

A Thoracic Spine Segmentation Technique for Automatic Extraction of VHS and Cobb Angle from X-ray Images

Ye-Eun Lee[†] · Seung-Hwa Han^{††} · Dong-Gyu Lee^{††} · Ho-Joon Kim^{†††}

ABSTRACT

In this paper, we propose an organ segmentation technique for the automatic extraction of medical diagnostic indicators from X-ray images. In order to calculate diagnostic indicators of heart disease and spinal disease such as VHS(vertebral heart scale) and Cobb angle, it is necessary to accurately segment the thoracic spine, carina, and heart in a chest X-ray image. A deep neural network model in which the high-resolution representation of the image for each layer and the structure converted into a low-resolution feature map are connected in parallel was adopted. This structure enables the relative position information in the image to be effectively reflected in the segmentation process. It is shown that learning performance can be improved by combining the OCR module, in which pixel information and object information are mutually interacted in a multi-step process, and the channel attention module, which allows each channel of the network to be reflected as different weight values. In addition, a method of augmenting learning data is presented in order to provide robust performance against changes in the position, shape, and size of the subject in the X-ray image. The effectiveness of the proposed theory was evaluated through an experiment using 145 human chest X-ray images and 118 animal X-ray images.

Keywords : Medical Image Processing, Cobb Angle, Image Segmentation, Thoracic Spine Segmentation

X-ray 영상에서 VHS와 콥 각도 자동 추출을 위한 흉추 분할 기법

이 예 은[†] · 한 승 화^{††} · 이 동 규^{††} · 김 호 준^{†††}

요 약

본 논문에서는 X-ray 영상에서 의료 진단지표를 자동으로 추출하기 위한 조직분할 기법을 제안한다. 척추질환이나 심장질환에 대한 진단지표로서, 흉추-심장 비율이나 콥 각도 등의 지표를 산출하기 위해서는 흉부 X-ray 영상으로부터 흉추, 용골 및 심장의 영역을 정확하게 분할하는 과정이 필요하다. 본 연구에서는 이를 위하여 계층별로 영상의 고해상도의 표현과 저해상도의 특징지도로 변환되는 구조가 병렬적으로 연결되는 형태의 심층신경망 모델을 채택하였다. 이러한 구조는 영상에서 세부 조직의 상대적인 위치정보가 분할 과정에 효과적으로 반영될 수 있게 한다. 또한 픽셀 정보와 객체 정보가 다단계의 과정으로 상호 작용되는 OCR 모듈과, 네트워크의 각 채널이 서로 다른 가중치 값으로 반영되도록 하는 채널 어텐션 모듈을 결합하여 학습 성능을 개선할 수 있음을 보인다. 부수적으로 X-ray 영상에서 피사체의 위치 변화, 형태의 변형 및 크기 변이 등에도 강인한 성능을 제공하기 위하여 학습데이터를 증강하는 방법을 제시하였다. 총 145개의 인체 흉부 X-ray 영상과, 총 118개의 동물 X-ray 영상을 사용한 실험을 통하여 제안된 이론의 타당성을 평가하였다.

키워드 : 의료영상처리, 콥 각도, 영상 분할, 흉추 분할

1. 서 론

인체의 흉부 X-ray 영상으로부터 판별하는 콥 각도(Cobb

angle)와 동물 X-ray 영상에서 측정하는 흉추-심장 비율(VHS: vertebral heart scale)은 각각 척추질환 및 심장질환을 진단하는 중요한 지표가 된다. 그런데 이들 지표를 자동으로 추출하기 위해서는 영상에서 흉추, 용골 및 심장 등 세부 조직의 영역을 정확하게 분할하고, 그 크기와 방향 및 구조를 분석하는 기술이 필요하다. 영상분할 기법은 전형적인 영상 처리 기법의 하나로서 오랜 기간 동안 축적된 다양한 방법론과 이론들이 알려져 있는데, 최근에는 딥러닝 기반 기법을 사용하여 영상의 영역을 자동으로 분할하는 방법들이 제안되고 있다[1-3]. 일반적인 영상 분할 기법으로서의 심층신경망 모델로는 U-Net[4], DeepLab[5] 등을 들 수 있으며, 의료영상

※ 이 논문은 과학기술정보통신부의 소프트웨어 중심대학 지원사업의 지원을 받아 수행하였음(2017-0-00130).

※ 이 논문은 2022년 한국정보처리학회 ASK 2022의 우수논문으로 "수정된 HRNet을 이용한 X-ray 영상의 흉추 분할 기법"의 제목으로 발표된 논문을 확장한 것임.

† 준 회 원 : 한동대학교 전산전자공학부 학사과정

†† 비 회 원 : 한동대학교 전산전자공학부 학사과정

††† 중신회원 : 한동대학교 전산전자공학부 교수

Manuscript Received : July 18, 2022

Accepted : August 29, 2022

* Corresponding Author : Ho-Joon Kim(hjkim@handong.edu)

분할 분야에서는, MRI 및 CT영상으로부터 인체의 조직구조를 분할하는 문제에서 비교사학습기법을 이용한 영상분할 기법[6], 측면 촬영한 X-ray 영상으로부터 요추와 경추를 판별하는 2단계 구조의 신경망 모델[7], 의료영상분할 문제에서 불완전한 데이터만으로 학습이 가능한 모델에 관한 연구[8], CT 영상에서 폐암조직을 분할하는 방법으로서 다중해상도 형태의 특징표현 관한 연구[9] 등이 발표되었다.

본 연구에서는 X-ray 영상으로부터 의료진단지표를 자동으로 산출하는 방법론으로서 조직구조 분할기법을 제시한다. 이는 수정된 HRNet 모델[10]을 기반으로 하는 딥러닝 기법으로 구현되는데, 저자의 이전 연구에서 콥 각도 산출 문제와 VHS 추출 문제에 적용한 기초연구 내용을 소개한 바 있다 [11, 12]. 본 연구에서 제안하는 의료영상 분할기법은 세부적으로 HRNet 모델의 기본 구조에 OCR(Object Contextual Representations) 모듈[13]과 채널 어텐션 모듈[14]을 추가한 형태로 구성되는데 이들은 각각 영상에서 각 픽셀의 정보와 객체 영역의 정보를 생성하고 이를 3단계의 과정으로 자기강조(self-attention)하는 과정과, 각 채널이 서로 다른 가중치로 반영되도록 하는 과정을 수행한다. 이러한 신경망 모델을 기반으로, 인체의 흉부 X-ray 영상으로부터 척추질환 진단지표인 콥 각도를 추출하는 문제와 동물 X-ray 영상으로부터 VHS를 추출하는 응용에 적용되는 영상 분할 문제를 고려하였다. 의료영상에서 조직구조를 분할하는 문제의 특징 중 하나로 영상 내에 나타나는 조직영역 간의 상대적인 위치 정보에 민감하다는 점을 들 수 있다. 이러한 특성을 고려하여 계층별로 처리되는 과정에서 고해상도의 표현이 유지되는 구조와 저해상도로 변환되는 구조가 병렬적으로 상호 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 제시한다.

2. VHS와 콥 각도 자동 산출 기법

2.1 VHS 지표 자동 산출 기법

본 연구에서는 우선 동물병원에서 개나 고양이와 같은 동물의 심장질환 진단을 위한 지표로서 VHS값을 자동으로 산출하는 문제를 고려하였다. 이 과정은 Fig. 1에 보인 바와 같이 조직분할 과정과 지표산출 과정의 2단계로 이루어진다.

Fig. 1에 보인 바와 같이 영상에서 흉추의 세부구조와 심장영역, 용골의 위치가 추출되면 이로부터 VHS값을 산출한다. 용골의 위치로부터 심장의 정점까지 최장 길이를 가지는 장축(L)을 찾고, 이에 수직 방향으로 길이가 최대인 단축(S)을 찾아 각각의 길이를 측정한다. 이어서 이 두 축에 대하여 각각 제4흉추의 위치를 기준으로 몇 개의 흉추를 포함하는지 측정된 값을 합산함으로써 VHS값이 결정된다.

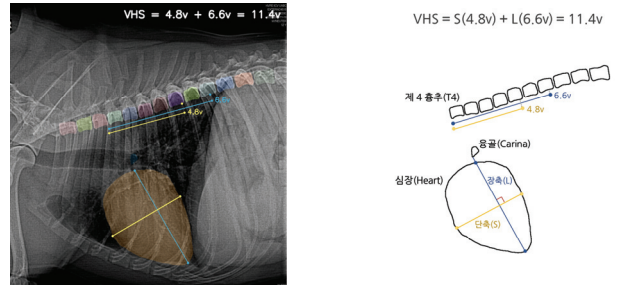


Fig. 1. An Example of VHS Measurement Process

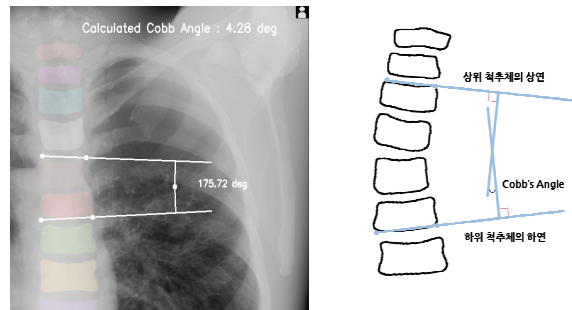


Fig. 2. An example of Cobb angle measurement process

2.2 척추질환 진단을 위한 콥 각도 자동 산출 기법

흉부 X-ray 영상으로부터 산출하는 콥 각도(Cobb angle)는 척추측만증 진단에 사용되는 지표로, 흉추 영역의 분할 결과로부터 그 값을 측정할 수 있다. Fig. 2에 보인 바와 같이, 먼저 만곡의 첨부(apex)를 기준으로 하여 가장 기울기가 큰 상위 척추체의 상연과 하위 척추체의 하연에서의 연장선을 추출한다. Equation (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(φ)는 임의의 두 쌍의 기울기 조합 (m_i, m_j) 이 이루는 각도 중 최댓값으로 산출하게 된다.

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

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

3. 조직구조 분할을 위한 딥러닝 모델

3.1 조직구조 분할 기법

앞 절에서 기술한 바와 같이 VHS 및 콥 각도 값을 산출하는 과정에서 그 정확도는 앞서 수행하는 흉추와 심장 영역 및 용골의 위치의 추출과정이 얼마나 정확한가에 좌우된다. 이에 본 연구에서는 이러한 영상 분할 과정에 초점으로 두어 그 방법론을 고찰하였다. 본 연구에서는 X-ray 영상에서 이와 같은 진단지표를 자동추출하는 과정에서 흉추의 세부 영역을 분할하는 방법으로서 HRNet 기반의 딥러닝 모델을 채택하였다.

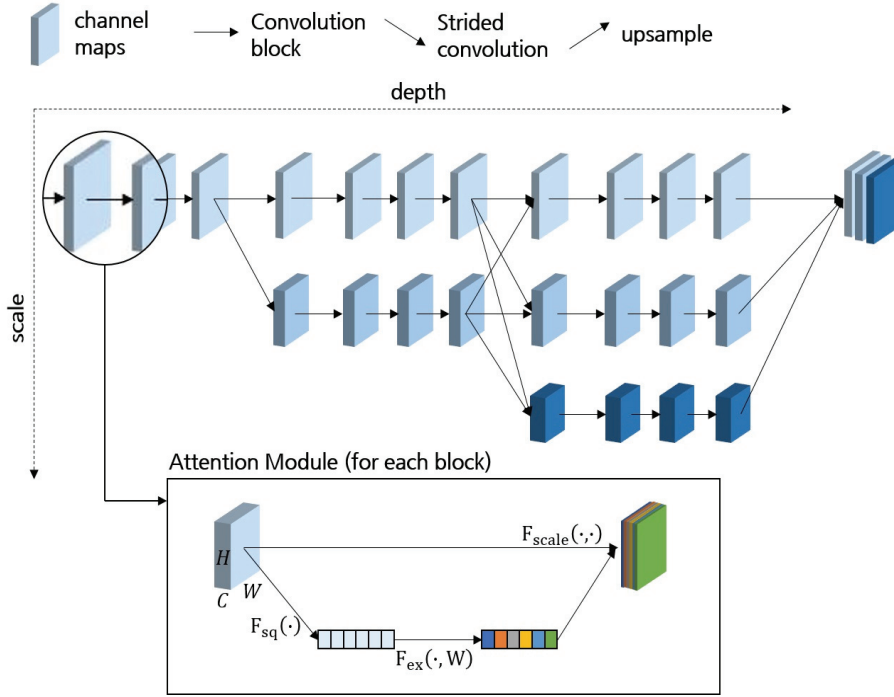


Fig. 3. The Deep-Learning Network for the Organ Segmentation

Fig. 3에 보인 바와 같이 HRNet은 기존의 인코더-디코더 기반 모델에 비해 고해상도 표현을 전체 연산 과정 동안 유지한다는 특징을 갖는다. 이는 기존의 모델에서 계층에 따라 점진적으로 해상도를 감소시키는 형태로 특징지도를 생성하고, 이를 다시 고해상도 정보로 복원하는 과정에서 픽셀 단위의 세부 정보가 소실될 수 있다는 단점을 보완할 수 있게 한다. 의료영상의 분할 문제에서 각 조직의 영역은 일정한 범위의 위치에 나타나게 되는데, 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 위치정보를 보존하고, 다양한 해상도에서 추출한 특징지도를 상호 결합하여 서로 다른 수용장의 크기를 가진 특징 정보를 활용하도록 함으로써 영상분할의 정확도를 높인다.

$$\phi_{n \times n}(\cdot) = \text{ReLU}(\text{BN}(\text{Conv}_{n \times n}(\cdot))) \quad (2)$$

$$\begin{aligned} \text{Bottleneck}(x) \\ = \text{ReLU}(\text{SE}(\text{BN}(\text{Conv}_{1 \times 1}(\phi_{3 \times 3}(\phi_{1 \times 1}(x))))) + x) \end{aligned} \quad (3)$$

$$\begin{aligned} \text{BasicBlock}(x) \\ = \text{ReLU}(\text{SE}(\text{BN}(\text{conv}_{3 \times 3}(\phi_{3 \times 3}(x))))) + x \end{aligned} \quad (4)$$

Equation (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 특징지도

를 추출한다. Equation (3)의 $\text{Bottleneck}(x)$ 함수는 $\phi_{1 \times 1}(\cdot)$ 연산을 이용해 채널 수를 감소시킨 후 $\phi_{3 \times 3}(\cdot)$ 연산으로 특징지도를 추출한다. 이어서 다시 $\text{Conv}_{1 \times 1}$ 연산을 통해 본래 채널 수로 증가시켜 정규화를 거친다. 해당 결과를 채널 어텐션 모듈($\text{SE}(x)$)에 적용하여 중요 채널을 강조한 후 잔차 학습을 적용하여 정류된 선형함수를 이용해 학습 초반의 특징지도를 추출한다. Equation (4)의 $\text{BasicBlock}(x)$ 함수는 연산을 통해 특징지도를 추출하여 컨볼루션 연산을 적용하는 과정이다. 이에 정규화를 거친 뒤 잔차학습을 적용하여 채널 어텐션을 수행한다. 이후 정류된 선형함수를 이용해 각 해상도에 따른 특징지도를 추출한다.

3.2 OCR 모듈 및 채널 어텐션 모듈 구조

본 논문에서는 기존의 전형적인 HRNet 모델의 기본 구조에 OCR(Object Contextual Representations) 모듈과 채널 어텐션 모듈을 결합한 모델을 구현하였다. OCR 모듈은 Fig. 4에 보인 바와 같이 한 픽셀의 라벨은 결국 해당 픽셀이 위치한 영역의 라벨이라는 사실에서 착안된 개념으로, 픽셀과 영역 사이의 관계를 통하여 단계적으로 픽셀의 라벨을 유추해 나간다. 그 과정은 총 3단계로 이루어지는데, 먼저 백본 신경망을 통해 얻은 특징지도를 바탕으로 클래스 수만큼의 개괄적인 객체 영역을 추출한다. 이는 주어진 K 개의 클래스에 대하여 각 픽셀이 클래스에 속하는 정도를 나타낸 K 개의 행렬로 표현된다. 다음으로 각각의 객체영역에 있는 픽셀의 값을 합하여 객체별 영역 표현을 추정한다. 주어진 클

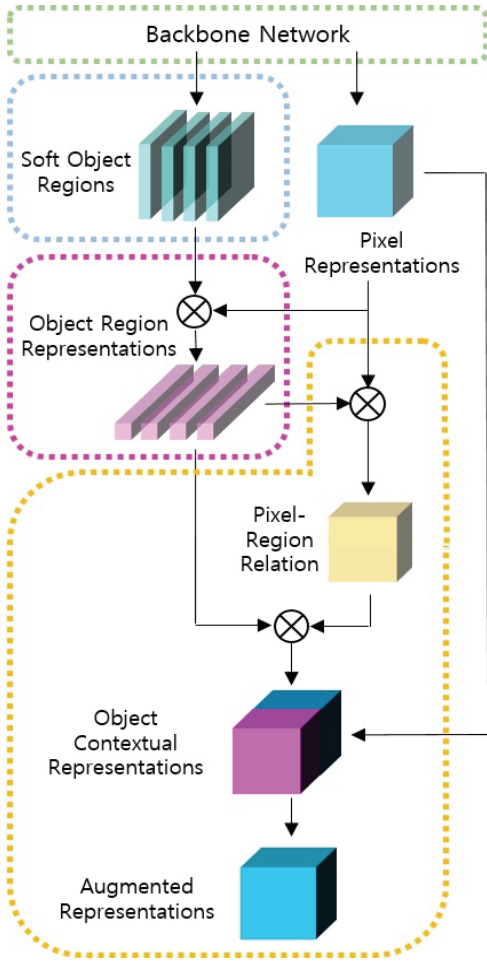


Fig. 4. The Structure of the OCR Pipeline

래스의 k 번째 객체별 영역 표현인 f_k 는 Equation (4)와 같이 입력 특징지도에 개괄적인 객체영역을 가중치로 변환한 $\text{softmax}(x_k)_i$ 를 적용하여 생성된다. 이어서 OCR 결과를 연산한 뒤, 이를 기존의 픽셀 정보와 총합하여 증강된 표현을 생성한다. Equation (5)의 y_i 는 i 번째 픽셀의 OCR 결과로, 객체별 영역 표현에 i 번째 픽셀과 k 번째 객체 간의 가중치인 $\text{softmax}(\phi_{1 \times 1}(x)^\top \cdot \phi_{1 \times 1}(f_k))_i$ 를 곱한 값으로 결정된다.

$$f_k = \sum_i \text{softmax}(x_k)_i \cdot x_i \quad (5)$$

$$y_i = \phi_{1 \times 1} \left(\sum_{k=1}^K \text{softmax}(\phi_{1 \times 1}(x)^\top \cdot \phi_{1 \times 1}(f_k))_i \cdot \phi_{1 \times 1}(f) \right) \quad (6)$$

이와 같은 과정은 기존 네트워크의 픽셀 분류가 해당 픽셀의 주변 정보만을 활용하여 결과를 도출했던 것과 달리, 픽셀에 해당하는 객체 정보를 이용하여 픽셀의 클래스 표현을 강화함으로써 분할과정의 성능을 개선할 수 있게 한다.

채널 어텐션은 여러 개의 특징지도가 서로 다른 중요도를 가지고 있다는 점을 활용하여 상대적으로 중요한 특징을 가진 채널은 강조하고 유용하지 않은 채널은 억제한다. 이를 위하여 먼저 전역 평균 풀링을 사용하여 각 채널을 대표하는 스칼라값을 구하고, 각 채널의 대푯값을 통해 채널 사이의 의존성을 계산한다. 두 개의 완전연결 계층에 활성화함수를 적용하고, 이를 대응하는 특징지도에 채널별 가중치가 부과되도록 재조정한다. 해당 네트워크는 계산 복잡도가 상대적으로 낮아 적은 오버헤드로도 성능을 높일 수 있으며, 기존의 다른 심층신경망에 유연하게 적용 가능하다는 장점이 있다. 특히 영상분할 과정에서 초기 추상적인 표현을 추출한 이후 픽셀 단위의 클래스 분류를 위한 구체적인 특징을 찾아내는 것에 효과적이다.

3.3 학습데이터 생성

심층신경망을 의료영상에 적용하는 응용은 충분한 양의 학습데이터를 확보하는 데에 어려움이 있다. 학습데이터의 부족으로 인한 과적합을 방지하고, 일반화된 성능을 보이기 위하여 학습데이터 증강기법의 필요성이 제기된다. 또 X-ray 영상에서 다양한 요인에 의해 피사체의 위치와 크기가 변화할 수 있는데, 이러한 변이에 강한 성능을 보일 수 있도록 데이터 증강기법을 적용하였다. 본 연구에서는 명암, 좌표, 스케일 등의 변이에 대하여 사전에 정의된 데이터 증강 연산을 일정 규칙에 따라 무작위로 선택하여 적용하는 방법으로 구현하였다. Fig. 5는 증강된 데이터 영상의 예시로, 좌표 이동 및 CLAHE를 적용한 예를 보였다.

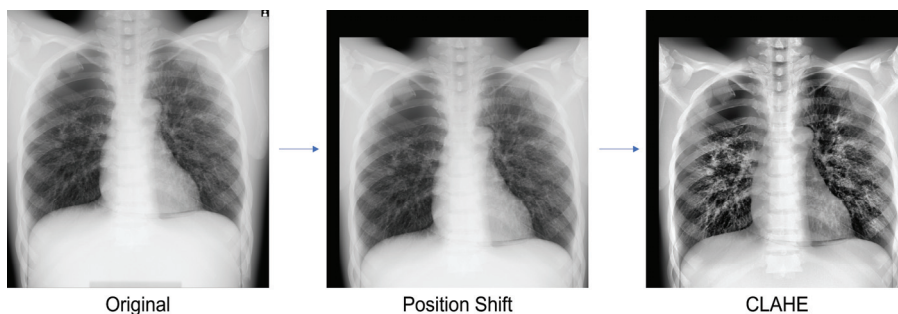


Fig. 5. Examples of the Data Augmentation Results

4. 실험 결과 및 고찰

제안된 이론의 타당성을 검증하기 위하여 인체 흉부 팬텀 X-ray 영상 65개와 NIHCC CXR[15] 영상 80개로 이루어진 총 145개의 인체영상 데이터와, 총 118개의 동물 X-ray 영상을 사용하였다. 인체 X-ray 영상 집합은 콥 각도 추출 문제에 적용하였고, 동물 흉부 X-ray 영상은 VHS 지표 값을 자동 추출하는 실험에 적용하였다. 실험에 사용된 영상데이터는 (1994×2430) 해상도의 12bit DICOM 형식의 영상과 (1024×1024) 해상도의 png 형식의 영상이다. 실험을 위하여 전체 영상 집합에서 65%의 영상을 무작위로 선택하여 학습에 사용하였고, 나머지 35%의 영상을 성능 테스트에 적용하였다. Fig. 6과 Fig. 7은 각각 콥 각도 측정과 VHS 산출 문제를 위한 학습데이터의 예이다. 그림에 보인 바와 인체의 흉부 X-ray 영상으로부터 총 12개의 흉추조직을 세분하여 분할하며, 동물 X-ray 영상에서는 VHS 산출에 필요한 13개의 흉추, 용골 및 심장 영역을 분할한다. 분할기법의 정량적 성능 평가를 위하여 파라미터 수, 연산량(GFlops: GPU-Floating point operations), 평균 합집합 대비 교차점 비율(mIoU: mean Intersection over Union) 등의 척도를 적용하였다.

첫 번째 실험에서는 기존의 알려진 인코더-디코더 기반 모델인 U-Net과 DeepLabv3를 사용한 경우와 본 논문에서 제안한 기법을 사용한 경우의 결과를 비교하였다. Fig. 8은 콥 각도 측정을 위한 영상분할 문제에 대하여, 제안된 모델을 사용한 경우와 기존의 모델의 결과를 파라미터의 수에 대비한 mIoU 척도 평가 변화로 비교한 것이다. 그림에 보인 바와 같이 파라미터의 양과 관계 없이 제안된 모델이 기존의 두 모델에 비하여 우수한 분할 성능을 보임을 알 수 있다.

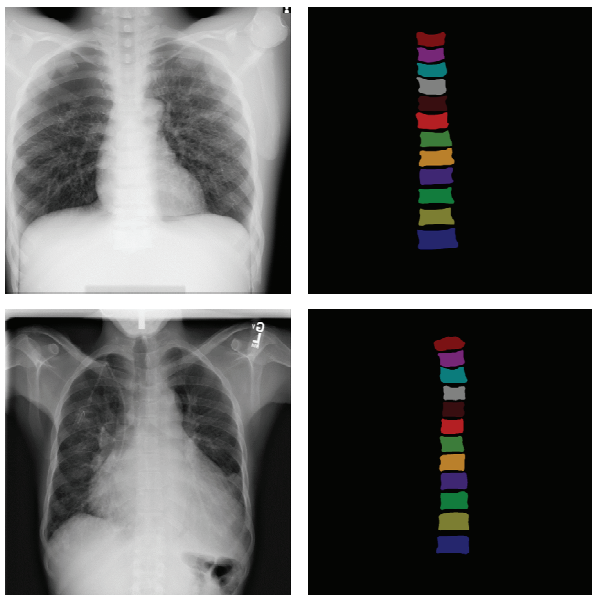


Fig. 6. Examples of the Training Data for Cobb Angle Measurement

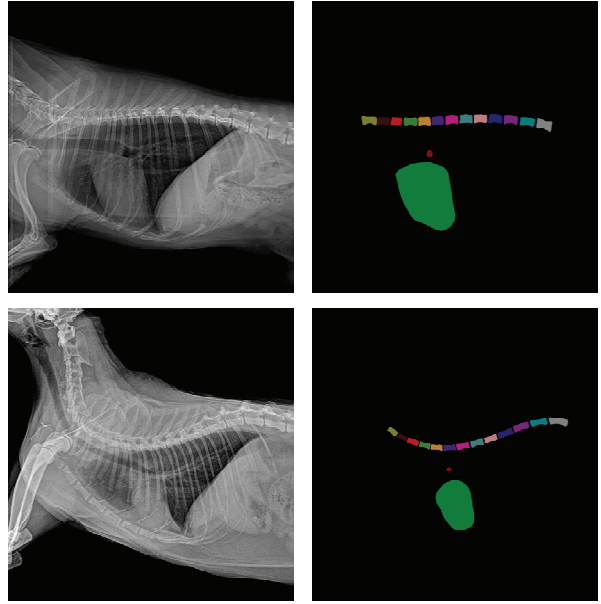


Fig. 7. Examples of the Training Data for VHS Measurement

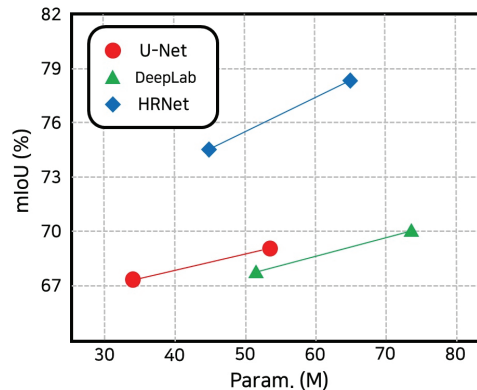


Fig. 8. Performance Comparison According to the Number of Parameters in the Thoracic Segmentation Experiments

Table 1은 VHS 지표값을 산출을 위한 영상 분할 실험 결과를 파라미터의 수, 연산량(GFlops)과 mIoU 척도를 구분하여 비교한 것이다. 표에 보인 바와 같이 제안된 모델이 기존의 두 모델에 비하여 파라미터 수 및 연산량 대비 우수한 성능을 보임을 알 수 있다. 이러한 결과는 고해상도 연산을 유지하고 저해상도의 특징추출을 병렬적으로 조밀하게 연결한 구조가 주어진 영상분할 문제에 효과적임을 시사한다.

두 번째 실험으로 본 논문에서 제안한 어텐션 모듈 및 OCR 모듈의 유용성을 평가하기 위하여 각 모듈의 적용 여부에 따른 실험 결과를 상호 비교하였다. Fig. 9에 보인 바와 같이 기존의 HRNet을 사용한 결과에 비해 OCR 모듈을 결합한 모델이 상대적으로 위치가 떨어진 객체에 대한 오분류 정도가 감소된 것을 확인할 수 있다. 또한 채널 어텐션모듈이 추가된

Table 1. Results of the VHS Measurement Experiments

Model	Parameter	GFlops	mIoU
U-Net	53.40M	207.03G	0.6021
DeepLabv3	73.09M	80.58G	0.4293
HRNet	65.85M	86.96G	0.7521

Table 2. Quantitative Evaluation Results of Using the OCR Module and the Attention Module

	HRNet	HRNet+OCR	HRNet+Att.
Mean IoU	0.6906	0.7522	0.7197
Mean Precision	0.7937	0.8430	0.8046
Mean Recall	0.8336	0.8699	0.8633
Mean F1-Score	0.8111	0.8543	0.8314

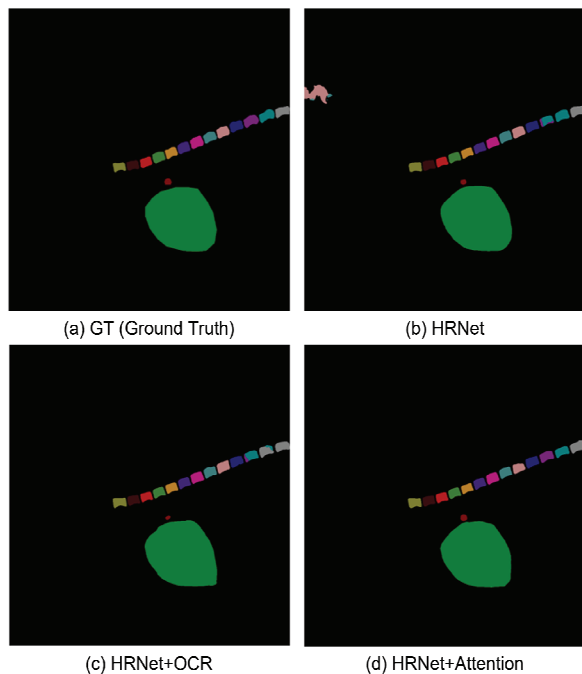


Fig. 9. Results of Using OCR Module and Attention Module

모델에서는 구체적인 세부 조직에 대한 분할의 정확도가 증가함을 보였다. 동일 실험에 대하여 4가지 종류의 정량적 지표를 사용하여 성능을 비교한 결과를 Table 2에 보였다. 표에 보인 바와 같이 OCR 및 어텐션 모듈을 추가한 모델이 기존의 HRNet에 비하여 4가지 척도를 모두 개선함을 보였다.

이러한 결과는 특징지도 생성 과정에서 객체의 배경영역 정보와 픽셀 간의 관계를 반영함으로써 분할 성능을 개선할 수 있게 하며, 영상의 분할 과정에서 상대적으로 중요한 특징을 가진 채널 및 객체영역 표현을 선별적으로 학습하는 구조가 의료영상에서의 조직분할 문제에 효과적으로 적용될 수 있음을 시사한다.

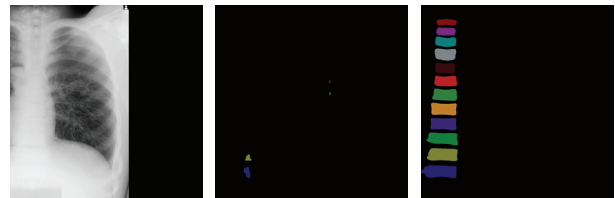


Fig. 10. Experimental Results Applying the Data Augmentation

세 번째 실험은 데이터 증강기법의 유효성에 대한 실험으로, 앞 절에서 기술한 것처럼 피사체의 위치 및 크기에 일정량의 변화를 준 영상에 대하여 분할 성능을 평가하였다. 이는 학습주기의 설정에 따라 증강연산을 무작위로 선택하여 적용하는 형식으로 동작하게 되는데, 이전 실험과 동일한 데이터 집합에 대하여 원 데이터 집합만을 사용한 경우와 매 반복마다 데이터 증강기법을 적용한 경우를 상호 비교하였다. 실험 결과, 테스트 데이터 집합에 대하여 데이터 증강이 적용된 경우의 mIoU는 0.713으로, 적용되지 않은 경우의 0.702보다 높은 정량적 수치를 기록하였다. 또한 Fig. 10에 보인 바와 같이 위치 변화가 발생한 입력 영상 (a)에 대하여, 원 데이터만을 사용한 경우의 결과는 (b)에 보인 것처럼 흉추분할에 실패했지만, 증강기법을 적용한 결과 (c)는 흉추의 세부 조직을 정확하게 분할하였다.

5. 결론

인체와 동물의 흉부 X-ray 영상에서 세부 조직의 위치는 일정 범위 내에서 각 영역의 상대적 위치에 종속적인 특징을 갖는다. VHS나 콧 각도를 자동으로 측정하는 문제를 위하여 흉추의 세부구조, 용골 및 심장의 영역을 정확하게 분할하기 위하여 이러한 의료영상의 특징을 잘 반영하는 모델을 고려할 필요가 있다. 본 연구에서 제안한 수정된 구조의 HRNet

모델은 분할의 위치 종속적인 특성을 잘 반영할 뿐만 아니라, OCR 및 어텐션 모듈의 특징을 추가하여 분할 성능을 개선하였다. 제안된 분할기법의 유용성을 평가하기 위하여 조직과 배경 영역의 정확한 구분을 측정하는 정량적 척도를 적용하였으며 기존의 딥러닝 기반 영상분할 모델과 성능을 비교하였다. 또한 진단과정에서 발생할 수 있는 영상 내 피사체의 형태 변화에도 안정적인 성능을 유지하기 위하여 위치 이동, 반전, 크기 변화, 명암 변화 등의 변이를 무작위로 적용하는 데이터 증강기법을 적용하였고, 그 유효성을 실험을 통해 평가하였다. 향후 연구에서는 간질환 진단을 위한 VLAS 지표 등 보다 다양한 진단지표를 고려하여 조직구조를 효과적으로 분할하는 방법론과 다양한 피사체에 대하여 일반화된 성능을 갖는 분할방법에 대한 연구가 필요한 것으로 판단된다.

References

- [1] J. Yu, D. Yang, and H. Zhao, "FFANet: Feature fusion attention network to medical image segmentation," *Bio-medical Signal Processing and Control*, Vol.69, pp.102-912, 2021.
- [2] Y. Hong, H. Pan, W. Sun, and Y. Jia, "Deep dual-resolution networks for real-time and accurate semantic segmentation of road scenes," *arXiv preprint arXiv:2101.06085*, 2021.
- [3] H. Yan, C. Zhang, and M. Wu, "Lawin transformer: Improving semantic segmentation transformer with multi-scale representations via large window attention," *arXiv preprint arXiv:2201.01615*, 2022.
- [4] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pp.234-241, 2015.
- [5] H. Wang, Y. Zhu, H. Adam, A. Yuille, and L.-C. Chen, "Max-deeplab: End-to-end panoptic segmentation with mask transformers," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5463-5474, 2021.
- [6] S. Bhattacharya, S. Shahnawaz, and A. Bhattacharya, "ClamNet: Using contrastive learning with variable depth Unets for medical image segmentation," *arXiv preprint arXiv:2206.05225*, 2022.
- [7] Y. Chen, Y. Mo, A. Readie, G. Ligozio, T. Coroller, and B. W. Papiez, "VertXNet: Automatic segmentation and identification of lumbar and cervical vertebrae from spinal X-ray images," *arXiv preprint arXiv:2207.05476*, 2022.
- [8] N. Tajbakhsh, L. Jeyaseelan, Q. Li, J. N. Chiang, Z. Wu, and X. Ding, "Embracing imperfect datasets: A review of deep learning solutions for medical image segmentation," *Medical Image Analysis*, Vol.63, p.101-693, 2020.
- [9] J. Jiang et al., "Multiple resolution residually connected feature streams for automatic lung tumor segmentation for CT images," *IEEE Transactions on Medical Imaging*, Vol.38, No.1, pp.134-144, 2019.
- [10] K. Sun et al., "High-resolution representations for labeling pixels and regions," *arXiv preprint arXiv:1904.04514*, 2019.
- [11] Y. Lee, D. Lee, J. Jeong, H. Kim, and H. Kim, "Thoracic spine segmentation of X-ray images using a modified HRNet," in *Proceedings of the Annual Spring Conference of Korea Information Processing Society Conference (KIPS) 2022*, Vol.29, pp.705-707, 2022.
- [12] S. Han, S. Hong, Y. Lee, D. Lee, K. Kim, and H. Kim, "A deep learning technique to automatically extract VHS from X-ray images," in *Proceedings of Korea Computer Congress 2022 (KCC 2022)*, Vol.29, pp.2075-2077, 2022.
- [13] Y. Yuan, X. Chen, and J. Wang, "Object-contextual representations for semantic segmentation," in *European Conference on Computer Vision (ECCV)*, pp.173-190, 2020.
- [14] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.7132-7141, 2018.
- [15] 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," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.2097-2106, 2017.



이 예 은

<https://orcid.org/0000-0003-4933-1081>

e-mail : leeyan413@naver.com

2020년 ~ 현 재 한동대학교

전산전자공학부 학사과정

관심분야 : 인공지능, 컴퓨터비전



한 승 화

<https://orcid.org/0000-0001-5446-8934>

e-mail : hans0525@naver.com

2020년 ~ 현 재 한동대학교

전산전자공학부 학사과정

관심분야 : 의료영상처리, 인공지능



이 동 규

<https://orcid.org/0000-0001-5187-1816>

e-mail : asas0409@nate.com

2018년 ~ 현 재 한동대학교

전산전자공학부 학사과정

관심분야 : 컴퓨터비전, 딥러닝



김 호 준

<https://orcid.org/0000-0002-8951-8441>

e-mail : hjkim@handong.edu

1995년 KAIST 전산학과(공학박사)

1987년 ~ 1991년 한국원자력연구소 연구원

2003년, 2010년 미국 캘리포니아주립대학교
방문교수

1996년 ~ 현 재 한동대학교 전산전자공학부 교수

관심분야 : 컴퓨터비전, 인공지능, 의료영상처리, 패턴인식