ONCOLOGY

UDC 616.711-007.55-071.3

Automatic scoliosis angle measurement using deep learning methods, how far we are from clinical application: A narrative review

D. Kh. I. Kassab[1], *I. G. Kamyshanskaya*[1], *A. A. Pershin*[2]

[1] St. Petersburg State University,
7–9, Universitetskaya nab., St. Petersburg, 199034, Russian Federation
[2] Federal Scientific Center of Rehabilitation of the Disabled named after G. A. Albrecht of the Ministry of Labour and Social Protection of the Russian Federation,
50, Bestuzhevskaya ul., St. Petersburg, 195067, Russian Federation

For citation: Kassab D. Kh. I., Kamyshanskaya I. G., Pershin A. A. Automatic scoliosis angle measurement using deep learning methods, how far we are from clinical application: A narrative review. *Vestnik of Saint Petersburg University. Medicine*, 2021, vol. 16, issue 2, pp. 85–94. https://doi.org/10.21638/spbu11.2021.202

In recent years, automatic measurement of scoliosis angle using deep learning (DL) techniques is being studied extensively. The objective of this study is to review and assess the clinical applicability of these new methods. A wide search for English and Russian literature was conducted, 13 studies were included. Although the results of many of the reviewed DL methods in measuring the angle of scoliosis are promising, their clinical implication is by far not possible. There is absence of consensus in many issues regarding these new methods (differences in architecture of the ANN, data set, principle of angle measurement and nature of the reported results). In order to successfully introduce these new methods into clinical practice, more comparative and prospective studies are needed. Also, a multidisciplinary team including technical and medical workers is needed.

Keywords: scoliosis, automated Cobb angle, artificial neural network, deep learning.

Introduction

The spine is the fundamental bone of the body to which all the other bones are connected. Normal position, shape and development of the vertebrae not only provide the body with normal erect symmetrical figure, but also it is essential to the development of other adjacent bones and internal organs. The most common spinal deformation is scoliosis. Scoliosis is an ancient disease that is known since the time of Hippocrates. Since that time pathophysiology, diagnosis, classification and treatment of the disease are being studied. However, all these aspects are continuously changing in order to better understand the disease and provide the patient with the best way of intervention.

Scoliosis is generally characterized by lateral deviation of the spine. However, the disease actually is a three-dimensional deformity with lordosis and vertebral rotation. In the most common and clinically important type of idiopathic scoliosis, which affects more often girls, the primary lesion lies in the sagittal plane, taking the form of lordosis [1]. Currently, the main method of evaluation of scoliosis is still that, which proposed by the American orthopedic surgeon John Robert Cobb. This angle is directedly correlated to all treatment decisions. Cobb angle is measured between the two tangents of the upper and lower endplates of the upper and lower end vertebrae, respectively. The Scoliosis Research Society (SRS) suggests that the diagnosis is confirmed when the Cobb angle is 10° or higher and axial rotation can be recognized. However, structural scoliosis can be seen with a Cobb angle under 10°[2]. There is a recognized measurement error of the usual manual method of about 5° and can reach 10°. This is mostly related to variances in choosing the end vertebrae. Also, the endplates usually are difficult to delineate. However, new computer-assisted measurement methods have lesser measurement errors, ranging from 1,2° to 3,6° [2]. In last years, artificial intelligence (AI) and machine learning applications (ML) are being studied extensively in many aspects of the medical field. Artificial neural network (ANN) is a ML technique that is similar to human neurons and synapse system. It can learn by analyzing training data, and then make a prediction when new data is put in [3]. Deep learning (DL) or deep neural network (DNN) is similar to ANN but consists of deeper layers. We can get better performance in prediction and recognition studies, when the layers become more complex. Convolutional neural network (CNN) is a special type of ANN that shows promising results in medical image analysis such as segmentation (including spinal and vertebral segmentation) (Fig. 1). Because of

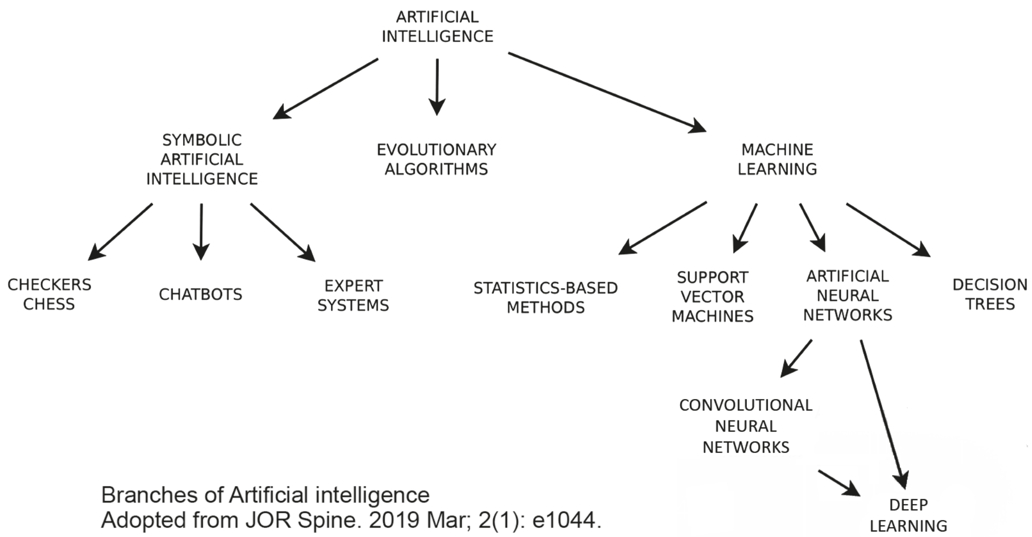

Fig. 1. Schematic overview of the main branches of artificial intelligence (AI), including machine learning (ML) methods which are having an impact on spine research [13]

the deep architecture of the network, large data set must be used for model training. For example, more than 1000 cases per class are needed to train deep learning architecture from scratch in classification [3]. However, using data augmentation or pre-trained network, around 100 cases per class could provide a reasonable outcome. Recently, different architectures of CNN are being proposed to analyze medical images. This article represents a review of the use of DL in measuring scoliosis angle. It was conducted to assess the clinical applicability of the newly proposed deep learning methods in scoliosis angle measurement.

Materials and methods

For this narrative review, we searched English literature in PubMed database, and Russian literature in e-library and CyberLeninka databases and forward and backward citations, using keywords deep learning and scoliosis or artificial neural network and Cobb angle. Only the studies that used a proposed deep learning method for the quantitative assessment of scoliosis were included. Studies of vertebral segmentation without measuring the angle were excluded. Also excluded, studies measuring the angle using other ML methods (such as vector regression). Actually, few studies available suggesting deep learning for measurement of scoliosis angle. The much more subject that has been studied is automatic spinal recognition, vertebral segmentation and 3D spinal reformation. We present the following article in accordance with the Narrative Review reporting checklist (available at: http://dx.doi.org/10.21037/atm-20-5495).

Results and discussion

13 studies, including 2 studies from Russian literature, met the criteria and were reviewed in this article. This review highlights certain main features of the proposed DL method in each study. Results were structured to represent these features and compare them. Those are type of NN that was used, measurement principle of Cobb angle, number of data used to train the network and to evaluate it (test data), range of angle if available and finally the results. Those data are summarized in table 1.

All the reviewed studies have similar broad lines in their approaches for measuring Cobb angle using deep learning techniques. Use of a certain data set of images to teach a proposed network that is based on CNN, to detect and segment either vertebrae or certain landmarks or the whole spine, and then measure the angle of inclination. The following paragraphs is a general discussion of the data represented in table 1, to highlight similarities or differences between the proposed methods from a clinical point of view.

Type of network

Most of the studies used networks that are based on CNN but with different architecture. Wang L. et al used a new Multi-View extrapolation net that consists of Joint-view net which simultaneously learns the features in AP and LAT images, Independent-view net which gets the independent pivotal landmarks in AP and LAT images separately and an inter-error correction net [4]. Few studies compared the ability of different types of

Table 1. **Deep learning proposed methods to measure the scoliosis angle**

Author	Title	Method to measure the angle	Type of ANN	Dataset	Results
Wang L. [4]	Accurate automated Cobb angles estimation using multi-view extrapolation net	Spinal land marks from AP/Lat X-ray to estimate the cobb angle	MVE Net: Joint-view net Independent-view net; Inter-error correction net	526 X-ray AP/lat cobb 0-96.33°	7,81 and 6,26 CMAE in AP and LAT angle respectively
Horng M. [5]	Cobb angle measurement of spine from X-ray images using convolutional neural network	Spinal isolation, vertebral detection and segmentation	U-Net, Res U-Net, and Dense U-Net	35 X-rays	MBR expert, novice, method −0,703 ± 12,552 −0,106 ± 13,582 −0,694 ± 12,091
Lein G. A. [6]	Automation analysis X-ray of the spine to objectify the assessment of the severity of scoliotic deformity in idiopathic scoliosis: a preliminary report	Angle measurement from the centerline of the spine	Mask (R-CNN); FCN ResNet; U-Net (Were compared)	300 X-rays children with IS	85% sensitivity 15% false results; X-ray bad quality or small angle<5
Wu H. [7]	Automated comprehensive adolescent idiopathic scoliosis assessment using MVC-Net	Spinal land marks from AP/Lat X-ray to estimate the cobb angle	MVC-Net Spinal Landmar-k Estimator net Cobb Angle Esti-mator network	526 X-ray bi planar images from 154 patients	CMAE 4,04°, 4,07° in AP and LAT Cobb angles, respectively
Kim C. K. [8]	Automation of spine curve assessment in frontal radiographs using deep learning of vertebral-tilt vector	Vertebral localization then vector tilt determination	Centroid-net, consists of: extraction net., initial prediction net, refinement net.	481 X-ray test 128 20 X-ray external dataset	CMAE 3,51 ± 3,89 for internal data set \ 4,11 ± 3,98 for external dataset
Padalko M. [9]	Automatic system for determining the angles of scoliotic deformity of the human spine	Spinal segmentation mask. measurement from the centerline	CNN, U-Net architecture	241 X-ray	—
Khanal B. [10]	Automatic Cobb angle detection using vertebra detector and vertebra corners regression	3 angles directly from landmarks (68 land mark)	CNN-based Faster-RCNN for vertebral detection Dense U-Net for land mark detection	609 spinal AP X-ray 98 test images	Error range 21,7%–33,3%

Watanabe K. [11]	An application of artificial intelligence to diagnostic imaging of spine disease: Estimating Spinal alignment from Moiré images	A curve fit 17 positions of vertebrae. 2 contact lines perpendicular to the curve used to measure the angle	CNN (Alexnet)	10,788 moiré & X-rays 0° to 55° 198 pair for testing	MAE VS the manual was 4,38 Normal 3,13 Mild 2,74 severe
Bernstein P. [12]	Radiographic scoliosis angle estimation: spline-based measurement reveals superior reliability compared to traditional Cobb method	Vertebral detection, measuring centers through perpendicular lines	Neural network (U-Net)	571 X-ray Cobb 20–40	Mean deviations below 0,5°
Galbusera F. [13]	Fully automated radiological analysis of spinal disorders and deformities: a deep learning approach	(3D) spinal anatomy from 78 landmarks	Simple seven-layer FCN	493 biplanar X-ray	Standard error 9,9°, mean difference —5,1°
Dubost F. [14]	Automated estimation of the spinal curvature via spine centerline extraction with ensembles of cascaded neural networks	Spinal Segmentation from vertebral landmarks. measuring from the centerline	Two cascaded CNN, U-Net architecture	609 X-ray scans 100 for testing	MAE 3,18 to 6,91 in the 3 curves
Pan Y. [15]	Evaluation of a computer-aided method for measuring the Cobb angle on chest X-rays	Segmentation of the spine and vertebral bodies	R-CNN	248 chest X-rays	MAE to manual was 3,32
Tu Y. [18]	Automatic measurement algorithm of scoliosis Cobb angle based on deep learning	Segmentation based method by calculating the slope of the curve	CNN for training Dense U-Net for segmentation	800 Xray 100 with scoliosis Test set 10 X-ray	Bias ranges from 1,0° to –5,4° (orthopedist reference)

Notes: FCN — fully convolutional networks; MVC — multi-view correlation; CMAE — circular mean absolute error; MVE — multi-view extrapolation; MAE — mean absolute error

CNN for detection and segmentation of the spine. Horng M. et al. compared U-Net, Res U-Net, and Dense U-Net. The segmentation results of the Residual U-Net were superior to the other two convolutional neural networks [5]. Another comparison was done between Mask (R-CNN), FCN (Fully Convolutional Networks), Res-Net and U-Net in a study by Lein G. A. et al, concluding that using U Net led to less error in spinal segmentation [6]. The availability of many types of DL network and the possibility to change the architecture of each network has made the choice of a gold standard method very difficult until now.

Measurement principle

Spinal isolation and vertebral detection are common initial steps for the automatic methods of measuring scoliosis angle, which can be divided roughly into two categories: segmentation-based methods and direct estimation methods [4]. Segmentation-based methods use vertebral segmentation to calculate the Cobb angle. Unfortunately, these methods are not robust because an accurate segmentation of the vertebra is extremely difficult owing to an unclear vertebral boundary in the radiographs, so the main disadvantage is the reliance on dedicated feature engineering and user bias [7]. Direct estimation methods attempt to extract the correlation between spine features (e. g., landmarks) from radiographs and the Cobb angle estimation without individual vertebral segmentation. However, these landmark-based methods also have disadvantages, because small errors in the landmarks can cause serious errors in the Cobb angle [8]. Another approach that may be considered as direct was proposed by Padalko M. et al., uses detection and segmentation of the whole spine with a mask, then determining its centerline. The angle is measured from the furthest detected points from the centerline [9]. Khanal B. et al used 68 land mark to detect 17 vertebrae, 12 thoracic and 5 lumbar (four corners of each vertebra). However, when the number of land marks in one image is not exactly 68, the program did not work well. Another limitation of their study is that the program falsely detected structures appearing similar to vertebra such as jaws, as skull and pelvis were cropped in the training data set [10]. Marking of the vertebrae also used by Watanabe K. et al. on Moiré images using a proposed Alexnet [11]. Kim K. C. et al. proposed a clinician friendly program, as they called it, that not only could measure the Cobb angle, but also could identify the end vertebrae. They combined centroid-net, that was used to localize and identify all thoracic and lumbar vertebrae and M-Net for vertebral-tilt field. They also compared the proposed method with other existing methods, which showed better results [8]. The same principle of using the center of the vertebrae to measure the angle was used by Bernstein P. et al. using neural network (NN), which has an architecture based on U-Net. They showed that even with poor quality images, the reliability of measuring scoliosis angle is increased by investigating the spinal curve as a whole and automating vertebral detection. Though the study included some drawbacks as the angle ranges and scoliosis spectrum were limited [12]·

Three studies combined multiple views to measure the angle, AP and lateral. Wang L. et al. proposed a Multiview exploration for accurate automated estimation of Cobb angle, using landmarks from AP and Lateral X-rays. Actually, another advantage of that study is the large dataset of 526 X-ray (plus augmentation) with different scoliosis severity [4]. Wu H. et al. proposed also a multi-view network (MVC-Net), for estimation of

Cobb angle using X-modules, land mark and Cobb angle estimators. They also used as, Wang et al. 526 X-ray (plus augmentation) [7]. Another study by Galbusera F. et al. used a 78 landmark from AP and Lat X-rays to measure sagittal balance and multiple spinopelvic parameters, including Cobb angle [13].

Data preparation and evaluation of the proposed method

Collecting a large number of images and generating the corresponding ground truth (or reference) is challenging and very time-consuming. Furthermore, the applicability of data augmentation which is an effective strategy to increase the size of the training data is inherently more restricted in medical imaging compared to other fields of research. Number of data set is a common limitation of almost all the reviewed studies that is most commonly in the range of hundreds. The Mask (R-CNN) was discarded in the preliminary test in the study of Lein G. A. et al. as it needed an extremely large data base [6]. In the same study, the authors stated that, there should be a minimum of 250 images to obtain an acceptable result, after using augmentation of data presented to the model. Dubost F. et al. and Khanal B. et al. used 609 data set and 100 test set which were available by Accurate Automated Spinal Curvature Estimation Challenge (AASCE) [10; 14]. Padalko M. et al., Pan Y. et al. and Horng M. et al. used small data sets 241,248 and 35 X-ray images respectively [5; 9; 15]. Moreover, regarding scoliosis, large total number of images is not enough to reflect the ideal data set. What is more important is the availability of minimum number of images that represent each degree of scoliosis (0–20°, 21–40°, 41–60°, >60°). Wang et al. used 526 X-rays equally divided into AP/LAT that included angle range from 0 to 96,33°. They also used dynamic data augmentation to increase the robustness of their model during training [4]. Watanabe K. et al. used 10788 pairs of Moire images and X-rays, 198 pairs for testing with angle range from 0 to 55° [11]. Data distribution in the study of Lein et al. revealed 15 % of the radiographs with grade I of scoliosis (angle up to 15–20°); 25 % with grade II (angle up to 40°); 45 % with grade III (angle up to 60°); and 15 % with grade IV (angle >60°) [6].

Reported results

Direct quantitative comparison of the reviewed studies cannot be made due to differences in the nature of the reported results. Good results can be seen by Wang et al. in their proposed method, the inter-error correction net was used for information fusion and better features in the final step of their proposed method. As a consequence, the combination of joint-view and independent-view nets can produce more accurate estimation than either net alone [4]. The authors also compared Symmetric Mean Absolute Percentage Error (SMAPE) with three other methods (the 1st two were excluded in this study for not meeting our criteria) on the same dataset, i.e., S2VR (Sun H. et al., 2017) was 37,08, BoostNet (Wu et al., 2017) was 41,35 and the MVC-Net (Wu et al., 2018) was 35,58 [7; 16; 17], their proposed model has achieved the most accurate result which was 18,95. Method of Galbusera F. et al. showed an excellent visual performance as the localizers in most cases were able to successfully capture the general shape of the thoracolumbar spine, even in patients with spinal instrumentation, no. matter which skeletal deformation is existed, however, the quantitative comparison with the ground truth, i.e., anatomical parameters extracted from the 3D reconstructions obtained with sterEOS, showed significant discrepancies. When measuring Cobb angle the standard error reaches 9,9° [13]. Kim C. K. el al. tested generalization ability of the proposed method by additional assessment of the Cobb angle measurement performance, using 20 frontal radiographs as external data set from a different hospital. The proposed method achieved small error with SMAPE 6,44 % and 6,87 % for the proposed M-net and U-Net, respectively [8]. Wu et al. who proposed MVC-network also shows accurate spinal land mark and Cobb angle measurement in AP & LAT. X-ray, with minimal differences in the mean of measured proximal thoracic (PT), major thoracic (MT) and thoracolumbar (TL) angles when compared to ground truth. However, significant differences of the minimum and maximum ranges were noted. For example, MT min-max distribution of AP X-ray, in the data set was 0,07-66 with a Mean 11,2 and in the proposed method the values were 1,75–41,9 and the mean was 10,1 [7]. Horng et al showed a detailed table representation of measurement results of a small data set (35) by their proposed method and comparison with an experienced observer and a novice observer [5]. Also, Tu Y. et al. presented their comparison results of only 10 automatically measured angles, ranges from 15,5° to 49,2° [18]. However, more organized detailed result representation could be about the accuracy of the proposed method for each degree of scoliosis. This was seen in the study by Watanabe et al., they reported mean absolute error of 4.38, 3.13 and 2.74 for normal, mild and sever scoliosis respectively where the same value for all the data set was 3.42. Such detailed data representation is more accurate for evaluation a certain method, though it is difficult to apply when the data set is large.

Conclusions

Although lately researchers actively are studying different methods and approaches for accurate assessment of scoliosis angle using DL with a lot of promising results, the best method is until now not defined. Which composition of ANN, whether or not to include more than one view or to use a 3D model, and what is the optimal quantity of learning data and tens of other questions still need consensuses in order to successfully introduce DL in clinical measurement of scoliosis angle. Actually, the mechanisms driving these methods are complicated and the basic principle appears as a black box to an external user [19]. Another issue is the importance of involving a multidisciplinary team of both technical and medical specialists like orthopedics or radiologists, not only during the steps of landmarking or manual measurement, but also in the very initial process of defining the objectives of the new method and at the final evaluation process. As evaluation and treatment of scoliosis patients depend on many factors, not only Cobb angle. Moreover, assessment of scoliosis severity is much way complex than other diseases that can be determined by one or multiple values or by certain range of values. Evaluation of such results by mean, for example, which has been seen in most of the studies reviewed in this article, cannot be an accurate representative value for all the grades of scoliosis, especially when the training and testing processes do not involve all the ranges of angles equally. It may be more accurate and clinically significant to represent each grade of scoliosis by a single mean, rather than representing the whole data set, as all these drawbacks will lead to under or overestimate the accuracy of the newly proposed method and even will be misleading. A third and very important object for future planning is the type of the studies conducted. There is a non-neglectable gap between the theoretical method and its clinical application no matter how effective this method is. The introduction of new techniques of measurement and evaluation may lead to change of an already known classifications or gradings. Thus prospective, comparative or case control like type of studies may be more beneficial to the active and successful integration of any new method into clinical practice.

References

1. Machida M., Weinstein S. L., Dubousset J. *Pathogenesis of idiopathic scoliosis*. Tokyo: Springer Nature, 2018. 124 p.
2. 2016 SOSORT guidelines: orthopedic and rehabilitation treatment of idiopathic scoliosis during growth. Scoliosis and Spinal Disorders, 2018. Available at: https://scoliosisjournal.biomedcentral.com/articles/10.1186/s13013-017-0145–8 (accessed: 20.03.2021).
3. Lee J. G., Jun S., Cho Y-W., Lee H., Kim G. B., Seo J. B., Kim N. Deep Learning in Medical Imaging: General Overview. *Korean J. Radiol.,* 2017, vol. 18, no. 4, pp. 570–584.
4. Wang L., Xu Q., Leung S., Chung J., Chen B., Li S. Accurate automated Cobb angles estimation using multi-view extrapolation net. *Med. Image Anal.*, 2019, vol. 58, 101542.
5. Horng M., Kuok C., Fu M., Sun Y. Cobb Angle Measurement of Spine from X-Ray Images Using Convolutional Neural Network. *Computational and Mathematical Methods in Medicine*, 2019, vol. 9, pp. 1–18.
6. Lein G. A., Nechaeva N. S., Mamedova G. M., Smirnov A. A., Statsenko M. M. Automation analysis X-ray of the spine to objectify the assessment of the severity of scoliotic deformity in idiopathic scoliosis: A preliminary report. *Pediatric Traumatology, Orthopaedics and Reconstructive Surgery*, 2020, vol. 8, no. 3, pp. 317–326.
7. Wu H., Bailey C., Rasoulinejad P., Li S. Automated comprehensive Adolescent Idiopathic scoliosis assessment using MVC-net. *Med. Image Anal.*, 2018, vol. 48, pp. 1–11.
8. Kim K. C., Yun H. S., Kim S., Seo J. K. Automation of Spine Curve Assessment in Frontal Radiographs Using Deep Learning of Vertebral-Tilt Vector. *IEEE,* 2020, vol. 8, pp. 84618–84630.
9. Padalko M. A., Orlov V., Naumov A. M., Nazarikov C. I., Aleksandrovich L. A. Automatic system to determine the angle of scoliotic deformation. *Fiziko-Matematicheskie Nauki*, 2019, vol. 3, pp. 55-68. (In Russian)
10. Khanal B., Dahal L., Adhikari P. Automatic Cobb Angle Detection using Vertebra Detector and Vertebra Corners Regression. computational methods and clinical applications for spine imaging 2019. *Lecture Notes in Computer Science*, vol. 11963, Springer, Cham, pp. 81–87.
11. Watanabe K., Aoki Y., Matsumoto M. An Application of Artificial Intelligence to Diagnostic Imaging of Spine Disease: Estimating Spinal Alignment from Moiré Images. *Neurospine,* 2019, vol. 16, no. 4, pp. 697–702.
12. Bernstein P., Metzler J., Weinzierl M. Radiographic scoliosis angle estimation: spline-based measurement reveals superior reliability compared to traditional COBB method. *JOR Spine,* 2021, vol. 30, no. 3, pp. 676–685.
13. Galbusera F., Casaroli G., Bassani T. Artificial intelligence and machine learning in spine research. *JOR Spine*, 2019, vol. 2, e1044.
14. Dubost F., Collery B., Renaudier A., Roc A., Posocco N., Niessen W., Bruijne M. Automated Estimation of the Spinal Curvature via Spine Centerline Extraction with Ensembles of Cascaded Neural Networks. *Computational Methods and Clinical Applications for Spine Imaging.* Shenzhen: Springer international publishing, 2019.
15. Pan Y., Chen Q., Chen T., Wang H., Zhu X., Fang Z., Lu Y. Evaluation of a computer-aided method for measuring the Cobb angle on chest X-rays. *Eur. Spine J.,* 2019, vol. 28, no. 12, pp. 3035–3043.
16. Sun H., Zhen X., Bailey C., Rasoulinejd P., Yin Y., Li S. Direct estimation of spinal cobb angles by structured multi-output regression. *International Conference on Information Processing in Medical Imaging*, 2017, pp. 529–540.
17. Wu H., Bailey C., Rasoulinejad P., Li S. Automatic landmark estimation for adolescent idiopathic scoliosis assessment using boost net. *Con Medical Image Computing and Computer Assisted Interventions,* 2017, pp. 127–135.

18. Tu Y., Wang N., Tong F., Chen H. Automatic measurement algorithm of scoliosis Cobb angle based on deep learning. *Journal of Physics: Conf. Series*, 2019, vol. 1187, iss. 4, 042100.
19. Chen K., Zhai X., Yang C., Sun K., Wang H., Li M. A narrative review of machine learning as promising revolution in clinical practice of scoliosis. *Annals of Translational Medicine*, 2021, vol. 9, no. 1. Available at: https://atm.amegroups.com/article/view/60113/html (accessed: 03.03.2021).

Received: April 20, 2021
Accepted: May 18, 2021

Authors' information:

Dima Kh. I. Kassab — PhD; Dimakk87@gmail.com
Irina G. Kamyshanskaya — PhD, Associate Professor; irinaka@mail.ru
Andrei A. Pershin — PhD; andrew.pershin@gmail.com