

# 수정된 HRNet을 이용한 X-ray 영상의 흉추 분할 기법

이예은<sup>1</sup>, 이동규<sup>1</sup>, 정지훈<sup>1</sup>, 김형규<sup>2</sup>, 김호준<sup>1</sup>

<sup>1</sup>한동대학교 전산전자공학부

<sup>2</sup>(주)제이피아이헬스케어

leeyan413@naver.com, asas0409@nate.com,

jihoojo03@naver.com, kimhk@jpi.co.kr, hjkim@handong.edu

## Thoracic Spine Segmentation of X-ray Images Using a Modified HRNet

Ye-Eun Lee<sup>1</sup>, Dong-Gyu Lee<sup>1</sup>,

Ji-Hoon Jeong<sup>1</sup>, Hyung-Kyu Kim<sup>2</sup>, Ho-Joon Kim<sup>1</sup>

<sup>1</sup>School of Computer Science and Electrical Engineering,

Handong Global University

<sup>2</sup>JPI Healthcare Co., Ltd

### 요 약

인체의 흉부 X-ray 영상으로부터 척추질환과 관련된 의료 진단지표를 자동으로 추출하는 과정을 위하여 흉추조직의 정확한 분할이 필요하다. 본 연구에서는 HRNet 기반의 학습을 통하여 흉추조직을 분할하는 방법을 고찰한다. 분할 과정에서 영상 내의 상대적인 위치 정보가 효과적으로 반영될 수 있도록, 계층별로 영상의 고해상도의 표현이 그대로 유지되는 구조와 저해상도의 특징 지도로 변환되는 구조가 병렬적으로 연결되는 형태의 심층 신경망 모델을 채택하였다. 흉부 X-ray 영상에서 콥 각도(Cobb's angle)를 산출하는 문제를 대상으로 흉추 분할을 위한 학습 방법, 진단지표 추출 방법 등을 소개하며, 부수적으로 피사체의 위치 변화 및 크기 변화 등에 강인한 성능을 제공하기 위하여 학습 데이터를 증강하는 방법론을 제시하였다. 총 145개의 영상을 사용한 실험을 통하여 제안된 이 론의 타당성을 평가하였다.

### 1. 서론

X-ray 영상에서 병변이나 질환을 자동으로 판별 하는 과정은 그 사전단계로 영상 내의 조직 영역을 정확하게 분할하는 과정이 필요하다. 본 연구에서는 흉부 X-ray 영상으로부터 척추질환 진단지표인 콥 각도(Cobb's angle)[1]를 추출하는 과정을 위한 기반 기술로서 영상 분할 기법을 고찰한다. 의료 영상에서 조직 구조를 분할하는 문제의 특징은 영상 내에 나타나는 조직 영역 간의 상대적인 위치 정보에 민감(position-sensitive)하다는 점이다. 이러한 특성을 고려하여 계층별로 처리되는 과정에서 고해상도의 표현이 유지되는 구조와 저해상도로 변환되는 구조가 병렬적으로 결합된 딥러닝 모델을 제시한다. 또한 실용적 환경을 고려한 성능 향상을 위하여 X-ray 진단 과정에서 발생할 수 있는 피사체의 위치와 크기 변화 등에 강인한 성능을 제공할 수 있도록 학습 데이터를 증강하는 방법론을 제시하였다.

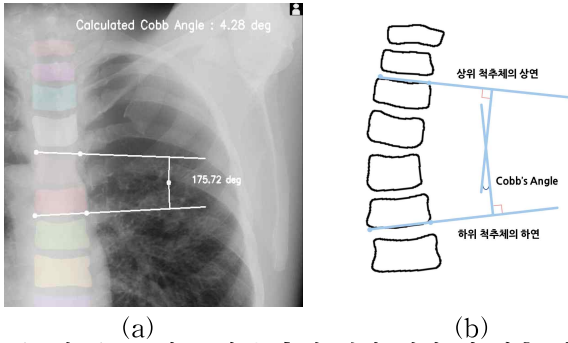
### 2. 콥 각도 산출을 위한 영상 분할 기법

흉부 X-ray 영상에서 판독되는 콥 각도(Cobb's angle)는 척추측만증 진단에 사용되는 지표로, 흉추 영역의 분할 결과로부터 그 값을 산출할 수 있다. (그림 1)에 보인 바와 같이, 먼저 만곡의 첨부(apex)를 기준으로 하여 가장 기울기가 큰 상위 척추체의 상연과 하위 척추체의 하연에서의 연장선을 고려한다. 식 (1)의  $N$ 을 만곡을 이루는 척추체의 총 개수,  $m_k$ 는 그 중  $k$ 번째 척추체의 기울기라 할 때, 두 연장선의 수직선이 이루는 콥 각도( $\varphi$ )는 임의의 두 쌍의 기울기 조합  $(m_i, m_j)$ 이 이루는 각도 중 최댓값으로 산출하게 된다.

본 연구에서는 X-ray 영상에서 이와 같은 진단 지표를 자동추출하는 과정에서 흉추의 세부 영역을 분할하는 방법으로서 HRNet[2] 기반의 딥러닝 모델을 채택하였다. (그림 2)에 보인 바와 같이 HRNet은 기존의 인코더-디코더 기반 모델에 비해 고해상도 표현을 전체 연산 과정 동안 유지한다는 특징을 갖는다.

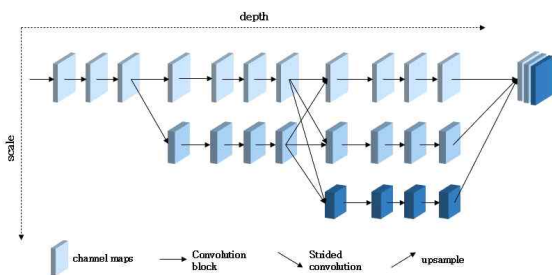
$$\varphi = \max_{i,j} \left| \tan^{-1} \left( \frac{m_i - m_j}{1 + m_i \cdot m_j} \right) \right| \quad (1)$$

$$(i,j) \in \{(a,b) | 0 \leq a \leq N-2, 2 \leq b \leq N \text{ and } b-a \geq 2\}$$



(그림 1) 굽 각도 자동 측정 결과 예시 및 산출 과정

이는 기존의 모델에서 해상도를 감소시키는 형태로 특징지도를 생성하고, 이를 다시 고해상도 정보로 복원하는 과정에서 픽셀 단위의 세부 정보가 소실될 수 있다는 단점을 보완할 수 있게 한다. 또한 의료 영상의 분할 문제는 각 조직의 영역이 일정 범위의 위치에 나타난다는 특성을 갖는데, 제안된 모델은 분할 과정에서 조직 영역 간 상대적 위치에 대한 정보를 효과적으로 반영할 수 있도록 한다. 즉 계층 간 처리 과정에서 고해상도의 표현을 유지한 채로 저해상도 컨볼루션을 통해 추출한 특징지도를 병렬적으로 연결하여 위치 정보를 보존한다. 또한 다양한 해상도에서 추출한 특징지도를 상호 결합하여 서로 다른 수용장의 크기를 가진 특징 정보를 활용하도록 함으로써 영상 분할의 정확도를 높인다.



(그림 2) 조직 구조 분할을 위한 딥러닝 네트워크의 구조

식 (2)의  $\phi_{n \times n}(\cdot)$  함수는 네트워크의 단위 블록에 사용되는 기본적인 연산으로  $(n \times n)$  컨볼루션 결과에 정규화를 거친 후 정류된 선형함수( $ReLU$ )를 이용해 특징지도를 추출한다. 식 (3)의  $Bottleneck(x)$  함수는  $\phi_{1 \times 1}(\cdot)$  연산을 이용해 채널 수를 감소시킨 후  $\phi_{3 \times 3}(\cdot)$  연산으로 특징지도를 추출한다. 이어서 다시  $Conv_{1 \times 1}$  연산을 통해 본래

채널 수로 증가시켜 정규화를 거친다.

$$\phi_{n \times n}(\cdot) = ReLU(BN(Conv_{n \times n}(\cdot))) \quad (2)$$

$$Bottleneck(x) = ReLU(BN(Conv_{1 \times 1}(\phi_{3 \times 3}(\phi_{1 \times 1}(x)))) + x) \quad (3)$$

### 3. 학습 데이터 생성 기법

심층 신경망을 의료 영상에 적용하는 응용에서 구현상의 제약은 충분한 양의 학습 데이터를 확보하기가 어렵다는 것이다. 학습 데이터의 부족으로 인한 과적합을 방지하고, 일반화된 성능을 보이기 위하여 학습 데이터 증강 기법의 필요성이 제기된다. 또 X-ray 영상에서 다양한 요인에 의해 피사체의 위치와 크기가 변화할 수 있는데, 이러한 변이에 강인한 성능을 보일 수 있도록 데이터 증강 기법을 적용하였다.

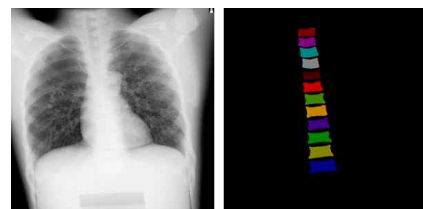
본 연구에서는 명암, 좌표, 스케일 등의 변이에 대하여 사전에 정의된 데이터 증강 연산들을 일정 규칙에 따라 무작위로 선택하여 적용하는 방법으로 구현하였다. (그림 3)은 증강된 데이터 영상의 예시로, 좌표 이동 및 CLAHE를 적용한 결과를 보였다.



(그림 3) 데이터 증강기법을 적용한 결과 예시

### 4. 실험 결과 및 고찰

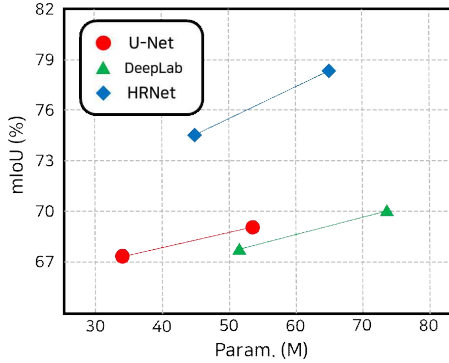
제안된 이론의 타당성을 검증하기 위하여 흉부 팬텀 모형을 촬영한 X-ray 데이터 집합 65쌍과 정면 흉부 X-ray 데이터인 NIHCC CXR[3] 80쌍, 총 145쌍을 학습 데이터셋으로 사용하여 실험을 수행하였다. 제시된 영상 데이터는  $1994 \times 2430$  12bit DICOM 형식의 영상과  $1024 \times 1024$  해상도의 png 형식의 영상이다.



(그림 4) 흉부 영상에서 척추 조직 분할 결과 예시

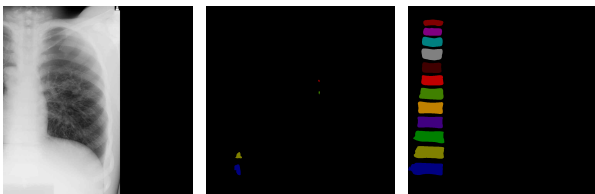
첫 번째 실험은 (그림 4)에 예시한 바와 같이 흉추 분할 기법에 대한 성능평가로서, 기존의 인코더-디코더 기반 모델인 U-Net[4], DeepLabv3[5]와의 성능을 상호 비교하는 실험을 수행하였다. 정량적

평가 척도로 평균 합집합 대비 교차점 비율(mIoU, mean Intersection over Union)을 적용하였는데, 실험 결과 (그림 5)에 보인 바와 같이 제안된 HRNet 기반 모델이 기존의 모델과 비교하여 파라미터 수 대비 정량적으로 우수한 성능을 보임을 알 수 있다.



(그림 5) 인체 영상 분할 성능 비교

두 번째 실험은 데이터 증강 기법의 유효성에 대한 실험으로, 피사체의 위치 및 크기에 일정량의 변화를 준 영상에 대하여 분할 성능을 평가하였다. 동일한 분할 문제에 대하여 원 데이터 집합만을 사용한 경우와 매 반복마다 증강 데이터 형식을 무작위로 선택하여 학습한 경우를 상호 비교하였다.



(그림 6) 데이터 증강 기법의 실험 결과 예시

테스트 데이터 집합에 대하여 데이터 증강이 적용된 경우의 mIoU는 71.34%로, 적용되지 않은 경우의 70.27%보다 높은 정량적 수치를 기록하였다. 또한 (그림 6)에 보인 바와 같이 위치 변화가 발생한 입력 영상 (a)에 대하여 원 데이터만을 사용한 경우, (b)에 보인 것처럼 흉추 분할에 실패했지만, 증강 기법을 적용한 결과 (c)에서는 흉추 영역의 정확한 분할에 성공하였다.

**5. 결론**

의료 영상으로부터 진단지표를 자동으로 산출하는 응용의 유용성은 조직 영역에 대한 분할 기법의 성능에 영향을 받는다. 전형적인 의료 영상에서 각 조직 영역의 위치가 일정한 형태로 나타나며 영역의 위치와 크기는 많은 경우 진단지표를 판단하는 기준이 된다. 이에 본 연구에서는 딥러닝 기반의 영상

분할 기법을 제시하였으며 그 유용성을 실험적으로 평가하였다. 또한 진단 과정에서 발생할 수 있는 영상 내 피사체의 형태 변화에도 안정적인 성능을 위하여 위치이동, 반전, 크기 변화, 명암 변화 등의 변이를 무작위로 적용하는 데이터 증강 기법을 적용하였다. 제안된 모델에 대하여 흉부 X-ray 영상으로부터 척추질환 진단지표의 하나인 굽 각도를 자동 산출하는 문제에 적용함으로써 실용적인 성능 향상을 시도하였다. 향후 연구에서는 보다 다양한 진단지표를 고려하여 조직 구조를 효과적으로 분할하는 방법론과 다양한 피사체에 대하여 일반화된 성능을 갖는 분할 방법에 대한 연구가 필요한 것으로 판단된다.

※ 본 연구는 과학기술정보통신부와 정보통신기술진흥 센터의 소프트웨어중심대학지원사업 (2017-0-00130)의 지원을 받아 수행되었음.

**참고문헌**

[1] Ming-Huwi Horng, Chan-Pang Kuok, Min-Jun Fu, Chii-Jen Lin, Yung-Nien Sun, "Cobb Angle Measurement of Spine from X-Ray Images Using Convolutional Neural Network", Computational and Mathematical Methods in Medicine, vol. 2019, Article ID 6357171, 18 pages, 2019. <https://doi.org/10.1155/2019/6357171>

[2] K. Sun, Y. Zhao, B. Jiang, T. Cheng, B. Xiao, D. Liu, Y. Mu, X. Wang, W. Liu, and J. Wang, "High-resolution representations for labeling pixels and regions," 2019, arXiv:1904.04514.

[3] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri and R. M. Summers, "ChestX-Ray8: Hospital-Scale Chest X-Ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 3462-3471, doi: 10.1109/CVPR.2017.369.

[4] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," Lecture Notes in Computer Science, pp. 234 - 241, 2015.

[5] L.-C. Chen, G. Papandreou, F. Schroff, and H. Adam, "Rethinking atrous convolution for Semantic Image segmentation," arXiv:1706.05587v3.