

VFSS에서의 흡인환자 검출을 돋기 위한 실시간 경추 세그멘테이션

김현일, 이고은*, 김보민*, 최상일, 이성재**

단국대학교 컴퓨터학과, 단국대학교 응용컴퓨터공학과*, 단국대학교 의학과**

gusdlf93@naver.com, ge971010@naver.com*, bim65@naver.com*, choisi@dankook.ac.kr,
rmlee@dankook.ac.kr**

Real-time Cervical Spine Segmentation to Help Detect Aspiration Patients in VFSS

HyunIl Kim, Goeun Lee*, Bomin Kim*, Sang-Il Choi, Seong Jae Lee**

Department of Computer Engineering, Dankook University

Department of Applied Computer Engineering, Dankook University*

Department of Rehabilitation Medicine, College of Medicine, Dankook University**

요약

본 논문에서는 흡인 장애를 겪고 있는 환자들의 VFSS(Videofluoroscopic Swallowing Study) 영상을 자동으로 진단하는데 도움을 주기 위해 세그멘테이션(Segmentation)을 이용해 경추를 검출하는 방법을 제안한다. CLAHE(Contrast-Limited Adaptive Histogram Equalization)로 밝기가 보정된 X-ray 영상으로부터 DNN(Deep Neural Network)학습에 사용할 경추의 GT(Ground Truth)를 생성하였으며, DNN을 이용한 세그멘테이션 모델중 Unet을 이용해 경추를 검출한다. Unet으로부터 예측된 경추의 ROI는 DBSCAN(Density-based Spatial Clustering of Applications with Noise)을 통해서 노이즈가 제거된다.

I. 서론

인구 고령화로 인해 인구당 환자수가 늘어남에 따라 의료인의 90%가 휴식시간을 보장받지 못하고 있는 실정이다.[1] 의료인의 부담을 줄여 주기 위해 인공지능 기술을 이용해 환자들의 병을 진단하는 사례들이 늘고 있다. 인공지능 기술은 특성상 환자들의 병을 진단하기 위해서 많은 양의 데이터들을 필요로 하지만, 개인정보를 중요시하는 의료분야의 특성상 많은 양의 데이터를 수집하기는 어렵다. 더불어 의료 영상의 경우 병의 진행 정도나 사람에 따라 많은 변형이 있을 수 있다. 특히 흡인장애를 겪고 있는 환자의 경우, 흡인이 일어나는 ROI(Region of Interest)은 VFSS 영상의 일부분에 불과해 인공지능이 의도치 않은 특징을 학습해서 환자와 정상인을 분류하도록 학습될 수 있으며 이렇게 학습된 모델은 일반화 성능이 떨어지게 된다. ROI의 설정은 인공지능이 영상 내에서 학습할 범위를 강제함으로써, 의도하지 않은 특징들로부터의 학습을 배제하고, ROI 밖에서 발생할 수 있는 거짓 양성(False Positive)을 줄이게 된다. 흡인장애는 식도와 기도사이에서 발생하기 때문에, ROI는 기도와 식도를 포함하는 영역으로 설정되어야 한다. 이 논문에서는 식도와 인접한 경추를 기준점으로 삼아 ROI 이미지 패치를 생성할 것을 제안한다.

II. 본론

실험에 사용된 데이터는 정상인 43 명, 흡인 장애 환자 28 명으로부터 촬영된 214 개의 VFSS 영상으로, 너무

어두워서 육안으로 관찰이 어려운 영상 17 개를 제외하고 197 개의 동영상을 대상으로 실험이 진행되었다. 동영상은 CLAHE[2]를 통해 밝기를 보정하였으며, 보정된 이미지의 어노테이션(Annotation)은 각 동영상에서 경추가 가장 선명하게 보이는 프레임을 한 장씩 선별하여 어노테이션을 진행하였다. 경추의 구성 요소 중 추체(Vertebral Body)는 VFSS 영상에서 관측이 잘되며, 사람마다 모양이 거의 유사하므로 어노테이션은 경추의 구성 요소 중 추체를 포함해야 한다. 이렇게 생성된 이미지 데이터 중 60%는 학습 데이터로 20%는 검증 데이터로, 나머지 20%는 테스트 데이터로 사용한다. 학습 데이터를 나누는 기준은 실험에 참여한 사람을 기준으로 나누도록 한다. 본 논문에서는 VFSS 영상으로부터 경추를 검출하기 위해서 DNN 세그멘테이션 모델 중 하나인 Unet[3]을 사용한다. Unet은 Convolution Encoder 와 Convolution Decoder로 구성된다. 이 실험에서 Convolution Encoder는 사전훈련된 vgg16[4]으로 구성되었으며, Convolution Decoder에서는 Encoding 단계에서 손실된 정보들을 복원해주기 위해서 Expanding Path에서 Contracting Path로 Feature를 Copy and Paste 하여 사용하였다. vgg16을 backbone으로 사용하는 Unet의 파라미터 개수는 2000 만개이다. 학습에 사용된 Unet의 네트워크 구조는 그림 1와 같다. 모델 학습에 사용된 Batch Size는 16, Optimizer는 Adam을 사용하였으며, 500 에폭동안 학습되었다. 이렇게 학습된 Unet의 세그멘테이션 성능은 표 1과 같다.

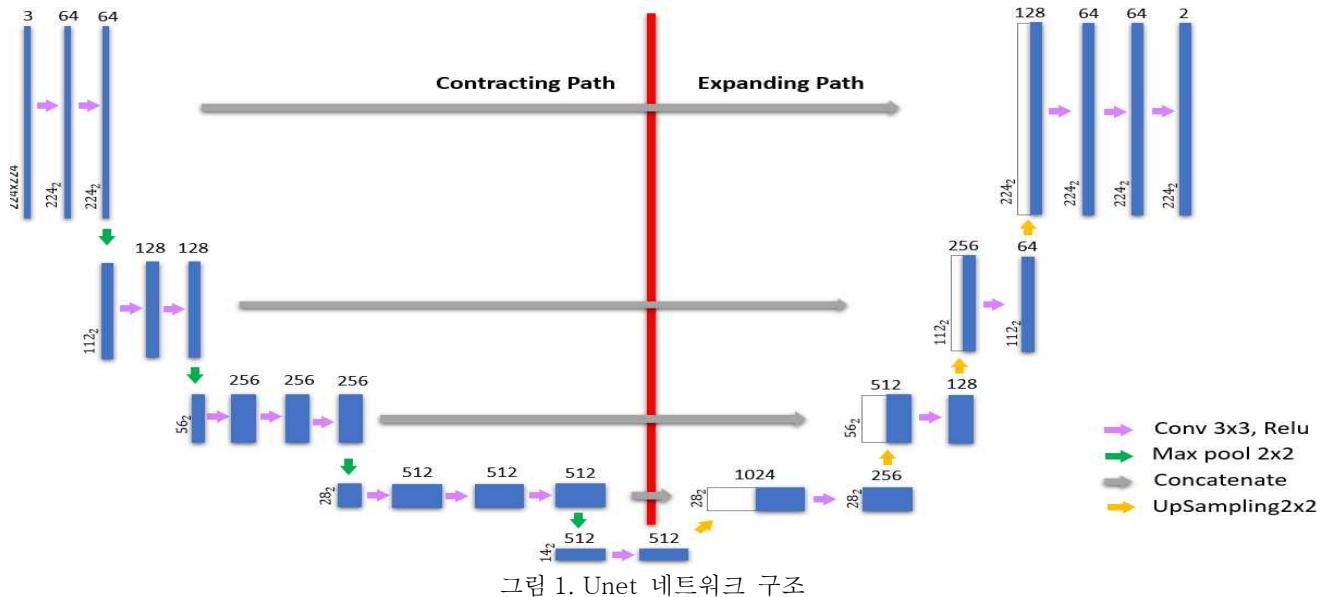


그림 1. Unet 네트워크 구조

표 1. Unet 세그멘테이션 성능

pixel acc	recall	precision	inference time
98.1%	90.7%	68.0%	0.01s

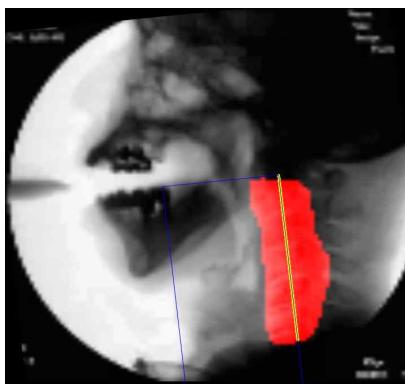


그림 2. 경추 세그멘테이션을 통해 생성된 이미지(빨간색 픽셀은 Unet 이 경추로 예상한 픽셀, 노란색선은 예측된 척주(vertebral column), 파란색 사각형은 ROI)

Unet(그림 1)을 통해 경추로 예측된 픽셀들은, PCA[5]를 통해 주성분 벡터를 뽑아낼 수 있고, 주성분 벡터는 척주(vertebral column)로 가정한다. 척주(vertebral column)는 ROI를 설정하기 위한 기준점이 되며, ROI는 척주 길이의 75%를 가로로, 영상의 바닥까지를 세로로하는 직사각형으로 휴리스틱하게 설정된다. 이렇게 생성된 ROI는 기도와 식도를 잘 포함하는 것을 알 수 있다. 그림 2는 경추 세그멘테이션을 통해 생성된 이미지 패치를 보여준다. 실시간으로 세그멘테이션 된 경추로부터 ROI를 생성하는 과정은 다음 링크에서 확인할 수 있다. (동영상 링크 : https://drive.google.com/file/d/1DzWrZDcxblvTvlRna9RCxc31RX9_cyRY/view?usp=sharing)

본 논문에서는 VFSS 영상으로부터 흡인환자를 자동으로 진단하기 전에, ROI를 생성하기 위하여 딥러닝 세그멘테이션 모델 중 Unet을 사용해 경추 검출을 진행한다. recall이 90.7%로 높은데 반해 precision이 68%로 성능이 낮은 이유는 학습이미지의 부족으로 인해 발생한 것으로, 추후 연구에서는 학습 데이터를 추가해서 False Positive를 줄이는 동시에 검출된 ROI를 통해 흡인 장애 환자를 자동으로 검출하는 인공지능 기술을 개발할 예정이다.

ACKNOWLEDGMENT

“본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW 중심대학 사업(2017-0-00091)과 글로벌 핵심인재양성지원사업의 연구결과로 수행되었음(No.2020-0-01463)”

참 고 문 헌

- [1] 2016 전국의사조사, Retrieved July., 15, 2020,https://www.dbpia.co.kr/pdf/pdfView.do?nodeId=NODE07370684&mark=0&useDate=&bookmarkCnt=0&ipRange=N&language=ko_KR portal/main/EachDtlPageDetail.do?idx_cd=2772”
- [2] K. Zuiderveld, “Contrast Limited Adaptive Histogram Equalization,” Chapter VIII.5, Graphics Gems IV. P.S. Heckbert (Eds.), Cambridge, MA, Academic Press, 1994, pp. 474– 485.
- [3] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” in Proc. Med. Image Comput. Comput.-Assisted Intervention, 2015, pp. 234– 241.
- [4] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” arXiv:1409.1556, 2014.
- [5] Jolliffe, I. T., Principal Component Analysis, 2nd edition, Springer, 2002.

III. 결론