

# 복부 CT 영상에서 신장 로컬 가이드 맵을 활용한 평균-교사 모델 기반의 준지도학습을 통한 신장 종양 분할

정희영<sup>0</sup>      김현진      홍헬렌\*

서울여자대학교 소프트웨어융합학과

{hyjeong, hyunjinkim, hlhong}@swu.ac.kr

## Kidney Tumor Segmentation through Semi-supervised Learning

### Based on Mean Teacher Using Kidney Local Guided Map in Abdominal CT Images

Heeyoung Jeong<sup>0</sup>      Hyeonjin Kim      Helen Hong\*

Department of Software Convergence, Seoul Women's University

#### 요약

부분신장절제술 전 수술 계획을 세우기 위해서는 신장 종양의 위치, 형태 및 수술 시 안전 마진 파악이 중요하므로 신장 종양을 정확히 분할하는 것이 필요하다. 그러나 신장 종양은 환자마다 위치 및 크기가 다양하며 소장과 비장 같은 주변 장기와 형태와 밝기값이 유사하여 신장 종양을 분할하는 것에 어려움이 있다. 본 논문에서는 레이블이 있는 데이터와 없는 데이터를 함께 사용하는 준지도학습 방법 중 하나인 평균-교사 모델을 활용하여 신장의 여러 위치에서 발생하는 작은 크기의 신장 종양을 분할하기 위해 신장 위치 정보를 가지는 신장 로컬 가이드 맵을 이용해 신장 종양에 집중하는 평균-교사 네트워크를 제안하고, 신장 종양의 크기에 따른 성능을 분석한다. 실험 결과, 제안 방법은 신장 주변에 존재하는 종양의 위치를 찾기 위해 신장 로컬 가이드 맵을 사용하여 신장의 국소 정보를 고려함으로써 75.24%의 F1-score를 보였다. 특히 분할이 어려운 작은 크기의 종양에 대한 과소분할을 개선하였으며 nnU-Net보다 적은 양의 레이블 데이터를 사용하여도 13.9% 높은 F1-score를 보였다.

#### Abstract

Accurate segmentation of the kidney tumor is necessary to identify shape, location and safety margin of tumor in abdominal CT images for surgical planning before renal partial nephrectomy. However, kidney tumor segmentation is challenging task due to the various sizes and locations of the tumor for each patient and signal intensity similarity to surrounding organs such as intestine and spleen. In this paper, we propose a semi-supervised learning-based mean teacher network that utilizes both labeled and unlabeled data using a kidney local guided map including kidney local information to segment small-sized kidney tumors occurring at various locations in the kidney, and analyze the performance according to the kidney tumor size. As a result of the study, the proposed method showed an F1-score of 75.24% by considering local information of the kidney using a kidney local guide map to locate the tumor existing around the kidney. In particular, under-segmentation of small-sized tumors which are difficult to segment was improved, and showed a 13.9%p higher F1-score even though it used a smaller amount of labeled data than nnU-Net.

\*corresponding author: Helen Hong / Department of Software Convergence, Seoul Women's University (hlhong@swu.ac.kr)

**키워드:** 신장 종양 분할, CT 영상, 준지도학습, 평균-교사 모델, 신장 로컬 가이드 맵

**Keywords:** Kidney Tumor Segmentation, CT Image, Semi-Supervised Learning, Mean Teacher Model, Kidney Local Guided Map

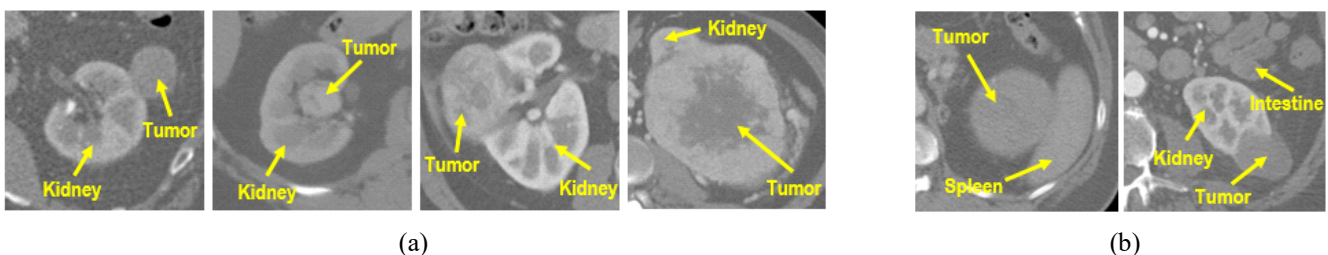
## 1. 서론

신장암은 전체 성인 종양 중 2~3%를 차지하며 전 세계적으로 43만건 이상의 신장 질환 사례가 발견되고 있다. 대부분 초기에 특별한 증상이 없어 증상이 나타났을 때 이미 질환이 악화된 경우가 많으므로 조기에 발견하여 수술적 치료를 하는 것이 중요하다[1-6]. 과거에는 신장과 신장 주위의 전이 조직이나 림프선을 절제하는 근치적 신장절제술이 일반적이었지만, 최근에는 신장을 최대한 보존하면서 종양만을 절제하는 부분신장절제술이 더 많이 수행되고 있다[7]. 부분신장절제술을 계획하기 위해서는 신장 종양의 위치, 형태 및 수술 중 주요 조직 보호를 위한 안전 마진(safety margin)을 파악하는 것이 중요하므로 신장 종양을 분할하는 것이 필요하다. 그러나 그림 1과 같이, 크고 작은 다양한 크기의 종양이 환자의 신장 내부, 외부, 경계 등 여러 위치에 존재하며, 소장과 비장 같은 주변 장기와 형태 및 밝기값이 유사하여 해당 기관으로 과대 분할(over-segmentation)과 이상치(outlier)가 발생할 수 있으므로 신장 종양을 분할하는 데 한계가 있다.

복부 CT 영상에서 신장 종양을 분할하기 위한 관련 연구는 일반적으로 지도학습(supervised learning)을 통해 수행되며 2019 Kidney and kidney Tumor Segmentation (KiTS19) Challenge에서 제공된 공개 데이터를 사용하여 성능을 평가한다. 복부 3D CT 영상에서 신장 종양의 대략적인 위치를 확인하기 위해 위치 식별 및 분할 단계로 이루어진 계단식 V-Net[8]과 중요한 특징 정보를 추출하여 주요 부분에 중점을 두는 Squeeze and Excitation(SE) 모듈과 같은 어텐션을 활용한 U-Net[9,10]을 통해 54.16~74.32%의 신장 종양 분할 성능을 보였으나 작은 크기의 종양 검출에는 한계가 있었다. 또한, 자가적응형 프레임워크를 기반으로 하여 네트워크의 파라미터를 자동으로 최적화하는 nnU-Net에 잔차연결(residual connection)을 적용하거나, 저해상도 3D 영상에서 신장 영역을 추출한 후 고해상도 3D 영상으로 업샘플링한 뒤

신장과 신장 종양으로 각각 분할하거나 신장 위치화와 신장과 신장 종양을 분할하는데 활용하는 계단식 네트워크를 통해 신장 종양에 대해 82.54%~85.09%의 개선된 성능을 보였다[11,12,13]. 이와 같은 딥 컨볼루션 기반의 지도학습을 통한 신장 종양 분할의 경우 레이블 데이터가 부족한 경우 성능 저하가 발생할 수 있으므로 충분한 레이블 데이터를 확보하는 것이 중요하다. 그러나 의료 영상에서 레이블링에 소요되는 시간과 비용이 높아 레이블 데이터 수집이 어려운 한계가 있다[14-18]. 이러한 레이블 데이터 부족 문제를 완화하기 위해 레이블이 있는 데이터와 레이블이 없는 데이터를 함께 활용하여 효과적으로 훈련하는 준지도학습(semi-supervised learning) 방법이 필요하다. 준지도학습은 훈련 시 레이블이 없는 데이터의 정보를 추가적으로 활용하여 네트워크의 일반화 성능을 향상시켜 적은 레이블 데이터 세트에서 발생할 수 있는 과적합(overfitting) 문제를 완화한다.

준지도학습 분야에서는 불확실성을 가중치로 활용하는 손실함수를 적용한 평균-교사모형을 제안하여 교사 모델이 더 정확하고 낮은 불확실성을 가진 예측을 생성하도록 하는 신장 분할 연구와, 레이블이 있는 데이터에서 입력 영상과 레이블 간의 분포를 계산하고 이를 기반으로 레이블이 없는 데이터에 대한 신뢰성 있는 의사 레이블(pseudo label)을 생성하는 3D U-Net을 제안하여 훈련 초기에 레이블이 있는 데이터에 과적합 되는 문제를 해결하기 위한 신장 분할 연구가 진행되었다[19,20]. 이외에도 신경교종과 비인두암종과 같은 종양에 대해서는 모델에서 레이블이 없는 데이터에 대한 높은 품질의 의사 레이블을 생성하기 위해 레이블이 있는 데이터에 적대적 손실(adversarial loss)을 적용하여 훈련 안정성을 높이거나, 교사 모델의 예측으로부터 불확실성을 계산하여 패널티를 부여하는 손실함수와 전역적 문맥 정보를 파악하기 위한 셀프-어텐션 (Self-Attention) 모듈을 사용한 평균-교사 모델[21,22]을 이용한 연구가 수행되었다. 그러나



**Figure 1** Characteristic of kidney on abdominal CT images: (a) Various location of kidney tumor, (b) Signal intensity similarity between other organs and kidney tumor.

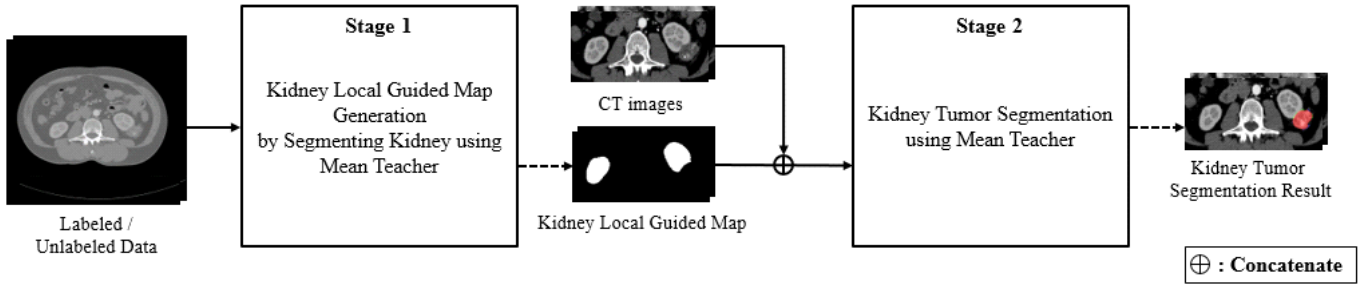


Figure 2 Proposed kidney tumor segmentation network.

신장 종양에 대한 연구 사례는 아직 없다.

따라서 본 논문에서는 레이블 데이터가 부족한 상황에서 작은 크기의 신장 종양을 효과적으로 검출하는 것을 목표로, 레이블이 없는 데이터를 함께 활용하여 신장의 위치 정보를 담은 신장 로컬 가이드 맵을 생성하고 이를 추가 입력 채널로 활용하여 신장 주변의 종양에 집중하여 분할하는 평균-교사 네트워크를 제안한다. 또한, 기본 U-Net과 신장 및 신장 종양 분할에서 우수한 성능을 보이는 nnU-Net과의 성능 비교를 수행하며, 종양 크기에 따라 4개의 그룹으로 분류하여 결과를 분석한다.

$$I_{new} = \begin{cases} 0 & \text{if } I < I_{min} \\ \frac{I - (I_{min}) \times 255}{I_{max} - I_{min}} & \text{if } I_{min} \leq I \leq I_{max} \\ 255 & \text{if } I > I_{max} \end{cases} \quad (1)$$

이 때,  $I$ 는 원본 영상의 밝기값이고,  $I_{min}$ 과  $I_{max}$ 은 각각 사용하고자 하는 밝기값 범위의 최소값과 최대값으로, 신장과 관련 없는 조직의 영향을 줄이기 위해 지방과 연조직 범위를 나타내는 -79HU와 304HU로 설정한다[11].

## 2. 제안 방법

본 제안 방법은 그림 2와 같이 두 단계의 평균-교사 모델 [23]로 구성되며, 신장의 위치 정보를 포함하는 신장 로컬 가이드 맵을 생성하는 신장 분할 단계와 신장 로컬 가이드 맵을 통해 신장 주변의 작은 크기 종양에 중점을 둔 신장 종양 분할단계로 이루어진다.

### 2.1 데이터 전처리

복부 CT 영상에서 신장이 상대적으로 작은 영역을 차지하기 때문에 신장의 지역적 문맥 정보에 집중하여 분할하기 위해 512x512 크기의 전체 복부 CT 영상을 사용하는 대신 신장 영역만을 포함하도록 영상을 자르고 이를 400x256 크기로 조정한다.

복부 CT 영상에서 신장 종양과 주변 조직의 밝기값이 유사한 경우, 다른 장기나 주변 조직을 신장 종양으로 판단하여 오분할이 발생할 수 있다. 따라서 그림 3과 같이 영상의 밝기값을 조정하여 신장 종양과 주변 조직이 잘 구별될 수 있도록 밝기값 정규화를 수행한다. 밝기값 정규화를 위해 식 (1)과 같이 영상의 밝기값을 최소-최대 정규화(min-max normalization)을 통해 0~255 사이로 정규화한다.

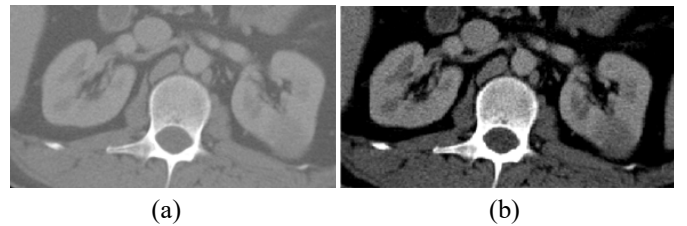


Figure 3 Effects of intensity normalization:

(a) Original image, (b) Min-max normalization.

### 2.2 신장 로컬 가이드 맵을 이용한 신장 종양 분할 네트워크

복부 CT 영상에서 신장 종양은 신장과 인접한 위치에서 발생하기 때문에 신장의 위치 정보를 활용하여 신장 종양의 위치를 파악할 수 있다. 먼저 평균-교사 모델을 사용하여 신장을 분할한 후, 분할 결과를 신장 로컬 가이드 맵으로 사용하여 종양 분할 단계에서 신장의 위치 정보를 제공한다.

평균-교사 모델은 그림 4와 같이 학생과 교사 모델로 구성되며, 레이블이 있는 영상과 레이블이 없는 영상을 모두 사용하여 준지도학습을 수행한다. 각 모델은 2D U-Net[24]을 백본 네트워크로 사용하며, 5개의 레이어로 이루어진 수축 경로(contracting path)와 확장경로(expanding path)를 가진다. 수축경로의 각 레이어에서는 3x3 컨볼루션 연산, 배치 정규화와 ReLU 활성화 함수를 적용하고, 2x2 맥스풀링 및 드롭아웃을 통해 특징맵을 생성한다. 확장경로의 각 레이어에

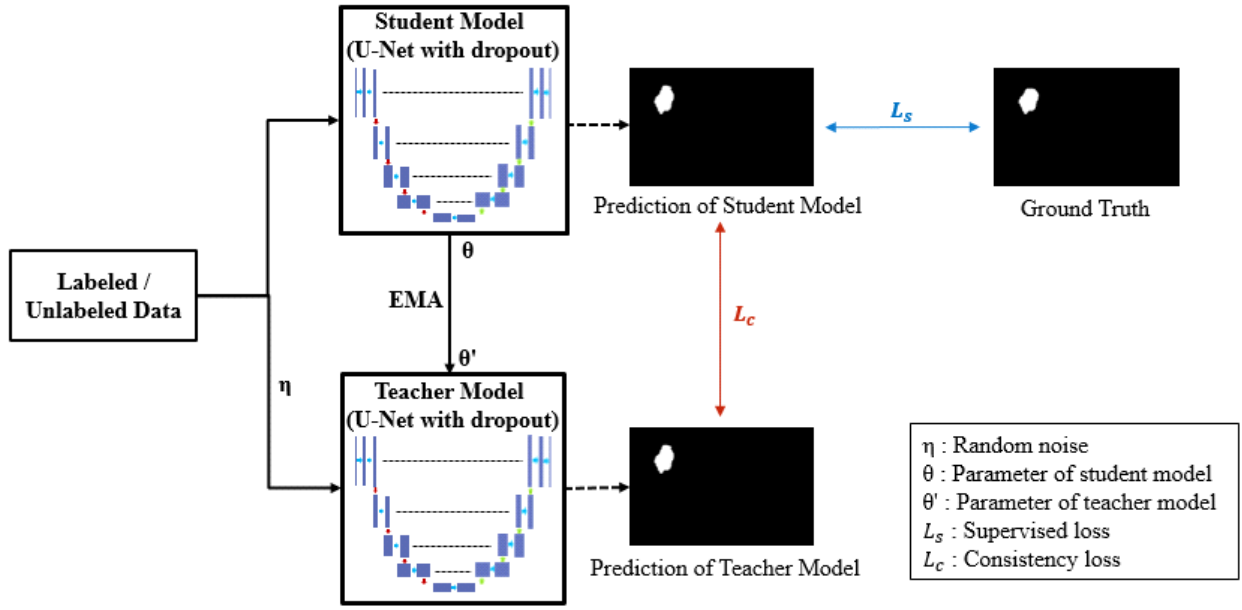


Figure 4 Mean teacher model framework.

서는 2x2 업샘플링을 거친 특징맵과 동일한 레이어에 위치한 수축경로의 특징맵을 연결연산(skip connection)을 통해 결합하고 3x3 컨볼루션 연산, 배치 정규화와 ReLU를 반복한다. 마지막 레이어는 예측을 위한 확률맵을 생성하기 위해 시그모이드 활성화 함수를 사용하는 1x1 컨볼루션 레이어로 구성된다. 네트워크의 일반화 성능 향상을 위한 추가적인 섭동(perturbation)을 가하기 위해 드롭아웃이 적용되며, 교사 모델의 경우 입력 영상에 랜덤 노이즈를 적용한다.

학생 모델은 경사하강법을 사용하여 파라미터가 갱신되어 훈련이 진행되고, 교사 모델의 파라미터는 식 (2)와 같이 모든 훈련 단계에서 학생 모델 파라미터와 이전 단계의 교사 모델 파라미터 간의 지수이동평균(exponential moving average)으로 갱신된다.

$$\theta_t \leftarrow \alpha \theta_{t-1} + (1 - \alpha) \theta_s \quad (2)$$

이 때,  $\theta_t$ 는 교사 모델의 파라미터이고  $\theta_{t-1}$ 는 이전 단계의 교사 모델 파라미터이고  $\theta_s$ 는 학생 모델의 파라미터이고  $\alpha$ 는 이동평균을 조절하는 하이퍼파라미터 값이고 0.99로 설정하였다.

이후 CT 영상과 신장 로컬 가이드 맵을 결합하여 생성한 두 개의 채널로 이루어진 영상을 평균-교사 모델의 입력 영상으로 사용하여 최종적으로 신장 종양을 분할한다. 신장 로컬 가이드 맵을 추가 입력 채널로 사용함으로써 신장의 위치 정보를 고려하며, 이를 통해 신장 주변 영역에 집중하여 작은 크기의 종양의 위치를 파악하는 데 도움이 될 수 있다.

## 2.3 손실함수

학습을 위한 전체 손실 함수로 식 (3)과 같이 지도 손실함수( $L_s$ )와 일관성 손실함수( $L_c$ )를 함께 사용하여 레이블이 있는 데이터와 없는 데이터를 학습에 함께 활용한다.

지도 손실함수는 레이블이 있는 데이터에 대한 학습을 수행하는 손실 함수로, 학생 모델이 출력한 예측값과 정답 레이블 간의 차이를 계산하여 모델이 예측한 결과가 실제 정답과 가까워지도록 모델을 학습시킨다. 지도 손실함수는 식 (4)와 같이 이항교차엔트로피(Binary cross entropy)와 다이스 손실(Dice loss)의 합으로 계산된다.

일관성 손실함수는 레이블이 있는 데이터와 레이블이 없는 데이터 모두에 대한 학습을 수행하는 손실 함수로, 학생 모델과 교사 모델이 출력한 예측값 간의 차이를 최소화하여 일관성을 유지하도록 모델을 학습시킨다. 일관성 손실함수는 식 (5)와 같이 평균제곱오차(Mean-squared error)를 사용하여 계산된다. 이러한 손실함수들을 통해 모델은 정답 레이블이 필요하지 않아 레이블이 없는 데이터에 대한 학습이 가능해진다.

$$L_t = L_s + L_c \quad (3)$$

$$L_s = \frac{1}{N} \sum_i^N [(-y_i \log(\hat{y}_i) - (1 - y_i) \log(1 - \hat{y}_i)) + (1 - \frac{2y_i \hat{y}_i}{y_i^2 + \hat{y}_i^2})] \quad (4)$$

$$L_c = \frac{1}{N} \sum_{i=1}^N (\tilde{y}_i - \hat{y}_i)^2 \quad (5)$$

이 때,  $y_i$ 는  $i$ 번째 데이터의 정답 레이블이고  $\hat{y}_i$ 은  $i$ 번째 데이터에 대한 학생 모델을 통해 출력된 예측값이고  $\tilde{y}_i$ 는  $i$ 번째 데이터에 대한 교사 모델을 통해 출력된 예측값이고  $N$ 은 전체 데이터 수이다.

### 3. 실험 및 결과

실험에서 사용된 데이터는 Minnesota Medical Center에서 제공한 KiTS19 Challenge 데이터로, 부분신장절제술 혹은 근처신장절제술을 받은 210명 환자의 수술 전 복부 CT 동맥기 영상이다. 이 영상의 해상도는 512x512이며, 슬라이스 장수는 19~1059장, 슬라이스 두께는 0.4375~1.041mm, 슬라이스 간격은 0.5~5mm이다.

실험 환경은 GeForce RTX 3090 그래픽 카드가 장착된 서버에서 파이썬 기반의 딥러닝 라이브러리인 Pytorch 1.10.1 버전을 사용하여 수행되었다. 네트워크 학습 시 에폭은 200, 학습률은 0.003, 배치 크기는 8로 설정하였으며 최적화를 위해 아담 옵티마이저를 사용하였다. 또한 학생과 교사 모델의 각 레이어에 [0.005, 0.1, 0.2, 0.3, 0.5]의 드롭아웃을 적용하였다. 실험은 홀드아웃(hold-out) 검증 방식을 사용하였으며, 훈련 데이터 169개와 테스트 데이터 41개로 나누고, 훈련데이터에서 레이블이 있는 데이터와 레이블이 없는 데이터의 비율을 조정하면서 실험을 진행했다.

본 제안 방법의 평가를 위해 2D U-Net[23]과 2D nnU-Net[25]을 사용한 지도학습 방법과 평균-교사 모델 (Mean Teacher, MT)과 신장 로컬 가이드 맵을 추가하여 신장위치 정보를 고려한 준지도학습 제안 방법을 비교하였으며, 레이블이 있는 데이터의 비율을 조정하여 네트워크 훈련에 사용하는 레이블 데이터 수에 따른 성능을 분석하였다. 신장 종양 크기 별 분할 성능을 평가하기 위해 종양의 크기를 기준으로 네 그룹으로 나누었다. 종양 크기가 2.3cm 미만인 경우 작은 크기 종양, 2.3cm 이상 4.3cm 미만인 경우 중간 크기 종양, 4.3cm 이상 7.4cm 미만인 경우 큰 종양, 7.4cm 이상인 경우 매우 큰 종양으로 설정했다[25]. 각 그룹 별 종양 크기의 평균 및 표준편차는 작은 크기 종양은  $1.8 \pm 0.2$ cm, 중간 크기 종양은  $3.1 \pm 0.5$ cm, 큰 종양은  $5.1 \pm 0.7$ cm, 매우 큰 종양은  $10.5 \pm 2.1$ cm로 나타났다.

본 제안 방법의 성능을 평가하기 위해 정량적 평가와 정성적 평가를 수행하였으며 정량적 평가에는 F1-score, 재현율(Recall), 정밀도(Precision)를 성능지표로 하여 제안 방법과 비교 방법 간의 분할 성능을 평가하였다. 또한 복부 CT 영상에서 신장 종양 영역은 배경 영역에 비해 상대적으로 적은 영역을 차지하므로 신장 종양과 배경의 비율이 불균형하므로 신장 종양 영역에 대한 성능을 균등하게 고려하기 위해 신장 종양과 배경의 정확도를 각각 계산한 후 평균을 취하는 Balanced accuracy를 추가적으로 평가하였다.

표 1은 레이블 데이터 비율에 따른 지도학습 방법과 준지도학습 방법의 신장 종양 분할 성능지표 결과이다. 지도학습으로 수행된 방법의 경우, 레이블이 있는 데이터 100%를 사용한 방법들에 비해 레이블이 있는 데이터를 25%만 사용한 방법들의 F1-score, 재현율, 정밀도 수치가 레이블 데이터의 비율이 줄어들면서 감소함을 나타냈다. 레이블이 있는 데이터 25%와 레이블이 없는 데이터 75%를 사용한 준지도학습 방법의 F1-score, 재현율, 정밀도는 26.23%, 27.50%, 34.66%로, 훈련 초기에 적은 레이블 데이터로 인해 과적합이 발생하고 일반화 성능이 낮아져 레이블이 없는 데이터를 충분히 활용하지 못해 과소 분할이 발생하여 낮은 성능을 보였다. 레이블이 있는 데이터 25%를 사용하고 신장 로컬 가이드 맵을 적용한 제안 방법의 경우 F1-score, 재현율, 정밀도 수치는 75.24%, 72.96%, 86.57%로 신장 로컬 가이드 맵을 사용하여 신장의 위치 정보를 고려함으로써 종양의 위치 파악에 도움이 되어 다른 비교 방법에서 발생한 작은 크기의 신장 종양에 대한 과소 분할 문제를 개선하고 가장 우수한 결과를 보였다. 또한, 제안 방법의 Balanced accuracy 측정 결과는 86.41%로 레이블이 있는 데이터 100% 사용한 nnU-Net에 비해 7.46%p 높은 성능을 나타내어 신장 종양 분할 정확도가 가장 크게 향상되었음을 보였다.

표 2는 신장 종양 크기 별 분할 성능 결과를 확인하기 위한 표이다. U-Net, nnU-Net과 같은 지도학습 방법 및 신장 로컬 가이드 맵을 사용하지 않은 준지도학습 방법인 MT의 경우, 종양의 위치를 파악하지 못해 복부 CT 영상에서 작은 크기 및 중간 크기 종양에 과소 분할이 발생하여 전반적으로 낮은 분할 성능을 보였다. 반면에, 신장 로컬 가이드 맵을 사용한 제안 방법의 경우 신장의 위치 정보를 고려하여 종양의 위치 파악에 도움이 되어 작은 크기 및 중간 크기의 종양에 대한 분할 정확도를 개선시켰으며, 큰 종양에서는 레이블이 있는 데이터 100%를 사용한 nnU-Net과 유사한 성능을 보였다. 그러나 신장 로컬 가이드 맵은 신장의 위치 정보만을 제공하기 때문에 종양이 신장에 비해 매우 크거나 신장 없이 종양만 존재하는 슬라이스의 경우 위치 정보가 제한되어 매우 큰 종양에서 레이블이 있는 데이터 100%를 사용한 nnU-Net에 비해 종양에 과소분할 경향을 보여 분할 성능이 낮아졌다.

그림 5는 복부 CT 영상에서 레이블 데이터 비율에 따른 지도학습과 준지도학습 방법의 종양 크기 별 신장 종양 분할의 정성적 결과이다. 그림 5(a)는 레이블이 있는 데이터 100%에서 지도학습을 기반으로 수행한 U-Net의 결과이고, 그림 5(b)는 레이블이 있는 데이터 25%에서 지도학습을 기반으로 U-Net을 수행한 결과이고, 그림 5(c)는 레이블이 있는 데이터 100%에서 지도학습을 기반으로 nnU-Net을 수행한 결과이고, 그림 5(d)는 레이블이 있는 데이터 25%에서

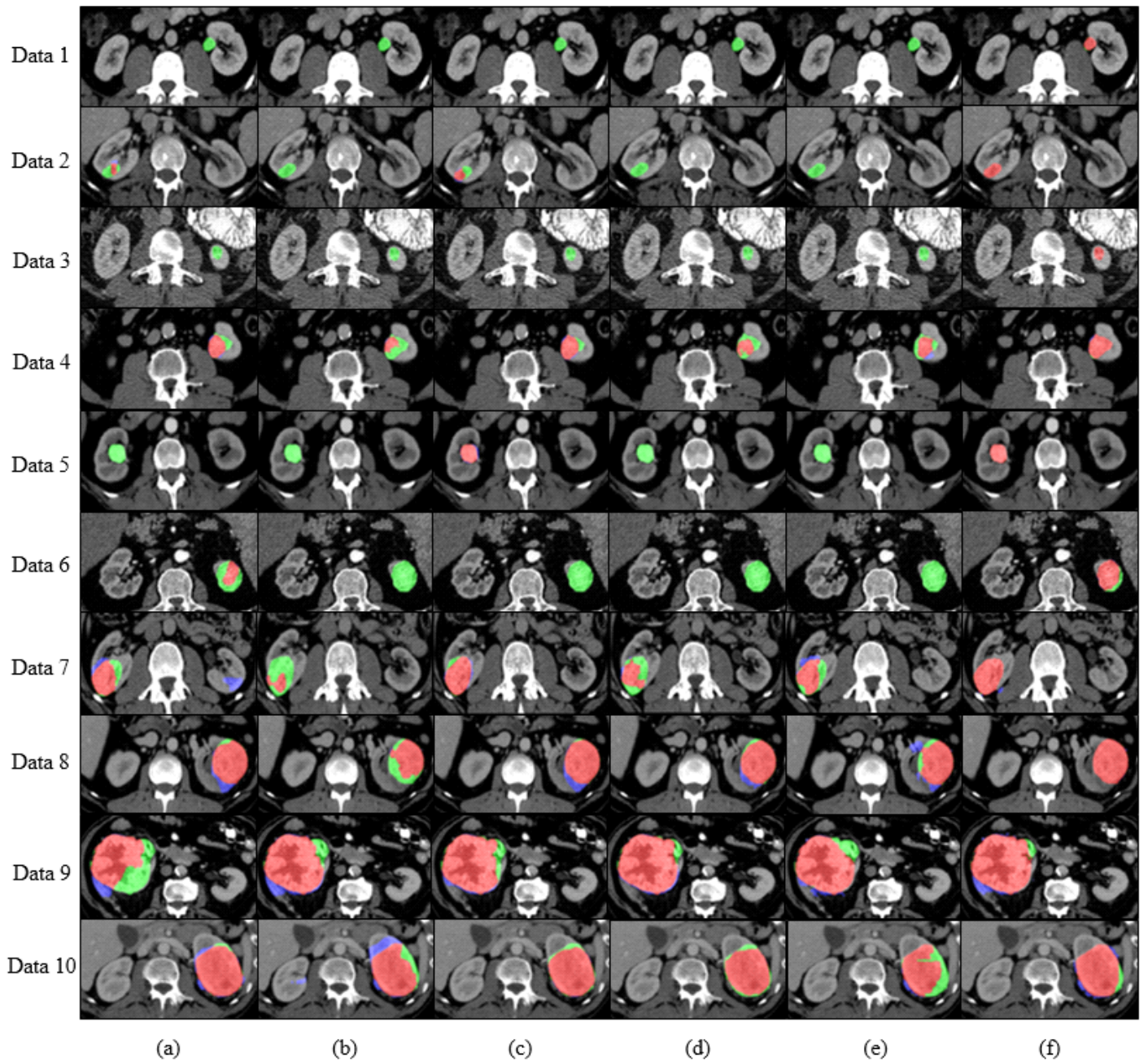
**Table 1 Performance evaluation of kidney tumor segmentation according to labeled data ratios in supervised and semi-supervised learning network**

Methods	Network	Ratio(%)		Metrics(%)			
		Labeled	Unlabeled	Balanced accuracy	F1-score	Recall	Precision
Supervised learning	U-Net[24]	100	0	69.7±16.0	43.4±33.0	39.7±32.3	61.3±39.2
		25	0	64.4±14.3	30.5±28.8	29.5±29.0	39.5±36.3
	nnU-Net[25]	100	0	78.9±17.7	61.2±34.4	58.0±35.6	76.9±34.1
		25	0	68.7±18.5	41.6±38.6	37.8±37.1	59.7±44.7
		25	75	63.4±13.2	26.2±27.2	27.5±26.7	34.6±35.0
Semi supervised learning	MT with Kidney Local Guided Map (Ours)	25	75	<b>86.4±14.7</b>	<b>75.2±26.8</b>	<b>72.9±29.4</b>	<b>86.5±24.6</b>

**Table 2 Performance evaluation of various-sized kidney tumor segmentation according to labeled data ratios in supervised and semi-supervised learning network by kidney tumor size**

Methods	Network	Ratio(%)		Balanced accuracy(%)			
		Labeled	Unlabeled	Small (tumor<2.3)	Medium (2.3≤tumor<4.3)	Large (4.3≤tumor<7.4)	Extremely Large (7.4≤tumor)
Supervised learning	U-Net[24]	100	0	57.7±12.7	62.3±10.5	74.7±13.9	84.8±12.4
		25	0	54.9±12.1	58.2±10.0	68.6±14.4	76.9±8.7
	nnU-Net[25]	100	0	62.7±14.4	70.3±14.2	87.8±13.5	<b>94.9±3.8</b>
		25	0	53.3±8.5	59.1±14.8	75.4±14.8	88.3±11.1
		25	75	56.4±10.9	60.6±12.0	63.1±12.8	74.0±10.1
Semi supervised learning	MT with Kidney Local Guided Map (Ours)	25	75	<b>88.74±16.0</b>	<b>88.9±14.6</b>	<b>88.2±13.4</b>	78.5±11.1

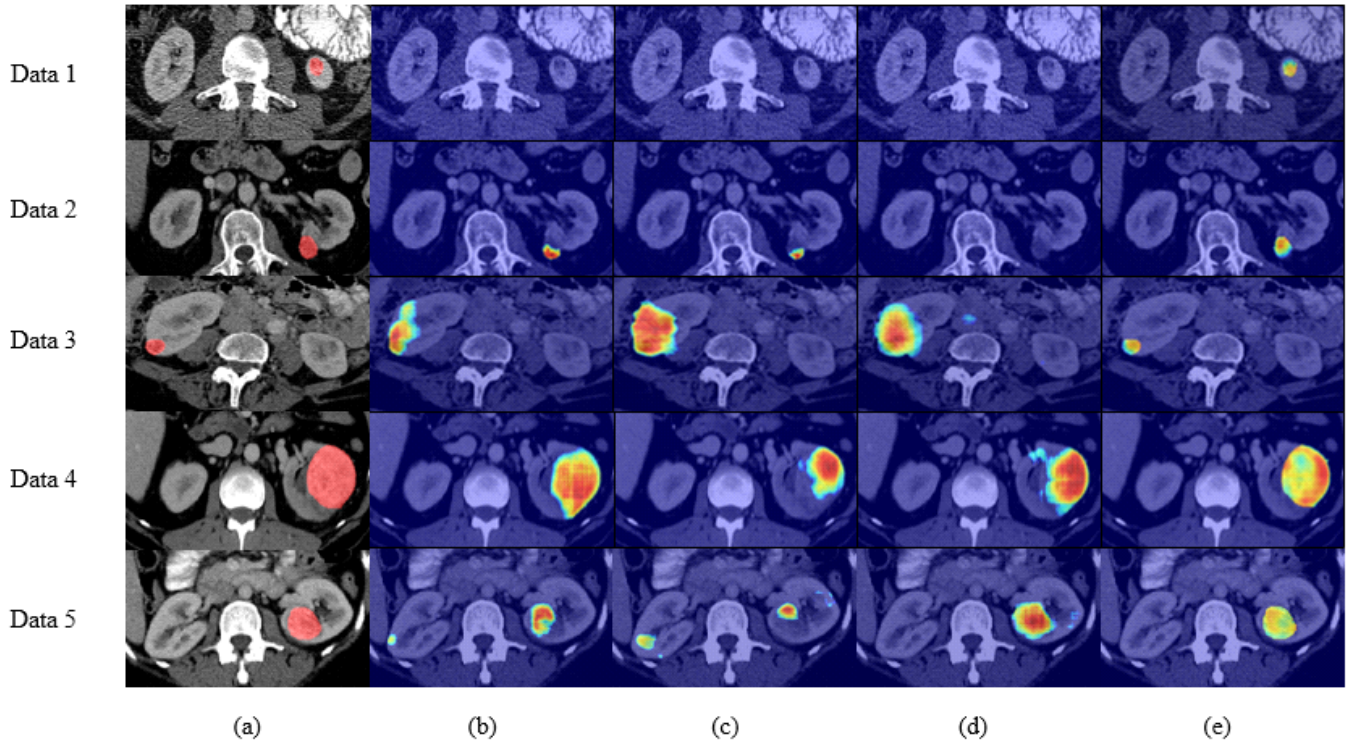




**Figure 5** Qualitative evaluation of various-sized kidney tumor segmentation results in supervised and semi-supervised learning network according to labeled and unlabeled data ratios. (a) U-Net using Labeled data 100%, (b) U-Net using Labeled data 25%, (c) nnU-Net using Labeled data 100%, (d) nnU-Net using Labeled data 25%, (e) MT using Labeled data 25% Unlabeled data 75% and (f) MT with Kidney Local Guided Map using Labeled data 25% Unlabeled data 75%. Row 1~3: small-sized kidney tumor, Row 4~6: medium-sized kidney tumor, Row 7~8: Large-sized kidney tumor, Row 9~10: Extremely large-sized kidney tumor. (Red: overlay, Green: under segmentation, Blue: over segmentation)

지도학습을 기반으로 nnU-Net을 수행한 결과이고, 그림 5(e)는 레이블이 있는 데이터 25%와 레이블이 없는 데이터 75%를 함께 사용하는 준지도학습의 평균-교사 모델을 수행한 결과이고, 그림 5(f)는 레이블이 있는 데이터 25%와 레이블이 없는 데이터 75%를 함께 사용하고 신장 로컬 가이드 맵을 적용한 평균-교사 모델을 수행한 결과이다. 작은 크기 종양 분할 결과인 데이터 1~3의 경우, 작은 크기의 신장 종양으로 인해 종양의 위치 파악이 어려워 (a)~(e) 방법

들은 적절한 분할을 수행하지 못했다. 그에 반해, 신장 로컬 가이드 맵을 사용한 (f) 방법은 종양 위치 파악에 도움을 주어 개선된 분할 결과를 보였다. 중간 크기의 종양 분할 결과인 데이터 4~6의 경우, 작은 크기의 신장 종양보다는 종양의 크기가 커져 (a)~(e) 방법들은 약간의 분할 개선을 보였지만 여전히 과소 분할 문제가 발생하였다. 반면에, 종양 위치 파악에 이점을 보이는 (f) 방법은 가장 좋은 분할 결과를 보였다. 큰 크기의 종양 분할 결과인 데이터 7과 데이터



**Figure 6 Grad-CAM visualization of kidney tumor segmentation results in supervised and semi-supervised learning network according to labeled and unlabeled data ratios. (a) Ground truth, (b) U-Net using Labeled data 100%, (c) U-Net using Labeled data 25%, (d) MT using Labeled data 25% Unlabeled data 75% and (e) MT with Kidney Local Guided Map using Labeled data 25% Unlabeled data 75%.**

8의 경우, 레이블이 있는 데이터 25%를 사용한 (b)와 (d) 방법들은 적은 레이블 데이터로 인해 과소 분할이 발생했으며, 레이블이 있는 데이터 100%를 사용한 (a)와 (c) 방법들은 더 많은 레이블 데이터를 활용하여 과소 분할이 개선되었다. 반면에, 레이블이 있는 데이터 25%를 사용하고 신장 로컬 가이드 맵을 사용하는 (f) 방법은 신장 주변 조직을 신장 종양으로 잘못 분할하는 과대분할 문제가 발생하지만 레이블 데이터를 적게 사용하면서도 비슷한 성능을 보여주었다. 매우 큰 크기의 종양 분할 결과인 데이터 9과 데이터 10의 경우, (f) 방법은 신장의 위치 정보만을 고려하기 때문에 매우 큰 종양에서는 과소 분할되는 경향을 보였다.

그림 6은 복부 CT 영상에서 레이블 데이터 비율에 따른 지도학습과 준지도학습 방법의 종양 크기별 신장 종양 분할의 Grad-CAM[27] 시각화 결과이다. 레이블이 있는 데이터 25%와 100%를 사용한 지도학습 방법의 경우, 신장 주위에 존재하는 작은 크기의 종양 위치를 제대로 파악하지 못하여 신장 종양의 일부에만 집중하거나 신장 영역까지 신장 종양 영역으로 분할하고 있는 것을 확인할 수 있다. 그러나 제안 방법은 신장의 위치 정보를 고려함으로써 적은 레이블 데이터 환경임에서도 다른 방법들에 비해 신장 종양 영역에 집중하여 분할하는 것을 확인할 수 있다.

#### 4. 결론

본 논문에서는 복부 CT 영상에서 신장 주위의 다양한 위치에 나타나는 작은 크기의 신장 종양 분할을 위해 신장 위치 정보를 활용하는 신장 로컬 가이드 맵을 사용하는 준지도학습 방법인 평균-교사 네트워크를 제안하였다. 또한, 신장 종양의 크기에 따른 성능을 분석하였다. 작은 크기의 신장 종양은 복부 CT 영상 내에서 종양의 위치를 파악하기 어려워 분할 시 과소 분할이 일어나는 경향이 있으나, 신장 로컬 가이드 맵을 이용하는 제안 방법은 신장의 위치 정보를 고려하여 종양 주변에 집중함으로써 종양 위치 파악을 개선하였고, 특히 작은 크기의 종양에 대한 분할 성능을 향상시켰다. 그러나 신장 로컬 가이드 맵은 신장의 위치 정보만을 제공하므로 종양에 비해 신장의 크기가 매우 작거나 종양만 존재하는 슬라이스의 경우에는 고려할 수 있는 신장 위치 정보가 부족해 성능에 제한이 있다. 향후 연구로는 종양의 전역적 문맥 정보를 고려하여 크기가 큰 신장 종양에 대한 분할 성능을 개선하고자 한다.

#### 감사의 글

본 연구는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원(No. RS-2023-00207947), 보건복지부의 재원으로 한국 보건산업진흥원의 보건의료기술연구개발사업 지원(HI22C1496) 및 서울여자대학교 학술연구비의 지원(2023-0112)을 받아 수행되었으며 이에 감사드립니다.



## References

- [1] H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram, I. Soerjomataram, A. Jemal, and F. Bray, "Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *A Cancer Journal for Clinicians*, vol. 71, no. 3, pp. 209-249, 2021.
- [2] Z. Du, W. Chen, Q. Xia, O. Shi, and Q. Chen, "Trends and projections of kidney cancer incidence at the global and national levels, 1990-2030: A Bayesian age-period-cohort modeling study," *Biomarker Research*, vol. 16, pp. 8, 2020.
- [3] S. Tangal, K. Önal, M. Yiğman and A.H. Haliloğlu, "Relation of neutrophil lymphocyte ratio with tumor characteristics in localized kidney tumors." *The New Journal of Urology*, vol. 13, no. 1, pp. 12-15, 2018.
- [4] M. Sun, F. Abdollah, M. Bianchi, Q. D. Trinh, C. Jeldres, R. Thuret, Z. Tian, S.F. Shariat, F. Montorsi, P. Perrotte, and P.I. Karakiewicz, "Treatment management of small renal masses in the 21st century: A paradigm shift." *Annals of Surgical Oncology*, vol. 19, pp. 2380-2387, 2012.
- [5] G. Yang, J. Gu, Y. Chen, W. Liu, L. Tang, H. Shu, and C. Toumoulin, "Automatic kidney segmentation in ct images based on multi-atlas image registration." in *Proc. of the 36th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society*, Chicago, IL, USA, pp. 5538-5541, 2014.
- [6] D.K. Kim, Y. Jang, J. Lee, H. Hong, K.H. Kim, T.Y. Shin, D.C. Jung, Y.D. Choi, and K.H. Rha, "Two-year Analysis for Predicting Renal Function and Contralateral Hypertrophy after Robot-assisted Partial Nephrectomy: A Three-dimensional Segmentation Technology Study." *International Journal of Urology*, Vol. 22, pp. 1105-1111, 2015.
- [7] J. Lee, H. Hong, K.H. Rha, "Automatic Segmentation of Renal Parenchyma using Graph-cuts with Shape Constraint based on Multi-probabilistic Atlas in Abdominal CT Images" *Journal of the Korea Computer Graphics Society*, 22(4) 11-19. doi: 10.15701/kcgs.2016.22.4.11, 2016.
- [8] H. Wei, Q. Wang, W. Zhao, M. Zhang, K. Yuan, and Z. Li, "Two-phase Framework for Automatic Kidney and Kidney Tumor Segmentation." 2019 Kidney Tumor Segmentation Challenge: KiTS19, 2019.
- [9] X. Xie, L. Li, S. Lian, S. Chen, and Z. Luo, "SERU: A cascaded SE-ResNeXT U-Net for kidney and tumor segmentation." *Concurrency and Computation: Practice and Experience* Volume 32, Issue 14, 2020
- [10] J. Wen, Z. Li, Z. Shen, Y. Zheng, and S. Zheng, "Squeeze-and-Excitation Encoder-Decoder Network for Kidney and Kidney Tumor Segmentation in CT images." *Kidney and Kidney Tumor Segmentation. KiTS 2021. Lecture Notes in Computer Science*, vol 13168. Springer, Cham. 2022.
- [11] F. Isensee, and K.H. Maier-Hein, "An attempt at beating the 3D U-Net." *ArXiv abs/1908.02182*, 2019.
- [12] X. Hou, C. Xie, F. Li, and Y. Nan, "Cascaded Semantic Segmentation for Kidney and Tumor." 10.24926/548719.002., 2019.
- [13] J. MA, "Solution to the Kidney Tumor Segmentation Challenge 2019." 10.24926/548719.005, 2019.
- [14] M.J. Willemink, W.A. Koszek, C. Hardell, J. Wu, D. Fleischmann, H. Harvey, L.R. Folio, R.M. Summers, D.L. Rubin, and M.P. Lungren, "Preparing medical imaging data for machine learning." *Radiology* 295, 2020.
- [15] S.Y. Huang, W.L. Hsu, R.J. Hsu, and D.W. Liu, "Fully Convolutional Network for the Semantic Segmentation of Medical Images: A Survey." *Diagnostics* 12, 2022.
- [16] Z. Yang, P. Xu, Y. Yang, and B.K. Bao, "A Densely Connected Network Based on U-Net for Medical Image Segmentation." *ACM Trans Multimed Comput Commun Appl* 17, 2021.
- [17] L. Liu, J.M. Wolterink, C. Brune, and R.N.J. Veldhuis, "Anatomy-aided deep learning for medical image segmentation: A review." *Phys Med Biol* 66, 2021.
- [18] D. Karimi, S.D. Vasylechko, and A. Gholipour. "Convolution-Free Medical Image Segmentation Using Transformers." *Lect Notes Comput Sci (Including Subser Lect Notes Artif Intell Lect Notes Bioinformatics)* 12901, 2021.
- [19] Y. Wang, Y. Zhang, J. Tian, C. Zhong, Z. Shi, Y. Zhang, and Z. He, "Double-Uncertainty Weighted Method for Semi-supervised Learning." *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2020.
- [20] J. Wang, and T. Lukasiewicz, "Rethinking Bayesian Deep Learning Methods for Semi-Supervised Volumetric Medical Image Segmentation," in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, LA, USA pp. 182-190, 2022.
- [21] Y. Tang, S. Wang, Y. Qu, Z. Cui, and W. Zhang, "Consistency and adversarial semi-supervised learning for medical image segmentation." *Computers in Biology and Medicine*, Volume 161, 2023.
- [22] L. Hu, J. Li, X. Peng, J. Xiao, B. Zhan, C. Zu, X. Wu, J. Zhou, and Y. Wang, "Semi-supervised NPC segmentation with uncertainty and attention guided consistency." *Knowledge-Based Systems*, Volume 239, 2022.

- [23] A. Tarvainen, and H. Valpola, "Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results." *Adv Neural Inf Process Syst* 2017-December, 2017.
- [24] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," *International Conference on Medical Image Computing and Computer-Assisted Intervention-MICCAI* 2015, Vol.9351, pp. 234-241, 2015.
- [25] F. Isensee, J. Petersen, A. Klein, D. Zimmerer, P. F. Jaeger, S. Kohl, J. Wasserthal, G. Koehler, T. Norajitra, S. Wirtk, and K.H. Maier-Hein, "nnU-Net: Self-adapting framework for u-net-based medical image segmentation." *arXiv preprint arXiv:1809.10486*, 2018.
- [26] E. Yang, C. K. Kim, Y. Guan, B. Koo, and J. Kim, "3D multi-scale residual fully convolutional neural network for segmentation of extremely large-sized kidney tumor." *Computer Methods and Programs in Biomedicine*, Volume 215, 2022.
- [27] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization," *2017 IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, pp. 618-626, 2017.

## 〈 저자 소개 〉

### 정 희 영

- 2021년 8월 서울여자대학교 소프트웨어융합학과 졸업(학사)
- 2021년 9월~현재 서울여자대학교 컴퓨터학과 소프트웨어융합전공 석사과정
- 관심분야 : 의료 인공지능, 딥러닝, 영상분할, 영상처리
- <https://orcid.org/0009-0001-2364-6459>



### 김 현 진

- 2014년 2월 서울여자대학교 멀티미디어학과 졸업(학사)
- 2016년 8월 서울여자대학교 정보미디어학과 졸업(석사)
- 2016년 9월~현재 서울여자대학교 컴퓨터학과 박사수료
- 관심분야 : 의료 인공지능, 딥러닝, 영상분할, 영상처리
- <https://orcid.org/0000-0003-0311-7530>



### 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>

