

# AutoSpine-Net: Spine Detection Using Convolutional Neural Networks for Cobb Angle Classification in Adolescent Idiopathic Scoliosis



Wahyu Caesarendra , Wahyu Rahmania , John Mathew, and Ady Thien

**Abstract** The Cobb angle measurement in adolescent idiopathic scoliosis is prone to inter- and intra-observer variations. This paper proposes a deep learning architecture for detection of spine vertebrae from X-ray images to automatically evaluate the Cobb angle, and to assess for the presence of scoliosis and severity of the curvature. The public AASCE MICCAI 2019 anterior–posterior X-ray image dataset was used for training and testing of the proposed convolutional neural network architecture. Seventeen vertebrae were detected from the input image to obtain 68 landmark features of the spine. The obtained landmarks were processed to measure the Cobb angle and to assess whether scoliosis was present. The severity of the curvature was further classified into mild, moderate and severe, if scoliosis was present. The results showed that the proposed algorithm has a classification accuracy of approximately 0.9 (90%). This architecture may be used as a tool to augment Cobb angle measurement in X-ray images of patients with adolescent idiopathic scoliosis in a real-world clinical setting.

**Keywords** Convolutional neural network · Deep learning · Scoliosis · Cobb angle · Vertebrae

---

W. Caesarendra (✉)

Faculty of Integrated Technologies, Universiti Brunei Darussalam, Jalan Tungku Link, Bandar Seri Begawan 1410, Brunei Darussalam  
e-mail: [wahyu.caesarendra@ubd.edu.bn](mailto:wahyu.caesarendra@ubd.edu.bn)

W. Rahmania

Department of Electrical Engineering, National Central University, Zhongli, Taoyuan, Taiwan

J. Mathew · A. Thien

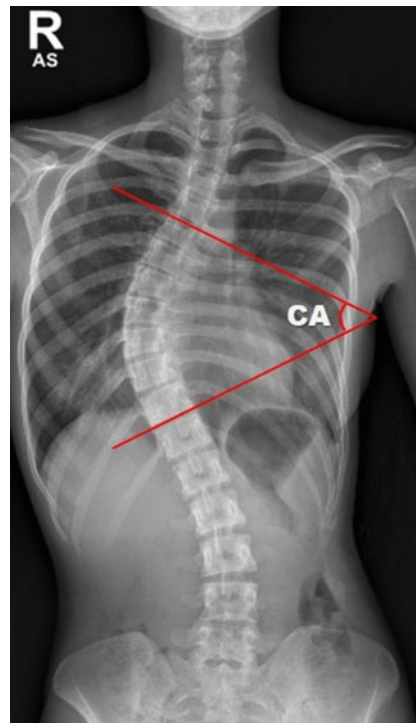
Department of Neurosurgery, Brunei Neuroscience, Stroke and Rehabilitation Centre, Pantai Jerudong Specialist Centre, Jerudong, Brunei Darussalam

## 1 Introduction

Adolescent idiopathic scoliosis (AIS) is an abnormal curvature, in which the spine curves abnormally from side to side and rotates. The Cobb measuring method is the gold standard used in quantifying the curve. The Cobb angle (CA) is measured from the most tilted vertebra (end vertebra) above and below the apex (most laterally placed vertebra) of the curve on radiographs taken either in the anterior–posterior or the posterior–anterior view on the coronal plane [1]. Generally, the manual procedure requires lines to be drawn onto a hardcopy of radiographic films and the angle between the two lines are measured using a protractor (Fig. 1).

In addition to being time consuming, CA measurement of spine is unreliable [2] and is associated with inter-observer and intra-observer variations [3]. Reported accuracies of measuring CA vary from  $2^\circ$  to  $11^\circ$  [4–6] with measurements differing up to  $5^\circ$  even with the same end vertebrae selected [4, 7]. The emergence of computerized digitization of radiographs has allowed for semi-automatic assessment of CA where the picture archiving and communications system (PACS) allows for an in-built function enabling the user to digitally draw the lines for the end vertebrae and the system automatically measures the CA. This has shown to have good reliability and

**Fig. 1** Measurement of Cobb angle (CA)



less variations compared to the manual method [6, 8]. However, this method is still dependent on the manual inputs from the user.

The availability of an accurate reproducible CA measurement method is important as the assessment of the CA is used to diagnose AIS and to guide decisions regarding curve progression as well as therapeutic options including surgical interventions. In addition, the method has been user-friendly and takes less time compared to the manual and semi-automatic methods.

The development of computer vision technologies [9], machine learning methods [10, 11], and deep learning methods [12–14] led to attempts for the transition from traditional and semi-automatic CA assessment to automated CA measurement. In this instance, the X-ray images were processed by means of computerized learning method to enable CA measurement and prediction. The various machine learning methods used in scoliosis clinical practice, including screening, diagnosis and classification, have previously been reported [11]. Vertebrae detection is an important stage to identify the landmarks of interest on the spine image for CA measurement. Studies on the vertebrae and spine detection based on machine learning methods [10, 11] and deep learning methods [12–14] have been previously reported. Bernstein et al. showed that a neural network (NN) can be applied to automatically train the vertebrae centroids detection [10]. However, automatic detection of the vertebrae in X-ray images can be difficult. Computer vision task is challenging in X-rays compared to computed tomography (CT) and magnetic resonance imaging (MRI) images due to multiple overlapping shadows of the ribs and pelvis, as well as, a differences in contrast between thoracic and lumbar vertebrae regions [15, 16]. Convolutional neural network (CNN) architecture can help overcome this problem [10]. Researchers in the field of deep learning (DL) have previously developed fully automated methods for Cobb angle measurement (Table 1). The methods had successfully detected and segmented the vertebrae and the spine. However, most of the studies

**Table 1** Brief summary of selected papers in CA measurement using deep learning methods

References	Image processing methods and the procedures	Method
[12]	<ol style="list-style-type: none"> <li>1. Vertebrae isolation and detection</li> <li>2. Drawing spinal curve line</li> <li>3. Vertebra rectangle segmentation</li> </ol>	CNN
[13]	<ol style="list-style-type: none"> <li>1. Usage of Moire screening system to obtain the curve</li> <li>2. Fitting the curve to the 17 position using cubic B-spline</li> <li>3. Calculate the two contact points of three lines perpendicular to the curve</li> <li>4. Define the middle vertebrae</li> <li>5. Calculate the angle</li> </ol>	CNN
[14]	<ol style="list-style-type: none"> <li>1. Use picture archiving and communication systems</li> <li>2. Consider the orientation of the overall spinal curve</li> <li>3. Identify the max angle between the superior perpendicular of the cranial vertebrae and the inferior perpendicular of the caudal vertebrae at the longitudinal central lines of the vertebral bodies</li> </ol>	R-CNN

were not collaborated with the clinicians and the results obtained were not compared with measurements made in the real-world setting.

In this paper, we propose a CNN for spine vertebra detection, CA evaluation and curve severity classification in AIS. The main objective is to automate and augment (1) the detection and the assessment of the CA, and (2) the confirmation of presence of scoliosis, based on standard spine X-ray images. The severity of the scoliosis is also classified into mild ( $10^{\circ}$ – $25^{\circ}$ ), moderate ( $>25^{\circ}$  to  $40^{\circ}$ ), and severe ( $> 40^{\circ}$ ).

## 2 Proposed Method

### 2.1 Datasets

The collection and labelling of spinal images were performed by the public AASCE MICCAI 2019 anterior–posterior X-ray images dataset [17]. The input images vary in sizes from  $359 \times 973$  to  $1427 \times 3755$ . Each image contains 17 vertebrae from the thoracic (upper spine) and lumbar (lower spine). The image input resolution is set to  $1024 \times 512$  for the algorithm development. A total of 962 images are used as follows, 481 images for training, 323 images for validation, and 158 images for testing are used. Each vertebrae is located by 4 corner landmarks. The ground-truth of the 68 landmarks or points in each image is provided by the dataset.

### 2.2 System Overview

The 50-layer ResNet [18] is used to classify 68 landmarks to obtain the corner offset of spine. This CNN consists of several convolutional layers that learn the local features of the images and generate the classifications. The proposed network (Fig. 2) includes

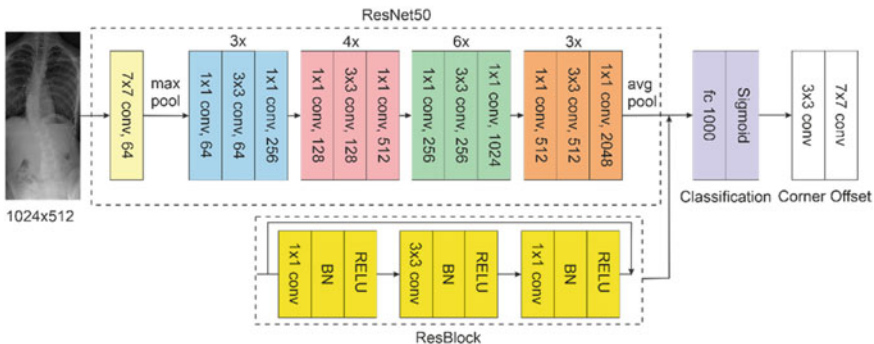


Fig. 2 Vertebrae detection network

pooling layers (average pool and max pool), classification, and corner offset. Combination of semantically similar features into a single feature reduces the dimensions of the extracted features and fully connected layers, and gives a final probability value for the class. Network depth has been previously shown to be beneficial to classification accuracy [19]. However, its performance can become saturated with resultant rapid decrease in performance as the network gained greater depth. This issue can be fixed by the ResNet framework [20] where a shortcut connection is added for every three convolution layers across the deep network. These shortcut connections performed identity mapping without additional parameters which can increase computational complexity. This simplification of network optimization during the training process enables ResNet to achieve a higher accuracy from deeper networks when performing image classification tasks.

The ResNet50 architecture is mainly composed of residual blocks (Fig. 2). Residual connection in ResNet architecture maintains connection to gain knowledge during training and speed up model training time by increasing network capacity. Batch normalization with ReLU activation is added for each convolutional layer. Bi-cubic interpolation is used as upscaling method. The skip connection technique is performed to exploit high-level semantic information and low-level fine details to improve model performance.

During the training process, a fine-tuning technique is applied to transfer the connection weights from the pre-trained model to our model and retrain the model to the current task. This model accepts an image as input and performs a fully connected layer as a final assessment. Finally, the model outputs the bounding box of each target object as well as the corresponding category label.

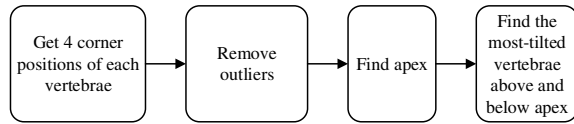
The X-ray images used contain 17 vertebrae, where each vertebrae has 4 corner landmarks: top-left, top-right, bottom-left and bottom-right. Therefore, each image has a total of 68 landmarks. The order of the landmarks is used to accurately localize the vertebrae, so that the slope of each can be known. The landmarks were separated into different groups to obtain an output feature map with a channel number of 68. Then, a heat map of the center point [21] is constructed to obtain a corner offset maps using a convolutional layers for landmark localization.

Landmarks of each corner of the vertebrae were obtained using the corner offset. The corner offset was obtained from the center of the heat map to the vertebrae margin using L1 loss to optimize the corner offset at the midpoint.

### ***2.3 Cobb Angle Measurement for Classification***

A review study on the classification of AIS is presented in [22]. The Author reviewed the clinical classification of AIS from a few previous studies. It mentioned that the classification provides a better and more reliable tool to assist surgeons in determining the appropriate method of treatment for certain curve pattern. In addition, with the developing methods in 3D reconstruction may be used as a basis classification for new therapeutic concepts [22].

**Fig. 3** CA measurement for classification



In this study, the application of CNN for AIS classification is presented. The steps to calculate the CA from the position of each corner found is presented in Fig. 3. After detecting an object on the X-ray image, detected bounding boxes are displayed on the spine. Boxes with a score of more than 0.5 were extracted. From the location of the detected boxes, the center point of each vertebrae is found to remove some outliers based on the anatomy of spine, where the adjacent vertebrae should not be far apart from each other. If the x-axis center of the detected bounding box is more than half the width of the box from the x-axis center of its two closest neighbors (top and bottom), the box is rejected as an outlier. Otherwise, the position of the box is reconsidered based on the position of the nearest boxes.

Following this, the depth of the curve at the found position of the corner box is calculated. For each of the two vertebrae, the distance between the bottom-left point of the upper box and the top-left of the lower box, and the bottom-right point of the upper box and the top-right of the lower box is calculated. The apex of the spinal is found as the deepest part of the curve.

For each box above the apex, the slope of each vertebrae is measured based on the position between top-left and top-right to detect the most-tilted vertebrae above the apex. For each box below the apex, the slope of each vertebrae is measured based on the position between bottom-left and bottom-right to detect most-tilted vertebra below the apex. The Cobb angle is then calculated as the angle of the intersection between two lines from the most-tilted vertebrae above the apex and most-tilted vertebrae below the apex.

### 3 Results

The datasets was trained on the RTX2060 GPU with Intel Core-i7 processor. Figure 4 shows the performance of the training dataset and the validation dataset when training the network. The models are initialized from the pre-trained weights on ImageNet. The network was trained with a learning rate of 0.0001 with Adam optimizer during training. The batch and epoch sizes are set as 2 and 100, respectively.

Figure 5 shows the result of the detection of vertebrae in the spine on the X-ray images. The condition of the patient's spine is classified as normal if the measured CA is less than  $10^\circ$ . For mild, moderate, and severe AIS, the CA measurements are  $10^\circ$  to  $25^\circ$ ,  $>25^\circ$  to  $40^\circ$ , and  $>40^\circ$ , respectively.

The performance metrics comparison of the four classes is summarized in Table 2. *Precision rate (PR)*, *Recall*, and *F1-measure* [23] can be computed as follows:

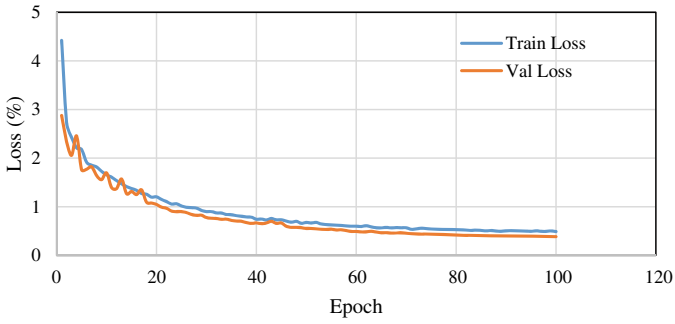


Fig. 4 Performance of the dataset in network training

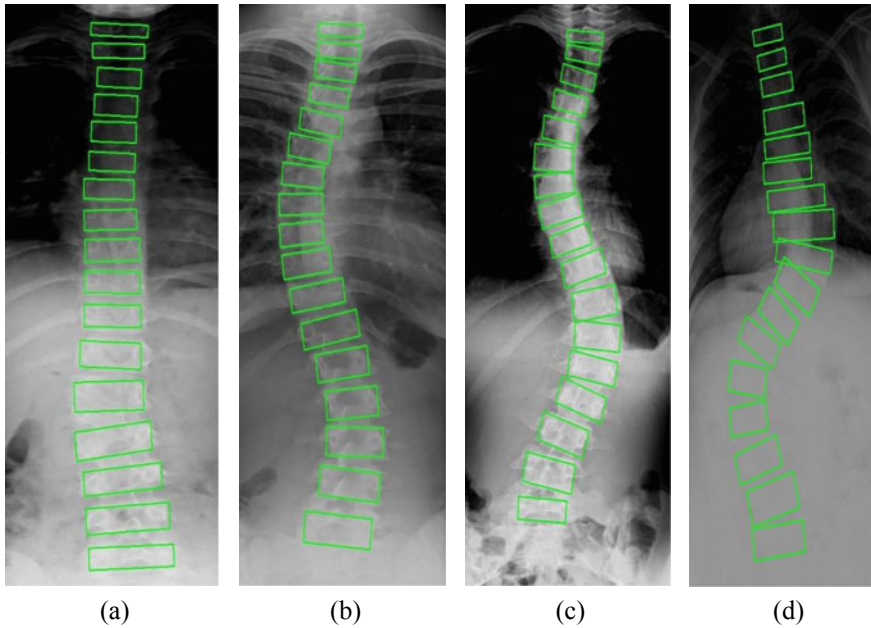


Fig. 5 Detection results: a normal, b mild, c moderate, and d severe

Table 2 Performance metrics comparison

Class	Accuracy results		
	PR	Recall	F1-measure
Normal	0.92	0.89	0.9
Mild	0.95	0.93	0.94
Moderate	0.94	0.89	0.91
Severe	0.9	0.8	0.84
Average	0.92	0.87	0.9

$$PR = \frac{TP}{TP + FP} \quad (1)$$

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

$$F1 - measure = 2 \times \frac{PR \times Recall}{PR + Recall} \quad (3)$$

where  $TP$  is true positive,  $FP$  is false positive, and  $FN$  is false negative.  $TP$  is the detected area of the vertebrae and corresponds to the associated class.  $FP$  is the detected area not associated with the vertebrae.  $FN$  is the area associated with the vertebrae that is not detected.

Severe AIS has the lowest  $PR$ ,  $Recall$ , and  $F1-measure$  compared to the other classes. The increase in the curvature made it difficult for the vertebrae to be detected in the area of the arch. In some cases, the vertebrae is completely undetected or misrepresented.

Mild AIS has the highest  $PR$ ,  $Recall$ , and  $F1-measure$  compared to other classes. The network can detect the vertebrae well as the spine is not too curved. In addition, the X-ray images have good lighting and contrast conditions for this class of spine in our dataset. Normal spine has lower accuracy than mild AIS as some image conditions for this class are not optimum. This resulted in a high number of  $FP$  and  $FN$  in the detection.

## 4 Discussion

The proposed architecture using CNN accurately detected the location of each of the 17 vertebrae in the spine X-ray. In addition to this, the bounding box was evaluated to be sufficient in its accordance with the vertebra positions. Its performance was accurate to provide the information needed to detect the superior and inferior end vertebrae, enabling the CA to be evaluated correctly.

The detection results also showed that the proposed architecture can be used to identify the vertebrae in X-ray images of different contrast and lighting conditions. Our test on several images with poor contrast and lighting conditions yielded good results. Importantly, CA measurements and curve classification were able to be accurately accomplished even when the detection process failed to identify one or two vertebrae. This is was a key part of the algorithm as X-ray images may come in different contrast and lighting qualities in the clinical setting, depending the severity of the curve as well as the patient's body habitus.

Previous studies using CNN [12–14] focus on vertebrae detection and measurement of CA under certain conditions but did not classify the severity of scoliosis. The method we proposed was able to measure CA from normal to severely scoliotic



spine (up to  $81^\circ$ ). This was an important step as severely abnormal curvatures were often difficult to detect.

There are some limitations to our study. In this proposed CNN, more errors in detection had occurred in images where the X-ray were of different sizes and when it involved larger areas from the neck to the hip which were not important landmarks for vertebrae detection. Further improvements with automatic image cropping to satisfy the conditions for optimal vertebrae detection is ongoing. Lastly, the results from this CNN were not validated against the clinicians' CA measurements (which remains the gold standard). This important final step will be crucial in confirming that this CNN will be capable in augmenting the specialist clinician's ability to accurately measure CA and may be used as a tool for non-specialist clinicians and nurses to assess CA in AIS patients.

## 5 Conclusions

A convolutional neural network for vertebrae spine detection, Cobb angle measurement and curvature severity classification in X-ray images of adolescent idiopathic scoliosis is proposed in this paper. The detection of vertebrae and classification had an accuracy of 0.9 (90%). Upon clinical validation, this architecture may be used as tool to augment Cobb angle measurement in X-ray images of patients with adolescent idiopathic scoliosis in a real-world clinical setting. A developed CNN method is also possible to be implemented for the real-time assessment or monitoring of scoliosis patients in the future [24].

**Funding** This project was supported by The AO Spine National Research Grant 2020 [AOSEA(R)2020-05].

## References

1. Cobb J (1948) Outline for the study of scoliosis. *Instr Course Lect* 5:261–275
2. Wu H, Bailey C, Rasoulinejad P, Li S (2018) Automated comprehensive adolescent idiopathic scoliosis assessment using MVC-Net. *Med Image Anal* 48:1–11
3. De Carvalho A, Vialle R, Thomsen L, Amzallag J, Cluzel G, le Pointe HD, Mary P (2007) Reliability analysis for manual measurement of coronal plane deformity in adolescent scoliosis. Are  $30 \times 90$  cm plain films better than digitized small films? *Eur Spine J* 16(10):1615–1620
4. Carman DL, Browne RH, Birch JG (1990) Measurement of scoliosis and kyphosis radiographs: intraobserver and interobserver variation. *J Bone Joint Surg* 72(3):328–333
5. Cheung J, Wever DJ, Veldhuizen AG (2002) The reliability of quantitative analysis on digital images of the scoliotic spine. *Eur Spine J* 11:535–542
6. Shea KG, Stevens PM, Nelson M (1998) A comparison of manual versus computer-assisted radiographic measurement intraobserver measurement variability for Cobb angles. *Spine* 23:551–555

7. Morrissy RT, Goldsmith GS, Hall EC, Kehl D, Cowies GH (1990) Measurement of the Cobb angle on radiographs of patients who have scoliosis. Evaluation of intrinsic error. *The J Bone Joint Surg Am* 72(3):320–327
8. Chockalingam N, Dangerfield PH, Giakas G, Cochrane T, Dorgan JC (2002) Computer-assisted Cobb measurement of scoliosis. *Eur Spine J* 11:353–357
9. Aroeira RM, de Las Casas EB, Pertence AE, Greco M, Tavares JM (2016) Non-invasive methods of computer vision in the posture evaluation of adolescent idiopathic scoliosis. *J Bodyw Mov Ther* 20(4):832–843
10. Bernstein P, Metzler J, Weinzierl M, Seifert C, Kisel W, Wacker M (2021) Radiographic scoliosis angle estimation: spline-based measurement reveals superior reliability compared to traditional COBB method. *Eur Spine J* 30:676–685
11. Chen K, Zhai X, Sun K, Wang H, Yang C, Li M (2021) A narrative review of machine learning as promising revolution in clinical practice of scoliosis. *Ann Transl Med* 9(1):1–16
12. Horng MH, Kuok PC, Fu M (2019) Cobb angle measurement of spine from X-ray images using convolutional neural network. *Comput Math Methods Med* 19(2019):6357171
13. Choi R, Watanabe K, Jingufi H (2017) CNN-based spine and Cobb angle estimator using Moire images. *IIEEJ Trans Image Electron Vis Comput* 5(2):35–144
14. Pan Y, Chen Q, Chen T (2019) Evaluation of a computer-aided method for measuring the Cobb angle on chest X-rays. *Eur Spine J* 28(12):3035–3043
15. Huang SH, Chu YH, Lai SH, Novak CL (2009) Learning-based vertebra detection and iterative normalized-cut segmentation for spinal MRI. *IEEE Trans Med Imaging* 28(10):1595–1605
16. Glocker B, Zikic D, Konukoglu E, Haynor DR, Criminisi A (2013) Vertebrae Localization in Pathological Spine CT via Dense Classification from Sparse Annotations. In: Mori K, Sakuma I, Sato Y, Barillot C, Navab N (eds) *Medical image computing and computer-assisted intervention—MICCAI 2013*. MICCAI 2013. Lecture Notes in Computer Science 8150. Springer, Berlin, Heidelberg, pp 262–270
17. AASCE—Grand Challenge Homepage. <https://aasce19.grand-challenge.org/>. Last accessed 30 Aug 2021
18. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: 2016 IEEE conference proceedings on computer vision and pattern recognition (CVPR). IEEE, Las Vegas, NV, USA, pp 770–778
19. Szegedy C, Liu W, Jia Y, Sermanent P, Reed S, Anguelov D, Erhan D, Vanhouche V (2015) Going deeper with convolutions. In: 2015 IEEE conference proceedings on computer vision and pattern recognition (CVPR). IEEE, Boston, MA, USA, pp 1–9
20. He K, Sun J (2015) Convolutional neural networks at constrained time cost. In: 2015 IEEE conference proceedings on computer vision and pattern recognition (CVPR). IEEE, Los Alamitos, CA, USA, pp 5353–5360
21. Yi J, Wu P, Huang Q, Qu H, Metaxas DN (2020) Vertebra-focused landmark detection for scoliosis assessment. In: 2020 IEEE 17th international symposium on biomedical imaging (ISBI). IEEE, Iowa City, IA, USA, pp 736–740
22. Ovadia D (2013) Classification of adolescent idiopathic scoliosis (AIS). *J Children’s Orthop* 7:25–28
23. Rahmaniar W, Wang W-J (2019) Real-time automated segmentation and classification of calcaneal fractures in CT images. *Appl Sci* 9(1):3011–3028
24. Jung JY, Bok SK, Kim BO, Won Y, Kim JJ (2015) Real-time sitting posture monitoring system for functional scoliosis patients. *Lecture Notes Electr Eng* 339:391–396