J Turk Spinal Surg 2024;35(2):49-54
DOI: 10.4274/jtss.galenos.2024.41736

SCOLIOSIS SCREENING FROM PLAIN RADIOGRAPHS INCLUDING CHEST AND ABDOMINAL X RAYS BY USING THE DEEP LEARNING METHOD: IS IT WORTH?

Alim Can Baymurat[1], Kemal Üreten[2], Tolga Tolunay[1], Gökhan Koray Gültekin[3], Muhammed Furkan Tosun[1], Muhammed Şakir Çalta[1], Alpaslan Şenköylü[1], Hakan Atalar[1]

[1]*Gazi University Faculty of Medicine, Department of Orthopaedics and Traumatology, Ankara, Turkey*
[2]*Ufuk University Faculty of Medicine, Department of Rheumatology, Ankara, Turkey*
[3]*Ankara Yıldırım Beyazıt University Faculty of Engineering and Natural Sciences, Department of Electrical and Electronics Engineering, Ankara, Turkey*

ABSTRACT

Objective: This study aimed to use deep learning techniques to discriminate different degrees of scoliosis on plain radiographs.

Materials and Methods: The study was performed on 1006 standing plain abdominal and chest radiographs (age range 10-18 years) obtained from the archive of our institution. The radiographs were divided into three groups according to the degree of scoliosis: normal (0-9°), mild (10-29°), and moderate/advanced (30° and above). The data were randomly selected and 15% were used for testing, 15% for validation, and the remaining 70% for training. Due to the limited data, the transfer learning (TL) method was used. Pre-trained VGG-16, ResNet-101, and GoogLeNet networks were used for TL. The original classifier was replaced with a new one. Geometric transformations of the radiographs were used for data augmentation. Rotation (-30, 30 degrees), translation (-30, 30 pixels), and scaling (0.9, 1.1 pixels) were applied to the images. The performance of the networks was evaluated using the performance parameters of accuracy, sensitivity, specificity, precision, and F1 score.

Results: Overall accuracy after testing the models was determined to be 90.1% for VGG-16, 86.1% for ResNet-101, and 85.5% for GoogLeNet. The accuracy, sensitivity, specificity, precision, and F1 score were 90.1%, 90.7%, 95.0%, 89.9%, and 90.1% for VGG-16, respectively. The VGG-16 values were determined to be higher than those of the ResNet-101 and GoogLeNet networks.

Conclusion: The results showed favorable results for deep TL methods in the assessment of normal, mild, and moderate/advanced scoliosis.
Keywords: Adolescent idiopathic scoliosis, deep learning method, pre-trained network.

INTRODUCTION

With advances in technology, artificial intelligence methods are now being applied in the field of medicine and in many areas of human life. Artificial intelligence studies are developing with increasing momentum. Recently, there have been many successful studies using convolutional neural networks (CNN), which is a deep learning (DL) method, especially in the field of image processing in medicine[1-3]. The layers in the CNN hidden layer mainly consist of the convolution layer, pooling layer, and fully connected layer. Feature extraction is performed in the convolution layers of the image transferred to the network operation through the input layer. In the pooling layer, the diversity of inputs is reduced. It is typically used as a fully connected layer classifier, and a softmax classifier is added to the output layer to calculate the prediction probability.

There are studies that have been conducted on different imaging methods of CNN in the field of medicine. CNNs are excellent at feature extraction, object detection, and classification, but a large amount of data is required to train the CNN model from the start. When there are not enough data, transfer learning (TL) is used to prevent overfitting and improve the performance of the model. For TL, some pre-trained networks developed from the imagent dataset are used. Some of these are AlexNet, VGG-16, GoogLeNet, ResNet-50, DenseNet, Inception, and MobileNet[4-6]. Another method applied to improve the performance of the model in DL methods is the data augmentation method. With this method, the number of data is artificially increased by

Address for Correspondence: Alim Can Baymurat, Gazi University Faculty of Medicine, Department of Orthopaedics and Traumatology, Ankara, Turkey
Phone: +90 312 202 22 33 **E-mail:** alimcanbaymurat@yahoo.com **Received:** 21.02.2024 **Accepted:** 21.03.2024
ORCID ID: orcid.org/0000-0002-0062-621X

©Copyright 2024 The Author. Published by Galenos Publishing House on behalf of Turkish Spine Society.
This is an open access article under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND) International License.

cropping, scaling, translating, and rotating the data, and these data are used for training the network[7].

Generally, scoliosis patients are referred to the orthopedist during school health screenings by chance or after the child's parents notice the patient's shoulder and trunk asymmetry. Sometimes, patients are referred to an orthopedist or spine surgeon after a curve in the spine is observed on the chest and outpatient abdominal radiographs taken in the hospital or health centers because of other complaints. While adolescent idiopathic scoliosis (AIS) patients who are diagnosed early with mild curvatures of the spine and start appropriate treatment are followed up and treated with conservative methods, surgical intervention is required in severe spinal curvatures that are overlooked or noticed late for any reason[8,9].

The application of DL methods has been extensively investigated across various medical disciplines. In addition, there have been studies on scoliotic deformities of the spine[10,11]. There have also been studies on the classification of scoliosis according to its degree[12-14]. DL studies of scoliosis in the literature have generally been conducted using computed tomography or scoliosis radiographs of the entire spine. There is no study in the literature on the effectiveness of the DL method in scoliosis screening from other plain radiographs including the spine, such as chest and/or standing abdominal radiographs, and grading the degree of curvature from these radiographs.

The aim of this study was to evaluate the effectiveness of DL methods in discriminating normal and scoliotic spines from chest and standing plain abdominal radiographs. To assess the ability of these methods to accurately classify scoliotic spines according to the degree of curvature.

MATERIALS AND METHODS

Approval for the study was obtained from the Gazi University Clinical Research Ethics (approval number: 02, date: 17.01.2023). In this retrospective study, posterior anterior chest radiographs and standing direct abdominal radiographs of patients aged 10-18 years who presented at our institution with any complaint between January 2021 and December 2022 were obtained from the picture archiving and communication systems (PACS) of our institution. A total of 20264 radiographs obtained from PACS were screened by four investigators. T1 and L1 vertebrae on chest radiography and T10 and L5 vertebrae on standing direct abdominal radiography were included in the study. Those who underwent spine surgery, rib cage surgery, implant-detected radiography, or radiography with incomplete spine images were excluded from the study. On the X-ray images used in the study, any identification such as hospital name, patient names, age, and date was removed, and the images were recorded as jpg files. The study included 1006 radiographs, which underwent assessment by four investigators for Cobb's measurement angles representing spinal curvatures. Subsequently, the precision of the measured Cobb angles was verified by an experienced spinal surgeon. The images were divided into three groups according to the degree of Cobb angles. Group 1 included cases with Cobb angles between 0° and 9°, Group 2 included Cobb angles between 10° and 29°, and Group 3 included cases with spinal deformity of 30° or more. There were 372 radiographs with Cobb angles between 0° and 9° (Group 1), 332 radiographs between 10° and 29° (Group 2), and 302 radiographs with Cobb angles of 30° or more (Group 3). Because the data used in the study was limited, the TL method was used. Pre-trained VGG-16, ResNet-101, and GoogLeNet networks were used for TL.

Data Processing Environment

The width and height of the PACS images were of different sizes and in jpg format. This study was conducted on a computer with a GeForce RTX2060 graphics processing unit, using MATLAB and Image Processing Toolbox. Each graph in the raw dataset was cropped using the rectangular area MATLAB® image cropping function, which includes the entire spine image. The hyperparameters used in the study were Optimizer; Stocastic gradient descent with momentum, mini batch size 16, dropout 0.5, initial learning rate 3e-4, and L2 regularization 0.004.

Data Preprocessing and Splitting

The data in this study were randomly selected and 15% were used for testing, 15% for validation, and the remaining 70% for training. Adjustment of parameters was made using validation data. Thus, the test and training data were not used during validation. No data augmentation was performed on the test and validation data. The number of images used in the splits, training, verification, and testing is shown in Table 1.

Transfer Learning, Data Augmentation

DL methods require a large amount of data for training. When there is insufficient data, data augmentation and the TL method are used. The data enhancement process is achieved by copying or converting images, such as sharpness, brightness, contrast change, and mirror symmetry. To improve the performance of the CNN model, online data augmentation by randomly flipping and rotating images was used. TL is the use of networks trained with ImageNet data, which contains many natural object images, for a new task. In this study, pre-trained VGG-16[12], ResNet-101[14], and GoogLeNet networks were used for TL. Therefore, the original classifier was replaced with a new classifier. Geometric transformations of radiographs were used for data augmentation, with rotation (-30, 30 degrees), translation (-30, 30 pixels), and scaling (0.9, 1.1 pixels) applied to the images.

Table 1. The number of images used for training, validation and testing

Cobb angle	Train	Validation	Test	Total
0-9 (normal)	261	54	57	372
10-29 (mild)	234	48	50	332
30 and above (moderate/severe)	213	44	45	302

Statistical Analysis

Because of the classification, the performances of the networks were evaluated with respect to accuracy, sensitivity, specificity, precision, and F1 score performance parameters. These performance measures were calculated using macroaverages from the confusion matrix obtained during the testing of the models.

$$\text{accuracy}=\frac{(TP+TN)}{(TP+TN+FP+FN)}$$

$$\text{sensitivity (recall)}=\frac{TP}{TP+FN}$$

$$\text{specificity}=\frac{TN}{TN+FP}$$

$$\text{precision}=\frac{TP}{TP+FP}$$

$$\text{F1 score}=\frac{2^*(\text{precision}^*\text{recall})}{(\text{precision}+\text{recall})}=\frac{2TP}{2TP+FP+FN}$$

TP: True positive; FP: False positive; TN: True negative; FN: False negative

RESULTS

The patients comprised 651 (64.7%) females and 355 (35.3%) males with a mean age of 14.61±3.89 years. Cobb angle values were within normal limits (0-9º) in 372 patients, mild scoliosis (10-29º) in 332 patients, and moderate/severe scoliosis (30º and above) in 302 patients (Table 2).

Three distinct groups of datasets were constructed based on the magnitude of the curves observed in the images: normal (Group 1), mild (Group 2), and moderate/severe (Group 3) (Table 1). The efficacy of the models was assessed by employing performance metrics, including accuracy, sensitivity, specificity, and precision. These metrics were computed using confusion matrices. The training process was repeated five times for each model. A summary of the performance of the pre-trained models is presented in Table 3. The overall accuracy of the VGG-16, ResNet-101, and GoogLeNet models tested was 90.1%, 86.1%, and 85.5%, respectively. The accuracy, sensitivity, specificity, precision, and F1 score results were 90.1%, 90.7%, 95.0%, 89.9%, and 90.1% in VGG-16, respectively, and were higher than the ResNet-101 and GoogLeNet results (Table 3). The confusion matrices obtained by testing the models are shown in Figure 1. Figure 2 shows the prediction results of six randomly selected samples during testing with VGG-16, first showing which classification was predicted, and then showing the percentage probability at which the predicted classification was predicted.

Table 2. Demographic characteristics and basic information of the patients

Mean age (year)	14.61±3.89
Gender F/M n (%)	651 (64.7%)/355 (35.3%)
Number of images (n)	1006
Cobb angle values	
Normal limits (0-9º)	372
Mild scoliosis (10-29º)	332
Moderate/severe scoliosis (30º and above)	302

F: Female; M: Male

DISCUSSION

The objective of this study was to assess the efficacy of DL techniques in detecting scoliosis from chest and standing plain abdominal radiographs, as well as accurately distinguishing various degrees of scoliosis based on the severity of spinal curvature. In addition, this study aims to gauge the potential of DL methods for scoliosis screening using chest and upright abdominal radiographs stored within healthcare institution image archives. TL was employed to implement the DL models using pre-trained VGG-16, ResNet-101, and GoogLeNet networks. These models were applied to the lung and standing abdominal radiographs used in the study. In the present study, the accuracy, sensitivity, specificity, precision, and F1 score results of VGG-16 with the DL method were 90.1%, 90.7%, 95.0%, 89.9%, and 90.1%, respectively, and it was found to be successful in both scoliosis detection and classification according to the degree of curvature.

Spine radiography is the first-line imaging method for the evaluation of patients with scoliosis. This study was conducted using radiographs with DL methods for the diagnosis and grading of scoliosis. Scoliotic deformities may be missed on chest or abdominal radiographs for any reason, especially in non-orthopaedic departments, because the physicians performing the evaluation do not have sufficient experience with scoliotic deformities. The success in scoliosis detection and grading with the DLmethod will help in scoliosis detection, grading, and referral to a specialist by physicians with limited knowledge of scoliosis disease. Previous studies have been conducted on DL methods in the diagnostic of musculoskeletal diseases. Successful results were obtained in radiological imaging studies where TL was applied with previously trained networks[15,16].

Table 3. Performance results obtained by testing the models

	VGG-16 (%)	ResNet-101 (%)	GoogLeNet (%)
Accuracy	90.1	86.1	85.5
Sensitivity	90.7	86.1	85.7
Specificity	95.0	93.0	92.6
Precision	89.9	86.1	85.6
F1 score	90.1	86.1	85.6

There are studies in the literature that have measured and predicted the Cobb angle of the scoliotic spine with the DL method using plain radiographs and computed tomography images[10,17,18]. Huang et al.[19] conducted a comparative analysis of CNNs and manual measurement of the Cobb angle in AIS. The findings revealed that the fully automated framework not only accurately identified vertebral borders, vertebral sequences, superior/inferior end vertebrae, and apical vertebrae but also successfully evaluated the Cobb angle for Proximal Thoracic, Main Thoracic (MT), and Thoracolumbar/Lumbar curves. In this study, the measurement of Cobb angle pain was excluded as a parameter for assessing spinal curvatures. Instead, the focus was on investigating the efficacy of DL methods in the detection of scoliosis and discrimination of the magnitude of curvature[19]. Fraiwan et al.[18] used deep TL to detect scoliosis and spondylolisthesis from patient X-ray images. In the study

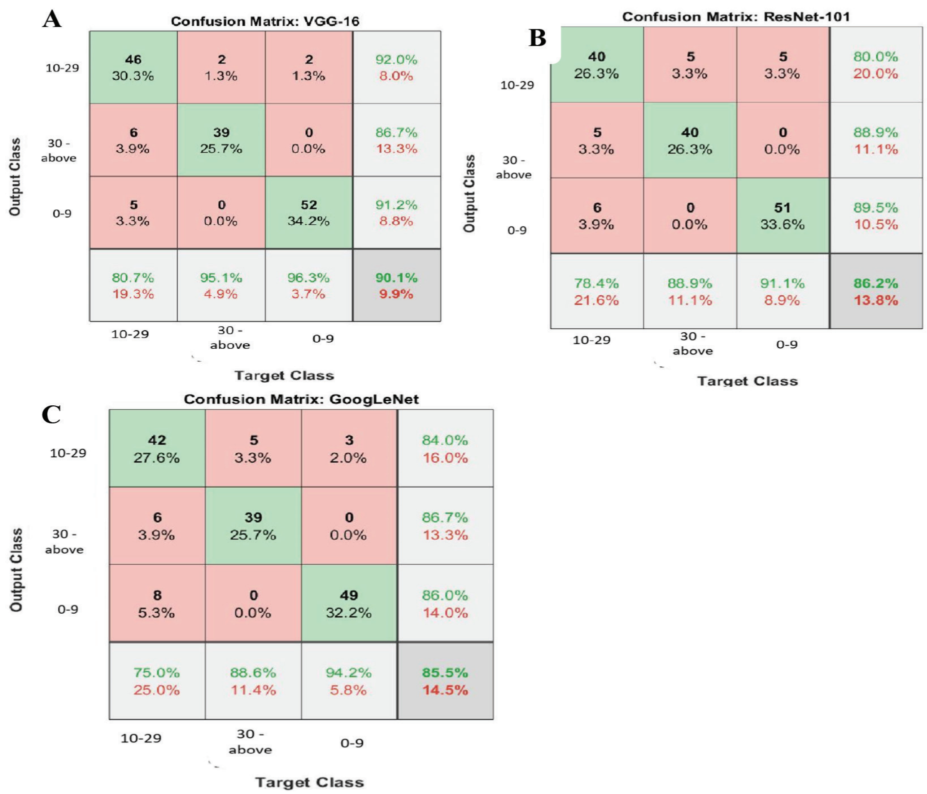

Figure 1. Confusion matrices of the pretrained models, A) VGG-16 model, B) ResNet-101 model, C) GoogLeNet model

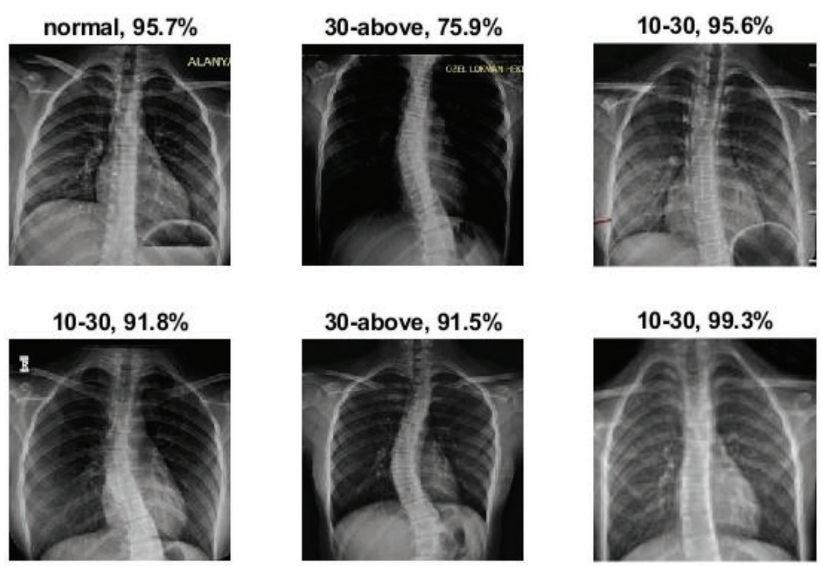

Figure 2. The prediction results of randomly selected six images during testing of the VGG-16 model

conducted on 188 scoliosis patients aged 5-35 years, successful results were obtained with an accuracy rate of 96.73% and 98.02%[18]. The current study is the first to focus on radiographs of adolescent patients aged 10-18 years, a population in which scoliosis commonly occurs. Furthermore, a substantial dataset of 1006 plain radiography images was analyzed, indicating a relatively extensive sample size. To mitigate the risk of overfitting, TL techniques were employed. The images were categorized into distinct groups based on the severity of curvature, as shown in Table 1. The evaluation of the VGG-16 model yielded promising results, with accuracy, sensitivity, specificity, precision, and F1 score values of 90.1%, 90.7%, 95.0%, 89.9%, and 90.1%, respectively.

Tavana et al.[20] performed a comparative analysis between a DL approach and a classical method for classifying the shape of spinal curvatures using a dataset of 1000 anterior-posterior spine radiographs. The study findings revealed that the pre-trained Xception model exhibited superior performance compared with the pre-trained MobileNetV2 network and the classical K nearest neighbors and support vector machine methods, specifically in discriminating between C- and S-shaped curvatures of the spine[20].

In a faster R-CNN and ResNet-based computer-aided scoliosis diagnosis study by Chen et al.[12], X-rays of the spine with normal, mild, moderate, and severe scoliosis were studied. For verification and analysis, the results were compared with the measurement results of the orthopedic surgeon. It has been reported that faster R-CNN and ResNet-based computer-assisted scoliosis diagnosis achieved successful results[12]. Chen et al.[12] used scoliosis radiographs including the entire vertebral column in their study. In the current study, using chest and standing abdominal radiographs, there were similar results, and it was seen that the DL method was successful in recognizing normal, mild, and moderate/severe scoliosis. However, chest and standing abdominal radiographs were not used in our study.

In this study, the reason for separating the images into three groups according to the degree of scoliotic curvatures of the spine was primarily because the treatment of patients with scoliosis varies depending on the size of the spinal curvature. While scoliotic curvature of the spine up to 10° is considered normal and does not require any follow-up and treatment, careful follow-up and appropriate treatment should be performed for errors where bone maturity has not been completed and in cases of spine curvatures of 10° and above. In this study, Cobb angles were considered normal at 0-9 degrees (Group 1) and did not require follow-up and treatment. Mild curvatures of 10-29 degrees (Group 2) are usually followed up radiologically. For curvatures of 30° and above (Group 3), brace treatment is applied, and surgical treatment is required in curvatures of 45°.

One of the key contributions of this study lies in the capacity of DL methods to discern different degrees of spinal curvature. The ability to classify scoliosis into normal, mild, moderate, and severe categories using DL techniques has significant implications for healthcare professionals. It allows for the provision of preliminary information to patients regarding their condition, guiding them toward appropriate treatment options, and facilitating referrals to specialists.

As a limitation of this study, due to limited number of radiological images, the radiographs were divided into only three groups (0-9 degrees: 1st group; 10-30 degrees: 2nd group and 30 degrees and above: 3rd group): 3rd group. In addition, not using the DL method in scoliosis measurements and the retrospective design of the study can also be considered limitations.

CONCLUSION

In conclusion, the results of this study show that DL methods are effective in detecting scoliosis in chest and standing plain abdominal radiographs and in classifying spinal curvature according to its magnitude. Also, the results of this study showed that deep TL methods may be useful in the evaluation of normal, mild and moderate/advanced scoliosis. The spine on X-ray images can be labeled and separated with DL methods as normal, mild, and moderate/advanced scoliosis. Thus, scoliosis patients requiring follow-up and treatment can be referred to orthopedics or spinal surgery. Furthermore, this study provides non-orthopedic physicians with limited expertise in spinal disorders with the opportunity to accurately assess the detection of scoliosis in a patient and the severity of spinal curvature. Thus, it offers the possibility of early detection of scoliosis and referral to the appropriate specialist, enabling early treatment. We also believe that this study will shed light on advanced scoliosis screening and statistical studies.

Ethics

Ethics Committee Approval: Approval for the study was obtained from the Gazi University Clinical Research Ethics (approval number: 02, date: 17.01.2023).

Informed Consent: Retrospective study.

Authorship Contributions

Surgical and Medical Practices: A.C.B., T.T., A.Ş., H.A., Concept: A.C.B., T.T., H.A., Design: M.F.T., A.Ş., Data Collection or Processing: K.Ü., G.K.G., M.F.T., M.Ş.Ç., Analysis or Interpretation: A.C.B., K.Ü., G.K.G., Literature Search: A.C.B., M.F.T., M.Ş.Ç., Writing: A.C.B., M.Ş.Ç.

Conflict of Interest: The authors have no conflicts of interest to declare.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

1. Üreten K, Maraş Y, Duran S, Gök K. Deep learning methods in the diagnosis of sacroiliitis from plain pelvic radiographs. Mod Rheumatol. 2023;33:202-6.

2. Atalar H, Üreten K, Tokdemir G, Tolunay T, Çiçeklidağ M, Atik OŞ. The diagnosis of developmental dysplasia of the hip from hip ultrasonography ımages with deep learning methods. J Pediatr Orthop. 2023;43:e132-7.
3. Yahara Y, Tamura M, Seki S, Kondo Y, Makino H, Watanabe K, et al. A deep convolutional neural network to predict the curve progression of adolescent idiopathic scoliosis: a pilot study. BMC Musculoskelet Disord. 2022;23:610.
4. Shen D, Wu G, Suk HI. Deep learning in medical image analysis. Annu Rev Biomed Eng. 2017;19:221-48.
5. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. Med Image Anal. 2017;42:60-88.
6. Wang J, Zhu H, Wang SH, Zhang YD. A review of deep learning on medical image analysis. Mobile Networks and Applications. 2021;26:351-80.
7. Shorten C, Khoshgoftaar TM, Furht B. Text data augmentation for deep learning. J Big Data. 2021;8:101.
8. Konieczny MR, Senyurt H, Krauspe R. Epidemiology of adolescent idiopathic scoliosis. J Child Orthop. 2013;7:3-9.
9. Fong DY, Lee CF, Cheung KM, Cheng JC, Ng BK, Lam TP, et al. A meta-analysis of the clinical effectiveness of school scoliosis screening. Spine (Phila Pa 1976). 2010;35:1061-71.
10. Zhao Y, Zhang J, Li H, Gu X, Li Z, Zhang S. Automatic cobb angle measurement method based on vertebra segmentation by deep learning. Med Biol Eng Comput. 2022;60:2257-69.
11. Zou L, Guo L, Zhang R, Ni L, Chen Z, He X, et al. VLTENet: A deep-learning-based vertebra localization and tilt estimation network for automatic cobb angle estimation. IEEE J Biomed Health Inform. 2023;27:3002-13.
12. Chen P, Zhou Z, Yu H, Chen K, Yang Y. Computerized-assisted scoliosis diagnosis based on faster R-CNN and ResNet for the classification of spine X-Ray images. Comput Math Methods Med. 2022;2022:3796202.
13. Yang D, Lee TTY, Lai KKL, Lam TP, Castelein RM, Cheng JCY, et al. Semi-automatic method for pre-surgery scoliosis classification on X-ray images using Bending Asymmetry Index. Int J Comput Assist Radiol Surg. 2022;17:2239-51.
14. He Z, Wang Y, Qin X, Yin R, Qiu Y, He K, et al. Classification of neurofibromatosis-related dystrophic or nondystrophic scoliosis based on image features using Bilateral CNN. Med Phys. 2021;48:1571-83.
15. Gao XW, Hui R, Tian Z. Classification of CT brain images based on deep learning networks. Comput Methods Programs Biomed. 2017;138:49-56.
16. Bressem KK, Vahldiek JL, Adams L, Niehues SM, Haibel H, Rodriguez VR, et al. Deep learning for detection of radiographic sacroiliitis: achieving expert-level performance. Arthritis Res Ther. 2021;23:106.
17. Ishikawa Y, Kokabu T, Yamada K, Abe Y, Tachi H, Suzuki H, et al. Prediction of cobb angle using deep learning algorithm with three-dimensional depth sensor considering the influence of garment in idiopathic scoliosis. J Clin Med. 2023;12:499.
18. Fraiwan M, Audat Z, Fraiwan L, Manasreh T. Using deep transfer learning to detect scoliosis and spondylolisthesis from X-ray images. PLoS One. 2022;17:e0267851.
19. Huang X, Luo M, Liu L, Wu D, You X, Deng Z, et al. The comparison of convolutional neural networks and the manual measurement of cobb angle in adolescent idiopathic scoliosis. Global Spine J. 2022:21925682221098672.
20. Tavana P, Akraminia M, Koochari A, Bagherifard A. Classification of spinal curvature types using radiography images: deep learning versus classical methods. Artif Intell Rev. 2023;56:13259-91.