Medical Image Analysis 58 (2019) 101542

Contents lists available at ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

Accurate automated Cobb angles estimation using multi-view extrapolation net

Liansheng Wang[a], Qiuhao Xu[a], Stephanie Leung[b,c], Jonathan Chung[b,c], Bo Chen[b,c,*], Shuo Li[b,c,*]

[a] Department of Computer Science, Xiamen University, Xiamen 361005, China
[b] Department of Medical Imaging, Western University, ON, Canada
[c] Digital Image Group, London, ON, Canada

ARTICLE INFO

Article history:
Received 26 December 2018
Revised 2 June 2019
Accepted 1 August 2019
Available online 9 August 2019

Keywords:
Scoliosis
Spinal curvature
Cobb angles
Multi-Task learning
Error estimation

ABSTRACT

Accurate automated quantitative Cobb angle estimation that quantitatively evaluates scoliosis plays an important role in scoliosis diagnosis and treatment. It solves the problem of the traditional manual method, which is the current clinical standard for scoliosis assessment, but time-consuming and unreliable. However, it is very challenging to achieve highly accurate automated Cobb angle estimation because it is difficult to utilize the information of Anterior-posterior (AP) and Lateral (LAT) view X-rays efficiently. We therefore propose a Multi-View Extrapolation Net (MVE-Net) that provides accurate automated scoliosis estimation in multi-view (both AP and LAT) X-rays. The MVE-Net consists of three parts: Joint-view net learning AP and LAT angles jointly based on landmarks learned from joint representation; Independent-view net learning AP and LAT angles independently based on landmarks learned from unique independent feature of AP or LAT angles; Inter-error correction net learning a combination function adaptively to offset the first two nets' errors for accurate angle estimation. Experimental results on 526 X-rays show 7.81 and 6.26 Circular Mean Absolute Error in AP and LAT angle estimation, which shows the MVE-Net provides an accurate Cobb angle estimation in multi-view X-rays. Our method therefore provides effective framework for automated, accurate, and reliable scoliosis estimation.

1. Introduction

Cobb angles (Cobb, 1948) are widely used for scoliosis diagnosis and treatment decisions. It is the most common and convenient standard for scoliosis measurement. Scoliosis is a structural, lateral, rotated curvature of the spine, which especially arises in children at or around puberty and leads to disability (Weinstein et al., 2008). Large studies have shown that the prevalence rate of scoliosis in children can be as high as 5.2%. This disease can also lead to complications from injury to the heart and lungs (Asher and Burton, 2006). Clinicians currently assess the Cobb angles to make treatment decisions. Therefore, it is essential to have a reliable way to measure Cobb angles.

However, existing manual scoliosis assessment in clinical practice is time-consuming and unreliable (Vrtovec et al., 2009; Pruijs et al., 1994). As shown in Fig. 1, the current standard for assessing scoliosis is to measure Cobb angles manually based on anterior-posterior (AP) and lateral (LAT) X-ray images. Clinicians manually measure the landmarks and choose the pivotal vertebrae using pivotal landmarks. This leads to the accuracy of the measurement unreliable since the results are affected by the selection of pivotal vertebrae and the bias of different observers. It is difficult to choose these vertebrae manually.

Existing segmentation-based methods and direct estimation methods for Cobb angles measurement still have limitations to achieve high accuracy.

Segmentation-based methods suffer from multiple error transmission since these methods firstly segment the required anatomical structures and then measure scoliosis based on the segmentation. Previous segmentation-based methods for scoliosis assessment such as filtering (Anitha et al., 2014; Zhang et al., 2010; Zhang et al., 2009), active contouring (Anitha and Prabhu, 2012), and physics models (Sardjono et al., 2013) locate the required vertebrae and calculate the Cobb angles. These methods are limited by some measurement disadvantages such as the selection of vertebrae and the bias of different users.

Direct estimation methods (Sun et al., 2017; Wu et al., 2017; Xue et al., 2017a; 2017b; Zhen et al., 2015; Wu et al., 2018) aim to obtain the relationship between medical images and clinical

* Corresponding authors at: Digital Image Group, London, ON, Canada.
E-mail addresses: bchen.chen@gmail.com (B. Chen), slishuo@gmail.com (S. Li).

https://doi.org/10.1016/j.media.2019.101542
1361-8415/

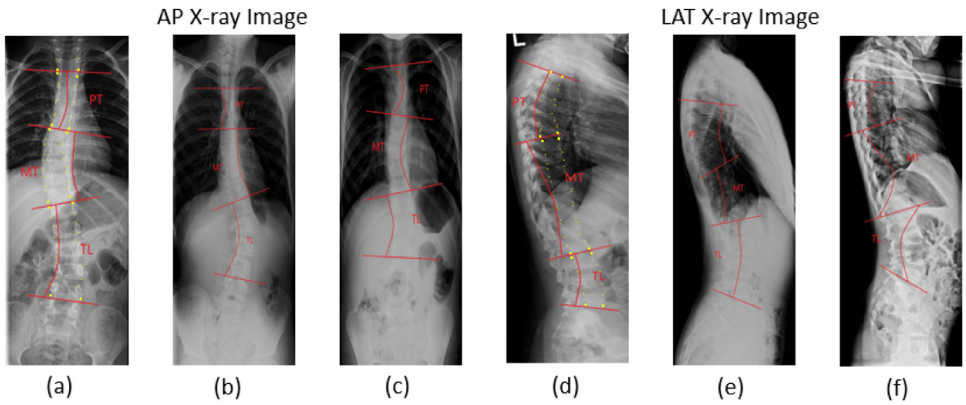

Fig. 1. Traditional methods use landmark (yellow points in (a) and (d), the pivotal landmarks have been magnified) to measure Cobb angles. It is challenging to measure Cobb angles due to high ambiguity and variability in scoliosis AP (a–c) and LAT (d–f) X-rays from different subjects since it is difficult to identify the most tilted vertebral endplates (red lines) on the X-ray images. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

measurements directly without the results based on segmentation. These methods have achieved great success recently in many clinical measurement fields such as cardiac volume measurement (Xue et al., 2017a; 2017b; Zhen et al., 2015) and scoliosis measurement in single view (Sun et al., 2017; Wu et al., 2017) and multi-view (Wu et al., 2018) images. Applied to scoliosis measurement, several initial attempts have been put forward in conference. S^2VR (Sun et al., 2017) achieved to build up the relationship between spinal landmarks and Cobb angles to improve the accuracy and robustness of the Cobb angle estimation. BoostNet (Wu et al., 2017) encoded and decoded the spinal features to remove deleterious outlier features, which achieved robust and accurate spinal landmarks. Despite their great effectiveness in single view estimations, these methods still cannot effectively handle multi-view X-ray images since the 3D spatial joint features between the AP and LAT views are not learned. The MVC-Net (Wu et al., 2018) takes into account the underlying relationship between AP and LAT X-rays, but it didn't consider the unique independent features the AP and LAT X-rays have, thus losing a certain precision.

High-precision calculation (Quarteroni et al., 2000) is a catalog of very popular methods in numerical mathematics for decreasing error. It can be used to achieve complex calculation with high accuracy. It can also be used to calculate complex expression numerically without plenty of formula derivation. For example, Laporta (2000) used a high-precision approximation to calculate a complex integration. High-precision calculation achieves this calculation. Stefano (2017) used high-precision calculation by calculating a physical constant for fast convergence and achieved great success. This technology has been proved to effectively improve the accuracy of the calculation. The extrapolation (Quarteroni et al., 2000) is one of the most useful methods in the high-precision calculation. It can be used to improve much popular numeric calculation methods such as trapezoidal method and Simpson method (Quarteroni et al., 2000). Mathematically, the extrapolation provides a combination function of the functional value at measured data to speculate the functional value, which makes the estimated value much more close to the ground truth. To achieve higher accuracy, the combination function depends on a series of rough measured data (Criminisi, et al., 2011). However, it is very time consuming to calculate the parameters of the combination function. Deep learning structure can effectively solve this problem by automatically and adaptively learning these parameters since it is much less time consuming using numeral approximation instead of complex precise calculation (Hinton et al., 2012; He et al., 2015).

In this paper, we propose a Multi-View Extrapolation Net (MVE-Net) to accurately estimate the Cobb angles from AP and LAT X-rays. Firstly, the MVE-Net directly estimate landmarks from two different ways: One uses a cross structure (joint-view net) to learn the relationship between both AP and LAT X-rays (joint features, the synchronous pivotal landmarks between AP and LAT images). The other uses a parallel structure (independent-view net) to learn the information of AP or LAT X-rays (unique independent features, the independent pivotal landmarks in AP and LAT images) to ignore the noise from the other. then both of the two ways use the landmarks to calculate Cobb angles. Secondly, since the extrapolation can leverage the two kinds of different pivotal landmarks, based on the two estimations through different ways, our MVE-Net applies extrapolation to adaptively offset the error of the two Cobb angles estimations by each other, then obtain an enhanced estimation. Incorporating a deep learning structure, the MVE-Net learns the combination of the two Cobb angles estimations adaptively (inter-error correction net), which further improves the accuracy and effectiveness of the calculation. The MVE-Net learns joint and independent features successfully and obtains a more accurate estimation.

Contribution In summary, our work contributes in the following aspects:

- The newly proposed MVE-Net achieves high accuracy in clinical Cobb angle measurement from AP and LAT X-rays.
- For the first time, the high-precision calculation (Quarteroni et al., 2000) is combined with deep learning, which leverages the first two networks, and provides an accurate estimation
- For the first time, a newly designed error controlled loss function is proposed, which achieves not only fast convergence but also high accuracy since the joint and independent features are combined effectively.

2. Proposed MVE-Net architecture

As shown in Fig. 2, the newly proposed MVE-Net consists of three parts: (1) A cross-learning joint-view net (Section 2.1) for learning joint features between AP and LAT angle, which simultaneously learns the features in AP and LAT images and gets the synchronous pivotal landmarks (Fig. 1) between AP and LAT images. (2) A parallel-learning independent-view net (Section 2.2) for learning unique independent features, which gets the independent pivotal landmarks in AP and LAT images separately as the basement of rough Cobb angle calculation. (3) An inter-error correction net (Section 2.3) to leverage the previous two calculations, offset the error through a combination function using extrapolation to

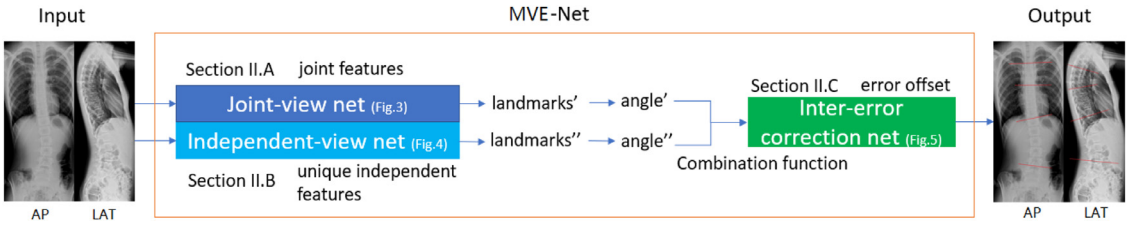

Fig. 2. Architecture of the MVE-Net for Cobb angle estimation. It contains (1) A joint-view net for learning joint features between AP and LAT angle, (2) An independent-view net for learning unique independent features, (3) An inter-error correction net to take advantage of the two former nets and increase the accuracy.

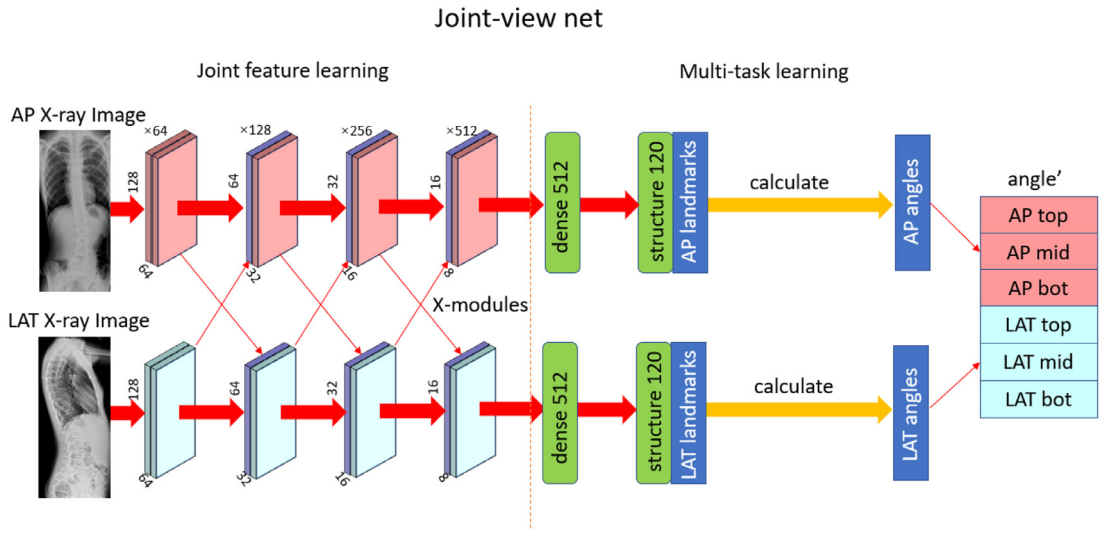

Fig. 3. The joint-view net uses a series of X-modules for learning joint features. It estimates landmarks form joint features and estimates Cobb angles using the landmarks.

combine the error complementarity between the two estimations, which increases the final estimation's accuracy. In our method, we use joint-view net and independent-view net to obtain two initial angles, then use an inter-error correction net based on extrapolation to improve the accuracy of the final estimation.

2.1. Joint-view net for synchronous pivotal landmarks

The joint-view net is designed to leverage the correlation between AP and LAT views in order to acquire robust spinal landmarks, and use those landmarks to calculate the Cobb angles. It consists of two convolution net parts: (1) Joint feature learning: a spinal landmark estimator network using a series of X-modules (Wu et al., 2018) to automatically learn common features between AP and LAT X-rays, and (2) Multi-task learning: a dedicated landmark estimator network that uses these joint AP/LAT X-ray features to regress landmarks for Cobb angles calculation. The X-module connections (McCloskey et al., 1989) are specifically designed to capture the underlying physical correlation of a structure using multi-view images, which are detailedly defined in the paper (Wu et al., 2018). The joint features can simultaneously learn the features in the AP and LAT images effectively, which improves the accuracy of landmark learning. The joint-view net architecture, illustrated in Fig. 3, captures the full extent of the spinal structure for robust Cobb angle estimation using the X-modules.

The loss function of the joint-view net (Eq. (4)) not only minimizes the error but also conforms to the spinal of our ground truth, it consists of two parts: a robust regression loss (Eq. (1)) and (a correction loss. (Eq. (4))

The robust regression loss uses the log of hyperbolic cosine as the objective function:

$$L'_{reg}(X^{\chi}, Y^{\chi}, \theta') = MEAN(log\ ch(M_{lm}(X^{\chi}; \theta') - Y^{\chi})) \quad (1)$$

Here χ represents for AP or LAT, X is the input image, M_{lm} is the estimated landmark, Y is the ground truth, θ' is the parameters of the joint-view net.

The correction loss uses the Pearson loss.

$$\rho' = \frac{MEAN(M_{lm}(X^{\chi}; \theta')Y^{\chi})}{STD(M_{lm}(X^{\chi}; \theta'))STD(Y^{\chi})} - \frac{MEAN(M_{lm}(X^{\chi}; \theta')MEAN(Y^{\chi}))}{STD(M_{lm}(X^{\chi}; \theta'))STD(Y^{\chi})} \quad (2)$$

Here MEAN is the element-wise arithmetic mean, and STD is the standard deviation of the predicted and ground truth landmark coordinates.

Since the output of Pearson coefficient r ranges from -1 (negative correlation) to 1 (perfect correlation), we rearranged it as follows in order to force the output to be between 0 (perfect correlation) and 1 (negative correlation):

$$L'_{cor}(X^{\chi}, Y^{\chi}, \theta') = \frac{1}{2}(1 - \rho') \quad (3)$$

The overall loss function of the joint-view net is thus defined as:

$$L_{jv}(X^{\chi}, Y^{\chi}, \theta') = L'_{cor}(X^{\chi}, Y^{\chi}, \theta') + \phi' L'_{reg}(X^{\chi}, Y^{\chi}, \theta') \quad (4)$$

where ϕ' is a scaling factor controlling the relative importance of the regression loss term, and $\phi' = 5$ in our experiments.

2.2. Independent-view net for independent pivotal landmarks

The independent-view net is designed to focus on the individual features of AP and LAT views in order to acquire individual spinal landmarks, and use those landmarks to estimate the Cobb angles. It consists of two convolution net parts: (1) Independent

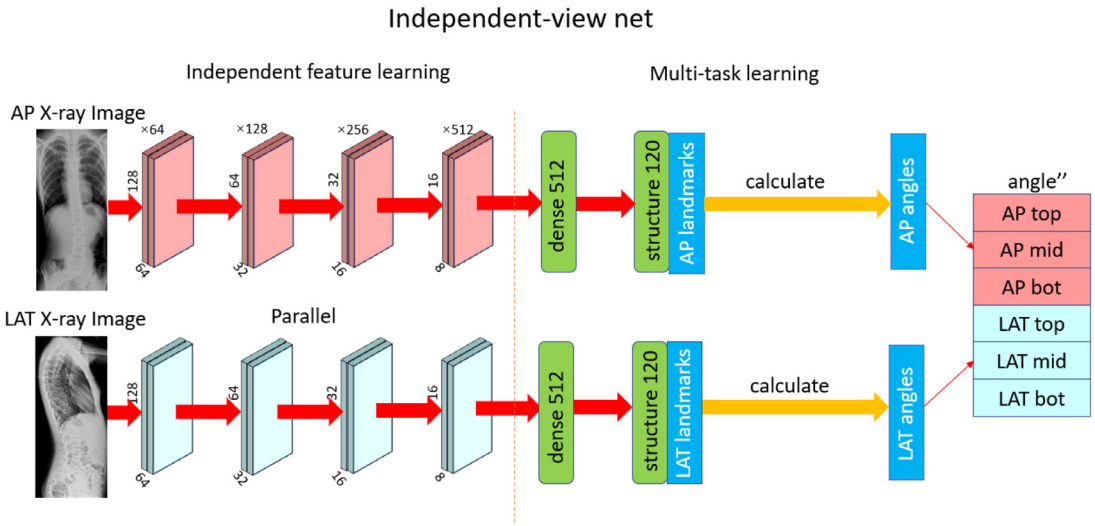

Fig. 4. The independent-view net learns independent features from AP and LAT X-rays in parallel. It estimates landmarks from independent features and estimates Cobb angles using the landmarks.

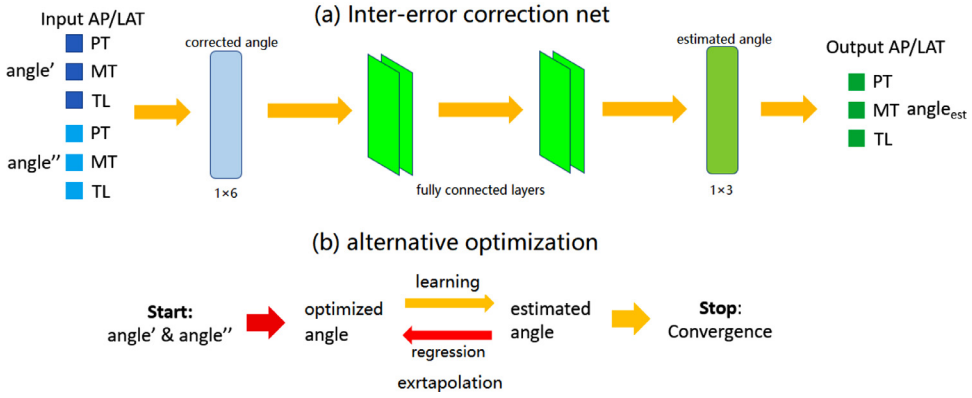

Fig. 5. The inter-error correction net (a) and the optimization of the loss function (b). The inter-error error correction net iteratively optimizes the corrected angle and the estimated angle alternatively to learn the optimization direction and step length, which can obtain a more accurate estimation. Theorem 1 proves the estimated angle has higher order accuracy.

feature learning: a spinal landmark estimator network using a series of convolution layers in parallel to automatically learn individual features of AP and LAT X-rays, and (2) Multi-task learning: a dedicated landmark estimator network that uses these individual AP/LAT X-ray features to regress landmarks for Cobb angles calculation. The independent features are the foundation of the landmark detecting. The independent-view net architecture, illustrated in Fig. 4, captures the single view of the spinal structure for accurate Cobb angle estimation using parallel learning.

Similarly, to minimize the error and conform to the spinal of our ground truth, the loss function of the independent-view net can be written as:

$$L_{iv}(X^{\chi}, Y^{\chi}, \theta'') = L''_{cor}(X^{\chi}, Y^{\chi}, \theta'') + \phi'' L''_{reg}(X^{\chi}, Y^{\chi}, \theta'') \tag{5}$$

where

$$L''_{reg}(X^{\chi}, Y^{\chi}, \theta'') = MEAN(log\ ch(M_{lm}(X^{\chi}; \theta'') - Y^{\chi})) \tag{6}$$

$$L''_{cor}(X^{\chi}, Y^{\chi}, \theta'') = \frac{1}{2}(1 - \rho'') \tag{7}$$

$$\rho'' = \frac{MEAN(M_{lm}(X^{\chi}; \theta'')Y^{\chi})}{STD(M_{lm}(X^{\chi}; \theta''))STD(Y^{\chi})} - \frac{MEAN(M_{lm}(X^{\chi}; \theta''))MEAN(Y^{\chi})}{STD(M_{lm}(X^{\chi}; \theta''))STD(Y^{\chi})} \tag{8}$$

θ'' is the parameters of the independent-view net. ϕ'' is a scaling factor controlling the relative importance of the regression loss termand, and $\phi'' = 5$ in our experiments.

2.3. Inter-error correction net using extrapolation

The inter-error correction net is designed for further improving the accuracy of the output of the joint-view net (*angle'*) and independent-view net (*angle''*). As shown in Fig. 5, the inter-error correction net uses a series of fully connected layers, alternating optimization (Section 2.3(a), Fig. 5(b)) and designed error controlled loss function (Section 2.3(b), Eq. (9)) to fit the extrapolation.

(a) In the alternative optimization, the inter-error correction net learns two extrapolated angles ($angle_{opt}$ and $angle_{est}$) for faster convergence by iteration: **Stage 1: Optimized angle regression** ($angle_{opt}$ Section 2.3.1) based the normalization of the error leverages the data's error, which provides a potential better correction as reference. The normalization makes sure our method is effective to calculate the Cobb angles on different scoliosis level (for example, low scoliosis level: less than 30°, and high scoliosis level: more than 45°) simultaneously; **Stage 2: Estimated angle learning** ($angle_{est}$ Section 2.3.2) based on the results of a combination function (Eq. (14)) using high-precision calculation improves the ac-

curacy of the final estimation. A two-stage alternating optimization scheme was used to train the optimized angle and the estimated angle of the inter-error correction net iteratively to ensure a synergistic effect when optimizing the two related tasks.

(b) To leverage the two extrapolated angles, CMEAN is defined in Eq. (10), the error controlled loss function of inter-error correction net is defined as follows:

$$Loss = \underbrace{CMEAN(angle_{est} - angle_0)}_{\text{estimated angle}} + \lambda \underbrace{CMEAN(angle_{opt} - angle_0)}_{\text{optimized angle}} \tag{9}$$

Here $angle_0$ is the ground truth.

The optimization of loss function consists of two parts, the optimized angle optimization and the estimated angle optimization. The optimized angle optimization obtains a potential better estimation, which effectively reduces the average error. The estimated angle optimization offsets the main error parts of the two estimations each other, which further improves the accuracy. The loss function leverages the optimized angle and the estimated angle, which improves the accuracy of the final estimation.

While the regression loss works well for Euclidean distance in general real space, it is not robust in dealing with compact space like circular quantities in argument angle space due to the phase-wrapping property of angles. For instance, the Euclidean distance of the principal argument angle between 5° and 355° is 350° in real space, but the actual difference is only 10° in circular space. We will therefore have to modify the loss function for circular quantities to alleviate this. To do so, we can consider performing arithmetics in polar coordinate space and then converting the result back to angular space using trigonometry.

The angle α can be decomposed as polar coordinates on the unit circle:

$$\alpha \rightarrow (x, y),\ \ x = cos(\alpha), y = sin(\alpha)$$

Given PT, MT and TL angles $angle = (\alpha_{PT}, \alpha_{MT}, \alpha_{TL})$, we can then use their circular mean as:

$$CMEAN(\alpha_{PT}, \alpha_{MT}, \alpha_{TL}) = arctan\left(\frac{\bar{y}}{\bar{x}}\right) \tag{10}$$

where

$$\bar{x} = \frac{1}{3}(cos(\alpha_{PT}) + cos(\alpha_{MT}) + cos(\alpha_{TL}))$$

$$\bar{y} = \frac{1}{3}(sin(\alpha_{PT}) + sin(\alpha_{MT}) + sin(\alpha_{TL}))$$

2.3.1. Optimized angle

The optimized angle based on the normalization ($angle_{norm}$) of the error is used to quantify the relationship between the angle and its error. In this paper, we use $angle_{norm}$ to normalize the local error, which leverages the scoliosis level discrepancy and shows the local characteristic of the loss function to learn the descending direction.

The normalization of the local error has a good correlation with the angle error, and it reflects the function information in a small neighborhood effectively. In order to achieve these characteristics, we formulate the $angle_{norm}$ as:

$$angle_{norm} = \frac{angle_0 - angle^{(i)}}{\int_{|x-angle^{(i)}|\le r} f_{angle^{(i)}}(|x - angle_0|)dx + \epsilon} \tag{11}$$

Here, $r = 5°$ is the radius of the neighborhood, and ϵ is a small constant avoiding zero denominators.

Based on the $angle_{norm}$, we can construct a function of angle as:

$$norm^{(i)} = \frac{angle - angle^{(i)}}{\int_{|x-angle^{(i)}|\le r} f_{angle^{(i)}}(|x - angle|)dx + \epsilon} \tag{12}$$

and then calculate the optimized angle as:

$$angle_{opt} = argmin(\lambda_{opt} norm' + (1 - \lambda_{opt}) norm'') \tag{13}$$

Here $\lambda_{opt} = \frac{|angle'|}{|angle' + angle''|}$.

The optimization of $angle_{opt}$ corrects the descent direction, which accelerates the convergence.

2.3.2. Estimated angle

The estimated angle is based on the results of the combination function using high-precision calculation. The $angle_{norm}$ is used to calculate the parameter of the combination function so that the main error parts of $angle'$ and $angle''$ can be offset by each other.

Following the extrapolation format, The estimated angle can be written as follows:

$$angle_{est} = \frac{K_2 * angle' - K_1 * angle''}{K_2 - K_1} - \frac{K_2 * K_1 * (angle'' - angle')}{K_2 - K_1} \tag{14}$$

Here, $K_i = f_i'(0)/\eta(angle^{(i)}), i = 1, 2$ is related to $angle_{norm}$, where

$$\eta(angle^{(i)}) = \int_{|x-angle^{(i)}|\le r} t^{(i)}(|x - angle_0|)dx + \epsilon \tag{15}$$

$t^{(i)}$ is the tolerate function, which indicates the net's acceptance of the error. In this paper, we use complex heaviside step function. f_i is function of $angle_{norm} - angle$ plot.

The general outline of the iterative 2-stage training scheme is summarized in Algorithm 1.

Algorithm 1 Iterative Angle Training.

```
Set initial of angle sequence angle = (angle', angle'')
repeat
    calculate angle_norm using Eq. 11
    update angle_opt using Eq. 13
    calculate angle_est using Eq. 14
    update angle using angle^(i) = angle_opt, i = 1 or 2
until Convergence
```

As Fig. 6 shown, the optimized angle and the estimated angle are alternatively updated and gradually tends to the ground truth, which shows the convergence of the inter-error correction net.

2.4. Theorem proof of the improved accuracy of our MVC-Net

The following Theorem 1 explains how we eliminate the low order error to achieve higher accuracy:

Theorem 1 (Our proposed estimation ($angle_{est}$) has higher order accuracy than the initial estimations ($angle'$ and $angle''$)). *Suppose the net Net for subscript 1 and 2, the calculated angle $angle'$, $angle''$, the norm-angle error $angle'_{norm}, angle''_{norm}$. When $angle'$ and $angle''$ are greater than or less than $angle_0$ at the same time, the $angle_{est}$ defined in Eq. (14) is an approximation of $angle_0$ (ground truth). And it has $|k(angle^{(i)}, Net)^* angle_{norm}(angle^{(i)})|$ improvement for $i = 1, 2$.*

Proof. For the definition as above, we have

$$angle_0 = angle^{(i)} + k(angle^{(i)}) * angle_{norm}(angle^{(i)})$$

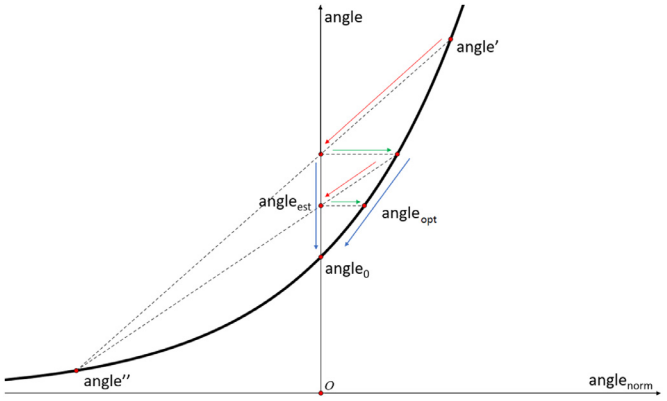

Fig. 6. The scheme of iteration in inter-error correction net. Using high-precision calculation, the optimization iteratively updates the norm-angle (red arrows) and angle itself (green arrows) to approach the ground truth, which shows the convergence of our method. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

By definition of $angle_{norm}$,

$$angle_{norm}(angle^{(i)}) = \lambda|angle_0 - angle^{(i)}| + \epsilon^{(i)}$$

Consider $\epsilon^{(i)}$ as 0, we can get

$$angle_0 = \frac{K_2 * angle' - K_1 * angle''}{K_2 - K_1} - \frac{K_2 * K_1 * (angle'' - angle')}{K_2 - K_1}$$

for $angle'$ and $angle''$ are greater than or less than $angle_0$ at the same time.

And we have

$$|angle_0 - angle^{(i)}| = |k(angle^{(i)}) * angle_{norm}(angle^{(i)})|$$

for $i = 1, 2$.

We have the similar conclusion for $angle_0$ is between $angle'$ and $angle''$.

Our inter-error correction net calculates the approximation directly while learning the K_i parameters since the approximation is a quasi-linear combination of $angle'$ and $angle''$. Therefore, we improve the accuracy of the estimation since we eliminate the low order error. □

3. Results and analysis

The MVE-Net has been validated on the spinal X-ray dataset with signs of scoliosis of varying extents. Extensive experiments show that our method with significant effectiveness, which can be practically used in clinical scoliosis analysis.

3.1. Data

Dataset Our dataset consists of 526 spinal X-ray images equally divided between AP and LAT views. These images were provided by local clinicians and all of them show signs of scoliosis to varying extent. Our dataset has covered different scoliosis levels, the range of the Cobb angle is distributed from 0 to 96.33°. The average resolution of the images is 0.26 mm/pixel. Our landmark ground truth consists of 60 landmarks per spinal images, which is the four corners of 15 thoracic and lumbar vertebrae. Each vertebrae was selected by an expert. We didn't select the Cervical vertebrae (vertebrae of the neck) since the cervical vertebrae are seldom involved in spinal deformity (Group, 2008). The ground truth of pivotal landmarks were determined by the ground truth of the landmarks. The Cobb angle ground truth was calculated by the ground truth of the pivotal landmarks.

Table 1
Mean error of the joint-view net, independent-view net, average (of the joint-view net and the independent-view net) and MVE-Net. Most of the time, the lowest error of our MVE-Net based on the improvement on the previous Nets indicates MVE-Net has higher accuracy.

Method	AP	
	CMAE(°)	SMAPE(%)
joint-view net	10.24	34.70
independent-view net	8.76	28.33
average	9.38	42.50
MVE-Net	**7.81**	**24.94**
Method	LAT	
	CMAE(°)	SMAPE(%)
joint-view net	16.00	40.17
independent-view net	13.15	29.98
average	14.56	45.75
MVE-Net	**6.26**	**11.90**

Data augmentation We use dynamic data augmentation to increase the robustness of our model during training. Our model has get better clinical applications since the common actual error such as Gaussian noise, rotation and shift have been resolved. In order to achieve this, we therefore augmented the data with:

(1) Randomly adding Gaussian noise directly to our images to simulate inherent noise.
(2) Randomly rotating the images up to 5° to allow for flexibility in rotation.
(3) Randomly shifting the images by 1% to encourage shift invariance.

Implementation details All networks are implemented in Keras using the Tensorflow backend. Training is implemented on four NVIDIA Tesla GPUs with a version of CUDA 8.0.

Performance metric For Cobb angle estimation, we used circular MAE (CMAE) and Symmetric Mean Absolute Error (SMAPE) to represent the relative error.

The circular MAE is defined as:

$$CMAE = \frac{1}{N}\sum_{i=1}^{N} CMEAN(M_{ang}(X^{\chi};\theta') - L^{\chi})$$

The SMAPE metric is defined as:

$$SMAPE = \frac{1}{N}\sum_{i=1}^{N} \frac{SUM(M_{ang}(X^{\chi};\theta') - L^{\chi})}{SUM(M_{ang}(X^{\chi};\theta') + L^{\chi})}$$

Here M_{ang} is our estimated Cobb angle and L is the ground truth.

3.2. Performance

As shown in Table 1, our model achieved a reputable CMAE of 7.81° in AP angle and 6.26° in LAT angle estimation and achieved a reputable SMAPE of 24.94% in AP angle and 11.90% in LAT angle estimation which proves our method is a useful clinical tool. Compared with only using joint-view net or independent-view net, our MVE-Net achieves higher accurate estimations. As Fig. 6 shown, this is contributed to our MVE-Net has successfully applied extrapolation and effectively offset the error of the joint-view net and the independent-view net each other. The joint-view net focuses on the joint feature between AP and LAT images. The independent-view net focuses on the pivotal landmarks. Out MVE-Net achieves to leverage the joint-view net and the independent-view net, and gives full play to their respective strengths, successfully improves the accuracy of the estimation. Fig. 7 shows the Cobb angle estimation has a high accuracy compared with the ground truth.

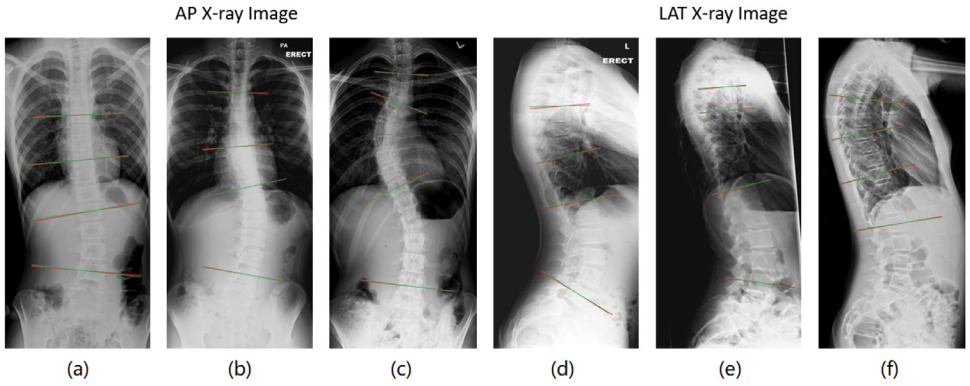

Fig. 7. The results of Cobb Angles estimation. The green lines are ground truth, and the red lines are our test performance. Our method overcomes huge variations and high ambiguities and achieves high accuracy in Cobb angles detection. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2
The comparison of the SMAPE in several methods. The manual-min method chooses the smaller average error value between joint-view net and independent-view net manually as a final result. Our MVE-Net has the lowest relative error indicates MVE-Net has the best performance.

Method	SMAPE(%)
S^2VR (Sun et al., 2017)	37.08
BoostNet (Wu et al., 2017)	41.35
MVC-Net (Wu et al., 2018)	35.85
manual-min	41.24
joint-view net	37.43
independent-view net	29.16
average	44.12
MVE-Net (ours)	**18.95**

Compared with other methods, our model has achieved the most accurate result. We compare with three other methods on the same dataset, i.e. S^2VR (Sun et al., 2017), BoostNet (Wu et al., 2017) and the MVC-Net (Wu et al., 2018) about the SMAPE, which shows the effectiveness of our method. The previous method didn't work well since it is a difficult problem to accurately calculate Cobb angles simultaneously on different scoliosis level. Our dataset has covered different scoliosis level (0–96.33°) comprehensively. Our MVE-Net has successfully handled this problem and effectively decreased the error. The extrapolation method leverages the difference among different scoliosis level, which constructs a more reliable network for Cobb angle estimation. Table 2 shows the results.

Fig. 8 shows the benefit of extrapolation. The extrapolation helps the inter-error correction net converge faster and achieves lower error since the extrapolation indicates the optimization direction and the step length. The leverage of the optimized angle and the estimated angle to a large extent correct the optimization direction, which speeds up the convergence. At the same time, it provides a more reliable reference value, which has determined more precise step length.

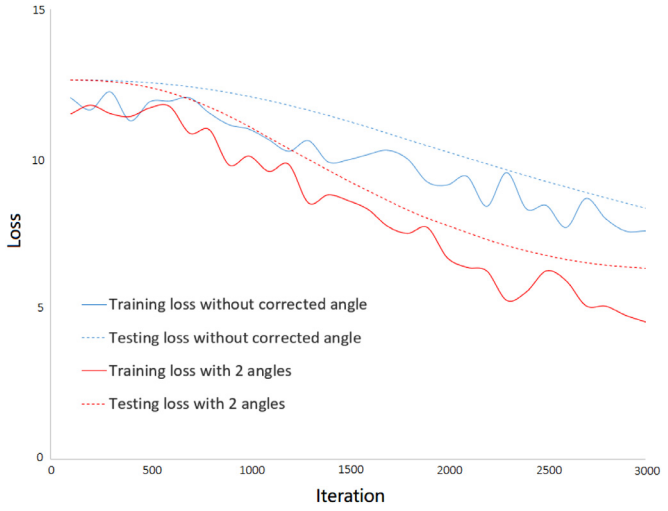

Fig. 8. Curves of estimation loss during the training and test procedures with (red lines) and without (blue lines) extrapolation calculated optimized angle. Extrapolation helps the network train faster and achieve lower estimation error for both training and test procedures. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.3. Analysis

The MVE-Net achieved the lowest CMAE of 7.81° in AP angle and 6.26° in LAT angle estimation on the unseen test set. This is due to the contributions of (1) the previous methods (joint-view net and independent-view net), which successfully learned joint and independent feature embeddings as is evident in the higher accuracy in images with noticeable variability and (2) the application of extrapolation, which faithfully increases the angles accuracy of joint-view net and independent-view net. The success of our method is further demonstrated by the more than 3 degrees on average of the circular mean absolute error, as well as the more rapid convergence compared to the conventional joint-view net and independent-view net DL model. Furthermore, our results demonstrate that the high performance can only be achieved by leveraging joint-view net and independent-view net using extrapolation iteratively, which indicates that reciprocal multi-task learning of both joint features and unique independent features is essential for achieving higher accuracy.

It is a difficult problem to accurately calculate Cobb angles simultaneously on different scoliosis level. In the proposed method, joint-view focuses on learning the common features (high-level features, lack of details) of two views, while each independent view alone contains only tiny unique information (low-level features, lack of semantic information) for each view. Therefore, we have an "inter-error correction net" for information fusion and better features in the final step of the proposed method. As a consequence, the combination of joint-view and independent-view nets can produce more accurate estimation than either net alone. It is the first time that the proposed MVE-Net has been used to solve this problem on such a wide distribution and get a more accurate estimation.

4. Conclusion

We proposed a automated Cobb angle estimation method for the scoliosis assessment using MVE-Net. The MVE-Net successfully leverages the joint features and independent features in multi-view X-ray images. Our MVE-Net achieved a high precision estimation of Cobb angles in both AP and LAT X-ray images in a large dataset of 526 X-ray images with different scoliosis level. Compared with other automated methods, extensive experiments show the high accuracy of our MVE-Net. Our accurate calculation method also has effective extensibility in other clinical applications for high precision estimations.

Declaration of Competing Interest

None.

Appendix A. Generalization to n-input situation

Our method can be generalized to not only two inputs ($angle'$ and $angle''$) but also n-inputs situation. If we have more inputs ($angle^{(i)},\ i = 1, 2, \ldots$), the inter-error correction net still can leverage these input information to get more accurate results. We formalize a more accurate approximation of $angle_0$ (ground truth) as follow:

$$angle_0 = angle^{(i)} + k(angle^{(i)}) * angle_{norm}(angle^{(i)}) + \mu(angle^{(i)}) * angle^2_{norm}(angle^{(i)}) + \epsilon^{(i)}$$

as we approximate the $\epsilon^{(i)}$ accurately.

Without loss of generality, we use three inputs ($angle'$, $angle''$ and $angle'''$) as example, we have the following corollary:

Corollary 1 (Inter-error correction net for n-inputs estimation (n=3 as example))**.** *Suppose we have the net Net for subscript 1, 2 and 3, the calculated angle $angle'$, $angle''$ and $angle'''$, the norm-angle error $angle'_{norm}$, $angle''_{norm}$ and $angle'''_{norm}$, and the parameter k defined the same as above. Assimilating the scale factor λ and μ to k, we define $K = k * \lambda$ and $M = k^2 * \mu$. We can find an approximation of $angle_0$ by the following equations*

$$angle_0 = angle' + k_1(angle') * angle'_{norm}(angle') + \mu_1(angle') * angle'^2_{norm}(angle')$$

$$angle_0 = angle'' + k_2(angle') * angle''_{norm}(angle') + \mu_2(angle') * angle''^2_{norm}(angle')$$

$$angle_0 = angle''' + k_3(angle') * angle'''_{norm}(angle') + \mu_3(angle') * angle'''^2_{norm}(angle')$$

the root $angle_{opt}$ is an approximation of $angle_0$

This equation is quadratic, but we can always eliminate the quadratic terms from two adjacent formulas to get a linear equation. Therefore the $angle_{opt}$ is still the linear combination of $angle'$, $angle''$ and $angle'''$. We can learn the linear coefficient with our method.

Due to the drawer principle, at least two of α are greater than or less than $angle_0$ at the same time. We may assume they are $angle''$ and $angle'''$. We can rewrite the equation as

$$\begin{aligned}0 = {} & angle'' - angle' \\ & + k_2(angle^{(i)}) * angle''_{norm}(angle^{(i)}) \\ & - k_1(angle^{(i)}) * angle'_{norm}(angle^{(i)}) \\ & + \mu_2(angle^{(i)}) * angle''^2_{norm}(angle^{(i)}) \\ & - \mu_1(angle^{(i)}) * angle'^2_{norm}(angle^{(i)})\end{aligned}$$

$$\begin{aligned}0 = {} & angle''' - angle' \\ & + k_3(angle^{(i)}) * angle'''_{norm}(angle^{(i)}) \\ & - k_1(angle^{(i)}) * angle'_{norm}(angle^{(i)}) \\ & + \mu_3(angle^{(i)}) * angle'''^2_{norm}(angle^{(i)}) \\ & - \mu_1(angle^{(i)}) * angle'^2_{norm}(angle^{(i)})\end{aligned}$$

we have

$$\begin{aligned}0 = {} & angle'' - angle' \\ & + K_2(angle^{(i)}) * |angle'' - angle_0| \\ & - K_1(angle^{(i)}) * |angle' - angle_0| \\ & + M_2(angle^{(i)}) * |angle'' - angle_0|^2 \\ & - M_1(angle^{(i)}) * |angle' - angle_0|^2\end{aligned}$$

$$\begin{aligned}0 = {} & angle''' - angle' \\ & + K_3(angle^{(i)}) * |angle''' - angle_0| \\ & - K_1(angle^{(i)}) * |angle' - angle_0| \\ & + M_3(angle^{(i)}) * |angle''' - angle_0|^2 \\ & - M_1(angle^{(i)}) * |angle' - angle_0|^2\end{aligned}$$

When these $angle^{(i)}$ are sufficiently close to $angle_0$, we have

$$\begin{aligned}0 = {} & angle'' - angle' \\ & + K * |angle'' - angle'| \\ & + M * |angle'' - angle'| \\ & * |angle'' + angle' - 2 * angle_0|\end{aligned}$$

$$\begin{aligned}0 = {} & angle''' - angle' \\ & + K * |angle''' - angle'| \\ & + M * |angle''' - angle'| \\ & * |angle''' + angle' - 2 * angle_0|\end{aligned}$$

or

$$\begin{aligned}0 = {} & angle'' - angle' \\ & + K * |angle'' + angle' - 2 * angle_0| \\ & + M * |angle'' - angle'| \\ & * |angle'' + angle' - 2 * angle_0|\end{aligned}$$

$$\begin{aligned}0 = {} & angle''' - angle' \\ & + K * |angle''' + angle' - 2 * angle_0| \\ & + M * |angle''' - angle'| \\ & * |angle''' + angle' - 2 * angle_0|\end{aligned}$$

Here

$$K = K_1(angle^{(i)}) = K_2(angle^{(i)}) = K_3(angle^{(i)})$$
$$M = M_1(angle^{(i)}) = M_2(angle^{(i)}) = M_3(angle^{(i)})$$

Since

$$angle_0 = angle^{(i)} + K|angle^{(i)} - angle_0| + M|angle^{(i)} - angle_0|^2$$

The expression of $angle_0$ will be simple.

Compared with the method in the theorem, the corollary method has higher accuracy. Since we use

$$angle_0 = angle^{(i)} + k(angle^{(i)}) * angle_{norm}(angle^{(i)}) + \mu(angle^{(i)}) * angle^2_{norm}(angle^{(i)})$$

instead of

$$angle_0 = angle^{(i)} + k(angle^{(i)}) * angle_{norm}(angle^{(i)})$$

with higher accuracy of $angle_0$

Besides using parabola, we can also use arc or some more complicated curves like higher degree polynomials curve or transcendental curves like

$$\begin{aligned} angle_0 = & \ angle^{(i)} + k(angle^{(i)}) * angle_{norm}(angle^{(i)}) \\ & + \mu(angle^{(i)}) * angle^2_{norm}(angle^{(i)}) \\ & + \nu(angle^{(i)}) * angle^3_{norm}(angle^{(i)}) \end{aligned}$$

or

$$\begin{aligned} angle_0 = & \ angle^{(i)} + k(angle^{(i)}) * angle_{norm}(angle^{(i)}) \\ & + \mu(angle^{(i)}) * angle^2_{norm}(angle^{(i)}) \\ & + \nu(angle^{(i)}) * \sqrt{angle_{norm}(angle^{(i)})} \end{aligned}$$

However, in practice, it is enough to use linear or quadratic functions. This is because as the power of the function increases, our method only adds to the precision of the remainder. At the same time, there is a systematic error in the measurement of Cobb's angle itself, and it does not make sense to pursue a single precision boost when our accuracy improvement has been less than this systematic error.

Therefore, our alternative error correction Net calculates the accurate approximation directly while learning the $K(angle^{(i)})$ parameters since the approximation is a linear combination of $angle'$ and $angle''$.

References

Anitha, H., Karunakar, A., Dinesh, K., 2014. Automatic extraction of vertebral endplates from scoliotic radiographs using customized filter. Biomed. Eng. Lett. 4, 158–165.

Anitha, H., Prabhu, G., 2012. Automatic quantification of spinal curvature inscoliotic radiograph using image processing. J. Med. Syst. 36, 1943–1951.

Asher, M.A., Burton, D.C., 2006. Adolescent idiopathic scoliosis: natural history and long term treatment effects. Scoliosis 1, 2.

Cobb, J., 1948. Outline for the study of scoliosis. Instr. Course Lect. 5, 261–275.

Criminisi, A., Shotton, J., Robertson, D., Konukoglu, E., 2011. Regression Forests for Efficient Anatomy Detection and Localization in CT Studies. Springer Berlin Heidelberg, Berlin, Heidelberg. 106–117.

Group, S.D.S., 2008. Radiographic measurement manual. Medtronic Sofamor Danek USA.

He, K., Zhang, X., Ren, S., Sun, J., 2015. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. CoRR abs/1502.01852. arXiv:1502.01852.

Hinton, G. E., Srivastava, N., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., 2012. Improving neural networks by preventing co-adaptation of feature detectors. CoRR abs/1207.0580. arXiv:1207.0580.

Laporta, S., 2000. High-precision calculation of multiloop Feynman integrals by difference equations. Int. J. Modern Phys. A 15 (32), 5087–5159.

McCloskey, M., Cohen, N.J., 1989. Catastrophic interference in connectionist networks: the sequential learning problem. Psychol. Learn. Motiv. 24, 109–165.

Pruijs, J.E.H., Hageman, M.A.P.E., Keessen, W., van der Meer, R., van Wieringen, J.C., 1994. Variation in Cobb angle measurements in scoliosis. Skelet. Radiol. 23, 517–520. doi:10.1007/BF00223081.

Quarteroni, A., Sacco, R., Saleri, F., 2000. Numerical Mathematics. New York, Springer-Verlag.

Sardjono, T.A., Wilkinson, M.H., Veldhuizen, A.G., van Ooijen, P.M., Purnama, K.E., Verkerke, G.J., 2013. Automatic cobb angle determination from radiographic images. Spine 38, 1256–1262.

Sun, H., Zhen, X., Bailey, C., Rasoulinejad, P., Yin, Y., Li, S., 2017. Direct estimation of spinal cobb angles by structured multi-output regression. In: International Conference on Information Processing in Medical Imaging, pp. 529–540.

Vrtovec, T., Pernus, F., Likar, B., 2009. A review of methods for quantitative evaluation of spinal curvature. Eur. Spine J. 18, 593–607.

Weinstein, S.L., Dolan, L.A., Cheng, J.C., Danielsson, A., Morcuende, J.A., 2008. Adolescent idiopathic scoliosis. Lancet 371, 1527–1537.

Wu, H., Bailey, C., Rasoulinejad, P., Li, S., 2017. Automatic landmark estimation for adolescent idiopathic scoliosis assessment using boostnet. In: Medical Image Computing and Computer Assisted Interventions, pp. 127–135.

Wu, H., Bailey, C., Rasoulinejad, P., Li, S., 2018. Automated comprehensive adolescent idiopathic scoliosis assessment using MVC-net. Med. Image Anal. 1–11.

Xue, W., Islam, A., Bhaduri, M., Li, S., 2017. Direct multitype cardiac indices estimation via joint representation and regression learning. IEEE Trans. Med. Imaging.

Xue, W., Nachum, I.B., Pandey, S., Warrington, J., Leung, S., Li, S., 2017. Direct estimation of regional wall thicknesses via residual recurrent neural network. The 25th Biennial International Conference on Information Processing in Medical Imaging (IPMI 2017).

Zhang, J., Lou, E., Hill, D.L., Raso, J.V., Wang, Y., Le, L.H., Shi, X., 2010. Computer-aided assessment of scoliosis on posteroanterior radiographs. Med. Biol. Eng. Comput. 48, 185–195.

Zhang, J., Lou, E., Le, L.H., Hill, D.L., Raso, J.V., Wang, Y., 2009. Automatic cobb measurement of scoliosis based on fuzzy hough transform with vertebral shape prior. J. Digit. Imaging 22, 463–472.

Zhen, X., Islam, A., Bhaduri, M., Chan, I., Li, S., 2015. Direct and simultaneous four-chamber volume estimation by multi-output regression. In: MICCAI. Springer, pp. 669–676.