

U-Net 을 이용한 의료 영상 이진 분할

임수빈¹, 김준석¹, 강미선²

¹순천향대학교 컴퓨터소프트웨어공학과 학부생

²순천향대학교 컴퓨터소프트웨어공학과 교수

0414subin@naver.com, gyk2043@gmail.com, ms.kang@sch.ac.kr

Binary Segmentation of Medical Images using U-Net

Subin Lim¹, Junseok Kim¹, Misun Kang¹

¹Dept. of Computer Software Engineering, Soonchunhyang University, Asan Korea

요 약

컴퓨터 비전 및 영상 처리 분야에서 정확하고 신뢰할 수 있는 분할을 달성하는 것은 중요하다. 특히 의료 분야에서는 분할의 부정확성이 부적절한 특징의 추출과 분류를 야기한다. 이 논문에서는 이중 어텐션 메커니즘을 U-Net 아키텍처에 통합하여 이진 영상 분할을 개선하는 방법을 제안한다. 인코더와 디코더 사이의 중요 접합점에 위치 및 채널 어텐션 모듈을 배치하여 기능 표현을 향상시켜, 벤치마크 실험에서 기존 방법보다 우수한 성능을 확인하였다. 접근 방식은 의료 영상에서의 이진 분할 작업에 대한 잠재력을 보여주며, 자동 분할을 통해 의료 전문가의 효율성을 높일 수 있다.

1. 서론

의료 분야에서 영상 내 정확하고 신뢰할 수 있는 관심영역 추출에 도달하는 것은 중요한 과제이다. 이 논문은 이중 어텐션 메커니즘을 통합한 새로운 U-Net 기반 접근법을 제안하여 의료 영상의 이진 분할 정확도를 향상시키는 방법을 제안한다. 이 방법은 위치 어텐션 모듈과 채널 어텐션 모듈을 사용하여 인코더와 디코더 사이의 연결점을 최적화하며, 이를 통해 특징 추출을 개선하고 강력한 분할 성능을 발휘할 수 있도록 한다. 제안된 방법은 안과 질환 및 피부 병변 데이터 세트를 대상으로 하였으며, 기존 방법과의 비교에서 우수한 성능을 확인하였다. 본 논문은 관련 연구, 방법, 실험 결과, 결론으로 구성되어 있다.

2. 관련 연구

컴퓨터 비전에서 분할 기술은 의료 영상, 자율 주행, 객체 인식 등 다양한 응용 분야에서 중요한 역할을 하고 있으며, 전통적인 임계값, 에지 검출, 영역 기반 방법은 간단하지만 복잡한 구조를 처리하는 데 한계가 있다. 최근 딥러닝 기반 아키텍처인 FCN[1], DeepLab[2], U-Net[3], SegNet[4] 등이 등장하면서 분할 작업의 정확성과 적응력이 크게 향상되고 있다. 특히, U-Net 은 지역 및 전역적 특징을 효과적으로 반영하

는 설계로 주목받지만, 디코더에서의 업-컨볼루션으로 인해 정확한 공간 정보가 손실될 수 있는 문제가 있다.

최근 의료 영상 분할 분야에서는 이 혁신을 이끌었습니다. 전통적인 방법론과 달리 CNN, U-Net, Res-Unet, DeepLab, Attention U-net 등의 딥러닝의 도입을 통해 분할 정확도를 크게 향상시켰다. 이러한 방법들은 종양 진단 등 중요한 의료 영역에서 효과를 입증했으며, 질병 진단 및 치료 계획에 실질적인 영향을 미치고 있다.

최근 연구에서는 딥러닝 모델의 성능을 높이기 위해 어텐션 메커니즘을 통합하는 접근이 주목받고 있다. 이 메커니즘을 모델이 입력 데이터에서 중요한 부분에 집중하도록 도와 영상 분할, 분류, 객체 감지 등에서 성능을 향상시킨다. 어텐션 U-Net[5]과 CBAM 어텐션 메커니즘을 활용해 분할 정확도를 높였으며[6], DA-Net 은 위치 어텐션과 채널 어텐션을 결합해 더욱 세분화된 정보를 효과적으로 사용하여 기능 표현을 개선하였다[7].

3. 제안된 방법

본 논문에서는 이진 영상 분할을 위한 듀얼 어텐션 U-Net 아키텍처를 제안한다. 이 모델은 의료 영상 분할의 정밀도를 높이기 위해 U-Net 아키텍처에 위치 어텐션과 채널 어텐션 모듈을 통합한다. 인코더와 디코더 사이에 이중 어텐션 모듈을 배치하여 중요한 공간 및 채널 정보를 학습하고 활용하도록 설계되었다.

이 접근 방식은 복잡한 세부 사항을 효과적으로 이용하고 분할 정확도를 향상시키는 데 기여한다.

4. 실험

본 연구에 사용된 두 가지 데이터는 눈 질환 데이터와 피부 병변 데이터이다.

A. 데이터 정보

1) 눈 질환 데이터

본 연구에 사용된 첫 번째 데이터는 Kaggle 에서 가져온 눈 질환 데이터이다. 1044 개의 눈 질환 이미지는 400×400 의 고정 해상도로 크기가 조정되며, 불룩한 눈, 백내장, 교차된 눈, 녹내장 및 포도막염의 다섯가지 그룹으로 분류된다. 본 연구에서는 해당 데이터에 수직 및 수평, 회전 기법을 통해 데이터 증강을 수행하였다.

2) 피부 병변 데이터

이 연구에 사용된 두 번째 데이터는 국제 피부 영상 협업(ISIC)의 피부 병변 데이터이다. 다양한 피부 병변 및 상태를 다루는 실측 레이블이 있는 2000 장의 영상이 포함되어 있고 각 영상은 256×256 의 고정 해상도로 표준화되어 있다.

B. 성능 평가

제안된 방법의 성능을 평가하기 위해서 기존 방법과 제안된 방법을 적용한 결과의 정확도를 정량적으로 비교하기 위해 픽셀 정확도, 다이스 계수, 평균 결합 교차점(mIoU) 및 민감도와 같은 평가 지표를 사용하였다. 픽셀 정확도는 이진 분할에서 정확하게 예측된 픽셀의 비율을 정량화하고, 평균 결합 교차점은 중첩을 측정하고 주사위 계수는 예측된 마스크와 지상 진실 사이의 유사성을 평가한다. 또한, 민감도는 이진 분할에서 긍정적인 인스턴스를 식별하는 데 모델의 효율성을 평가한다.

C. 구현

제안된 모델은 80%의 데이터를 훈련 목적으로 사용하여 훈련을 수행하였으며, 나머지 20%는 테스트를 위해 사용되었다. 훈련 프로세스는 아담 옵티마이저와 초점 손실함수를 사용하여 학습률 $1e-5$, 학습률 0.001, 배치 크기 8 로 50 개의 에포크를 정의했고, 과적합을 방지하기 위해 인코더 블록에 0.2 의 탈락률을 적용하였다. 출력을 정규화 하기 위해 Sigmoid 분류기를 사용하였다.

5. 실험 결과

우리가 제안한 방법은 눈 질환 데이터에서 좋은 성능을 발휘했고 피부 병변 데이터에서도 기존의 접근 방식과 비교하여 다른 모든 모델보다 성능이 뛰어난 것을 확인할 수 있었다.

이 평가는 이러한 테스트 데이터 내에서 적절한 기능을 정확하게 분류하고 세분화하는 모델의 효과를 탐구하는 것을 목표로 한다. 본 논문에서 제시한 모델은 각각 픽셀 정확도 98.19%, 97.21%로 정확한 픽셀 수준 분류를 보여주었다.

6. 결론

본 논문에서 제안한 방법이 픽셀 수준의 신뢰할 수 있는 이진 분할 방법으로 보여지며, 다른 의료 영상 데이터에도 확장 적용한다면 의료 영상 분석 기술 발전에 도움을 줄 수 있을 것이라 기대된다.

감사의 글

본 연구는 2024 년 과학기술정보통신부 및 정보통신기획평가원의 SW 중심대학사업의 연구 결과로 수행되었음(2021-0-01399)

참고문헌

- [1] J. Long, E. Shelhamer, and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," 2015, arXiv:1411.4038 [cs.CV].
- [2] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 40, no. 4, pp. 834-848, April 2018. DOI: 10.1109/TPAMI.2017.2699184.
- [3] N. Siddique, S. Paheding, C. P. Elkin, and V. Devabhaktuni, "U-Net and Its Variants for Medical Image Segmentation: A Review of Theory and Applications," IEEE Access, vol. 9, pp. 82031-82057, 2021. DOI: 10.1109/ACCESS.2021.3086020.
- [4] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation," 2016. Available: <https://doi.org/10.48550/arXiv.1511.00561>
- [5] O. Oktay, J. Schlemper, L. L. Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y. Hammerla, B. Kainz, B. Glocker, and D. Rueckert, "Attention U-Net: Learning Where to Look for the Pancreas," arXiv e-prints, Apr 2018, arXiv:1804.03999. DOI: 10.48550/arXiv.1804.03999.
- [6] S. Woo, J. Park, J. Lee, and I. S. Kweon, "CBAM: Convolutional Block Attention Module," in Computer Vision – ECCV 2018, edited by V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss, Springer International Publishing, Cham, 2018, pp. 3-19. DOI: 10.1007/978-3-030-01234-2_1.
- [7] J. Fu, J. Liu, H. Tian, Y. Li, Y. Bao, Z. Fang, and H. Lu, "Dual attention network for scene segmentation," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019, pp. 3146-3154.