

합성곱 신경망을 이용한 요추 특징점 검출에 관한 연구

A Study on Lumbar Vertebrae Landmark Detection using Convolutional Neural Networks

장준수¹

Jun-Su Jang¹

요약

본 연구에서는 요추 추체의 특징점을 추출하기 위한 합성곱 신경망 기반의 검출 방법을 제안한다. 추체 특징점 검출은 Cobb 각도, 척추전만각, 추체 전위 여부 등의 다양한 임상지표를 도출하기 위한 가장 기본적인 부분이다. 특징점 검출을 위해서 coarse-to-fine 접근 방법의 두 단계 네트워크 구조를 제시한다. 첫 번째 단계 네트워크는 전체 요추 특징점 검출을 강인하게 수행하는데 초점을 두었고, 두 번째 단계 네트워크는 각 추체별로 특징점 위치의 정확도를 높일 수 있도록 학습되었다. 두 네트워크 모두 ResNet-50을 수정하여 직접 좌표값을 출력할 수 있는 회귀 모델로 설계되었다. 제안한 검출 구조의 성능을 확인하기 위해서, 1,000장의 요추부위 측면 X-ray 영상을 활용하여 실험을 수행하였다. 실험 결과로, 검출 성공률 99.7%, 특징점의 위치 오차 4.54 ± 3.00 의 우수한 성능을 보였다. 제안한 특징점 검출 방법은 기존의 다른 추체 분석 연구에 적용될 수 있으며, 전체 시스템의 성능 향상에 기여할 수 있을 것이다.

핵심어 : 요추 진단, 추체 특징점 검출, 딥러닝, 합성곱 신경망

Abstract

In this study, a detection method based on convolutional neural networks is proposed to extract the lumbar vertebrae landmarks. Vertebrae landmark detection is the most fundamental part to derive various clinical indices such as Cobb angle, vertebral lordosis, and vertebral dislocation. For landmark detection, we present a two-stage network structure of a coarse-to-fine approach. The first stage network focused on robustly detecting all lumbar landmarks, and the second stage network was learned to increase the accuracy of the locations of landmarks for each vertebra. Both networks were designed as regression models that can directly output coordinates of landmarks by modifying ResNet-50. In order to confirm the performance of the proposed detection structure, an experiment was performed using 1,000 lumbar lateral X-ray images. As the result, the success rate of detection was 99.7%, and the position error of the landmarks was 4.54 ± 3.00 . The proposed landmark detection method can be applied to other previous vertebral analysis studies, and contribute to improving the performance of the entire system.

Keyword : Lumbar vertebrae diagnosis, Vertebrae landmark detection, Deep learning, Convolutional neural networks

¹ Future Medicine Division, Korea Institute of Oriental Medicine, Daejeon, Korea [Principal Researcher]

e-mail: junsu.jang@kiom.re.kr

* 본 연구는 한국한의학연구원 “AI 한의사 개발을 위한 임상 빅데이터 수집 및 서비스 플랫폼 구축(KSN2012110)” 과제의 지원을 받아 수행되었습니다.

Received(August 18, 2020), Review Result(1st: August 27, 2020), Accepted(September 4, 2020), Published(September 30, 2020)



© 2020 The Authors. Published by NCISS.
This is an open access article licensed under the Creative Commons Attribution-NonCommercial 4.0 International License.
To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

1. 서론

인공지능이 발전함에 따라서 해당 분야의 전문가가 하던 작업을 인공지능이 전문가 못지않게 혹은 전문가보다 더 뛰어나게 수행하는 시대가 왔다. 특히 깊은 신경망(Deep Neural Networks, DNN)으로 대표되는 현대의 인공지능은 시각 인지 분야에서 뛰어난 성능을 입증하고 있는데, 의료 영상 분석에 있어서도 DNN 기반의 수많은 자동화 알고리즘들이 연구되고 있다 [1][2].

단순 방사선 사진(X-ray)을 이용한 척추 구조의 자동 분석은 오랫동안 연구되어 온 분야인데, 전문가의 수작업에서 발생할 수밖에 없는 정확성, 반복성 문제를 보완하기 위해서 자동화된 알고리즘이 필수적으로 필요하다. 척추 구조 분석에 관련된 연구는 추체(vertebrae)의 특징점을 검출하는 연구 [3][4], 추체의 영역을 분할하는 연구 [5][6], 추체의 크기, 각도 등 다양한 임상 지표를 추출하는 연구 [7][8] 로 나뉘 볼 수 있다. 추체의 특징점 검출, 영역 분할을 수행한 후, 목표로 하는 임상 지표를 추출하는 순차적 접근이 일반적이거나, 최근에는 그 구분을 명시적으로 하지 않거나, 하나의 네트워크 구조에 융합하여 임상 지표를 자동으로 추출하는 연구도 활발히 진행되고 있다 [9].

위에 나열한 연구들 중에서도 추체의 특징점을 검출하는 연구는 자동화 알고리즘의 가장 기본이 되는 부분인데, 그 중요성은 다음과 같이 볼 수 있다. 첫째, 시각적으로 해석하기 쉬운 방법으로 임상에서 직관적으로 활용할 수 있다. 임상에서 활용하는 지표는 대부분 추체의 상연, 하연과 평행한 직선을 긋고, 이 직선으로 정의되는 각도를 활용한다. 대표적으로 Cobb 각도 [10], 척추전만 각 [11] 등이 있다. 추체의 자세를 표현하는 직선을 결정하기 위해서 2개의 점 위치를 지정하는 것이 전문가의 진단과정과 가장 유사하며, 결과를 해석하는데 있어서도 가장 직관적이다. 각도 측정을 위해서 영역 분할 알고리즘을 활용하는 경우가 많은데 [6], 영역 분할이 이루어진 후에도 최종적으로 하나의 각도를 결정하기 위해서는 2개의 점을 결정하는 과정이 필요하다.

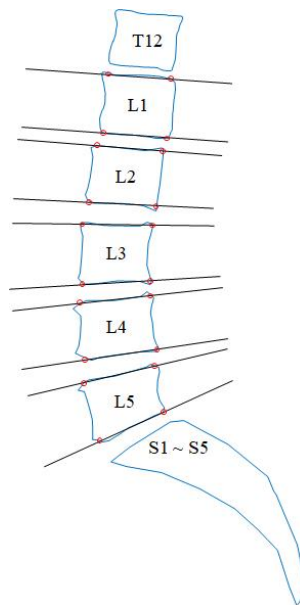
둘째, 영역 분할이나 임상 지표 추출 연구에 요소기술로 적용될 수 있다. 영역 분할 연구에서 초기 관심영역(ROI) 설정이 필요한데, ROI 설정을 위해서 추체의 중심점 혹은 4개의 꼭지점 검출을 활용한다. 특징점 검출의 정확도에 따라 전체 시스템의 성능이 결정될 수 있으며, ROI 설정이 실패할 경우, 이어지는 영역 분할이나 지표 추출 부분이 전혀 동작하지 않을 수도 있다.

본 연구에서는 DNN 중 영상 인식에서 가장 널리 활용되고 있는 합성곱 신경망(Convolutional Neural Networks, CNN)을 이용하여 요추 추체의 특징점을 자동 검출하는 방법을 제안한다. 두 단계의 CNN 구조를 활용하여 검출의 강인성과 정확도를 높이는 시도를 하였으며, 각 단계에 활용한 CNN 구조는 거의 동일하여 시스템의 효율성도 확보하였다. 각 추체별로 학습된 특징점 검출기의 성능을 테스트하여 결과를 제시한다.

2. ResNet-50 기반 요추 특징점 검출

2.1 요추 특징점 검출 문제 정의

정상 요추의 경우 [그림 1]과 같이 L1 ~ L5 의 5개의 추체로 이루어져 있는데, L1 의 위쪽에는 흉추의 마지막 추체인 T12 가 위치하고, 아래쪽으로는 천추 S1 이 위치한다. X-ray 영상에서 관측 되는 추체의 형태를 사각형에 근사한다고 볼 때 사각형의 꼭지점 4개를 특징점으로 정의할 수 있다. 추체의 각도를 측정하기 위해서는 추체의 상연 혹은 하연의 꼭지점 2개를 선택하여 직선을 결정하면 된다.



[그림 1] 요추 특징점 정의

[Fig. 1] Definition of lumbar vertebrae landmarks

추체를 대표하는 4개의 점은 다양한 지표를 계산하기 위한 가장 기본적인 요소이다. 추체 하나 당 4개의 특징점을 지정하면 Cobb 각도, 척추전만각 뿐만 아니라, 추체의 전위 여부 등 다양한 진단에 활용할 수 있다.

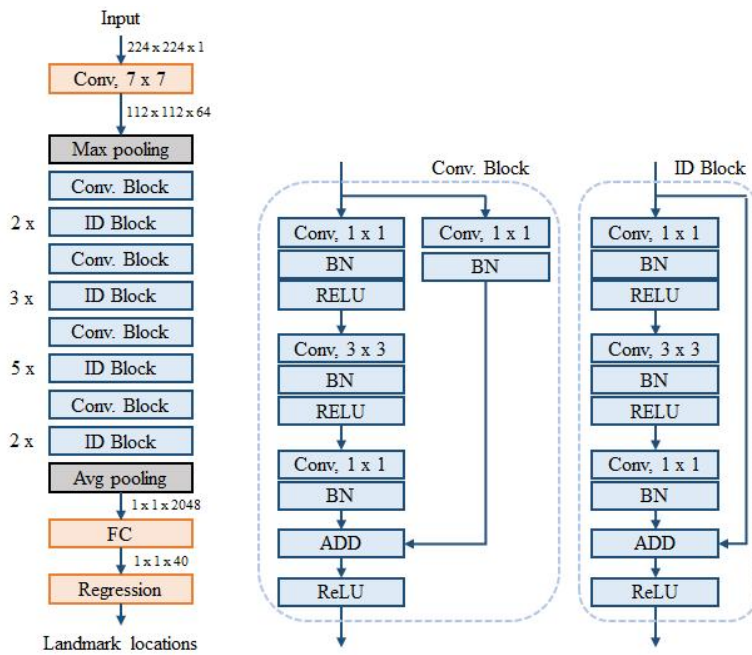
본 논문에서는 요추 부위를 촬영한 X-ray 영상에서 5개의 요추 추체(L1 ~ L5)를 검출하고, 추체 하나 당 4개의 특징점으로 총 20개의 특징점을 찾는 자동화된 알고리즘을 제시한다. 요추 L5와 흉추 T12는 형태적으로 구분이 쉽지 않으므로, 전신 척추가 모두 포함되지 않은 요추 부위 영상 입력



에서는 S1을 기준으로 5개 요추를 정의하는 것이 타당하다. 임상적으로는 요추화(lumbarization), 천추화(sacralization)로 인해 요추가 5개가 아닌 경우가 있으나, 본 연구에서는 모두 정상 요추로 가정하였다. 따라서 S1을 기준으로 위쪽으로 5개의 추체를 요추로 정의하였다.

2.2 특징점 검출을 위한 ResNet-50 기반 두 단계 네트워크

요추 특징점 검출을 위해서 CNN 모델중 다양한 응용분야에서 뛰어난 성능이 입증된 ResNet-50을 활용하였다 [12]. ResNet-50은 컬러 영상 입력에 대해서 영상 분류를 목적으로 만들어진 네트워크로, 입력은 3 channel 컬러영상이고 출력은 분류층(classification layer)을 가진다. 본 연구에서는 입력이 1 channel인 X-ray영상을 이용하고, 요추의 특징점 좌표 (x, y)를 출력하는 구조가 필요하므로, [그림 2]와 같이 ResNet-50 구조를 수정하여 사용하였다. 즉, 입력 및 바로 이어지는 7 x 7 합성곱 layer는 1 channel에 맞게 변경되었고, 해상도는 224 x 224로 유지하였다. 출력부의 경우에는 분류층 대신 특징점들의 좌표 출력을 위한 회귀층(regression layer)로 구성하였다. 즉, 네트워크는 224 x 224 해상도의 흑백 X-ray 영상 입력에 대해서 (x, y) 좌표들을 출력하는 구조이다. 원본 ResNet-50에서 수정된 부분은 [그림 2]에서 주황색으로 나타내었다.

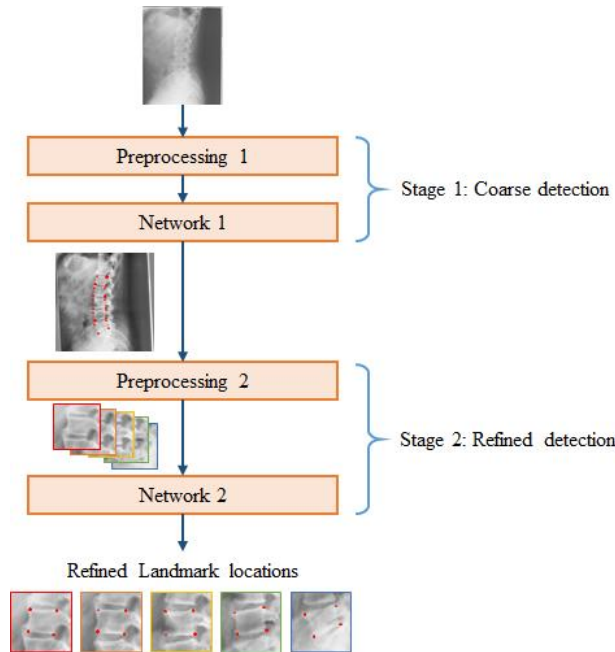


[그림 2] 특징점 검출을 위해 수정된 ResNet-50 구조

[Fig. 2] Modified version of ResNet-50 for landmark detection

일반적으로 객체 검출 방법은 ROI 선택을 위해서 영상 공간을 탐색하는 구조를 활용하는데 [13], [그림 2]의 구조는 특징점 좌표값을 직접 네트워크가 출력하는 구조이다. 이는 일반적인 객체 검출 문제와 달리 특정부위 진단을 위한 X-ray 영상에서는 입력 영상에서 추체들의 크기와 위치 변화가 한정적이기 때문에, 후보 영역을 탐색하는 구조를 사용하지 않고 비교적 단순한 구조의 네트워크를 활용하기로 한다. 하지만 요추 진단을 위한 X-ray 영상에서도 5개의 추체를 포함한 ROI의 위치 및 스케일 변화가 존재하므로 하나의 단순한 네트워크만으로는 검출의 강인성과 정확성을 모두 만족하기는 어렵다. 따라서 본 연구에서는 coarse-to-fine 접근 방법으로 두 단계 네트워크 구조의 검출기를 제안한다.

[그림 3]에 두 단계 네트워크 구조를 나타내었다. 첫 번째 단계는 전체 요추 특징점 검출을 위한 네트워크이다. 입력 영상을 기준 해상도인 224 x 224로 변환하고, 대비 제한 적응 히스토그램 평활화(Contrast-Limited Adaptive Histogram Equalization, CLAHE) [14] 를 Preprocessing 1에서 수행한다. Network 1은 5개의 추체 특징점 좌표를 출력하는 구조로서 [그림 2]에서 설명한 수정된 ResNet-50이다.



[그림 3] 요추 특징점 검출을 위해 제안한 두 단계 네트워크

[Fig. 3] Proposed two-stage network for lumbar vertebrae landmark detection

두 번째 단계는 개별 요추 특징점 검출을 위한 네트워크이다. 전체 요추 검출 네트워크에서 찾아진 각 추체 대표점 위치를 기준으로 ROI 를 설정하고 세밀하게 대표점 위치를 검출한다.



Preprocessing 2에서는 5개 추체에 대한 ROI를 설정하고 각 ROI를 224 x 224 영상으로 변환한다. 또한 CLAHE를 적용하여 Network 2에서 정교한 검출 작업을 수행한다. L1~ L5 까지 추체의 모양이 유사하므로, 추체 번호를 구분하지 않고 동일한 네트워크를 반복하여 사용한다. Network 2는 1개의 추체에 대한 특징점의 좌표 4개를 출력하므로 [그림 2]의 구조에서 출력 dimension만 8로 변경되고 동일한 네트워크 구조를 가진다.

3. 실험 결과

3.1 데이터

제안하는 알고리즘을 검증하기 위해서 미국 National Library of Medicine에서 제공하는 NHANES II(the second National Health and Nutrition Examination Surveys) 의 요추 X-ray 데이터를 활용하였다 [15][16]. NHANES II 영상 데이터는 아날로그 X-ray 영상을 디지털화 한 것으로 최근의 연구들 [3-11] 에서 활용한 영상에 비해서 명암, 잡음 등의 조건이 열악하며, 따라서 검출 알고리즘의 강인성과 정확도를 평가하기에 난이도가 높은 데이터이다.

전체 요추 특징점 검출에 활용한 데이터는 총 1,000장이며 수작업으로 특징점 참값(Ground Truth, GT)이 준비되었다. 원본 영상의 해상도는 2048 x 2487 인데, 좌우 패딩을 추가하여 2487 x 2487 영상으로 만든 후, 224 x 224 의 기준 해상도로 스케일 조절을 하였다.

데이터 개수를 증강하기 위해서 ± 25 pixel 범위에서 평행이동, $\pm 3^\circ$ 의 회전 및 0.9 ~ 1.1 배의 크기 변환 등을 통하여 총 15,000장의 영상을 생성하였다. 이 중에 13,500장은 학습 영상으로 1,500장은 테스트 영상으로 활용하였다.

또한, 개별 요추 특징점 검출을 위해서는 각 추체의 특징점 좌표를 이용하여 정사각형 형태의 ROI를 설정하였다. 특징점 4개의 평균으로 ROI의 중심점을 계산하였고, ROI의 한 변의 길이는 중심점과 중심점에서 가장 멀리 떨어져 있는 특징점의 길이를 기준으로 3.3배 크기로 설정하였다. 최종적으로 ROI를 224 x 224 의 기준 해상도로 정규화 하였다. L1 ~ L5 추체의 형태는 유사하므로 따로 객체 분류를 나누지 않고, 하나의 객체로 데이터를 준비하였다. 전체 요추 특징점 검출을 위한 데이터와 동일하게 데이터 증강을 통하여 만들어진 총 15,000장의 영상 중에서, 13,500장을 학습에 1,500장을 테스트에 사용하였다.

3.2 네트워크 학습 및 결과 분석

네트워크의 학습을 위해서 모멘텀을 사용한 확률적 경사하강법(stochastic gradient descent with momentum)을 사용하였고, [표 1]에 학습 파라미터를 정리하였다. Network 1과 Network 2 모두 동일

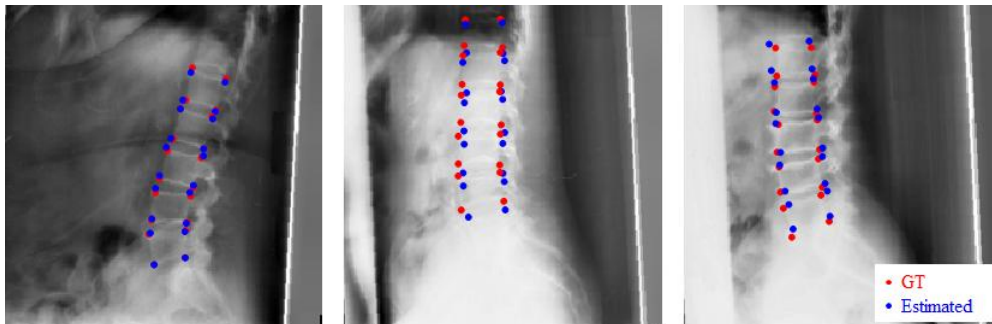
한 파라미터 조건으로 학습하였다. 매 epoch 마다 학습 데이터 순서를 섞어서 사용하였으며, learning rate는 $10^{-5} \sim 10^{-3}$ 범위에서 가변적으로 사용하였다.

[표 1] 네트워크 학습 파라미터

[Table 1] Parameters for network training

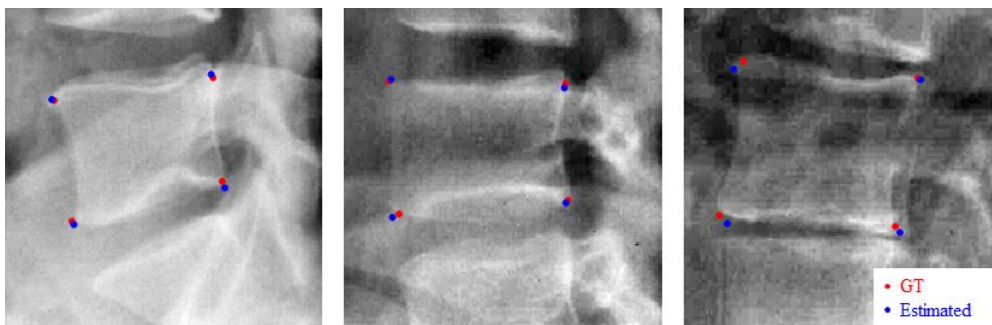
batch size	4
learning rate	$10^{-5} \sim 10^{-3}$
momentum	0.9
number of epoch	100

[그림 4]는 테스트 데이터에 대한 Network 1의 결과 예시를 보여준다. 붉은색 점이 참값이며, 푸른색 점이 검출 결과이다. Network 1에서는 Network 2를 학습하는데 사용된 영상과 유사한 위치와 크기로 ROI를 설정해 줄 수 있으면 된다. ROI의 중심점 위치 오차가 10 pixel 이하일 때를 성공적인 검출로 정의했을 때, 검출 성공률은 99.7% 이었다.



[그림 4] Network 1을 이용한 초기 특징점 검출 결과 예시

[Fig. 4] Examples of coarse landmark detection result using Network 1



[그림 5] Network 2을 이용한 정제된 특징점 검출 결과 예시

[Fig. 5] Examples of refined landmark detection result using Network 2



테스트 데이터에 대한 Network 2의 정제된 검출 결과 예시를 [그림 5]에 나타내었다. Network 1의 초기 검출 결과보다 위치 정확도가 개선되는 것을 시각적으로 확인할 수 있다.

각 요추 및 전체 특징점 위치 오차의 평균 및 표준편차를 [표 2]에 정리하였다. L2에서 가장 작은 오차 3.89 ± 2.48 를 보였으며, L5에서 가장 큰 오차 5.41 ± 3.38 가 관측되었다. 전체 특징점의 위치 오차는 4.54 ± 3.00 로 나타났다. 오차는 모두 224 x 224 해상도 기준의 픽셀 거리이다.

[표 2] 특징점 위치 오차

[Table 2] Position errors of landmarks

L1	L2	L3	L4	L5	전체
4.94 ± 3.60	3.89 ± 2.48	4.22 ± 2.48	4.25 ± 2.61	5.41 ± 3.38	4.54 ± 3.00

4. 결론

본 연구에서는 요추 추체의 특징점 검출을 위한 두 단계 CNN 구조를 제안하였다. 두 단계의 네트워크는 모두 특징점의 좌표값을 출력하도록 ResNet-50 기반의 회귀 모델로 설계되었다. 첫 번째 단계에서는 전체 5개의 요추 위치를 동시에 강인하게 찾는데 초점을 두었고, 두 번째 단계에서는 각각의 요추별로 특징점 위치의 정확도를 향상할 수 있도록 하였다. 이러한 coarse-to-fine 접근 방법이 잘 동작됨을 NHANES II 데이터에 적용하여 검증하였으며, 검출 성공률 99.7%, 특징점의 위치 오차 4.54 ± 3.00 픽셀의 우수한 결과를 얻었다.

특징점 검출을 위해서 입력 영상의 해상도를 고정한 상태에서 하나의 네트워크를 활용하는 것 [8] 보다, 두 단계로 이루어진 구조를 활용하는 것이 유리하다. 첫 번째 단계 네트워크에서는 초기 검출(coarse detection)을 강인하게 할 수 있으며, 두 번째 단계 네트워크에서는 상대적으로 고해상도의 영상에 대해 정밀한 검출(refined detection)을 수행하는데 집중할 수 있다.

추체의 특징점을 검출하는 기술은 여러 가지 임상지표를 추출하기 위한 가장 기본적이며 범용적으로 사용할 수 있는 기술이다. 빈번히 사용되는 임상지표인 Cobb 각도의 경우에 각도를 정의하는 두 특징점으로부터 각도를 계산하지 않고 여러 특징을 융합하거나 [8], 직접 각도값을 추정하는 방법이 더 정확할 수 있다 [9]. 하지만 Cobb 각도뿐만 아니라 척추전만각, 추체 전위 등 다양한 목적에 사용할 요소기술로서 특징점 추출의 중요성은 여전히 높다고 볼 수 있다. 특징점 추출 기술은 각도에 관련된 지표에 모두 사용이 가능하며, 영역 분할을 위한 시드 점(seed point)를 제공할 수도 있다. 제안하는 특징점 추출 방법이 기존의 추체 분석 연구에 적용된다면 전체 알고리즘의 성능 향상에 기여할 수 있을 것으로 생각된다.

References

- [1] M. Kim, J. Yun, Y. Cho, K. Shin, R. Jang, H. J. Bae, N. Kim, "Deep learning in medical imaging", *Neurospine*, vol. 16, no. 4, December 2019, pp. 657-668, doi: 10.14245/ns.1938396.198.
- [2] C. Qin, D. Yao, Y. Shi Z. Song, "Computer-aided detection in chest radiography based on artificial intelligence: a survey", *BioMedical Engineering OnLine*, vol. 17, no. 113, August 2018, doi: 10.1186/s12938-018-0544-y.
- [3] H. Wu, C. Bailey, P. Rasoulinejad, S. Li, "Automatic landmark estimation for adolescent idiopathic scoliosis assessment using boostNet", *Proceedings of the Medical Image Computing and Computer Assisted Intervention (MICCAI)*, September 11-13, 2017, Quebec City, Quebec, Canada, pp. 127-135, doi: 10.1007/978-3-319-66182-7_15.
- [4] S. Ebrahimi, L. Gajny, W. Skalli, E. Angelini, "Vertebral corners detection on sagittal X-rays based on shape modelling, random forest classifiers and dedicated visual features", *Computer Methods in Biomechanics and Biomedical Engineering Imaging & Visualization*, vol. 7, no. 2, May 2018, pp. 132-144, doi: 10.1080/21681163.2018.1463174.
- [5] S. Al-Arif, K. Knapp, G. Slabaugh, "Fully automatic cervical vertebrae segmentation framework for X-ray images", *Computer Methods and Programs in Biomedicine*, vol. 157, April 2018, pp. 95-111, doi: 10.1016/j.cmpb.2018.01.006.
- [6] M. H. Beak, J. G. Jeong, Y. J. Kim, K. G. Kim, J. Y. Jeon, "An Automated Vertebra Segmentation model based on Deep learning and an Application to Cobb angle Measurement based on Spine X-ray", *Journal of Next-generation Convergence Information Services Technology*, vol. 9, no.1, March 2020, pp. 1-9, doi: 10.29056/jncist.2020.03.01.
- [7] H. Wu, C. Bailey, P. Rasoulinejad, S. Li, "Automated comprehensive adolescent idiopathic scoliosis assessment using MVC-net", *Medical Image Analysis*, vol. 48, May 2018, pp. 1-11, doi: 10.1016/j.media.2018.05.005.
- [8] L. Wang, Q. Xu, S. Leung, J. Chung, Bo Chen, S. Li, "Accurate automated Cobb angles estimation using multi-view extrapolation net", *Medical Image Analysis*, vol. 58, December 2019, 101542, doi: 10.1016/j.media.2019.101542.
- [9] K. C. Kim, H. S. Yun, S. Kim, J. K. Seo, "Automation of Spine Curve Assessment in Frontal Radiographs Using Deep Learning of Vertebral-Tilt Vector", *IEEE Access*, vol. 8, May 2020, pp. 84618-84630, doi: 10.1109/ACCESS.2020.2992081.
- [10] H. Anitha, G. K. Prabhu, "Automatic quantification of spinal curvature in scoliotic radiograph using image processing", *Journal of Medical Systems*, vol. 36, no. 3, June 2012, pp. 1943-1951, doi: 0.1007/s10916-011-9654-9.
- [11] B. H. Cho, D. Kaji, Z. B. Cheung, I. B. Ye, R. Tang, A. Ahn, O. Carrillo, J. T. Schwartz, A. A. Valliani, E. K. Oermann, V. Arvind, D. Ranti, L. Sun, J. S. Kim, S. K. Cho, "Automated Measurement of Lumbar Lordosis on Radiographs Using Machine Learning and Computer Vision", *Global Spine Journal*, vol. 10, no. 5, August 2020, pp. 611-618. doi: 10.1177/2192568219868190.



- [12] K. He, X. Zhang, S. Ren, J. Sun, “Deep Residual Learning for Image Recognition”, Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, June 26-July 1, 2016, Las Vegas, Nevada, USA, pp. 770-778, doi: 10.1109/CVPR.2016.90.
- [13] J. R. R. Uijlings, K. E. A. van de Sande, T. Gevers, A. W. M. Smeulders, “Selective search for object recognition”, International Journal of Computer Vision, vol. 104, no. 2, April 2013, pp. 154-171, doi: 10.1007/s11263-013-0620-5.
- [14] Z. Karel, “Contrast Limited Adaptive Histogram Equalization”, in Graphic Gems IV, P. S. Heckbert, Eds., San Diego, CA, USA: Academic Press Professional, 1994, 474-485.
- [15] National Library of Medicine, “Download Digitized NHANES II X-ray Films”, nlm.nih.gov, <https://www.nlm.nih.gov/databases/download/nhanes.html>, (accessed December 11, 2018).
- [16] Z. Xue, S. Rajaraman, R. Long, S. Antani, G. Thoma, “Gender Detection from Spine X-Ray Images Using Deep Learning”, 2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS), June 18-21, 2018, Karlstad, Sweden, pp. 54-58, doi: 10.1109/CBMS.2018.00017.