

3D Spine Shape Estimation from Single 2D DXA

Emmanuelle Bourigault, Amir Jamaludin, and Andrew Zisserman

Visual Geometry Group, Department of Engineering Science,
University of Oxford, Oxford, UK
`emmanuelle@robots.ox.ac.uk`

Abstract. Scoliosis is currently assessed solely on 2D lateral deviations, but recent studies have also revealed the importance of other imaging planes in understanding the deformation of the spine. Consequently, extracting the spinal geometry in 3D would help quantify these spinal deformations and aid diagnosis. In this study, we propose an automated general framework to estimate the 3D spine shape from 2D DXA scans. We achieve this by explicitly predicting the sagittal view of the spine from the DXA scan. Using these two orthogonal projections of the spine (coronal in DXA, and sagittal from the prediction), we are able to describe the 3D shape of the spine. The prediction is learnt from over 30k paired images of DXA and MRI scans. We assess the performance of the method on a held out test set, and achieve high accuracy¹.

Keywords: 2D-3D Symbiosis · Scoliosis · MRI · DXA

1 Introduction

The standard procedure to examine the spine for the presence of scoliosis is using antero-posterior (AP) X-rays and measuring the angle between the most tilted vertebrae [7]. Scoliosis typically affects growing children and proper diagnosis of scoliosis requires multiple follow-up scans. As such, Dual-energy X-ray (DXA) scans, with its lower radiation dose than X-rays is quickly becoming an acceptable alternative [12, 13, 18, 23]. However, this still does not solve the fundamental issue of scoliosis diagnosis; it is essentially a 3D disorder but the focus is on the 2D lateral shift of the spine. Imaging this disorder on a Magnetic Resonance Imaging (MRI) is a good alternative, and multiple studies have shown that there is indeed more useful information that can be extracted from MRIs for scoliosis [4, 8, 11]. MRI however is more expensive and requires more time for one single scan compared to X-ray/DXA.

In this paper, our objective is to obtain the 3D patient-specific spine shape from a 2D DXA. We explore whether this is even possible given that 3D information is “lost” in the projection of a 2D DXA scan. We show that it is possible to learn to infer this “lost” information by leveraging a paired imaging set of DXA scans with corresponding MRIs taken at roughly the same time. The MRI

¹ Our code is available at <https://www.robots.ox.ac.uk/~vgg/research/dxa-to-3d/index.html>

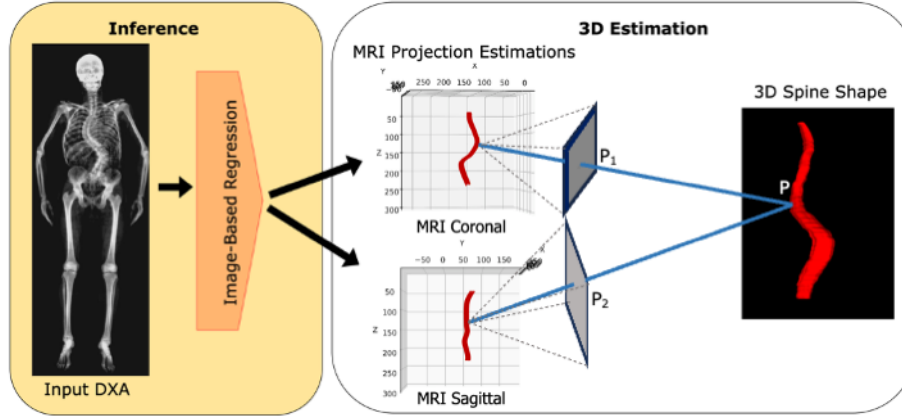


Fig. 1. Inference. Given a DXA scan, the model predicts the coronal and sagittal projections of the 3D spine. Once these two orthogonal views of the spine are obtained, the 3D spine can be reconstructed. A detailed explanation of the procedure is given in the Supplementary Material.

provides the 3D geometry of the spine, and a network can be trained to map the 3D spine from the 2D DXA. Through these two orthogonal curves, coronal and sagittal, we can recover the 3D geometry of the spine.

1.1 Related Work

To date, the vast majority of scoliosis research has focused on 2D shape analysis of the spine or part of the 3D spine [5, 12, 13, 14, 18, 23, 24]. Limitations of 2D spine analysis arise particularly in classifying 3D curve shape. Indeed, deviations are not limited to the coronal plane, they include twisting of the spine in multiple directions [20], and the importance of other planes, e.g. for axial rotations, is well recognized [10]. Recently, there has been a growing interest in the community on EOS imaging with simultaneous acquisition of coronal and sagittal views producing 3D spine shape [19]. This enables, for example, a better evaluation of the effect of brace surgical treatment [2, 9].

There are several methods similar to ours; the most similar work is by [17] which works on DXA scans and predicts the 3D model of the spine using statistical shape models (SSM). Other works use biplanar X-rays either with SSM [1, 3, 6] or contour matching [26]. In our work, we directly estimate the 3D spine from a single 2D DXA. We leveraged the paired MRI attached to each DXA scan in our dataset to learn the 3D estimate. Our 3D regression involves only a single pass through a feed-forward network.

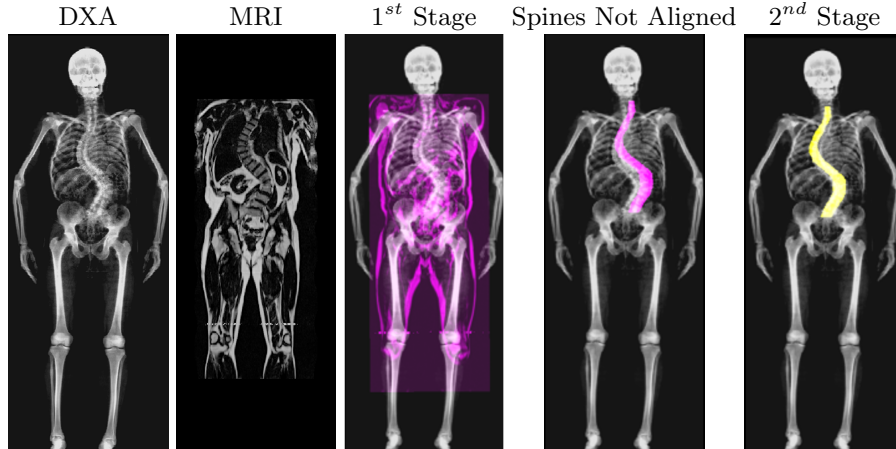


Fig. 2. DXA to MRI two-Stage Alignment. The two scans are iteratively aligned using a three parameter planar transformation. From left to right: DXA scan; original coronal projection of MRI scan (not-aligned to the DXA); overlay of MRI aligned to DXA after the image-level alignment first stage; overlay of segmented spines after the first stage; overlay of spine segmentation after the spine-level alignment second stage.

2 3D From 2D

Our method for estimating 3D spine geometry from a single 2D DXA is simple. We essentially learn to regress 2D curves; first by directly regressing the curve of the spine extracted from the 2D DXA itself (this is a coronal projection of the 3D spine), and then by predicting the sagittal projection of the 3D curve extracted from a paired MRI. This is only feasible through the use of a large-scale public dataset, consisting of *paired* whole-body DXA and MRI of the same subjects [22]. The challenges of this problem and the solutions we propose are discussed separately in their respective sections: (i) image-level alignment of modalities, and spine-curve alignment in Section 2.2, (ii) how to directly regress 2D curves from the DXA scan in Section 2.3.

2.1 Problem definition

The problem consists of regressing 2D spine curves obtained from the projections of 3D MRI to 2D from a DXA image using a single model (see Figure 3). This can be separated into two separate regression problems, namely: (i) the regression of the coronal or AP curve, and (ii) the regression of the sagittal or lateral curve. We define three sets of points $P_i = \{x \in \mathbb{R} \mid 0 \leq x \leq 209\}$, $i = a, b, c$ for the singular central and two lateral curves of the spine respectively. These define the segmentation and mid-curve of the spine in the 2D projections. Our objective is to regress these curves (in the coronal and sagittal planes) from the DXA scan.

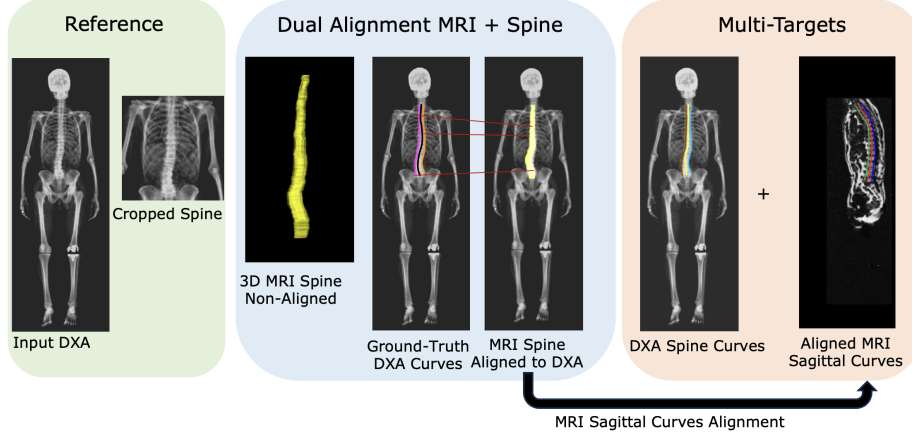


Fig. 3. Model Learning. The cropped DXA is the input to the regression network. The regression targets are the DXA curve, and the sagittal curve (projected from the 3D MRI spine). The alignment for the sagittal curve to DXA is obtained from the alignment of the coronal projection of the 3D MRI to the DXA. Six curves are regressed: the centerline of the spine as well as the left and right boundaries of the segmentation, for both the coronal and sagittal views.

2.2 Modality Alignment

The paired images, the 2D DXA scans and the 3D MRIs, do not come registered. Hence, the first step we take is to align these two modalities. As we are interested in inferring the 3D information from 2D images; we register the 3D MRI and the extracted spine curve to the 2D DXA. The alignment proceeds iteratively in two stages: (i) a rough image-to-image alignment of the two imaging modalities followed by, (ii) a finer alignment of the extracted segmentation/curve of the spine from the MRI to the segmentation/curve of the spine from the DXA. See Figure 2.

For the rough alignment, we use a pipeline proposed by [25] which finds the best 2D transformation to align the 3D MRI (via its coronal projection) to the 2D DXA. The 2D transformations involve three parameters (a rotation angle θ , and two translations). To compute the transformation, 10 rotation angles are sampled in the range $[-2, 2]$ degrees, and the translation is obtained by convolving CNN spatial-feature maps of the DXA and MRI, and selecting the point of maximum response.

For the second stage, we align the spine curves between the two modalities. Again, the 2D transformation consists of a rotation and translation. We use keypoint matching sampled along the spine contour of the DXA, and compute the transformation that minimises the mean squared error (MSE).

It is possible that the person changed position too much between the DXA and MRI scans, and it is not possible to align the spines due to deformations. To check for this we measure the overlap of the spine segmented in the DXA scan

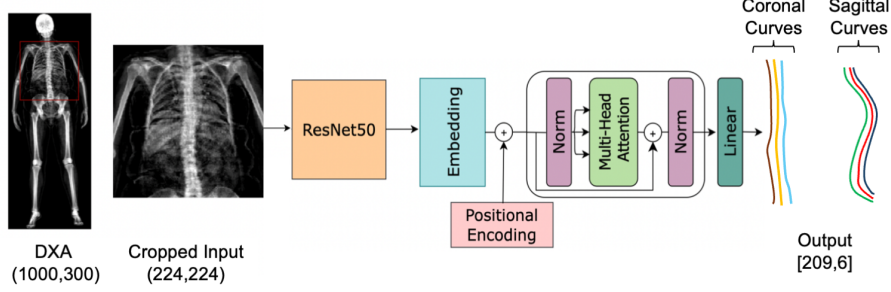


Fig. 4. Image-Based Regression of Coronal and Sagittal Spine Curves. We use a ResNet50, pre-trained on Imagenet-21k, with a transformer layer to regress the spine curves $(x_{(1,2,3)}(z), y_{(1,2,3)}(z)), z \in [0, 209]$ for left, center and right curves.

with the projection of the 3D spine segmented in the MRI (see Figure 2). We apply a threshold for filtering out the badly aligned scans: if the Intersection-over-Union (IoU) of spine masks is below 70%, then they are discarded. A total of 14,065 paired DXA-MRIs are rejected by this test which is 28.8% of the data.

2.3 Learning the Regressor

Encoder. Our regressor consists of a multi-scale feature extractor coupled with a transformer layer to learn the long-range dependencies more effectively (see Figure 4). The initial part of the architecture consists of an image feature extractor with ResNet50 pre-trained on ImageNet-21k. The feature map of the penultimate convolutional block is extracted as $7 \times 7 \times 2048$ and fed into a standard transformer layer adding the positional encoding, which is then average pooled, and flattened to a 2048-dimensional feature. The regression head consists of linear layer to predict the output curve points from the transformer output vector.

Loss. Our loss is the L_1 difference between the target and predicted spine curve points: $L = \sum_{i=1}^n |x^i - \hat{x}^i|$, where n is the number of samples, x^i are the ground-truth spine points, and \hat{x}^i the predicted spine points at a given z^i .

We regress three curves for each view (coronal (x) and sagittal (y)): the central curve of the spine as well as the left and right curves bounding the segmentation. The central and lateral curve points are $(x_1, x_2, x_3, y_1, y_2, y_3)$, where 1 and 3 are either the right/left or anterior/posterior curves depending on the plane projection, and 2 is the central curve of the spine.

3 Dataset & Implementation Details

The UKBiobank is a publicly available dataset of 48,384 full-body Dixon MRI paired with DXA scans [22]. The UK Biobank MRIs are resampled to be isotropic

and cropped to a consistent resolution ($700 \times 224 \times 224$). The dataset is split into 80:10:10 for training (27,816), validation (3,477), and testing (3,477) after filtering the non-aligned scans by the dual alignment procedure explained in section 2.2. Our training-validation-test is balanced in scoliosis cases, each containing 20% of scoliosis cases. Our model takes as input a cropped DXA image of the whole spine. This is achieved using a 224×224 cropping window of the spine given spine segmentation from [5].

Obtaining the 2D spine from the DXA images. We obtain the spine centroids and 2D spine mask segmentation in DXA which involves spine segmentation using pseudo-labelling in an active learning framework [5, 13]. This is used as the target for the coronal view regressor.

Obtaining the 3D spine from the MRIs. We used a segmentation network to obtain the 3D whole spine curve in the MRI (3D centroids and segmentation) [4] trained on adjacent axial slices ($n-1, n, n+1$) to limit the loss of depth information. These labels are used for training our sagittal regression model.

3.1 Implementation Details

We use the Bottleneck Transformer from [21] with a ResNet-50 encoder for our regressor; it regresses points for different projections of the 3D spine. For curve regression, all DXA inputs and target spine curves are normalised to fixed height for spine ranging from 0 to 209 pixels. We train our model for 500 epochs. We use five-fold cross-validation, where we repeat validation on 5 stratified folds. The batch size is set to 16, optimizer is Adam [15] with $\beta = (0.9, 0.999)$, and the learning rate is initially set to $1e^{-4}$ with decay every 200 epochs. We used one 32GB Tesla V100 GPU. To reduce overfitting, we employ two techniques, using dropout with a probability of $p = 0.3$ and we employ a regularizer to the L1 loss. We also use different augmentation techniques with cropping, image contrast and random Gaussian noise in training.

4 Results

4.1 Evaluation Metrics

We measure the performance of our model using the mean absolute error $MAE = \frac{1}{n} \sum_{z=1}^n |y_z - \hat{y}_z|$, and the mean of the relative error $RE = \frac{1}{n} \sum_{z=1}^n \frac{|y_z - \hat{y}_z|}{\hat{y}_z}$ between the predicted points \hat{y}_z and the ground truth $y_z, z \in [0, 209]$. The mean of RE is always between 0 and 1, the lower the relative error the better. To assess the accuracy of whole spine predicted masks obtained from the area between the lateral curves compared to project ground-truth masks obtained from segmentation of the MRI, we use the 2D Intersection-over-Union, IoU .

4.2 Spine Curves Estimation and Robustness

The DXA and Coronal MRI spine curves are predicted at sub-pixel level precision with better precision and lower variance for the combined ResNet50+Transformer

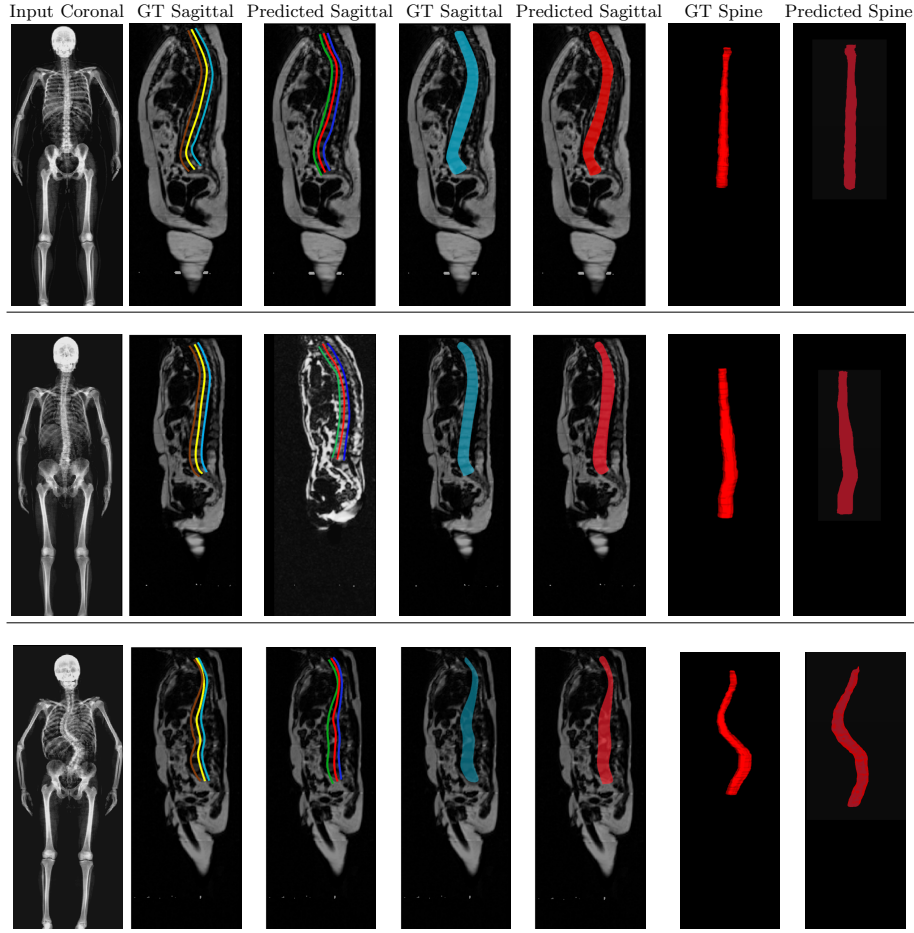


Fig. 5. Qualitative Results of DXA to MRI Sagittal Curve Generation on the Test Set. We show from left to right, the input DXA, the ground-truth sagittal curves, predicted sagittal curves, ground-truth spine mask, predicted spine mask overlayed on MRI sagittal slice ($n=112$), the ground-truth 3D spine, and its prediction using Gaussian rendering. The severity of the scoliosis increases from top to bottom.

framework ($MAE = 0.66 \pm 0.21$, $RE = 0.084 \pm 0.02$, $IoU = 89.6$) compared to ResNet50 or ViT (see Table 1). We show that using residual blocks with a transformer layer, our model can reliably estimate not only the coronal MRI projection from a single DXA but any spine curve projections along a 360° rotation about the vertical axis such as the sagittal projections ($MAE = 1.65 \pm 0.6$, $RE = 0.21 \pm 0.08$, $IoU = 86.8$), see Table 1). Our model is also able to capture the pattern of curves with straighter sagittal curves for more severe cases of scoliosis (see Figure 5).

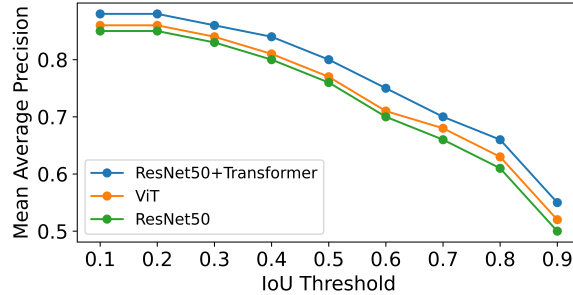


Fig. 6. Mean Average Precision of Predicted Spine Masks for Different IoU Thresholds. IoU thresholds start from 0.1 to 0.9 following COCO dataset [16], with steps of 0.1. We show that our model *ResNet50+Transformer* surpasses baselines for 3D spine shape estimation.

3D evaluation. We also evaluate the performance of our model to reconstruct 3D spines from 2D DXAs. We effectively have a segmentation of the spine bounded by the lateral curves from our pipeline. We use an untouched set of 150 MRIs from the UKBiobank manually annotated spines from [4] to measure the IoU. In Figure 6, we show the segmentation performance of predicted spines in the orthogonal views. We also measure the average deviation of spine curve point-wise along the spine as a metric for 3D error. Our best model *ResNet50 + Transformer* achieves a 3D average spine curve deviation of 1.42 voxels or 3.12mm and a 3D spine mask IoU of $83.8 \pm 1.1px$. See Figure 6 for visualising the mean average precision at different IoU thresholds.

Discussion. One probable reason that the model is able to infer a sagittal view from coronal DXA is that it has access to the entire (cropped) DXA scan, and there are at least two cues it can use: (i) the different intensities of the imaged bones of the spine give information about their angle and depth; and (ii) the position of the rib cage in the image depends on the spine, and so indirectly the 2D layout of the ribs in the image gives information about the spine.

5 Conclusion

Our model is able to give a patient-specific representation of the spine in 3D from a single DXA scan. We show that our method is effective in capturing the intricacies of the 3D spine. As such, this work has the potential to assist in the diagnosis and screening of scoliosis and other spinal disorders. Our primary focus for future work includes investigating confidence prediction for the sagittal curve.

Configuration	Model	Spine Curves Absolute Error			Spine Curves Relative Error			Spine Mask
		Mean	Median	SD	Mean	Median	SD	IoU
DXA	ResNet50	0.71	0.70	± 0.26	0.08	0.076	± 0.03	91.4
	ViT	0.69	0.67	± 0.24	0.073	0.071	± 0.02	91.9
	ResNet50 + Transformer	0.58	0.57	± 0.18	0.062	0.057	± 0.01	92.3
Coronal MRI	ResNet50	0.81	0.79	± 0.3	0.11	0.094	± 0.05	88.3
	ViT	0.78	0.78	± 0.28	0.091	0.083	± 0.03	88.9
	ResNet50 + Transformer	0.66	0.64	± 0.21	0.084	0.079	± 0.02	89.6
Sagittal MRI	ResNet50	3.63	3.29	± 1.2	0.41	0.39	± 0.14	83.6
	ViT	2.99	2.73	± 1.1	0.38	0.36	± 0.11	84.1
	ResNet50 + Transformer	1.65	1.58	± 0.6	0.26	0.21	± 0.08	86.8

Table 1. Performance of Image-based Models for Spine Curves Regression.

For a given input DXA, we predict curves from the DXA target 1st block or curves from the corresponding MRIs 2nd and 3rd blocks. For each target, we predict 3 curves (center + 2 laterals) for each block. The 2nd and 3rd blocks is a single multi-view model which outputs coronal and sagittal curves. The 1st column specifies the model configuration with output curve modality (DXA for baseline, coronal and sagittal MRI), the 2nd column the different models, the 3rd column and the 4th column for absolute error and relative error respectively in terms of mean, median and standard deviation of spine curves predicted and reference curves in pixels. The far right column shows the 2D IoU of the masks bound by the lateral spine curves.

Acknowledgments. This work was supported by the SABS: R³ CDT, University of Oxford (EP/S024093/1), by the EPSRC Programme Grant Visual AI (EP/T025872/1), and the Novartis-BDI Collaboration. This work has been conducted using the UK Biobank resource (application number 17295).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

Bibliography

- [1] Aubert, B., Vazquez, C., Cresson, T., Parent, S., de Guise, J.A.: Toward automated 3d spine reconstruction from biplanar radiographs using cnn for statistical spine model fitting. *TMI* **38**(12), 2796–2806 (2019)
- [2] Aurélien Courvoisier, R.V., Skalli, W.: Eos 3d imaging: assessing the impact of brace treatment in adolescent idiopathic scoliosis. *Expert Review of Medical Devices* **11**(1), 1–3 (2014)
- [3] Benameur, S., Mignotte, M., Parent, S., Labelle, H., Skalli, W., de Guise, J.: 3d/2d registration and segmentation of scoliotic vertebrae using statistical models. *Computerized Medical Imaging and Graphics* **27**(5), 321–337 (2003)
- [4] Bourigault, E., Jamaludin, A., Clark, E., Fairbank, J., Kadir, T., Zisserman, A.: 3d shape analysis of scoliosis. In: *MICCAI Workshop on Shape in Medical Imaging* (2023)
- [5] Bourigault, E., Jamaludin, A., Kadir, T., Zisserman, A.: Scoliosis measurement on DXA scans using a combined deep learning and spinal geometry approach. In: *MIDL* (2022)
- [6] Clogenson, M., Duff, J.M., Luethi, M., Levivier, M., Meuli, R., Baur, C., Henein, S.: A statistical shape model of the human second cervical vertebra. *IJCARS* **10**, 1097–1107 (2015)
- [7] Cobb, J.: Outline for the study of scoliosis. *Instr Course Lect AAOS* **5**, 261–275 (1948)
- [8] Donzelli, S., Poma, S., Balzarini, L., Borboni, A., Respizzi, S., Villafañe, J., Zaina, F., Negrini, S.: State of the art of current 3-d scoliosis classifications: a systematic review from a clinical perspective. *JNER* **12**, 91 (10 2015)
- [9] Dubousset, J., Ilharreborde, B., Huec, J.C.L.: Use of eos imaging for the assessment of scoliosis deformities: application to postoperative 3d quantitative analysis of the trunk. *European Spine Journal* **23**, 397–405 (2014)
- [10] Illés, T.S., Lavaste, F., Dubousset, J.: The third dimension of scoliosis: The forgotten axial plane. *OTSR* **105** **2**, 351–359 (2019)
- [11] Illés, T.S., Tunyogi-Csapó, M., Somoskeőy, S.: Breakthrough in three-dimensional scoliosis diagnosis: significance of horizontal plane view and vertebra vectors. *European Spine Journal* **20**, 135–143 (2010)
- [12] Jamaludin, A., Kadir, T., Clark, E., Zisserman, A.: Predicting spine geometry and scoliosis from dxa scans (2023), <https://arxiv.org/abs/2311.09424>
- [13] Jamaludin, A., Kadir, T., Clark, E.M., Zisserman, A.: Predicting scoliosis in dxa scans using intermediate representations. In: *CSI@MICCAI* (2018)
- [14] Jamaludin, A., Lootus, M., Kadir, T., Zisserman, A.: Automatic intervertebral discs localization and segmentation: A vertebral approach. In: *CSI@MICCAI* (2016)
- [15] Kingma, D., Ba, J.: Adam: A method for stochastic optimization. In: *ICLR* (2015)

- [16] Lin, T., Maire, M., Belongie, S.J., Bourdev, L.D., Girshick, R.B., Hays, J., Perona, P., Ramanan, D., Doll'ar, P., Zitnick, C.L.: Microsoft COCO: common objects in context. CoRR **abs/1405.0312** (2014), <http://arxiv.org/abs/1405.0312>
- [17] López Picazo, M., Magallón Baro, A., Del Río Barquero, L.M., Di Gregorio, S., Martelli, Y., Romera, J., Steghöfer, M., González Ballester, M.A., Humbert, L.: 3-d subject-specific shape and density estimation of the lumbar spine from a single anteroposterior dxa image including assessment of cortical and trabecular bone. TMI **37**(12), 2651–2662 (2018)
- [18] Ng, P.T.T., Straker, L., Tucker, K., Izatt, M., Claus, A.: Advancing use of dxa scans to quantitatively and qualitatively evaluate lateral spinal curves, for preliminary identification of adolescent idiopathic scoliosis. CTI **112** (03 2023)
- [19] Rehm, J., Germann, T., Akbar, M., Pepke, W., Kauczor, H.U., Weber, M.A., Spira, D.: 3d-modeling of the spine using eos imaging system: Inter-reader reproducibility and reliability. PLoS ONE **12** (2017)
- [20] Rockenfeller, R., Müller, A.: Augmenting the cobb angle: Three-dimensional analysis of whole spine shapes using bézier curves. Computer Methods and Programs in Biomedicine **225**, 107075 (08 2022)
- [21] Srinivas, A., Lin, T.Y., Parmar, N., Shlens, J., Abbeel, P., Vaswani, A.: Bottleneck transformers for visual recognition. CVPR (2021)
- [22] Sudlow, C.L.M., Gallacher, J.E., Allen, N.E., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P., Green, J., Landray, M.J., Liu, B.C., Matthews, P.M., Ong, G., Pell, J.P., Silman, A.J., Young, A., Sprosen, T., Peakman, T.C., Collins, R.: Uk biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Medicine **12** (2015)
- [23] Taylor, H.J., Harding, I., Hutchinson, J., Nelson, I., Blom, A., Tobias, J.H., Clark, E.M.: Identifying scoliosis in population-based cohorts: development and validation of a novel method based on total-body dual-energy x-ray absorptiometric scans. Calcified tissue international **92**, 539–547 (2013)
- [24] Windsor, R., Jamaludin, A., Kadir, T., Zisserman, A.: A convolutional approach to vertebrae detection and labelling in whole spine MRI (2020)
- [25] Windsor, R., Jamaludin, A., Kadir, T., Zisserman, A.: Self-supervised multi-modal alignment for whole body medical imaging. In: MICCAI (2021)
- [26] Zhang, J., Lv, L., Shi, X., Wang, Y., Guo, F., Zhang, Y., Li, H.: 3-d reconstruction of the spine from biplanar radiographs based on contour matching using the hough transform. IEEE TBME **60**(7), 1954–1964 (2013)