

뇌-종양 영역 분할 개선을 위한 계층적 3차원 어텐션 U-Net 모델

노연우, 노원종*, 이상웅

가천대학교, *한림대학교

nyw0207@gachon.ac.kr, *wonjong.noh@hallym.ac.kr, slee@gachon.ac.kr

Hierarchical 3D Attention U-Net For Improving Brain-Tumor Segmentation

Yeonwoo Noh, Wonjong Noh*, Sang-Woong Lee

Gachon University, *Hallym University

요약

본 논문은 뇌-종양 부위를 자동으로 찾아내기 위한, 어텐션 U-Net 기반의 새로운 머신러닝 학습 모델을 제시한다. 기존 연구와 다르게, U-Net에서 정보를 디코딩을 할 때, 계층적 어텐션을 적용하여, 거시적으로 추상화된 이미지 정보와 미시적으로 추상화된 이미지 정보를 모두 이용하도록 설계하였으며, 이 때 적용되는 어텐션 필터는 공간축면과 채널 축면에서 어텐션 기능을 동시에 지원하는 3차원 필터를 적용하도록 설계하였다. 이를 통해, 정상 영역의 학습 데이터가 대부분을 차지하는 데이터 불균형 상황에서도, 종양 영역에 대한 세그멘테이션의 기능이 향상되도록 하였다. 기존 cascaded-CNN [2], U-Net [3], 어텐션 U-Net [4] 연구들과의 비교 실험을 통해, 제안하는 학습 모델은 dice, 민감도, 특이도 지표가 평균적으로 각각 11%, 7%, 12% 향상됨을 확인하였다.

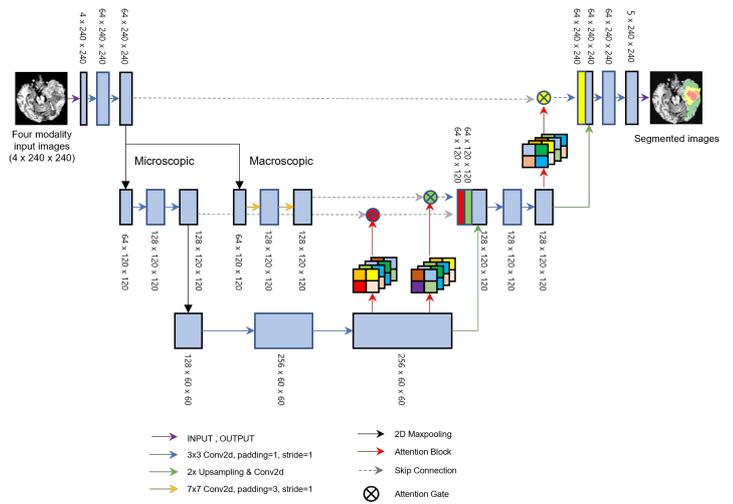
I. 서론

신경교종(Glioma)은 사람에게 가장 흔한 유형의 뇌종양이며, 정확한 종양의 세그멘테이션은 다중 모드 자기 공명 영상(MRI)에서 다양한 모양과 모양으로 인해 가장 어려운 의료 영상 분석 작업 중 하나이다. 이러한 뇌 종양을 수동으로 분할하려면 많은 의학적 전문 지식이 필요하고 시간이 많이 걸리며 인적 오류가 발생하기 쉽다. 또한 수동 프로세스는 일관성과 재현성이 부족하여 결과에 부정적인 영향을 미치고 궁극적으로 잘못된 예 후 및 치료로 이어질 수 있다. 최근 심층 신경망(Deep Neural Network, DNN) 알고리즘의 급속한 발전은, 컴퓨터 기반 자동 의료영상 분석에 큰 가능성을 보여주고 있다. 예를 들어, CNN(Convolutional Neural Networks)의 빠른 발전은 이미지 분류 또는 분할에서 기계학습모델이 인간 수준의 성능에 접근하거나 능가 하고 있다. 훈련된 인간 수준의 성능을 능가할 수 있는 심층 신경망(DNN) 모델을 얻는 것은 임상에게 질병 감지, 치료 계획 및 모니터링에 대한 보다 정확하고 신뢰할 수 있으며 표준화된 접근 방식을 가능하게 할 것이다 [1].

뇌종양 세그멘테이션을 위한, 최근의 대표적인 DNN 연구로는, cascaded-CNN을 이용한 모델 [2], U-Net 기반의 인코더-디코더 아키텍처 모델 [3], transformer의 개념을 적용한 어텐션 U-Net 모델 [4]들이 제시되었다. 본 연구에서는 위 대표적인 연구들의 성능을 개선하기 위하여, 현재 최고 수준인 어텐션 U-Net 모델을 기반으로 하는 새로운 향상된 모델을 제시하고, 실험을 통해 성능향상을 입증하였다.

II. 제안하는 이중-계층적 뇌-세그멘테이션 방식

본 논문에서 제안하는 '계층적 3D 어텐션 U-Net (계층적 3D 어텐션 U-Net)' 세그멘테이션 모델은 다음 [그림 1]과 같다. 기존 연구들과 비교하여, 첫 번째 특징은, 어텐션 U-Net에서 종양의 영역을 세부적으로 세그멘테이션할 수 있는 능력을 강화하기 위해, large-scale의 추상화된(압축된) 정보와 small-scale로 추상화된(압축된) 정보를 모두 이용하는 방식을 이용하였다.



[그림 1] 제안하는 계층적 3D 어텐션 U-Net 학습 모델

이를 위해, 거시적인 영상 압축은 7x7 convolution 필터를 사용하였으며, 미시적인 영상 압축은 3x3 convolution 필터를 적용하였다. 위와 같은 학습 모델 구조는, 정상 영역에 대한 데이터가 상대적으로 훨씬 많이 존재하는, 데이터 불균형 상황에서, 정상 영역의 데이터에 의해 학습에 bias가 생기는 것을 막아주는 역할을 할 수 있게 된다. 두 번째 특징으로는, 3차원 어텐션 필터를 적용하였다. 공간(spatial-wise) 및 채널(channel-wise) 측면에서 모두 어텐션 U-Net이 적용된다는 것이다. 이를 통해, 뇌 영상의 다양한 입력 modality에 대하여 어텐션을 적용할 영역을 좀 더 적응적으로 그리고 유연하게 결정하는 것이 가능해진다.

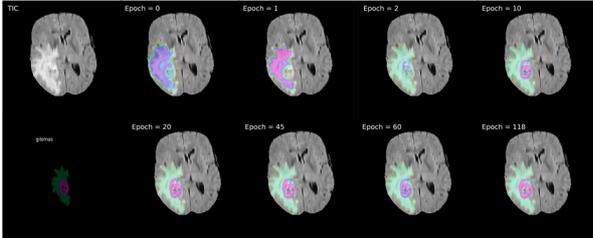
III. 실험 결과

본 실험은, 데이터로 BRAT2013의 공개 데이터셋을 사용하였으며, 최적

화 방법으로 확률경사강하법(SGD)가 사용되었고, 교차엔트로피 함수가 손실함수로 사용되었다. 본 논문에서 제안하는 방식은 아래와 같다.

(1) 어텐션 계수 학습 및 영향

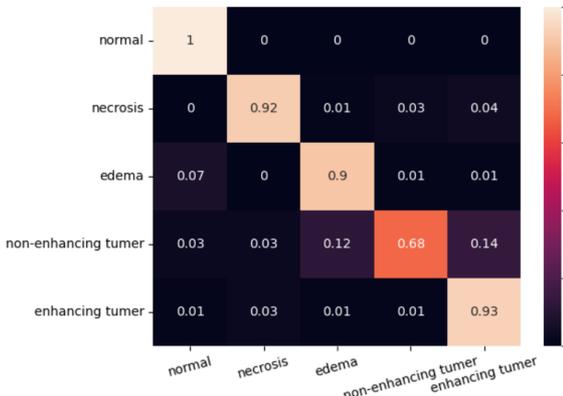
본 연구에서 제안하는 기계학습 모델의 학습 과정을 보여 주고 있다. [그림 2]에서는 어텐션에 의해, 종양의 영역이 점점 더 명확하게 학습되고 있음을 확인하였다.



[그림 2] 세그멘테이션 학습

(2) Confusion matrix를 측정

각 necrosis, enhancing tumor, non-enhancing tumor, edema로 이루어지는 종양 영역에 대하여, 각 세부 영역에 대한 예측 분포도를 [그림 2]에서 확인하였다. 각 영역에 대하여, 68% - 93% 사이의 예측 정확도를 확보하였으며, non-enhancing tumor에 대한 예측도가 가장 낮았는데, 이는 학습 데이터가 부족해서 생긴 현상으로 해석이 된다.



[그림 3] 세부 종양 영역에 대한 예측 분포 행렬

(3) 비교 실험

본 실험에서는, 종양 영역에 대해, 아래와 같이 정의된 Dice, 민감도, 특이도를 통하여, 성능을 측정한다.

- Dice : $2 * |P_1 \cap T_1| / (|P_1| + |T_1|)$
- 민감도 : $|P_1 \cap T_1| / |T_1|$
- 특이도 : $|P_0 \cap T_0| / |T_0|$

여기서, P는 모델이 예측한 값을 나타내며, T는 실제값을 의미한다. 좀 더 구체적으로, P₁과 P₀는 종양 영역에 대해, 예측된 레이블이 positive와 negative 값을 갖는 voxel들을 의미한다. T₁과 T₀는 종양 영역에 대해, truth 레이블이 positive와 negative 값을 갖는 voxel들을 의미한다. 평가 지표로서, 'Dice' 는 실제 종양 영역과 예측된 종양 영역을 직접 비교하여, 겹치는 영역이 차지하는 비율을 의미하며, 민감도는 실제 종양 영역을 종양 영역으로 올바르게 분류한 비율을 의미하고, 특이도는 실제로 종양이 아닌 영역을 종양이 아닌 것으로 올바르게 분류한 비율을 의미한다. 본 논문에서는, 또한, 뇌의 종양 영역을 아래와 같이 중분류하여, 각각의 영역에

대한 예측 정확도를 평가하였다.

- Enhancing := necrosis + enhancing tumor
- Core := necrosis + enhancing tumor + non-enhancing tumor
- Complete := necrosis + enhancing tumor + non-enhancing tumor + edema

본 논문에서 제안하는 방식을 기존의 대표적인 연구들([2]-[4])과 비교 평가하였으며 (표 1), 비교군의 평균 대비, 기존 U-Net 구조에 제안하는 3D 어텐션 필터만 적용했을 때는, Dice, Specificity, Sensitivity가 각각 11%, 7%, 11% 향상되었음을 확인하였으며, 거시적 영상과 미시적 영상을 동시에 이용하는 '계층적 3D 어텐션 U-Net'을 적용했을 때는, Dice, Specificity, Sensitivity가 각각 12%, 7%, 13% 향상되었음을 확인하였다.

[표 1] 성능 비교

지표 방법론	Dice			특이도			민감도		
	Enhancing	Core	Complete	Enhancing	Core	Complete	Enhancing	Core	Complete
Cascaded CNN [2]	0.880	0.790	0.730	0.890	0.790	0.680	0.870	0.790	0.800
U-Net [3]	0.680	0.647	0.737	0.999	0.999	0.999	0.585	0.509	0.630
어텐션 U-Net [4]	0.879	0.844	0.886	0.999	0.999	0.999	0.855	0.834	0.921
3D 어텐션 U-Net	0.896	0.874	0.908	0.999	0.999	0.999	0.869	0.847	0.893
Proposed	0.912	0.892	0.919	0.999	0.999	0.999	0.888	0.870	0.910

IV. 결론

본 논문에서는 효과적인 뇌-종양 세그멘테이션을 위한 새로운 어텐션 U-Net 기반 모델을 제안하였다. 뇌 종양의 세그멘테이션 학습에 방해가 되는 데이터 불균형 현상을 효과적으로 해결하기 위해, 거시적으로 추상화된 영상 정보와 미시적으로 추상화된 영상 정보를 모두 이용하면서, 공간 측면과 채널 측면에서 어텐션 기능을 동시에 지원하는 3차원 어텐션 필터를 이용하는 '계층적 3D 어텐션 U-Net' 모델을 설계하였다. 기존 cascaded-CNN, U-Net, 어텐션 U-Net 연구들과의 비교 실험을 통해, 세그멘테이션 관련 주요 성능 지표가 크게 향상됨을 확인하였다. 향후 연구로서, 종양 세그멘테이션에서 데이터 불균형을 효과적으로 해결하기 위해, 최근의 확산(diffusion)기반 데이터 생성 모델을 추가적으로 활용하는 U-Net 모델을 연구하고자 한다.

참고 문헌

[1] M. FuFutrega, M., Milesi, A., Marcinkiewicz, M., Ribalta, 'Optimized U-Net for Brain Tumor Segmentation,' Lecture Notes in Computer Science, vol 12963, Springer, Cham, 2022

[2] Mohammad Havaei, Axel Davy, David Warde-Farley, Antoine Biard, Aaron Courville, Yoshua Bengio, Chris Pal, Pierre-Marc Jodoin, Hugo Larochelle, 'Brain Tumor segmentation with Deep Neural Networks,' Medical Image Analysis, vol 35, pp. 18-31, 2017

[3] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. 'U-net: Convolutional networks for biomedical image segmentation,' International Conference on Medical image computing and computer-assisted intervention, Springer, Cham, 2015.

[4] Ozan Oktay, Jo Schlemper, Loic Le Folgoc, Matthew C. H. Lee, Mattias P. Heinrich, Kazunari Misawa, Kensaku Mori, Steven G. McDonagh, Nils Y. Hammerla, Bernhard Kainz, Ben Glocker, Daniel Rueckert, 'Attention U-Net: Learning Where to Look for the Pancreas,' CoRR abs/1804.03999, 2018