

# deep learning 기반의 CT 영상 multi class bone segmentation

임희주<sup>1</sup>, 신태훈<sup>1</sup>, 최문형<sup>2</sup>, 정승은<sup>2</sup>, 남윤호<sup>2,3\*</sup>

<sup>1</sup>이화여자대학교 휴먼기계바이오공학부

<sup>2</sup>가톨릭대학교 은평성모병원 영상의학과

<sup>3</sup>가톨릭대학교 서울성모병원 영상의학과

## Multi class bone segmentation of CT image using deep learning

Heejoo Lim<sup>1</sup>, Taehoon Shin<sup>1</sup>, Moon-Hyung Choi<sup>2</sup>, Seung-eun Jung<sup>2</sup>, Yoonho Nam<sup>2,3\*</sup>

<sup>1</sup>Ewha Womans University, Division of Mechanical and Biomedical Engineering

<sup>2</sup>Department of Radiology, Eunpyeong St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea

<sup>3</sup>Department of Radiology, Seoul St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea  
gmlwn97@ewhain.net

### Abstract

Bone segmentation of CT images plays an important role in clinical diagnosis. Although deep learning algorithms have been adopted in many fields of body segmentation problems, manual segmentations are still inevitable for constructing the ground truth data. Here, we present an automatic bone segmentation approach which minimizes users' manual interventions. We obtained the ground truth data from femur images by using thresholding and convolutional neural network. U-net was used for training dataset and was proceeded in four different environment: (1)multi segmentation, (2)single segmentation, (3)pretrained multi segmentation, and (4)pretrained single segmentation. . The dice coefficient of the model was calculated as 0.945, 0.945, 0.941, 0.932, respectively of each of the four cases.

### I. 연구 배경

CT image에서의 bone segmentation은 bone metastasis detection과 같은 질병 진단에 중요한 역할을 하지만 clinical routine에서의 CT scan은 종종 낮은 radiation dose에서 행해져서 영상 quality 가 낮다는 어려움이 있다 [1]. 단순한 thresholding값으로 인한 segmentation은 intensity가 낮은 조직(bone marrow)은 포함시킬 수 없게 때문에 cortical bone과 bone marrow를 모두 보여줄 수 있는 automated segmentation의 방법이 필요하다. 또한 bone 만을 대상으로 하는 single class segmentation의 경우 조영제 투여 시 bone 뿐만 아니라 다른 부가적인 vessel이 같이 잡힐 수 있기 때문에 추후 어려운 task를 위해 multi-class segmentation을 해주는 것이 더 바람직하다. 최근 Automated abdominal segmentation of CT방법[2]이 제시되었으나 여전히 ground truth data는 manual segmentation이 필요하다는 점이 한계로 남아있다. 본 연구에서는 manual segmentation을 최소화한 automated multi class bone segmentation 방법을 제시한다.

### II. 연구 방법

#### 1. 데이터셋

91명의 peripheral CT image와 65명의 abdominal CT image가 연구에 이용되었다. Unet 데이터셋은 54명(83%)의 train dataset, 11명(17%)의 validation dataset, 이루어져 있으며 총 4번의 cross validation을 거쳤다. 제한된 dataset으로 인해 모든 데이터를 3-D에서 2-D로 변환해주어 최종적으로 3088개의 train data와 646개의 validation data가 모델에 쓰이게 되었다.

#### 2. ground truth segmentation

Manual segmentation의 최소화를 위해 우선적으로

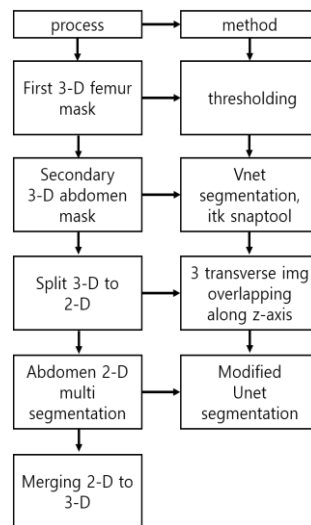


그림 1. bone segmentation의 순서도

91명의 peripheral CT의 thigh 영역에 해당하는 slice들에서 thresholding값을 주어 3-D bone mask를 얻었다. Thigh 영역을 선택한 이유는 thresholding만으로 femur영역을 비교적 쉽게 구분해 낼 수 있기 때문이다. 이후 V-net을 이용하여 femur mask로부터 3-D abdomen에 대한 second mask를 획득하였다. V-net 네트워크의 train data로는 72명, validation data로는 19명을 사용하였으며, test data로는 10명의 복부 이미지를 사용하였다. 이 때 각각의 input data들은 interpolation과 random crop을 이용하여 512x512x32\*6의 size와 0.5x0.5x5mm<sup>3</sup>의 spatial resolution을 갖도록 조정해주었다. 네트워크의 convolution kernel size는 기존의 Vnet과 동일하게 3x3x3을 사용하였다[3]. 이로부터 얻어진 secondary mask는 femur image를 ground truth label로 사용했기 때문에 bone뿐만 아니라 aorta, kidney와 같은 부가적인 장기들도 포함한다. 따라서 ITK-SNAP toolbox를 이용해 bone mask와 다른 장기들이 서로 다른 intensity 값을 갖도록 manual segmentation을 해주었다. Semi-automated segmentation을 통해 얻은 bone, kidney, vessel, Liver에 대한 3-D mask는 이후 transverse section을 기준으로 2-D로 변환되었다. Image는 320 x 320의 크기를 갖도록 resample되었는데 이때 RGB 채널별로 서로 다른 인접한 transverse plane image가 들어갔다. 하나의 이미지

채널 당 3개의 인접한 이미지 정보가 담겨 있기 때문에 이후 U-net을 통해 얻어진 2-D output image를 다시 3D로 재구성할 때 더 자연스러운 이미지를 얻을 수 있다

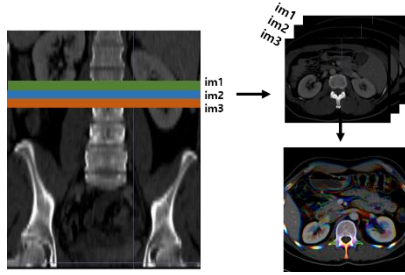


그림 2. 3-D image의 2-D로의 변화. 각 채널별로 z축의 정보를 갖도록 했다.

### 3. 네트워크

ResNet-34 encoder로 하는 변형된 U-net 딥러닝 모델을 pytorch 및 fast.ai library 를 이용하여 구현하였으며, 제한된 GPU memory 를 고려하여 16비트로 최적화 연산을 진행하였다. 네트워크는 encoder부분인 ResNet34와 batch norm, ReLu 그리고 convolution으로 이루어져 있으며  $10^{-3}$ 의 weight decay 값과 32의 batch size를 갖는다. 총 25번의 epoch으로 1 cycle policy에 따라 model을 fitting해주었다[4]. 이때 encoder 네트워크의 초기 parameter는 ImageNet data로 pretrained된 것과 random initialization을 시켜준 것을 모두 테스트하여 비교하였다. 이후 2D output image를 환자별로 이미지 slice들을 쌓아 3-D 이미지로 변환 시켜주었다. 우리는 변환된 이미지에서 bone(class=1)에 해당되지 않는 pixel 값들을 0으로 초기화한 후 Dice Loss를 주어 계산했는데 이는 우리가 관심있는 class의 loss 값을 살펴보기 위함이다. 이후 아래 도식처럼 각 cross validation 단계마다 산출한 loss값들의 평균을 가지고 최종 모델의 loss값을 도출하였다.

$$\frac{1}{N} \sum_i^N (1 - 2 \times \frac{\sum output_i mask_i}{\sum output_i^2 + \sum mask_i^2 + eps})$$

### III. 연구 결과

Unet 모델을 5명의 실험 data에 적용한 dice coefficient 결과는 다음 그래프와 같다.

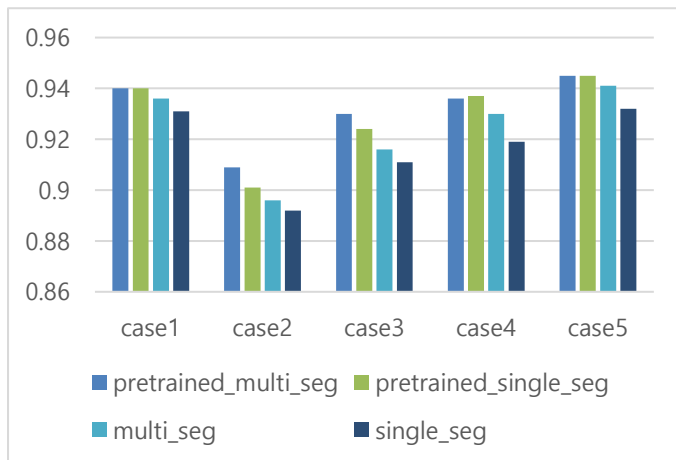


그림 3. 실험 data에 따른 dice coefficient 결과값. Transfer learning을 하지 않은 어려운 task일수록 multi class segmentation이 더 좋은 성능을 보인다.

Case 5에 관해 ImageNet database로 transfer learning을 거친 model의 dice coefficient(0.945)값은 그렇지 않은 model의 dice loss(0.941)의 99.5%로 parameter들이 미리 학습된 network가 더 높은 성능을 보인다. Pretrained된 model의 경우 single class segmentation과 multi class segmentation 모두 비슷한 accuracy를 보여주었지만, random initialized된 model의 경우 대부분의 data에 대해 multi class로 segmentation을 해준 output image가 그렇지 않았을 때보다 더 높은 accuracy를 얻었다. Case 2와 Case 5의 output image는 아래 그림과 같다.

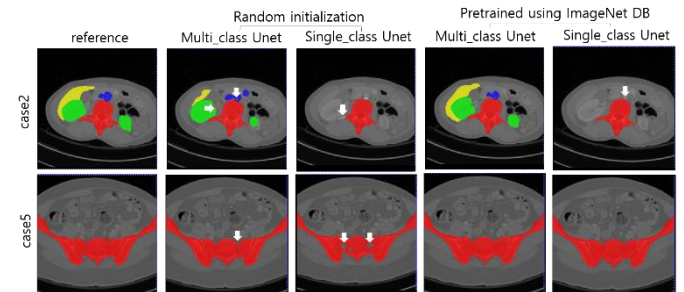


그림 4. Case 2와 Case 5에 대한 output image. 각각의 model에 대해 잘못 예측한 부분은 화살표로 표시되어 있다.

Case 5의 경우 transfer learning의 유무와 상관없이 단일 class segmentation이 multi class segmentation보다 bone marrow 영역을 잘 segment하지 못했음을 알 수 있다.

이 모델은 transfer learning을 시켜준 경우 한 개의 class로 segmentation했을 때와 거의 유사한 정확도를 갖는 한계점이 있다. 또한 65명이라는 적은 수의 data를 사용했기 때문에 overfitting되었을 가능성을 무시할 수 없다. 그러나 우리는 조영제를 투여한 CT image나 질병을 갖고 있는 환자 image 등 좀더 다양하고 많은 train data로 학습이 된다면 multi class segmentation이 단일 class로 학습시킬 때보다 더 좋은 결과를 가져올 것이라 예상된다. 현재 개발된 네트워크를 바탕으로 bone metastasis를 판단하는 알고리즘 연구가 진행중이다.

### IV. Acknowledgements

본 연구는 한국연구재단의 지원을 받아 수행됨 (NRF-2019R1F1A1058872)

### V. 참고문헌

- [1] André Kelin, and Jan Warszawski, "Automatic bone segmentation in whole-body CT images", Vol 14, No.1, p.21-29, 2018
- [2] Alexander D. Weston, and Panagiotis Korfiatis, "Automated Abdominal Segmentation of CT Scans for Body Composition Analysis Using Deep Learning", Vol 290, No.3, 2018
- [3] Fausto Milletari, and Nassir Navab, "V-net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation", 2016
- [4] Leslie N. Smith, "Cyclical Learning Rates for Training Neural Networks", 2017