Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Fourier-inspired Automatic Cobb Angles Measurement for Adolescent Idiopathic Scoliosis

Yishan Zhong
Georgia Institute of Technology
Felipe Giuste
Emory University
Wenqi Shi
The University of Texas Southwestern Medical Center
Ben Tamo
Georgia Institute of Technology
Micky Nnamdi
Georgia Institute of Technology
Benoit L Marteau
Georgia Institute of Technology
Kaiwen Wang
Georgia Institute of Technology
Steven W Hwang
Shriners Hospitals for Children - Philadelphia
Selina Poon
Shriners Hospitals for Children - Pasadena
Henry J Iwinski
Shriners Hospitals for Children Medical Center - Lexington
May D Wang
maywang@gatech.edu

Georgia Institute of Technology

Article

Keywords:

Posted Date: April 23rd, 2025

DOI: https://doi.org/10.21203/rs.3.rs-6010694/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Fourier-inspired Automatic Cobb Angles Measurement for Adolescent Idiopathic Scoliosis

Yishan Zhong[1], Felipe Giuste[2], Wenqi Shi[3], J. Ben Tamo[1], Micky C. Nnamdi[1], Benoit L. Marteau[1], Kaiwen Wang[4], Steven W. Hwang[5], Selina Poon[6], Henry J. Iwinski[7], and May D. Wang[1,2]

[1]School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA
[2]Department of Biomedical Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA
[3]O'Donnell School of Public Health at UT Southwestern Medical Center. Dallas, TX. 75390
[4]Department of Computer Science, Georgia Institute of Technology, Atlanta, GA 30332 USA
[5]Shriners Children's Philadelphia, Philadelphia, PA 19140 USA
[6]Shriners Children's Pasadena, Pasadena, CA 91105 USA
[7]Shriners Children's Lexington, Lexington, KY 40508 USA
[*]maywang@gatech.edu

ABSTRACT

Measuring Cobb angles in adolescent idiopathic scoliosis (AIS) patients using anteroposterior (AP) radiographs is time-consuming and prone to variability due to subjective vertebrae selection. To address this, we propose an automated model for Cobb angle measurement using a Fourier-inspired method. The model starts by estimating the spinal curvature spectrum and then predicts the Cobb angles. We use a VGG-16 transfer learning model to regress the concatenated real and imaginary components of the spine's spectral representation. This feature is used to reconstruct the curvature through an inverse Fourier transform for improved model explainability. Simultaneously, it serves as input to a projection head for Cobb angle estimation. Evaluating on the MICCAI 2019 SpineWeb dataset, our model shows a Pearson correlation of 0.9382 for curvature estimation and a symmetric mean absolute percentage error (SMAPE) of 12.25% for Cobb angle measurement, performing similarly to other top models. By treating spinal curvature prediction as a frequency domain analysis, our method offers more accurate curvature shapes than spatial domain approaches. Additionally, applying zero padding improves spectral resolution and reduces SMAPE in Cobb angle estimation. Spectral prediction consistently outperforms spatial prediction in both curvature and Cobb angle measurement.

Introduction

Scoliosis is a condition that describes an abnormal lateral curvature of the spine. Scoliosis has been shown to hamper patients in conducting daily activities[1,2]. or adolescents with scoliosis, the condition also carries significant psychological implications, exacerbating the physical challenges. Patients may experience distress or social rejection[3–5]. As a result, efficient diagnosis and assessment are crucial. One commonly used qualitative assessment is Adam's forward bend test; however, its reliability has been questioned[6].

Accurate characterization of spinal deformities is essential for assessing patients with adolescent idiopathic scoliosis. Among the various measurements, Cobb angles are most frequently used. Typically measured from anterior-posterior (AP) radiographs, Cobb angles are composed of three individual angle measurements: proximal-thoracic (PT), main-thoracic (MT), and thoracic-lumbar (TL). This measurement requires identifying four key vertebrae as landmarks, drawing perpendicular lines to the spinal curvature, and calculating the angles between adjacent lines to form the Cobb angle. Larger Cobb angles indicate more severe spinal tilt.

Assessing Cobb angles can be labor-intensive and time-consuming. Studies have reported that clinicians typically spend between two and five minutes per radiograph[7]. Moreover, accurate measurement requires extensive training and experience. Even with expertise, the results are often subject to substantial interobserver and intraobserver variability[8,9]. In other words, even among well-trained clinicians, measurements of the same radiograph can vary significantly. Furthermore, measurements made by the same clinician on the same radiograph may lack reproducibility.

An automated measurement pipeline could effectively address both sources of variability. The pipeline must provide reliable and accurate estimations of the three Cobb angles based on an anterior-posterior radiograph. In this paper, we present a novel approach utilizing Fourier analysis. We hypothesize that spinal curvatures in scoliosis patients exhibit distinct patterns in the

frequency domain. Therefore, rather than performing regression in the spatial domain, our model regresses to the concatenated real and imaginary components of the spinal spectrum.

Our pipeline consists of three key steps. In the first step, we generate a segmentation of the spinal column from the radiograph using a U-Net model[10]. The input channels for this step include the segmentation, the grayscale radiograph after histogram equalization, and the radiograph's gradients, which highlight the bone edges. In the second step, we leverage transfer learning with the VGG-16 model to perform regression on the concatenated spectrum of the spinal curvature. The final step involves a projection head that regresses the three Cobb angles. We evaluate our approach on the MICCAI 2019 competition dataset, achieving a Symmetric Mean Absolute Percentage Error (SMAPE) of 12.25%, which is competitive with the top results reported.

To support the clinical adoption of our pipeline, we have developed a web application that allows clinicians to upload a radiograph and receive the estimated Cobb angles within two seconds. To further enhance interpretability and build clinician trust, the application enables users to visualize the three input channels used by the model, as well as the radiograph overlaid with the predicted spinal curvature. This curvature is generated by performing an inverse Fourier transform of the predicted spectrum back into the spatial domain.

We summarize our contribution as:

- Using spectral analysis to address the problem of Cobb angle measurement shows promising results. Regression of spine curvature in the frequency domain yields a higher correlation, lower Hausdorff distance, and reduced Kullback-Leibler divergence compared to regression in the spatial domain. This reduction in curve deviation leads to more accurate Cobb angle measurements.
- Downsampling in the spatial domain combined with zero padding of the spectrum yields varying performance in curvature estimation. By downsampling by a factor of 4 and applying zero-padding with the same factor, our approach achieved a Symmetric Mean Absolute Percentage Error (SMAPE) of 12.25%.
- Integrating the grayscale radiograph, its segmentation, and its gradient into a model originally developed for RGB images enhanced performance.
- Our web application, which integrates our AI-assisted measurement pipeline, has the potential to enhance clinical efficiency and facilitate rapid scoliosis screening in teenagers. The visualizations not only improve interpretability but also support evidence-based clinical decision-making.

Results

Curvature model performance

The primary metric is Pearson correlation, because the difference is most significant and less fluctuated during training than Hausdorff and KL divergence. The first few rows in **Table. 1** shows the correlations, Hausdorff distances, and KL divergences with a downsampling factor of 4. This gives an initial spatial size of 256. Before transforming into the frequency domain, the series is modified in three ways: no padding applied, padding to twice the size, and padding to four times the size.

Downsample factor	Domain of analysis	Zero padding	Pearson corr.	Spearman corr.	Hausdorff dist.	KL div.
4	**Frequency**	256 unpad	0.9185	0.8967	99.10	0.02874
		256 pad to 512	0.9166	0.8950	111.37	0.02513
		256 pad to 1024	0.9382	0.9277	105.50	0.02283
	Space	256	0.9164	0.9025	106.54	0.02306
2	**Frequency**	512 unpad	0.8866	0.9101	112.83	0.02652
		512 pad to 1024	0.9342	0.8959	131.10	0.01667
	Space	512	0.9149	0.8921	115.74	0.01691
No downsampling	**Frequency**	1024 unpad	0.9171	0.9017	119.92	0.01143
	Space	1024	0.9093	0.8774	119.62	0.01270

Table 1. The spatial and spectral domain curvature prediction with a few setups. (1) the downsampling factor is 4. (i.e. Across the 1024 height we take 256 equidistant points). (2) the downsampling factor is 2. (i.e. Across the 1024 height we take 512 equidistant points). (3) no downsampling is applied.

Similarly, the process is repeated with a downsampling factor of 2 and without downsampling entirely. The results are listed in the remaining rows in **Table. 1**. A visualization of the predicted spectrum is shown in **Figure 1**.

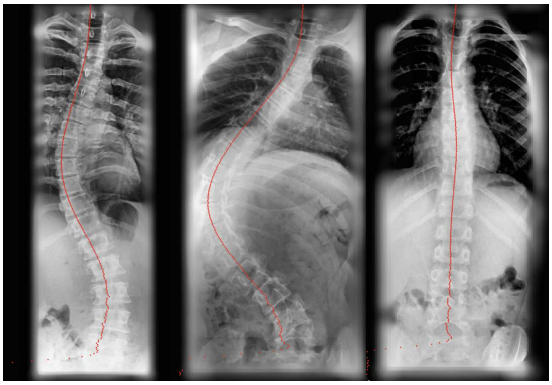


Figure 1. The red curves are the predicted spinal curvature in the frequency domain, with downsampling by a factor of 4 and subsequent zero padding to 1024. Qualitatively, they show strong agreement along the center of the spine.

Through head-to-head comparisons between various setups, the curve fitting with the spectral method is more accurate than that with the spatial method. In addition, with the same spectral sequence length, the more downsampling and padding applied, the higher correlation are the curves. For example, the three setups with 1024 spectral lengths are: (1) downsampling with a factor of 4 and zero padding four times the length, (2) downsampling with a factor of 2 and zero padding twice the length, and (3) no downsampling and no zero padding. Among the three, the first setup obtained the best curvature prediction. Likewise, the same trend can be observed with 512 spectral length.

Cobb angle model performance

The ground truth values of the Cobb angles are provided for all radiographs in the testing dataset. The SMAPEs between the ground truth and the AI-driven measurements are shown in **Figure. 2**. Similar to the results on curvature, it suggests consistently better performances in the frequency domain than in the spatial domain. In addition, a setup resulting in more accurate curvature tends to obtain lower SMAPE as expected.

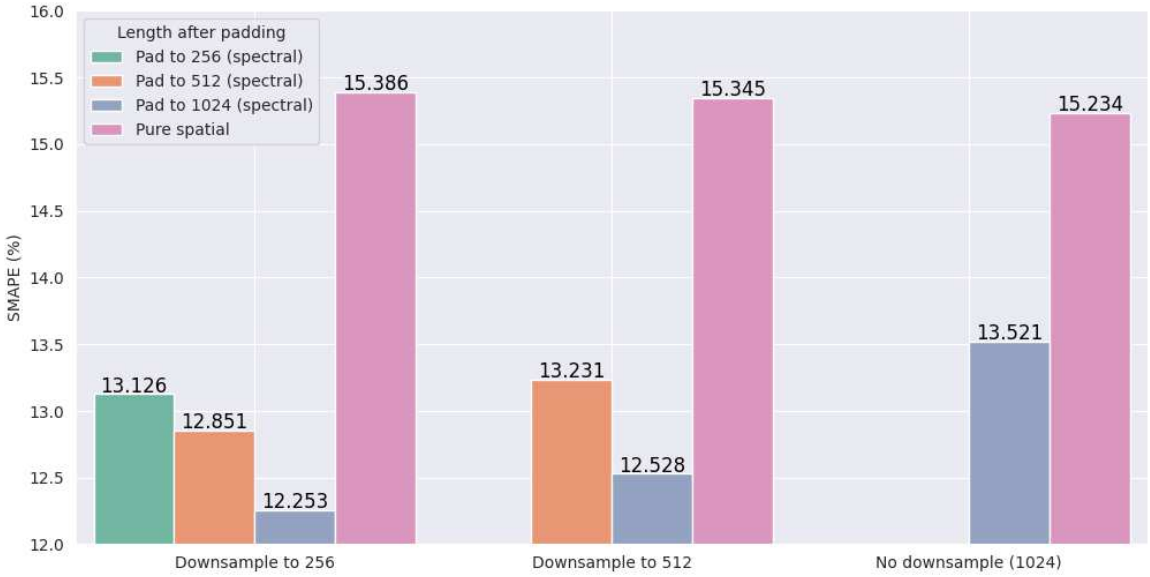


Figure 2. The SMAPE of Cobb angles of different setups. The performance is evaluated on the testing dataset containing 129 images. A similar yet more pronounced trend is observed as the case of curvature prediction. The spectral analysis is superior to spatial analysis. Also, given the same initial length, zero padding lowers the SMAPE.

Ablation study of the three channels

While fixing the downsampling and padding to 256 and 1024 respectively, ablation study is conducted to determine the necessity to include three channels. Noted in **Table. 2**, we observe considerable drops in SMAPE if any of the channels is repeated three

times rather than combining all three. This strongly indicates that the three channels achieved a synergistic effect.

Input channels	SMAPE
Original + Spinal segmentation + Gradient	**12.253%**
Radiograph only	14.649%
Spinal segmentation only	17.431%
Gradient only	30.644%

Table 2. The synergistic effect of having three channels results in higher SMAPE than any of the grayscale radiograph, segmentation, and gradient map alone.

Authors	Methods	SMAPE
Khanal et al.	Bounding box + Landmark	25.69%
Chen et al.	Landmark	23.59%
Wu et al.	Landmark	23.44%
Reddy et al.	Segmentation + Landmark	15.85%
Chen et al.	Segmentation + Landmark	12.76%
Ours	**Spectral**	**12.25%**
Shao et al.	Segmentation	10.31%

Table 3. Our Fourier-inspired approach achieves a competitive result on the MICCAI 2019 dataset. The 10.31% SMAPE is achieved on finetuning a large segment anything model (SAM).

Discussion

Previous efforts on Cobb angle measurements fall into three categories: edge-based, segmentation-based, and landmark-based approaches.

For edge-based or segmentation-based approaches, typically a polynomial curve fit on the centers of the vertebrae can obtain the spinal curve can be fit to estimate the angles. Kusuma[11] cropped each radiograph into numerous horizontal stripes. With Canny edge detection, the regions around the vertebrae were highlighted. The centers of the vertebrae correspond to the centroids of the edges. They achieved an absolute angular deviation of 5.86°. Canny edge detection can fail on radiograph due to noise. Mukherjee et al.[12] experimented with four denoising methods before edge detection to obtain the dominant edge of the spine. They used the Hough transform to extract the slopes for angle computation. They achieved an absolute deviation of 1.27°. The rise of segmentation architectures offers plenty of breakthrough as well. Zhang et al.[13] incorporated reinforcement learning to generate accurate vertebra-level masks. They first segmented the edge of the spine with a UNET architecture. The predicted segmentation then went through feed forward layers to directly predict the three Cobb angles. Shao et al.[14] finetuned the segment anything model (SAM)[15] and achieved an SMAPE of 10.31%.

Surprisingly, landmark-based methods could also involve segmentation as a preliminary step. The key sign of landmark-based method is the extraction of corners or centers of the bounding box or segmented area. For example, Yi et al.[16] used the conventional heatmap prediction with the UNET architecture, combined with a variant of focal loss to calculate the center of vertebral bodies. Chen et al.[17] adapted the feature pyramid network (FPN) common in Masked-RCNN for the task and achieved SMAPE of 12.76%. Similarly, Khanal et al.[18] used Mask-RCNN to identify and crop individual vertebrae and a separate CNN predicts the coordinates of the vertebral corners of the identified patch. The coordinates relative to each patch were remapped to the global radiograph-level coordinates. Outlying coordinates were filtered, and two curves were fit to the left and right predicted landmarks, which enabled geometry-based angle calculation.

A common challenge after landmark extraction is to obtain pairs that correspond to the same vertebra. Kim et al.[19] presented vertebral tilt field to regularize the predicted heatmaps and matched the left and right corners of the same vertebra into the same pair. The matching process minimized possible cross-over among all possiblities connecting the predicted landmarks. Combining global and local information, Reddy et al.[20] took a global-local network approach, where a U-NET predicting the heatmap of vertebral centers was masked by the global path to predict the location. The local U-NET and global paths remained separate until the last step, when the information was merged. Chen et al.[21] proposed model to maximize joint information sharing between direct Cobb angle prediction and indirect angle prediction through curvature prediction. The two objective regularized the training and guaranteed the angles predicted by the two approaches agree. Due to the radiograph variations, boostrapping has been shown effective. Wu et al.[22] proposed a BoostNet architecture that used a CNN approach to directly

predict the coordinates of the four corners of the vertebrae. Data samples that reproduce poorly on a shallow autoencoder were considered outliers and were discarded. Two autoencoder-based filtering units were sandwiched between the CNN and the feed forward layers to generate the coordinates. The authors reported that the addition of filtering units reduces the mean standard deviation of the predicted coordinates by more than 10 fold. They achieved an SMAPE of 23.44%.

Achieving a SMAPE of 12.25% positions our work as highly competitive on the same dataset. The SMAPE values reported in other studies are summarized in **Table. 3**. Our pipeline introduces spectrum prediction as an intermediate task between the input radiograph and the output Cobb angles. This intermediate supervision offers two key advantages. First, it helps regularize the problem and bridge the significant disparity between the input and output dimensionalities. Second, the predicted spectrum enables visualization of the corresponding spine curvature through an inverse Fourier transform, providing users with a means to interpret the deep learning-based pipeline.

Notably, the work by Shao et al.[14] slightly outperforms our model. However, their transformer-based approach, an adaptation of SAM, may be less feasible for clinical translation in resource-constrained environments due to its model complexity. In contrast, while our simpler Fourier-inspired model results in a higher percentage error, it remains compact and more suitable for practical deployment.

Our experiments reveal two key findings. First, spectral prediction outperforms spatial prediction in accurately predicting Cobb angles. We reason that the frequency spectrum more effectively captures generalizable patterns of spinal curvature among scoliosis patients. In contrast, demographic factors, such as patient height, contribute to positional variations across radiographs, making regression in the noisy spatial domain more challenging. Second, applying zero padding before the Fourier transform enhances both spectral and angular predictions. Zero padding in the spatial domain not only standardizes the sequence length across the pipeline, but also refines the frequency spectrum[23,24]. As illsutrated in (**Supplementary Figure S1**), the same effect is found with padding the spinal curve. Higher-resolution spectra align more closely with the ground truth curvature, providing the projection head with richer information for more accurate Cobb angle measurements. However, excessive padding may have detrimental effects. It elongates the spectral sequence, increases the spectral feature size, and introduces additional trainable parameters. With our chosen downsampling and padding parameters, the model successfully mitigates the risk of underfitting while benefiting from a more refined spectrum. Upon further analysis, the extent of zero padding appears to have a minimal impact on curvature prediction. For example, starting with a sequence length of 256, Pearson correlations remain largely unchanged across the three padding configurations. However, the differences in SMAPE are more pronounced, likely due to the increased input dimension of the projection head, which yields a higher-definition spectrum.

The pipeline has several areas for potential improvement. For instance, the gradient channel may still include image boundaries before the widening step. To better suppress these edges, a smoother blur could be applied. However, the abrupt transition from near-white pixels (particularly around the pelvis) to the added black boundary would require a large blurring radius to soften the transition, which could inadvertently obscure important details in the radiograph. This issue stems from the unstandardized dimensions of the dataset. In the future, a larger dataset consisting of radiographs captured with standardized equipment could mitigate this problem. Moreover, the prototype application may face challenges in real-world settings due to the fact that all training data currently come from scoliosis patients. As a result, if the model is used for screening healthy patients, it may bias towards detecting a tilted spine. This limitation is inherent to the spinal curvature approach when compared to landmark-based methods. However, with a sufficiently large and diverse dataset, the Fourier method could potentially capture the spectral features of the spine effectively for both healthy individuals and scoliosis patients.

To support the clinical adoption of our pipeline, we have developed a prototype application (**Supplementary Figure S2**) that enables rapid Cobb angle measurements. Designed for ease of use, the application requires minimal technical skills or medical expertise, making it ideal for resource-limited settings. Users can upload a radiograph in PNG or JPEG format, and the application automatically calculates and displays the spinal curvature along with the three Cobb angles. Additionally, the segmentation of the spine and the measured curvature are shown, offering visual feedback for verification and quality assurance. We foresee this deployment as a valuable tool for quick, quantitative scoliosis screening, offering a more objective alternative to the qualitative Adam's bend test.

Methods

Dataset

This study uses the MICCAI 2019 competition dataset (AASCE SpineWeb). The dataset contains an existing split of 483 training and 129 testing anteroposterior (AP) radiographs. The ground truth labels, available for all 612 images, are the three Cobb angles measured in degrees: proximal thoracic (PT), main thoracic (MT), and thoracic lumbar (TL). The vertebral landmarks of all images are also provided, which consist of the four corners of the middle 7 vertebrae. The labeled vertebrae vary from image to image. **Supplementary Figure S3** shows some example images from the dataset, and **Supplementary Figure S4** shows the histograms of the three Cobb angle measurements of the training data.

The histograms of ground truth Cobb angles indicate the presence of unusual outliers among the 483 training images. Therefore, the two images corresponding to over 140 degrees PT and over 70 degrees in TL are removed from the training set. Within the remaining 481 training images, images were split into a training set and a validation set with a 7:3 ratio.

Pipeline overview

The complete pipeline is illustrated in **Figure 3**. The intermediate supervision of the spectral information of curvature is represented by the concatenation of the real and imaginary parts. The first half of the pipeline obtains the spectral prediction of the spinal curvature. The VGG-16 is adopted as the basis of transfer learning, and it ingests the grayscale radiograph, the edges, and the segmentation as the three channels.

There are two branches in the second half. First, a feed-forward net (FFN) projection head is used to do regression on the three Cobb angles. Second, the predicted curvature in the spatial domain is plotted, using inverse Fourier transform of the spectra. This encourages explainability and qualitative validation.

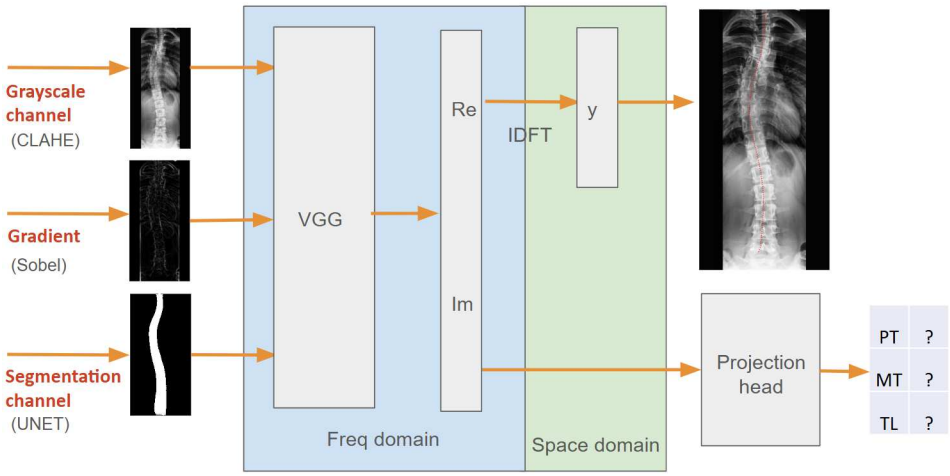


Figure 3. The original VGG-16 model and our transfer learning model. To efficiently utilize the three channels, the model takes in the grayscale radiograph, the image after edge detection, and the U-NET segmented image. The output of the transfer learning model is the spectrum of the spinal curvature. Then the model splits into two branches. One for the angles measurement and the other for qualitative visualization via inverse Fourier transform for clinical explainability.

Preprocessing

Standardization of the radiographs is achieved as follows. First, all image-mask pairs are padded with black margins to convert into an aspect ratio of 1:2. Offline augmentation is used to enlarge the dataset by 128 times. Random rotation by at most 10°, random translation, random brightness, and random contrast are applied. Each augmented image is cropped at the center and rescaled to a dimension of 1024x512.

To obtain the spectra of the spinal curvatures for intermediate supervision, binary spinal masks are generated with the help of the landmark annotations of the four corners of vertebrae. Splines are fit through the annotations along the left and right edges. The top and bottom of the spine are manually identified.

The spinal curvature consists of the midpoints between the two splines. The resulting midpoint series can be viewed in such a way that the x-axis matches the height of the image, and the y-axis is across the width. With an image height of 1024, a series of 1024 y values is sufficient to describe the spinal curvature. Then, the series is downsampled by a factor of 4, and 768 zeros are added to the end so the size of the resulting series remains 1024. In other words, Fourier transform is applied to this downsampled and zero-padded series, as shown in **Figure 4**. The result serves as the annotation for the spectral training.

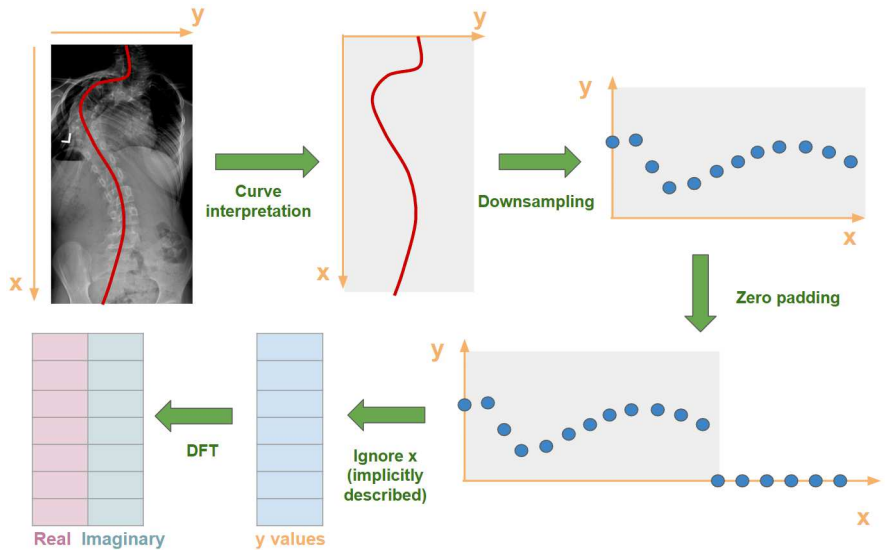


Figure 4. The process of preparing the curvature spectrum for supervision is shown. The curvature is treated as a series of y values, where y is along the width of the image. The series is first downsampled and then padded with zeros before applying the Fourier transform. Note that the x values are implicit because they are equidistant along the heights of the image. The illustration does not reflect the exact values of downsampling and padding.

Three input channels

The three adapted input channels of the VGG-16 transfer learning model are the grayscale image, the gradient image, and the segmentation image. Each of the three channels is by itself grayscale, and together they complement each other to better utilize the three channels of the VGG-16 architecture.

Grayscale

The grayscale channel is the radiograph after contrast-limited adapted histogram equalization (CLAHE). This technique is widely used in solving other problems involving X-ray to emphasize the bone structures.[25,26]

Segmentation

The segmentation channel is the predicted binary segmentation of the spinal region **Supplementary Figure S5**. The segmentation is obtained using a very small U-NET model[10], which comprises of five convolution blocks with 4, 16, 64, 16, 4 channels, respectively. The 64-channel block was the bottleneck of the U-NET architecture. During training, binary cross entropy loss is used to minimize the differences between the masks. Also, a DICE score (Equation 1) is adopted as the evaluation metric.

$$DICE = \frac{2(X_{pred} \cap X_{gt})}{X_{pred} + X_{gt}} \tag{1}$$

The model is trained for 50 epochs with a batch size of 16, and an Adam optimizer with a learning rate of 1×10^{-4}. To accommodate the simplicity of our U-NET model, the input radiograph is resized to a quarter of its original width and height.

Gradient

The gradient channel is generated in three steps **Supplementary Figure S6**. Firstly, the Sobel operator and LaPlacian filters are separately used to find the edges in the radiograph. The two resulting images are pixel-wise added. As a side effect, this operation may also highlight the edges of the image due to the previous resizing step. Secondly, the pixels are normalized to fit the range from 0 to 255. Finally, a gamma curve (exponent of 2) is applied to transform the pixel values to further highlight the bone edges.

Transfer learning model and curvature estimation

Model

The model is adapted from the VGG-16 backbone[27]. It is initialized with pretrained parameters, all of which can be updated during the training. After the average pooling step, a new projection head replaced the original layers. The projection head has three layers: 10000, 4000, and the length of the concatenated spectrum. Alternatively, in the case of prediction in the space domain, the output dimension is the number of sampled spinal coordinates. Mean square error loss and Adam optimizer with learning rate of 3×10^{-5} are used.

Evaluation

All curvature evaluations are conducted in the spatial domain regardless the specific setups. In other words, using the spectral method, an inverse Fourier transform is first applied to obtain the spatial y values. The following metrics are used to evaluate the deviation between the predicted series of points from the actual spinal curvature.

- Pearson correlation (Equation 2). It measures the agreement between two series. If all spatial y values in the ground truth are placed along the x-axis, whereas the predicted spatial y values are along the y-axis, all points in a perfect prediction would aligned along a line with no noise.

$$r = \frac{\sum(X_{pred_i} - \bar{X_{pred}})(X_{gt_i} - \bar{X_{gt}})}{\sqrt{\sum(X_{pred_i} - \bar{X_{pred}})^2 \sum(X_{gt_i} - \bar{X_{gt}})^2}} \tag{2}$$

- Spearman correlation (Equation 3). Like Pearson, it is also a measure of correlation. However, Spearman emphasizes the ordering of the series. d is the differences in the ranking, and n is the number of points in the series.

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)} \tag{3}$$

- Kullback-Leibler divergence (Equation 4). KL divergence measures the difference between two probability distributions.

$$KL(P \parallel Q) = \sum_{i \neq j} p_{ij} \log \frac{p_{ij}}{q_{ij}} \tag{4}$$

  The problem of comparing two curves is translated into comparing two distributions. For each x coordinate, there are two y values associated with the predicted y and ground truth y. This is analogous to two probabilities of the same class. In the original context, p refers to the true distribution, and q is the approximate distribution. Now, they become the ground truth curvature and the predicted curvature. Before KL divergence analysis, the y values of the sequence are normalized. From the implementation perspective, all y coordinates less than or equal to 0 are converted into 1×10^{-8} before normalization.

- Hausdorff distance. Unlike the previous three metrics, the Hausdorff distance takes into account both the x and the y coordinates. For every predicted coordinate, the nearest ground-truth coordinate based on Euclidean distance is determined. The Hausdorff distance is the greatest deviation among all coordinates in the series.

Cobb angles estimation

The angle prediction model comprises a simple projection head on top of the predicted spectrum (or spatial y values). The output is a three-dimensional vector indicating the three Cobb angles (PT, MT, and TL). We used symmetric mean absolute percentage error (SMAPE) for evaluation (Equation 5). SMAPE is the standard evaluation in the MICCAI competition.[28]

$$SMAPE = \frac{1}{N}\sum_{i=1}^{N} \frac{\sum_j^{PT,MT,TL} \left| X_{pred_j} - X_{gt_j} \right|}{\sum_j^{PT,MT,TL} (X_{pred_j} - X_{gt_j})} \tag{5}$$

In the equation, N is the number of images, and j covers the three angles. *pred* refers to prediction and *gt* denotes ground truth. The resulting fraction is converted to a percentage value.

Author contributions statement

Y. Z. designed the methodology and performed the algorithm implementation. F. G., W. S., and B. M. proposed significant improvements to the methodology. K. W. developed the clinical application with the trained model. B. T., M. N., S. H., S. P., H. I., and M. W. provided critical revisions to the manuscript. All authors read and approved the final manuscript.

Additional information

The author(s) declare no competing interests.

Data availability statement

The code for the model is available at https://github.com/zhongys97/Fourier-inspired-scoliosis-assessment

Acknowledgements

This research has been supported by a Wallace H. Coulter Distinguished Faculty Fellowship, a Petit Institute Faculty Fellowship, and research funding from Amazon and Microsoft Research to Professor May D. Wang. We extend special thanks to Dr. Michael Wattenbarger for providing us with valuable clinical insights into scoliosis in the pediatric population.

References

1. Downs, J. *et al.* Impact of scoliosis surgery on activities of daily living in females with rett syndrome. *J. Pediatr. Orthop.* **29**, 369–374 (2009).
2. Ishikawa, Y. *et al.* Activities of daily living and patient satisfaction after long fusion for adult spinal deformity: a retrospective study. *Eur. Spine J.* **28**, 1670–1677 (2019).
3. MacLean Jr, W. E., Green, N. E., Pierre, C. B. & Ray, D. C. Stress and coping with scoliosis: psychological effects on adolescents and their families. *J. Pediatr. Orthop.* **9**, 257–261 (1989).
4. Eliason, M. J. & Richman, L. C. Psychological effects of idiopathic adolescent scoliosis. *J. Dev. & Behav. Pediatr.* **5**, 169–172 (1984).
5. Sanders, A. E. *et al.* Clinically significant psychological and emotional distress in 32% of adolescent idiopathic scoliosis patients. *Spine deformity* **6**, 435–440 (2018).
6. Côté, P., Kreitz, B. G., Cassidy, J. D., Dzus, A. K. & Martel, J. A study of the diagnostic accuracy and reliability of the scoliometer and adam's forward bend test. *Spine* **23**, 796–802 (1998).
7. Kamal, S. A., Sarwar, M., Razzaq, U. A. *et al.* Effective decision making for presence of scoliosis. *Int. J. Biol. Biotechnol.* **12**, 317–328 (2015).
8. Pruijs, J., Hageman, M., Keessen, W., Van Der Meer, R. & Van Wieringen, J. Variation in cobb angle measurements in scoliosis. *Skeletal radiology* **23**, 517–520 (1994).
9. Kuklo, T. R. *et al.* Reliability analysis for manual adolescent idiopathic scoliosis measurements. *Spine* **30**, 444–454 (2005).
10. Ronneberger, O., Fischer, P. & Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, 234–241 (Springer, 2015).
11. Kusuma, B. A. Determination of spinal curvature from scoliosis x-ray images using k-means and curve fitting for early detection of scoliosis disease. In *2017 2nd International conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, 159–164 (IEEE, 2017).
12. Mukherjee, J., Kundu, R. & Chakrabarti, A. Variability of cobb angle measurement from digital x-ray image based on different de-noising techniques. *Int. J. Biomed. Eng. Technol.* **16**, 113–134 (2014).
13. Zhang, D., Chen, B. & Li, S. Sequential conditional reinforcement learning for simultaneous vertebral body detection and segmentation with modeling the spine anatomy. *Med. Image Analysis* **67**, 101861 (2021).
14. Shao, Y. & Li, S. Adaptive loss engine for x-ray segmentation (alexs) for scoliosis intervention: assess digital segmentation and angle approximation. In *Medical Imaging 2024: Image-Guided Procedures, Robotic Interventions, and Modeling*, vol. 12928, 228–234 (SPIE, 2024).
15. Kirillov, A. *et al.* Segment anything. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 4015–4026 (2023).
16. Yi, J., Wu, P., Huang, Q., Qu, H. & Metaxas, D. N. Vertebra-focused landmark detection for scoliosis assessment. In *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, 736–740 (IEEE, 2020).
17. Chen, C. *et al.* Automating cobb angle measurement for adolescent idiopathic scoliosis using instance segmentation. In *2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 1–5 (IEEE, 2024).

18. Khanal, B., Dahal, L., Adhikari, P. & Khanal, B. Automatic cobb angle detection using vertebra detector and vertebra corners regression. In *International Workshop and Challenge on Computational Methods and Clinical Applications for Spine Imaging*, 81–87 (Springer, 2019).

19. Kim, K. C., Yun, H. S., Kim, S. & Seo, J. K. Automation of spine curve assessment in frontal radiographs using deep learning of vertebral-tilt vector. *IEEE Access* **8**, 84618–84630 (2020).

20. Reddy, P. K. *et al.* Anatomical landmark detection using deep appearance-context network. In *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 3569–3572 (IEEE, 2021).

21. Chen, B. *et al.* An automated and accurate spine curve analysis system. *IEEE Access* **7**, 124596–124605 (2019).

22. Wu, H., Bailey, C., Rasoulinejad, P. & Li, S. Automatic landmark estimation for adolescent idiopathic scoliosis assessment using boostnet. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 127–135 (Springer, 2017).

23. Cormack, G. D., Blair, D. A. & McMullin, J. N. Enhanced spectral resolution fft for step-like signals. *IEEE transactions on instrumentation measurement* **40**, 34–36 (1991).

24. Li, R., Xuan, J. & Shi, T. A frequency estimation method for two adjacent frequencies. *Measurement* **157**, 107614 (2020).

25. Umri, B. K., Akhyari, M. W. & Kusrini, K. Detection of covid-19 in chest x-ray image using clahe and convolutional neural network. In *2020 2nd international conference on cybernetics and intelligent system (ICORIS)*, 1–5 (IEEE, 2020).

26. Ikhsan, I. A. M., Hussain, A., Zulkifley, M. A., Tahir, N. M. & Mustapha, A. An analysis of x-ray image enhancement methods for vertebral bone segmentation. In *2014 IEEE 10th International Colloquium on Signal Processing and its Applications*, 208–211 (IEEE, 2014).

27. Simonyan, K. & Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556* (2014).

28. Wang, L. *et al.* Evaluation and comparison of accurate automated spinal curvature estimation algorithms with spinal anterior-posterior x-ray images: The aasce2019 challenge. *Med. Image Analysis* **72**, 102115 (2021).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryFile.docx