, S. 2017. Automatic landmark estimation for adolescent idiopathic scoliosis assessment using BoostNet. In *MICCAI*, 127–135. Springer.
- Wu, H.; Bailey, C.; Rasoulinejad, P.; and Li, S. 2018. Automated comprehensive adolescent idiopathic scoliosis assessment using MVC-Net. *Medical Image Analysis*, 48: 1–11.
- Wu, W.; Liang, J.; Du, Y.; Tan, X.; Xiang, X.; Wang, W.; Ru, N.; and Le, J. 2014. Reliability and reproducibility analysis of the Cobb angle and assessing sagittal plane by computer-assisted and manual measurement tools. *BMC Musculoskeletal Disorders*, 15(1): 1–8.
- Zhang, C.; Wang, J.; He, J.; Gao, P.; and Xie, G. 2021. Automated vertebral landmarks and spinal curvature estimation using non-directional part affinity fields. *Neurocomputing*, 438: 280–289.
- Zhang, J.; Lou, E.; Hill, D. L.; Raso, J. V.; Wang, Y.; Le, L. H.; and Shi, X. 2010. Computer-aided assessment of scoliosis on posteroanterior radiographs. *Medical & biological engineering & computing*, 48(2): 185–195.
- Zhu, J.-Y.; Park, T.; Isola, P.; and Efros, A. A. 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *ICCV*, 2223–2232.