

흉부 CT 영상에서 듀얼 윈도우 영상을 활용한 트랜스포머 기반 폐암 분할

이민진^{1,○}

김세희^{1,○}

홍헬렌^{1,*}

¹서울여자대학교 소프트웨어융합학과

(minjin, 526sally, hlhong)@swu.ac.kr

Transformer-based Lung Cancer Segmentation Using Dual Window Images in Chest CT Images

Min Jin Lee^{1,○}

Se Hee Kim^{1,○}

Helen Hong^{1,*}

¹Department of Software Convergence, Seoul Women's University

요 약

폐암은 크기, 모양, 위치가 다양하다. 더욱이 폐 실질에 위치한 고립된 폐암은 경계가 명확한 반면, 비슷한 강도 값을 가진 주변 구조에 붙은 종양은 경계가 불분명하여 종양 분할에 어려움이 있다. 이러한 문제를 해결하기 위해 본 논문에서는 흉부 CT 영상에서 종격동과 흉벽 등에 부착된 폐암의 경계 불명확성을 해결하기 위해 듀얼 윈도우 영상을 활용한 트랜스포머 기반의 폐암 분할 방법을 제안한다. 폐암 특징을 잘 표현하는 두가지 윈도우 설정을 적용하여 폐 윈도우 영상과 종격동 윈도우 영상을 생성한다. 제안 방법은 SegFormer를 백본으로 사용하며 각각의 영상을 서로 다른 인코더의 입력으로 사용하는 이중 인코더로 다중 스케일의 특징맵을 생성하고, 디코더에서 특징맵 결합 및 채널 주의 블록을 통해 주요 특징에 가중치를 부여하는 방식을 적용한다. 실험을 통해 전체적인 폐암에서 제안 방법이 DSC와 재현율에서 각각 75.74%, 80.44%로 가장 우수한 성능을 보였으며, 폐암의 위치 별 분할 결과에서 제안 방법이 가장 우수한 재현율을 보였다.

Abstract

Lung cancer varies in size, shape, and location. Furthermore, isolated tumors located in the lung parenchyma have clear boundaries, whereas tumors attached to surrounding structures with similar intensity values have indistinct boundaries, making tumor segmentation challenging. To address these issues, we propose a transformer-based lung cancer segmentation method using dual-window images to resolve the boundary ambiguity of lung cancers attached to structures such as the mediastinum and chest wall in chest CT images. To effectively represent the characteristics of lung cancer, we generate lung window and mediastinal window images using two window settings. The proposed method utilizes SegFormer as the backbone and employs a dual-encoder architecture, where each image is input into a different encoder to generate multi-scale feature maps. In the decoder, feature map concatenation and channel attention blocks are applied to assign weights to the key features. Through experiments, the proposed method demonstrated the best performance in overall lung cancer segmentation, achieving a DSC of 75.74% and a recall of 80.44%. Additionally, in the segmentation results by the location of the lung cancer, the proposed method achieves the highest recall.

[○] These authors contributed equally to this work.

^{*} corresponding author

*corresponding author: Helen Hong/Seoul Women's University(hlhong@swu.ac.kr)

키워드: 폐암 분할, 트랜스포머기반, 듀얼 윈도우 영상, 듀얼 인코더, SegFormer, CT 영상

Keywords: Lung cancer segmentation, Transformer-based, Dual window image, Dual encoder, SegFormer, CT image

1. 서론

폐암은 전 세계적으로 암으로 인한 사망의 주요 원인 중 하나로, 국내에서도 암 사망률 중 가장 높은 비중을 차지하고 있다[1]. 폐암의 병기 결정, 치료 계획 수립 및 치료 효과 모니터링을 위해 주로 흉부 컴퓨터 단층 촬영(Computed Tomography, CT) 영상을 활용하여 폐암의 크기, 면적 및 체적을 측정한다[2,3]. 이러한 크기, 면적 및 체적을 정확하게 측정하기 위해서는 임상 의에 의한 수동 분할이 필요하지만, 이는 시간이 많이 소요될 뿐 아니라 임상 의마다 종양 경계를 설정하는 방식이 다를 수 있어 같은 데이터에 대해서 결과가 다르게 나올 수 있어 일관성 있는 결과를 얻기 어렵다. 따라서 흉부 CT 영상에서 자동으로 폐암 영역을 분할하는 것이 필요하다.

그림 1 은 흉부 CT 영상에서 나타나는 폐암의 특징을 나타내며, 폐암은 크기나 형태가 다양하며, 위치에 따라 폐실질(lung parenchyma)에 명확히 나타나거나 폐 흉벽(chest wall), 흉강(chest cavity), 종격동(mediastinum), 폐혈관 등에 부착되어 있어 주변 구조물과의 경계를 정확하게 구분하기 어렵다. 이러한 이유로 폐암 분할의 자동화에 어려움이 있다.

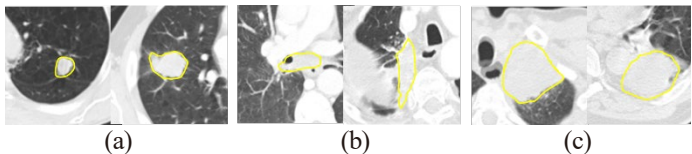


Figure 1. Examples of lung cancers with various sizes, shapes and locations on chest CT images: (a) isolated cancers, (b) mediastinum attached cancers, (c) chest wall attached cancers. The lung cancer boundary is represented in yellow.

폐암 분할 연구들은 컨볼루션 신경망(Convolutional Neural Network, CNN) [4,5]을 활용하고 있으며, 다양한 크기의 폐암을 다루기 위해 다중 스케일을 고려하거나 주의 메커니즘(attention mechanism)을 추가한 네트워크 구조를 제안하고 있다 [6,7,8,9]. 또한, 트랜스포머 기반의 네트워크는 기존 CNN 방식에 비해 전역적인 문맥 정보를 고려할 수 있어, 여러 의료 분할 연구들에 사용되고 있다. Zhang[10]은 CNN 병목 구간(bottleneck)에 트랜스포머(Transformer)를 적용한 TransUNet[11] 기반의

네트워크를 제안하여 폐암 분할을 시도하였다. 그러나 트랜스포머를 기본 네트워크로 사용한 폐암 분할 연구는 드물다.

본 논문에서는 흉부 CT 영상에서 흉벽과 종격동 등에 부착된 폐암의 경계 불명확성을 해결하기 위해 듀얼 윈도우 영상을 활용한 트랜스포머 기반의 폐암 분할 방법을 제안한다. 첫째, 폐암 특징을 잘 표현하는 두 가지 윈도우 영상을 생성하고 이를 입력으로 사용하여 각 인코더에서 다중 스케일의 특징 맵을 생성한다. 둘째, 디코더는 이들 다중 스케일 특징 맵을 결합하고, 채널 주의(channel attention) 블록을 사용하여 주요 특징을 선별함으로써 분할 정확도를 향상시킨다. 셋째 TCIA 의 NSCLC-Radiomics 환자 데이터셋을 사용하여 제안 방법 및 단일 네트워크 방식과 비교하여 전체 폐암과 폐암의 위치에 따른 분할 결과를 제시한다.

2. 제안방법

2.1 데이터 전처리

CT 영상은 -1024~3096 까지의 넓은 밝기값 범위를 가지기 때문에 임상 의들은 판독하려는 목표 장기에 최적화된 밝기값을 0~255 범위의 회색조로 변환하기 위해 밝기값 윈도우 설정을 사용한다. 일반적으로 폐암을 관찰하기 위해서는 폐 윈도우 설정을 사용하는데 폐실질 내부에 위치한 고립된 폐암은 명확하게 나타나지만, 폐 흉벽, 흉강, 종격동, 폐 혈관에 부착된 폐암은 주변 구조물과 유사한 밝기값을 가져 경계를 구분하는 데 어려움이 있다. 이러한 문제를 해결하기 위해 일반적인 폐 윈도우 설정과 함께 종격동 윈도우 설정도 함께 사용한다.

그림 2와 같이 폐 윈도우 설정은 윈도우 폭(Window width) 1500HU(Hounsfield Unit), 윈도우 레벨(Window level)을 -600HU으로 설정되며 폐실질 내부에 독립적으로 존재하는 폐암을 명확하게 나타낼 수 있다. 종격동 윈도우 설정은 윈도우 폭 50HU, 윈도우 레벨 350HU로 설정되며 폐 윈도우에 비해 좁은 윈도우 폭으로 HU 범위를 줄여 대조를 강화하고 주변 구조물과의 경계를 더 뚜렷하게 관찰할 수 있다. 이후, 폐암의 중심 좌표를 기준으로 폐암을 포함하는 영역을 128 x 128 크기

로 잘라 최종 입력 영상을 생성한다.

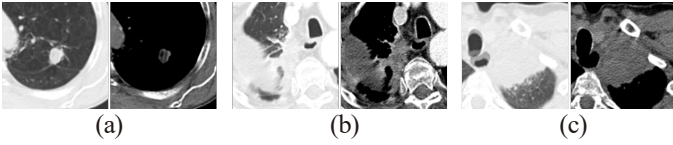


Figure 2. The generated images using lung window(left) and mediastinum window(right) settings: (a) isolated cancer, (b) mediastinum attached cancer, (c) chest wall attached cancer

2.2 기본 네트워크 구조

본 연구에서 제안하는 이중 인코더 방식의 분할 네트워크는 SegFormer[12]를 백본 네트워크로 사용한다. 그림 3와 같이 SegFormer는 트랜스포머 계열 중 380만 개(3.8M)의 파라미터를 가지는 가벼운 네트워크로 제안 방법처럼 이중 인코더로 파라미터의 증가가 필수적인 방법에서 좀 더 효율적으로 사용할 수 있는 장점을 가진다.

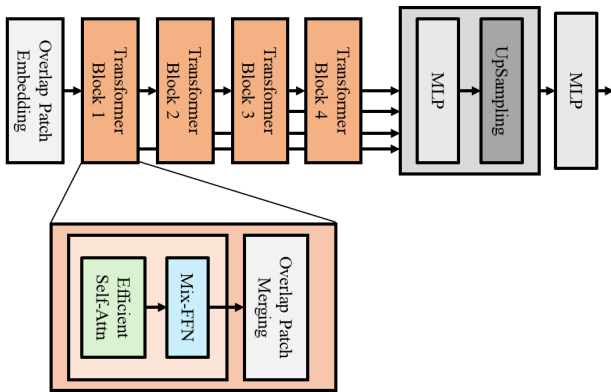


Figure 3. Architecture of the SegFormer network used as the baseline.

SegFormer의 구조는 계층적 특징 맵(feature map)을 생성하는 4개의 레이어로 구성된 인코더(encoder)와 각 레이어에서 생성된 특징 맵들의 채널 수와 공간 크기를 동일하게 맞추고 최종적인 분할 마스크를 생성하는 단순한 다층 퍼셉트론(Multilayer perceptron: MLP)과 업샘플링 과정의 디코더로 구성된다. SegFormer의 입력 패치는 기존 비전 트랜스포머 방법인 ViT[13]의 입력 패치보다 작은 4x4 크기의 작은 패치를 입력으로 사용한다. 이는 좀더 작은 영역에 대한 국소적인 문맥을 파악할 수 있는 장점이 있으며, 여러 스케일의 패치 크기 사용을 통해 국소적, 전역적 문맥을 함께 학습시킬 수 있다. 또한, 기존의 공간적인 정보를 담기 위한 포지셔널 인코딩을 사용하지 않음으로써 복잡성이 감소되어 가벼운 네트워크 모델의 효율을 높일 수 있는 장점을 가진다.

각 인코더의 트랜스포머 블록은 효율적인 셀프 어텐션(Efficient Self-Attention), 믹스 피드 포워드 네트워크(Mix-FFN) 및 오버랩 패치 머징(Overlapped Patch Merging) 과정으로 구성된다. 효율적인 셀프 어텐션 단계에서는 기존 비전 트랜스포머에 비해 늘어난 패치들의 연산량을 줄이기 위해, 셀프 어텐션에서 사용되는 키(key)와 값(value)의 시퀀스 길이를 일정 비율로 줄일 수 있도록 선형 레이어를 적용하여 연산량을 감소시키고, 이를 통해 시간 복잡도를 줄일 수 있다. 믹스 피드 포워드 네트워크 단계에서는 부족한 공간적인 정보를 보완하기 위해 선형 레이어를 한 번 수행한 후, 활성화 함수를 지나기 전에 3x3 컨볼루션 연산을 추가로 수행하여 공간적인 정보를 추가할 수 있도록 한다. 마지막 단계인 오버랩 패치 머징 단계에서는 패치들을 결합해 새로운 패치로 변환하는 과정을 수행한다.

디코더는 인코더에서 생성된 다양한 스케일의 특징 맵을 결합하여 결과를 생성하는 과정이다. 인코더에서 입력 받은 각 레이어의 특징 맵들의 채널 차원과 공간 크기를 동일하게 만들기 위해 MLP와 업샘플링(upsampling)을 수행한 후 각 레이어의 특징 맵들의 결합을 수행한다. 이후, 결합된 특징 맵은 최종적으로 선형 레이어를 거쳐 분할 마스크를 생성한다.

SegFormer는 이러한 네트워크 구조를 통해 분할 성능을 유지하면서도 연산량을 크게 줄일 수 있다. 또한 높은 수준의 특징 맵과 낮은 수준의 특징 맵을 모두 고려하기 때문에 다양한 스케일의 특징 맵을 효과적으로 활용하여 가벼운 네트워크임에도 불구하고 우수한 분할 성능을 보여준다.

2.3 듀얼 윈도우 기반 분할 네트워크 구조

폐암의 다양한 경계의 특징들을 반영하기 위하여 전처리 과정에서 생성된 폐 윈도우 영상과 종격동 윈도우 영상을 모두 활용하여 분할을 수행한다. 이를 위해 각각의 영상을 서로 다른 인코더의 입력으로 사용하는 이중 인코더 기반 네트워크를 제안하며, 최종적으로 하나의 디코더에서 이들 특징 맵을 결합하여 분할을 수행한다.

그림 4(a)는 제안된 네트워크의 구조를 나타내며, 기본 구조는 이전 절에서 설명한 SegFormer의 구조를 따른다. 폐 윈도우 영상과 종격동 윈도우 영상은 각각 SegFormer의 인코더에 입력으로 사용된다. 각 SegFormer 인코더는 다중 스케일 구조로 구성되어 있으며, 폐 윈도우 영상을 입력으로 사용하는 상위 인코더에서는 자기-어텐션(self-attention)

매커니즘을 통해 영상 내 모든 요소 간의 관계를 학습하여 문맥적 특징을 추출한다. 하위 인코더에서는 종격동 윈도우 영상을 입력으로 하여 동일한 과정을 통해 문맥적 특징을 추출하며, 이를 통해 각 인코더에서 다중 스케일의 문맥적 특징을 독립적으로 추출한다. 이후, 동일한 스케일에서 생성된 폐 윈도우 특징 맵과 종격동 윈도우 특징 맵은 채널 주의 블록의 입력으로 사용되고 통합되어진다.

그림 4(b)는 두 특징 맵들을 통합하는 채널 주의 블록의 구조를 나타낸다. 먼저, 두 특징맵들에 채널 결합을 수행한 후, 기존 채널 수를 유지하기 위해 1×1 크기 커널의 컨볼루션 연산을 통해 채널 수를 반으로 줄인다. 이후, 분할에 중요한 영향을 미치는 채널을 강조하기 위해 채널 주의 모듈을 적용한다. 채널 주의 모듈은 전역 최대 풀링(global max pooling) 및 전역 평균 풀링(global average pooling)을 병렬적으로 수행하고, 이를 공유된 다층 퍼셉트론에 통과시킨 후, 두 결과를 더한 뒤 시그모이드 함수를 거쳐 각 채널의 가중치를 계산한다. 마지막으로, 해당 가중치를 특징 맵의 값과 원소 곱(element-wise multiplication)하여 반영한다. 이후, SegFormer 디코더와 동일한 과정을 통해 최종 분할 마스크를 생성한다.

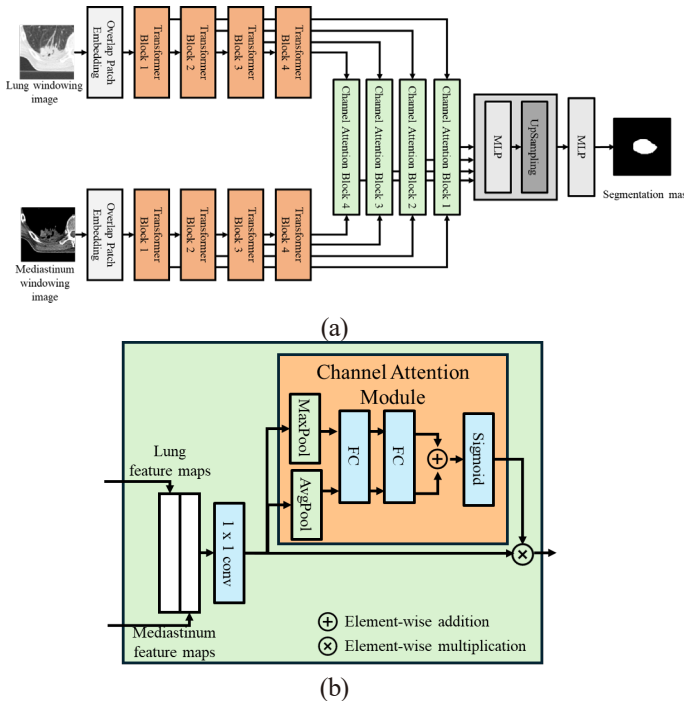


Figure 4. Overview of the proposed Dual-window SegFormer network for lung cancer segmentation: (a) the proposed network with a dual-encoder and channel attention mechanism, (b) the architecture of the channel attention block

3. 결과

실험에서 사용한 데이터는 암 영상 아카이브에서 제공하는 공공 데이터셋인 NSCLC-Radiomics 환자 데이터셋[14]으로, 총 422 명의 폐암 CT 영상으로 구성되어 있다. 이 중 39 명의 선암(Lung Adenocarcinoma, LUAD) 과 113 명의 폐 편평 세포암(Lung Squamous Cell Carcinoma, LUSC) 데이터를 선별하여 총 142 개의 데이터를 사용하였다. 폐암의 직경은 1.36cm 에서 14.63cm 까지 다양하며, 평균 크기는 5.58cm 이고 표준 편차는 2.69cm 이다. 해당 데이터셋은 2004 년 9 월 27 일부터 2014 년 1 월 1 일까지 Biograph 40, SOMATOM Sensation 10, SOMATOM Sensation 16, SOMATOM Sensation Open (Siemens Healthineers) 및 XiO (CMS Imaging, Inc) 장비를 통해 촬영된 흉부 CT 영상으로, 각 슬라이스는 512×512 픽셀 해상도에 평면 내 해상도는 0.97mm, 슬라이스 두께는 3mm 이다.

제안한 네트워크를 학습하고 평가하기 위해, 142 개의 데이터를 학습, 검증 및 테스트 데이터를 각각 98 개, 24 개, 30 개로 나누었고, 5-겹 교차 검증을 수행하여 모든 데이터에 대하여 실험을 수행하였다. 또한 학습 과정에서 데이터 부족으로 인한 과적합을 방지하기 위해 데이터 증강(Data augmentation) 기법을 적용하였으며, -20 도에서 20 도 사이의 임의의 회전(Random rotation), -20 픽셀에서 20 픽셀 사이의 임의의 이동(Random translation), 0.8 배에서 1.2 배 사이의 임의의 스케일(Random scaling)을 적용하였다. 학습 데이터에 대해 무작위로 3 번의 데이터 증강을 적용하여 학습 데이터를 3 배로 증가시켰다. 전체 실험은 Windows 10 (64-bit) 운영체제에 NVIDIA GeForce RTX 2080 Ti 그래픽 카드를 장착한 PC 에서 CUDA 12.2 버전의 GPU 환경에서 파이썬(Python) 3.9 및 파이토치(Pytorch) 2.2 라이브러리를 사용해 진행하였다. 하이퍼 파라미터로는 배치(Batch) 크기를 8, 학습률(Learning rate)을 $1e-4$ 로 설정하였으며, 이진 교차 엔트로피(BCE, Binary Cross Entropy) 손실 함수와 Adam 옵티마이저를 사용하였다. 학습은 50 에폭으로 설정하였으나, 검증 손실이 20 에폭 동안 최소값 보다 감소하지 못하면 조기 종료(Early stopping)하도록 설정하였다.

제안 방법의 성능은 정성적 및 정량적 평가로 분석하였다. 정성적 평가는 폐암 분할 결과 영상의 육안 평가를 수행하였고, 정량적 평가는 다이스 유사 계수(Dice Similarity Coefficient(DSC)), 재현율(Recall), 정밀도(Precision), 균형 정확도(Balanced accuracy(Balanced Acc))를 기준으로 평가하였다. 두 영상을 사용하는 이중 인코더와 채널 주의

블록을 사용하는 제안 방법(Dual-SegFormer)와 비교하기 위하여, 제안 방법의 기본 네트워크인 SegFormer 에 단일 영상인 폐 윈도우 영상과 종격동 윈도우 영상을 각각 적용단일 네트워크(Single-SegFormer) 방식과 두 영상을 고려할 수 있도록 단일 네트워크의 확률 결과값을 평균하여 융합하는 후기 융합 방법(late fusion)에 대한 3 가지를 선정하였다.

그림 5 는 흉부 CT 영상에서 폐암 분할 결과를 나타낸다. 상단 두 줄은 고립된 폐암들을 나타내며, 하단 네 줄은 종격동이나 흉벽에 부착된 폐암을 나타낸다. 고립된 폐암의 경우, 폐 윈도우 영상을 사용하는 Single-SegFormer 방법의 결과와 제안된 방법의 결과는 유사하나, 종격동 윈도우 영상을 사용하는 Single-SegFormer 방법은 과소 분할이 많이 발생하였다. 이는 종격동 윈도우 영상에서 폐암 영역이 폐 윈도우 영상 보다 더 작게 나타나는 원본 영상(Figure 5(a), Figure 5(b))의 특성과 일치한다. 반면, 종격동이나 흉벽에 부착된 폐암의 경우 폐 윈도우 영상을 사용하는 Single-SegFormer 방법이 더 많은 과소 분할이 발생했으며, 이는 경계가 명확하지 않은 부위에서 나타난 현상이다. 이에 비해 제안된 방법은 고립된 폐암과 다른 구조물에 부착된 폐암 모두의 경우에서 두 가지의 Single-SegFormer 방법보다 더 안정적인 분할 결과를 보여주었다. 이는 제안 방법이 두 가지의 인코더로부터 폐 윈도우 영상과 종격동 윈도우 영상에서 폐암의 분할에 주요한 각각의 특징들을 잘 포착하였고, 디코더의 결합 과정에서 채널 주의 블록을 사용하여 좀더 적합한 특징맵에 가중치를 부여할 수 있었기 때문에, 다양한 위치에서 다른 특징을 보이는 폐암에 대하여 안정적인 분할 결과를 보일 수 있었다.

Table 1 은 전체 폐암에 대한 분할 성능을 보여준다. 제안된 방법은 정밀도를 제외한 모든 지표에서 가장 우수한 성능을 보였다. 특히 재현율에서는 폐 윈도우와 종격동 윈도우 영상을 각각 입력으로 사용한 Single-SegFormer 방법보다 각각 5.76%p 와 8.46%p 향상되었다. Single-SegFormer 방법에서는 폐 윈도우 영상을 사용한 결과가 종격동 윈도우 영상 보다 더 높은 분할 성능을 보였으며, 이는 종격동 윈도우 영상을 사용한 결과에서 과소 분할이 더 많이 발생했음을 의미한다. 후기 융합 방식은 전체적으로 가장 낮은 성능을 보였으나, 정밀도에서는 가장 높은 성능을 보였는데, 이는 과소 분할이 많이 발생했기 때문이다.

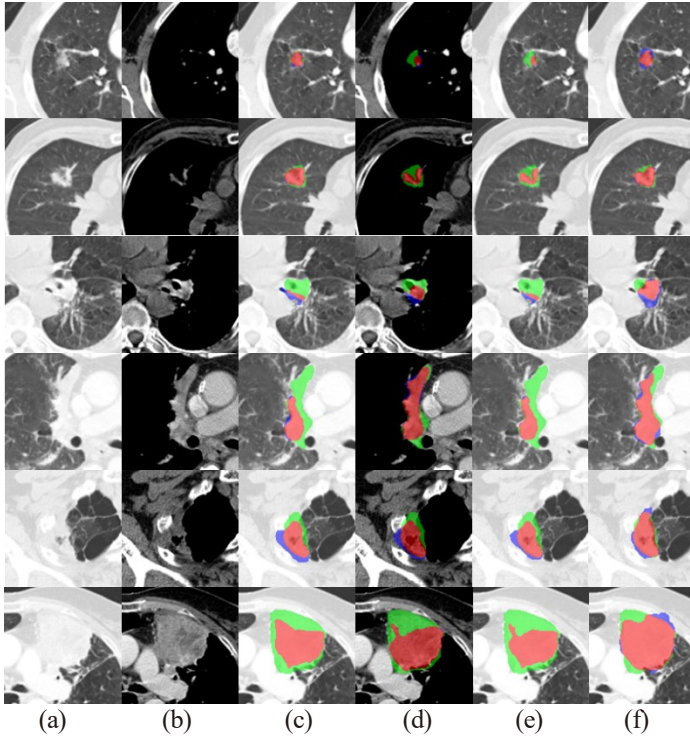


Figure 5. Quantitative evaluation of lung cancer segmentation (a) lung window image, (b) mediastinum window image, (c) Single-SegFormer using the lung window image, (d) Single-SegFormer using the mediastinum window image, (e) late fusion, (f) Dual-SegFormer (Ours) (Red : overlapping areas with ground-truth, Green: under-segmented areas, Blue: over-segmented areas)

Table 1. Performance evaluation of lung tumor segmentation for the entire lung tumors. Mean and standard deviations are provided, with the highest values highlighted in bold.

	(%)			
Methods	DSC	Recall	Precision	Balanced Acc
Single-SegFormer (lung window)	75.37 (16.04)	74.68 (18.04)	79.98 (17.99)	86.97 (8.90)
Single-SegFormer (mediastinum window)	74.53 (16.07)	71.98 (18.61)	81.41 (17.48)	85.50 (9.15)
Late fusion	73.43 (17.60)	67.05 (20.18)	86.47 (15.52)	83.21 (10.00)
Dual-SegFormer (Ours)	75.74 (13.93)	80.44 (15.38)	74.90 (17.37)	89.48 (7.58)

Table 2 는 폐암의 위치에 따른 성능 결과를 나타낸다. 고립된 폐암과 종격동이나 흉벽에 부착된 폐암으로 나누어 분석한 결과, 고립된 폐암에서는 제안된 방법이 재현율과 균형 정확도에서 가장 우수한 성능을 보였던 반면, 종격동이나 흉벽에 부착된 폐암에서는 제안된 방법이 DSC, 재현율, 균형 정확도 모두에서 가장 좋은 성능을 보였다. 이는 제안된 방법이 고립된 폐암보다 부착된 폐암에서 더 효과적임을

확인할 수 있었다. 또한, Single-SegFormer 방식들끼리의 비교에서 고립된 폐암의 경우 폐 윈도우 영상을 사용한 Single-SegFormer 방법의 결과가 더 높은 DSC 성능을 보였고, 부착된 폐암의 경우 종격동 윈도우 영상을 사용한 Single-SegFormer 방법의 결과가 근소한 차이로 더 나은 성능을 보였다.

Table 2 Performance evaluation of lung cancer segmentation into two groups according to their location type. Mean and standard deviations are provided, with the highest values highlighted in bold. (%)

Methods	DSC	Recall	Precision	Balanced Acc
*** Isolated cancers ***				
Single-SegFormer (lung window)	76.79 (17.37)	76.56 (17.91)	81.80 (20.48)	88.04 (8.84)
Single-SegFormer (mediastinum window)	74.65 (18.06)	71.18 (19.42)	83.62 (19.38)	85.40 (9.60)
Late fusion	74.56 (18.36)	67.57 (20.29)	88.87 (17.87)	83.67 (10.08)
Dual-CASegFormer (Ours)	76.63 (15.73)	82.75 (16.12)	74.92 (19.46)	90.98 (7.92)
*** Attached cancers ***				
Single-SegFormer (lung window)	74.43 (15.30)	73.29 (17.84)	78.89 (16.50)	85.88 (8.74)
Single-SegFormer (mediastinum window)	74.55 (15.15)	72.32 (18.47)	80.25 (16.46)	85.48 (9.05)
Late fusion	72.65 (17.60)	66.60 (20.38)	84.94 (14.50)	82.85 (10.08)
Dual-SegFormer (Ours)	75.24 (13.07)	78.54 (14.84)	75.35 (16.04)	88.31 (7.28)

4. 결론

본 논문에서는 흉부 CT 영상에서 다양한 크기와 유사한 밝기값을 가진 주변 구조물이 존재하는 폐암의 분할 성능을 향상시키기 위한 네트워크를 제안하였다. 제안된 네트워크는 트랜스포머 기반의 SegFormer 구조를 백본으로 사용하였으며, 고립된 폐암과 주변 구조물에 부착된 폐암을 모두 효과적으로 분할하기 위해 두 가지 윈도우 설정을 적용한 영상을 생성하였다. 각각의 영상에서 문맥적 특징을 추출하기 위해 두 개의 인코더를 사용하여 별도의 다중 스케일의 특징

맵을 생성하고, 디코더에서 동일한 스케일의 특징 맵을 결합한 후, 채널 주의 블록을 통해 주요 특징에 가중치를 부여하는 방식을 적용하였다.

실험 결과, 전체적인 분할 성능과 폐암 위치 별 분할 결과에서 제안된 네트워크는 모든 위치에서 우수한 재현율을 기록하였으며, DSC 지표에서도 좋은 성능을 보였다. 특히 고립된 폐암에 비해 종격동이나 흉벽에 부착된 폐암에서 더 큰 성능 향상을 확인할 수 있었다. 향후 연구 방향으로는 폐암의 공간 정보를 활용한 다중-뷰 기반의 3 차원 폐암 분할 네트워크로 확장하여, 폐암 분할 성능을 개선하고자 한다.

감사의 글

본 연구는 보건복지부의 재원으로 한국 보건산업진흥원의 보건의료기술연구개발사업 지원 (HI22C1496), 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원(No. RS-2023-00207947) 및 서울여자대학교 학술연구비의 지원(2024-0221)을 받아 수행되었습니다.

References

- [1] National Cancer Information Canter 2022. Proportion of deaths from major cancers. <https://cancer.go.kr/lay1/S1T645C646/contents.do>
- [2] P.D. Mozley, C. Bendtsen, B. Zhao, L.H. Schwartz, M. Thorn, Y. Rong, and A.J. Buckler, “Measurement of tumor volumes improves RECIST-based response assessments in advanced lung cancer,” Translational oncology, 5(1), pp. 19-25, 2012.
- [3] S.A. Hayes, M.C. Pietanza, D. O’Driscoll, J. Zheng, C.S. Moskowitz, M.G. Kris, and M.S. Ginsberg, “Comparison of CT volumetric measurement with RECIST response in patients with lung cancer,” European journal of radiology, 85(3), pp. 524-533, 2016.
- [4] O. Ronneberger, P. Fischer, and T. Borx, “U-net: Convolutional networks for biomedical image segmentation,” Medical Image Computing and Computer-Assisted Intervention, pp 234-241, 2015.
- [5] V. Badrinarayanan, A. Kendall, and R. Cipolla, “Segnet: A deep convolutional encoder-decoder architecture for image segmentation,” IEEE transactions on pattern analysis and machine intelligence, 39(12), pp 2481-2495, 2017.
- [6] J. Yang, B. Wu, L. Li, P. Cao, and O. Zaiane, “MSDS-UNet: A multi-scale deeply supervised 3D U-Net for automatic segmentation of lung tumor in CT,” Computerized Medical Imaging and Graphics 92 101957, 2021.

- [7] S.F. Banu, M.M.K. Sarker, M. Abdel-Nasser, D. Puig, and H.A. Raswan, "Aweu-net: an attention-aware weight excitation U-net for lung nodule segmentation," *Applied Sciences*, 11(21), 10132, 2021.
- [8] J. Lee, M.J. Lee, B.S. Kim, and H. Hong, "Automated lung tumor segmentation robust to various tumor sizes using a consistency learning-based multi-scale dual-attention network," *Journal of X-Ray Science and Technology*, pp.1-14, 2023.
- [9] M.K. Lee, Y.S. Oh, and H. Hong, "Enhanced Lung Cancer Segmentation with Deep Supervision and Hybrid Lesion Focal Loss in Chest CT Images," *Journal of the Korea Computer Graphics Society*, 30(1), 11-17, 2024.
- [10] F. Zhang, Q. Wang, E. Fan, N. Lu, D. Chen, H. Jiang and Y. Yu, "Enhancing non-small cell lung cancer tumor segmentation with a novel two-step deep learning approach," *Journal of Radiation Research and Applied Sciences*, 17(1), 100775, 2024.
- [11] J. Chen, Y. Lu, X. Luo, E. Adeli, Y. Wang, L. Lu and Y. Zhou, "TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation," *arXiv preprint arXiv:2102.04306*, 2021.
- [12] E. Xie, W. Wang, Z. Yu, A. Anandkumar, J.M. Alvarez, and P. Luo, "SegFormer: Simple and efficient design for semantic segmentation with transformers." *Advances in neural information processing systems*, 34, pp.12077-12090, 2021
- [13] A. Dosovitskiy, "An image is worth 16x16 words: Transformers for image recognition at scale," *arXiv preprint arXiv:2010.11929*, 2020.
- [14] H.J.W.L. Aerts, L. Wee, E. Rios Velazquez, R.T.H. Leijenaar, C. Parmar, P. Grossmann, S. Carvalho, J. Bussink, R. Monshouwer, B. Haibe-Kains, D. Rietveld, F. Hoebers, M.M. Rietbergen, C.R. Leemans, A. Dekker, J. Quackenbush, R.J. Gillies, P. Lambin, "Data From NSCLC-Radiomics (version 4) [Data set]," *The Cancer Imaging Archive*, 2014.

〈 저자 소개 〉

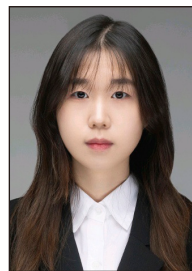
이 민 진

- 2007년 2월 서울여자대학교 컴퓨터학과 졸업(학사)
- 2016년 8월 서울여자대학교 컴퓨터학과 졸업(박사)
- 2016년 9월~현재 서울여자대학교 소프트웨어융합학과 초빙강의교수
- 관심분야 : 의료 인공지능, 딥러닝, 영상 분할 및 분석
- <https://orcid.org/0000-0002-6773-1364>



김 세 희

- 2024년 8월 서울여자대학교 소프트웨어융합학과 졸업(학사)
- 관심분야 : 의료 인공지능, 딥러닝, 컴퓨터 비전
- <https://orcid.org/0009-0000-7539-142X>



홍 헬 렌

- 1994년 2월 이화여자대학교 전자계산학과 졸업(학사)
- 1996년 2월 이화여자대학교 전자계산학과 졸업(석사)
- 2001년 8월 이화여자대학교 컴퓨터학과 졸업(박사)
- 2001년 9월~2003년 7월 서울대학교 컴퓨터공학부 BK 조교수
- 2006년 3월~현재 서울여자대학교 소프트웨어융합학과 교수
- 관심분야 : 의료 인공지능, 딥러닝, 영상처리 및 분석
- <https://orcid.org/0000-0001-5044-7909>

