# 3D inference of the scoliotic spine from depth maps of the back

Nicolas Comte[1,2], Sergi Pujades[2], Aurélien Courvoisier[3,4], Olivier Daniel[3,4], Jean-Sébastien Franco[2], François Faure[1], and Edmond Boyer[2]

[1] Anatoscope, 38330, Montbonnot-Saint-Martin, France
[2] Inria, Université Grenoble Alpes, CNRS, Grenoble INP, Laboratoire Jean Kuntzmann, 38000 Grenoble, France
[3] TIMC, University Grenoble Alpes, CNRS, UMR 5525, VetAgro Sup, Grenoble INP, CHU Grenoble Alpes, 38000 Grenoble, France
[4] Grenoble Alps Scoliosis and Spine Center, Grenoble Alps University Hospital, Bvd de la Chantourne, CEDEX 09, 38043 Grenoble, France

**Abstract.** Recent advances combining outer images and deep-learning algorithms (DLA) show promising results in the detection and the characterization of the Adolescent Idiopathic Scoliosis (AIS). However, these methods are providing a limited 2D characterization while scoliosis is defined in 3D. In this study we propose an inference method that takes as input a depthmap of the back of a person and outputs the 3D shape estimation of the thoracolumbar spine. Our DLA method predicts 3D vertebrae positions with an average 3D error of 7.1mm (std: 4.7mm). From the predicted 3D positions, scoliosis can be located and estimated with a mean absolute error (MAE) of 5.5° (std: 6.2°) in the frontal plane. Moreover, sagittal alignments can be estimated with a MAE of 6.4° (std: 5.5°) in kyphosis and 8.3° (std: 6.8°) in lordosis. In addition, our non-ionizing approach can detect scoliosis with an accuracy of 89%.

**Keywords:** spine, adolescent idiopathic scoliosis, depthmaps

## 1 Introduction

Adolescent idiopathic scoliosis (AIS) is a progressive disease mostly affecting young people and evolving during the period of growth [5]. The main recommendation is to use braces to stop scoliosis progression with medical monitoring until the end of the adolescence. Diagnosis and follow-up are usually performed with radiographs by measuring the Cobb Angle at the scoliotic curvature on the coronal plane. This method is highly accurate but it involves multiple radiation exposures which increase the risk of cancer [3][6]. Therefore, there is a strong interest to examine spines using non-invasive and non-ionising methods, especially with young people [10].

In the recent years, deep learning methods based on Convolutional Neural Networks (CNN) showed promising results from outer images of the back. Yang et al 2019 [16] trained CNN models to classify RGB images of the back according to the scoliosis severity. Kokabu et al 2021 [20] extended the Adam's

forward bend test by the inference of the severity of scoliosis, *i.e.* the main Cobb angle, from depthmaps. Watanabe et al. 2019 [17] predicted the coronal spinal alignments from Moiré images and estimate the cobb angle along the predictions. However, these methods are providing a limited 2D characterization of the deformities while scoliosis is inherently a three-dimensional condition [4].

Our hypothesis is that the information present in a depth map of the back of a person has a very strong correlation with the underlying 3D shape of the spine, i.e. the 3D location of each vertebra. Thus, a model taking as input a depthmap could predict the 3D shape of the spine. The goal of this study is to evaluate such a strategy and its accuracy by comparing the predictions made by the proposed CNN model with real-life measurements of spine positions and curvatures.

## 2 Data collection and processing

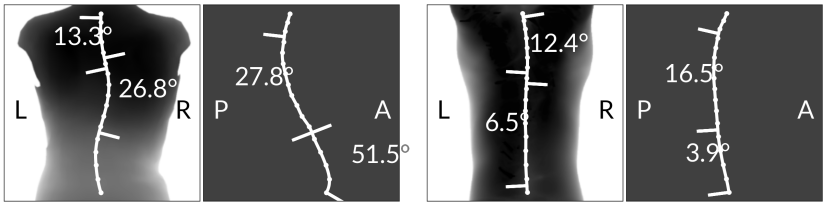
To validate our hypothesis and approach we use two different datasets. One acquired at Grenoble Hospital (GH) and the publicly available NMDID dataset [19].

**GH dataset** The study was approved by two Ethical Committees. The GH dataset was collected and published in Courvoisier et al 2019 [18]: CECIC Rhône-Alpes-Auvergne, Clermont-Ferrand, IRB 5891. The IRB for recent cases is CPP Ile de France 2 on the 07/20/2020: n○ ID RCB: 2020-A01071-38. All parents and patients received an information letter.

Two types of data were collected from 32 patients with AIS (7-16 y. old, 81% females) with no treatment: a 3D surface scan of the back or the torso with an optical scan (ScanGogh II Vorum Research Corporation or the Occipital Structure Sensor Mark II), and the corresponding 3D spine reconstruction obtained from biplanar X-rays with the EOS Imaging systems [7].

As the surface scan and the EOS images were not performed simultaneously, the patient pose can vary between both acquisitions. We therefore added an additional step to adjust the skin surface pose to match the EOS pose. We use a 3D kinematic anatomic model in Sofa [8] which we personalize to the back surface. We then deform the pose of the avatar to match the visual envelope extracted from the biplanar EOS radiographs. The obtained avatar has thus the skin surface in the same pose as the pose observed during the radiographs.

**NMDID dataset.** The GH dataset is composed of AIS patients only. To explore asymptomatic AIS cases we included subjects from the NMDID [19] database, which provides full-body CT-scans of *ex vivo* subjects. We included 89 cases (15-30 y. old, 32% females, 7% with scoliosis), which had 17 thoracolumbar vertebrae and a Body Mass Index (BMI) under $30kg/m^2$. A selection on BMI was needed since the corpses are in supine position, and the back is flatten during the CT-acquisition. High BMI bodies had completely flat backs with no shape information.




(a) GH (b) NMDID

Fig. 1: Dataset examples. For each subject we have a depthmap, the 3D vertebrae locations and the measured angles. Coronal (left) and sagittal (right) views.

From the selected cases, we automatically segmented the thoracolumbar spine vertebrae using Meng et al. 2023 method [22] and the skin with a binary segmentation approach on each slice of the volumetric image. The results were then turned into meshes using the marching-cubes algorithm [2].

**Depthmap generation.** From the meshes of the backs of the patients we rendered depthmaps using an orthographic camera model. For the NMDID subjects we directly used the segmented skin, whereas for the GH subjects, we used the skin of the repositionned avatar matching the 3D spine shape in the radiographs.

**Spine 3D characteristics.** From the 3D vertebra model positions obtained from the EOS Imaging system [7] and the automatic CT segmentation [22], we computed the 3D spine characteristics. Using the method by Choi et al. [12] based on cubic B-Splines, we computed the scoliosis severity by selecting the maximum absolute Cobb angle value. Kyphosis and lordosis were obtained with splines fitted in the sagittal plane between T04-T12 and L01-L05 respectively. As we do not model S01 we computed the lordosis Cobb angle using the L05 centroid.

**Final dataset.** From the NMDID and GH datasets we reconstructed 121 pairs of 3D spines and surfaces of the back: 89 from the NMDID (6 annotated with scoliosis) and 32 from GH (all with AIS). The input dataset we consider is therefore composed of 45% females aged between 7 and 30 years old. 38 cases with scoliosis are included in our analysis with different type of curvatures (simple, double curvatures, 10-65°). Fig. 1 illustrates one case of each dataset with the depthmap and the 3D characterization of the associated spine.

## 3 Method

Our method takes as input a depth map of the back of a patient and outputs the 3D location of the thoraco-lumbar vertebrae. To ease the spine prediction we

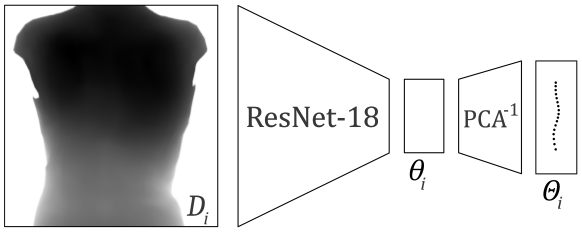



Fig. 2: An input depthmap $D_i$ is fed into the Resnet that predicts the latent PCA representation $\theta_i$. From $\theta_i$ the 3D vertebrae coordinates $\Theta_i$ are computed.

use a low dimensional representation of the spine based on Principal Component Analysis (PCA). Fig. 2 presents the overview of the method.

**PCA spine representation.** To ensure that the predicted individual vertebrae create a consistent spine shape, we use a low dimensional representation based on PCA[1]. The learned PCA space can produce a compressed latent representation of the spine data, while preserving most of its information and variability. Two main advantages follow. The prediction of the $n_\theta = 20$ PCA coefficients is less complex than the prediction of each of the 17 individual vertebrae 3D locations ($17 \times 3 = 51$). In addition, the compressed latent representation enforces regularization over the estimated spine locations.

**Architecture.** We use ResNet-18 [11] model as our backbone CNN architecture, which has been shown to be efficient for regression tasks. It encodes the depth image (224×224, 1 channel) into the aforementioned latent PCA representation of the spine, which is in turn easily decoded into the 3D vertebrae positions. We train the network to minimize an L2-loss between the vertebrae coordinates predicted by the neural model, and the annotated coordinates of the training examples. The model is trained using an Adam optimizer with PyTorch [15] with 2000 epochs and a learning rate of 1e-4.

**Data Processing.** To structure the dataset and ease the learning task we process the data. Following Choi et al. [12] the depthmaps on the spine are centered and cropped according to its length, then resized to a $224 \times 224$ resolution. Depth pixel observations are further normalized between $[-1, 1]$, and the antero-posterior positions of the vertebrae are defined with respect to a median vertebra (T08). This 3D transformation allows a simpler representation of the spine while keeping the vertebrae alignment information. It differs from [14] and [21] which are regressing absolute positions.

We augment the dataset applying a random set of different transformations, such as mirroring and rotating the torso in 3D. From the 121 rendered depthmaps, the data augmentation raises the amount of images to 10,890.

## 4 Evaluation

We start by validating the accuracy of the PCA spine representation. Then we evaluate the vertebrae location accuracy by computing the 3D distances between the ground-truth and the predictions as well as 2D distances in the coronal and sagittal planes. We evaluate the spine curvature accuracy by computing the Mean Absolute Error (MAE) in three angles: the main Cobb angle [12], the kyphosis (T04-T12) and the lordosis (L01-L05). As the 3D predicted spine of our method can also be used for spine classification (main Cobb angle $\geq 10°$) we compute binary classification metrics: sensitivity, specificity, positive predictive value, accuracy and AUC. All our quantitative results are computed using a stratified 20-fold cross-validation (CV) on the full dataset. All metrics are computed by averaging over all test sets.

### 4.1 PCA dimensional reduction.

A first analysis is dedicated to the PCA and how well the resulting representation is able to encode the vertebra positions. For each fold, the PCA model is learned from the training set and evaluated by comparing the PCA reconstructions of the spines of the test set with their true X-rays reconstructions. Our analysis shows that with 20 components, the spines can be reconstructed with a Mean Absolute Error (MAE) in 3D vertebra location of 0.92mm (std: 0.75 mm) and an MAE in scoliosis severity of 2.71deg (std: 2.49). These values are close to the annotation accuracy ($< 1$ mm and $< 5°$).

### 4.2 3D spine prediction accuracy

**Vertebra locations.** MAE of the vertebrae location predictions are reported in Table 1. Our model is able to predict the 3D vertebra positions with an average 3D error below the cm. These predictions allows the characterization of the spinal alignments (scoliosis, kyphosis, lordosis) with automated measurements [12]. Visuals of the predictions are presented in Fig. 3.

**Spine 3D characteristics.** Table 1 reports the angle errors between the predicted spine shapes and the ground truth ones. The last three lines of the Table 3) are provided for the comparison with the literature. Note how our approach is the only one predicting kyphosis and lordosis. Numerical results are provided for informative comparison on the range of values, as the considered datasets are different, both in size and population.

**Classification.** Classification results are reported in Table 1. Our method is able to discriminate cases with scoliosis with a sensitivity of 64% and a specificity of 99%. We report a positive predictive value of 95%, an accuracy of 89% and an AUC of 90%. Fig. 4 shows the predicted severity against the ground truth one with the $10°$ classification thresholds. Let us note how several miss-classifications

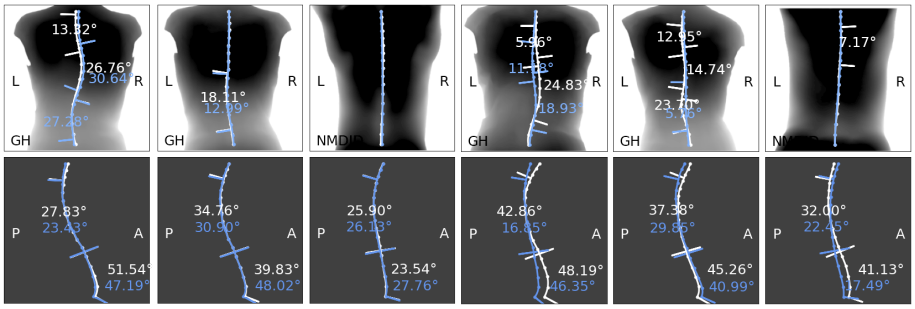



Fig. 3: Selection of results. White: ground-truth; blue: predictions. Top row: coronal view for each case, bottom row: sagittal view. GH: institution; NMDID: New Mexico Decedent Image Database; L: Left; R: Right; P: Posterior; A: Anterior.

Table 1: Spine prediction accuracy. Vertebrae errors in mm of the 3D locations (3D) and 2D coronal (Cor.) and sagittal (Sag.) projections. Angle errors in degrees in the scoliosis severity (Sco. sev.), kyphosis (Kyph.) and lordosis (Lord.). Scoliosis classification ($\geq 10°$) sensitivity (Sens.) and specificity (Spec.).

| | Locations (mm) | | | Characteristics (°) | | | Classification | |
|---|---|---|---|---|---|---|---|---|
| Study | 3D | Cor. | Sag. | Sco. sev. | Kyph. | Lord. | Sens. | Spec. |
| Ours | 7.07 (4.69) | 4.51 (2.99) | 5.60 (4.62) | 5.46 (6.19) | 6.44 (5.49) | 8.26 (6.85) | 0.64 | 0.99 |
| [12] | × | 5.4 (3.5) | × | 3.42 (2.64) | × | × | NA | NA |
| [16] | × | × | × | × | × | × | 0.88 | 0.84 |
| [20] | × | × | × | [4.4 - 4.7] | × | × | 0.99 | 0.42 |

arise near the 10° threshold and the underestimation of moderate and severe scoliosis cases. In conclusions we provide leads to improve these predictions.

### 4.3 Ablation studies

To further understand the proposed approach we performed two complementary experiments.

Recent work by Klarqvist et al. 2022 [23] has shown that it is possible to estimate body composition from solely body silhouettes. We consider the analogous case for the prediction of the 3D spine: instead to use a depthmap as input we consider the case where only the binary silhouette of the back surface is observed.

In addition, we also experiment by training and testing on the two different datasets (GH and NMDID) to see if the imbalance of scoliotic patients in the datasets has an effect on the predictions.

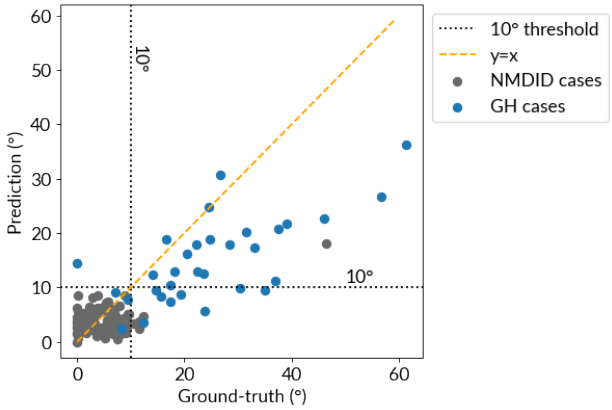



Fig. 4: Comparison between predicted scoliosis severity and ground-truth measurements.

Table 2: Comparison with binarized depth information, i.e. silhouette (silh). MAE (with std.) on locations in 3D, coronal (Cor.) and sagittal (Sag.) plane and the estimated scoliosis severity. Detection is evaluated sensitivity (Sens). and specificity (Spec.).

| Type | 3D (mm) | Cor. (mm) | Sag. (mm) | Severity (°) | Sens. | Spec. |
|---|---|---|---|---|---|---|
| silh | 10.02 (6.96) | 4.91 (3.34) | 8.55 (7.22) | 6.34 (7.86) | 0.52 | 0.98 |
| depth | 7.07 (4.69) | 4.51 (2.99) | 5.60 (4.62) | 5.46 (6.19) | 0.64 | 0.99 |

**Predictions from binary images.** We tested the impact of depth information by training our model to reconstruct the 3D positions on binary images that show only the silhouette of the torso. The results in Table 2 illustrate that model using depth maps achieves overall better performance. Depth is informative for the network. It is interesting to note the capability of the sihouette network to reconstruct scoliosis curvatures with a severity error of 6.34°. The biggest difference is obtained in the sagittal plane alignments, which is coherent with the lack of depth information in the binary silhouette image.

**Different datasets.** The inclusion of NMDID allows to validate the approach on asymptomatic cases. However, it could have a negative impact on the prediction of the scoliotic cases by adding bias. We thus evaluate the impact of the NMDID dataset on the network regression performance. We train solely on GH (scoliotic data) and on both GH and NMDID (scoliotic + asymptomatic) data and report the results on Table 3. Including the NMDID cases in the training set does not introduce any bias in the prediction performances (first two lines) on GH subjects and confirms that the inclusion of spines without scoliosis from

deceased subjects in supine position does not deteriorate the predictions over living patients acquired with the regular protocol.

Table 3: MAE in locations in mm and standard-deviations compared with the reported values in the literature.

| Study | Train | Test | 3D | Coronal | Sagittal |
|---|---|---|---|---|---|
| Ours | GH | GH | 8.81 (5.19) | 5.87 (3.73) | 6.76 (4.97) |
| Ours | GH+NMDID | GH | 8.80 (5.46) | 5.90 (3.73) | 6.74 (5.26) |
| Ours | GH+NMDID | NMDID | 6.45 (4.21) | 4.01 (2.49) | 5.19 (4.30) |
| Ours | GH+NMDID | GH+NMDID | 7.07 (4.69) | 4.51 (2.99) | 5.60 (4.62) |

## 5 Conclusion

In this work we presented an approach to predict the 3D spine shape of a patient from a depth map of the back. This depthmap can be obtained with depth sensors, which are usually inexpensive, portable, making them a promising cost-effective, non-ionizing approach to quantify scoliotic deformities. Our approach provides accurate 3D predictions of the vertebrae locations and allows to compute relevant anatomic curvatures along the spine as well as a fast scoliosis diagnosis.

Our presented work predicts 3D vertebrae locations, but does not consider today their individual 3D orientation. It has been studied that the axial vertebra rotation contains information related to the early detection of scoliosis [9] and its evolution [13]. Our approach could be adapted in the future to also predict the orientations of the individual vertebrae. To that end, the PCA representation should be reconsidered, as a linear approach might not be well suited to capture the 3D rotations.

Another venue of improvement of the presented work would be to add more subjects with moderate and severe scoliosis in the training pipeline. As shown in Fig. 4, our model tends to underestimate severe deviations due to the low number of such cases in our dataset. To make this possible, as well as to foster future research in this direction, we make our trained model and training code available for research purposes.

## References

1. Pearson, K.: LIII on lines and planes of closest fit to systems of points in space. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 2(11), 559?572 (1901). doi:10.1080/14786440109462720

2. Lorensen, W.E., Cline, H.E.: Marching cubes: A high resolution 3d surface construction algorithm. ACM SIGGRAPH Computer graphics 21(4), 163?169 (1987). doi:10.1145/37401.37422
3. Hoffman, D.A., Lonstein, J.E., Morin, M.M., Visscher, W., Harris III, B.S.H., Boice, J.D. Jr.: Breast Cancer in Women With Scoliosis Exposed to Multiple Diagnostic X Rays. JNCI: Journal of the National Cancer Institute 81(17), 1307?1312 (1989). doi:10.1093/jnci/81.17.1307
4. Stokes, I.A.: Three-dimensional terminology of spinal deformity. A report presented to the Scoliosis Research Society by the Scoliosis Research Society Working Group on 3-D terminology of spinal deformity. Spine (Phila Pa 1976) 19(2), 236?248 (1994).
5. Weinstein, S.L., Dolan, L.A., Cheng, J.C., Danielsson, A., Morcuende, J.A.: Adolescent idiopathic scoliosis. The Lancet 371(9623), 1527?1537 (2008). doi:10.1016/s0140-6736(08)60658-3
6. Ronckers, C.M., Doody, M.M., Lonstein, J.E., Stovall, M., Land, C.E.: Multiple Diagnostic X-rays for Spine Deformities and Risk of Breast Cancer. Cancer Epidemiol Biomarkers Prev 1 17 (3), 605?613. doi:10.1158/1055-9965.EPI-07-2628
7. Humbert, L., De Guise, J.A., Godbout, B., Skalli, W.: 3D reconstruction of the spine from biplanar x-rays using parametric models based on transversal and longitudinal inferences. Medical Engineering & Physics 31(6), 681?687 (2009). doi:10.1016/j.medengphy.2009.01.003
8. Dicko, A.H., Liu, T., Gilles, G., Kavan, L., Faure, F., Palombi, O., Cani, M.P.: Anatomy Transfer. ACM Transactions on Graphics, ACM SIGGRAPH ASIA, 32(6), (2013). doi:10.1145/2508363.2508415
9. Courvoisier, A., Drevelle, X., Dubousset, J., Skalli, W.: Transverse plane 3D analysis of mild scoliosis. European Spine Journal 22(11), 2427?2432 (2013). doi:10.1007/s00586-013-2862-x
10. Knott, P., Pappo E., Cameron, M., deMauroy, J.C., Rivard, C., Kotwicki, T., Zaina, F., Wynne, J., Stikeleather, L., Bettany-Saltikov, J., Grivas, T.B., Durmala, J., Maruyama, T., Negrini, S., O'Brien, J.P., Rigo, M.: SOSORT 2012 consensus paper: reducing x-ray exposure in pediatric patients with scoliosis. Scoliosis (2014). doi:10.1186/1748-7161-9-4
11. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770?778 (2015). doi:10.1109/CVPR.2016.90
12. Choi, R., Watanabe, K., Jinguji, H., Fujita, N, Ogura, Y., Demura, S., Kotani, T., Wada, K., Miyazaki, M., Shigematsu, H., Aoki, Y.: CNN-based spine and Cobb angle estimator using moire images. IIEEJ Trans. Image Electron. Vis. Comput. 5, 135?144 (2017). doi:10.11371/tievciieej.5.2_135
13. Skalli, W., Vergari, C., Ebermeyer, E., Courtois, I., Drevelle, X., Kohler, R., Abelin-Genevois, K., Dubousset, J.: Early Detection of Progressive Adolescent Idiopathic Scoliosis: A Severity Index. Spine (Phila Pa 1976) 42(11), 823?830 (2017). doi:10.1097/BRS.0000000000001961
14. Nerot, A., Skalli, W., Wang, X.: Estimation of spinal joint centers from external back profile and anatomical landmarks. Journal of Biomechanics 70, 96?101 (2018). doi:10.1016/j.jbiomech.2017.11.013
15. Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Köpf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., Chintala, S.: Pytorch: An imperative style, high-performance deep learning

library (2019). In: Proceedings of the 33rd International Conference on Neural Information Processing Systems, Vancouver, BC, Canada, 8026?8037 (2019). doi:10.48550/arXiv.1912.01703
16. Yang, J., Zhang, K., Fan, H., Huang, Z., Xiang, Y., Yang, J., He, L., Zhang, L., Yang, Y., Li, R., Zhu, Y., Chen, C., Liu, F., Yang, H., Deng, Y., Tan, W., Deng, N., Yu, X., Xuan, X., Lin, H. (2019).: Development and validation of deep learning algorithms for scoliosis screening using back images. Communications Biology 2, 390 (2019). doi:10.1038/s42003-019-0635-8
17. Watanabe, K., Aoki, Y., Matsumoto, M.: An application of artificial intelligence to diagnostic imaging of spine disease: Estimating spinal alignment from moiré images. Neurospine 16(4), 697?702 (2019). doi:10.14245/ns.1938426.213
18. Courvoisier, A., Nesme, M., Gerbelot, J., Moreau-Gaudry, A., Faure, F.: Prediction of brace effect in scoliotic patients: blinded evaluation of a novel brace simulator - an observational cross-sectional study. Eur. Spine J. 28, 1277?1285 (2019). doi: 10.1007/s00586-019-05948-9
19. Edgar, H.J.H., Daneshvari Berry, S., Moes, E., Adolphi, N.L., Bridges, P., Nolte, K.B.: New Mexico Decedent Image Database. Office of the Medical Investigator, University of New Mexico (2020). doi:10.25827/5s8c-n515.
20. Kokabu, T., Kanai, S., Kawakami, N., Uno, K., Kotani, T., Suzuki, T., Tachi, H., Abe, Y., Iwasaki, N., Sudo, H.: An algorithm for using deep learning convolutional neural networks with three dimensional depth sensor imaging in scoliosis detection. Spine Journal 21, 980?987 (2021). doi:10.1016/j.spinee.2021.01.022
21. Caturano, S., Kaiser, M., Bertsch, M., Bassani, T., Taylor, W.R., Ćuković, S.: Prediction of the 3d spinal alignment from external shape of the back in ais patients using regression model. IEEE 21st Mediterranean Electrotechnical Conference (MELECON), 1147?1152 (2022). doi:10.1109/MELECON53508.2022.9843065
22. Meng, D., Boyer, E., Pujades, S.: Vertebrae localization, segmentation and identification using a graph optimization and an anatomic consistency cycle. Comput. Med. Imaging Graph. 107, 102235 (2023). doi:10.1016/j.compmedimag.2023.102235
23. Klarqvist, MDR et al:Silhouette images enable estimation of body fat distribution and associated cardiometabolic risk. NPJ digital medicine 5(1), 105 (2022). doi: 10.1038/s41746-022-00654-1