

Pancreas Segmentation using Multiple Concatenated U-Net Model for Medical Image Systems

의료 영상 시스템에서 다중 연쇄 U-Net 모델을 이용한 췌장 영역 분할 탐지

심사용 작성시 저자정보는 삭제하여 올려주시기 바랍니다.

Abstract – The pancreas, which is small in size and superimposed on other organs, is difficult to accurately locate on a medical image. In this paper, we propose a multiple concatenated U-net(MCU-net) model that can more accurately segmentation the position of the pancreas in an abdominal image through machine learning. The MCU-net model is a model that concatenated multiple U-net structures in the final layer with various sizes of patches, which perform better than the original U-net model. Experiments showed that the multiple concatenated U-net model had 91.6% accuracy, about 2% higher than the original U-net model.

Key Words : Pancreas segmentation, Image processing system, Deep learning, Artificial intelligence

1. Introduction

인체 장기 중 췌장은 치사율과 위험성이 높은 병변이 자주 발견되기에 의료 영상을 통해 정확한 진단이 필요한데, 크기가 작고 위장의 뒤쪽 십이지장과 연결된 위치적 특징 등으로 인해 CT와 같은 영상에서 그 위치와 모양을 정확히 찾기 어렵다. 이를 해결하기 위해 최근에는 CNN(Convolutional Neural Networks) 모델 기반 딥러닝 시스템을 이용해 췌장을 분할(segmentation)하고 있으나, 폐와 같은 다른 장기가 98% 이상으로 잘 분할 [1]되는 반면 췌장은 아직 90%를 넘지 못하고 있다 [2]. 따라서 본 논문에서는 다양한 패치 크기의 U-net [3] 구조를 다중 연쇄 결합(concatenation)하여 췌장 분할 성능을 향상시킨 MCU-net 모델을 제안한다.

2. Related Work

췌장 분할을 위한 초기 연구는 통계적 모델을 쓰거나 다중 아틀라스(multi-atlas) 기법을 사용했다 [4]. 그런데 아틀라스 기법은 분할하고자 하는 대상 이미지가 훈련했던 아틀라스의 형태와 다른 경우 분할 정확도가 낮아지는 한계점이 있다 [5]. 반면, 기계 학습의 CNN 모델을 발전시킨 U-net [3], Attention U-net[6], RSTN[7]은 췌장 분할 정확도가 아틀라스 기법에 비해 모두 크게 향상되었다. 입력 이미지로부터 특정 기관 및 세포를 분할하는 U-Net 모델은 U자 형태의 네트워크를 구성한다. 이때, 수축 경로(contracting path)과 확장 경로(expanding path)를 통해 이미지의 사소한 부분은 배제되고 특징이 되는 부분은 유지 전달한다. U-net 모델을 기반으로 한 Attention U-net[6]

모델은 기존 U-net 시스템에 비해 특징에 집중하는 개념인 Attention Gate (AG)를 추가한 모델로, U-net에서 이미지의 특징 정보를 게이트를 통해 선택적으로 유지한다. AG로 훈련된 모델들은 특정 작업에 유용한 두드러진 특징을 강조하면서 입력 이미지의 관련 영역을 억제하는 방법을 암묵적으로 학습한다. 이로써 모델 민감도와 예측 정확도를 높였고, 패치 단위로 영역을 탐지해서 훈련 속도 또한 빠른 편이다. 이외에도 R-CNN과 같이 2단계 학습을 진행하는 RSTN[7]은 신장, 비장, 췌장 등의 몸의 작은 장기들을 분리하는 방법을 제시했다. 이 방법은 첫 번째 단계에선 찾고자 하는 장기의 위치를 찾고, 두 번째 단계에서는 장기 식별을 수행하여 영역을 분할한다. 그러나 RSTN은 지나치게 긴 학습 시간을 요구한다. 따라서 본 논문에는 다양한 패치 크기를 가지는 Attention U-net 모델을 연쇄 결합하여 췌장 분할 성능을 향상하는 MCU-net(Multiple Concatenated U-net) 모델을 제안한다.

3. Proposed Method

제안하는 MCU-net 모델은 서로 다른 패치 크기를 가지는 U-net 모델을 연쇄 결합한 구조이다. 탐색이 끝난 곳을 다시 탐색하는 필터 슬라이딩(filter-sliding) 기반 특징 탐색 방식과 달리 패치 기반 탐색 방식은 지정된 간격인 스트라이드(strides)만큼 건너뛰면서 입력데이터를 순회하기 때문에 CNN과 달리 중복된 영역에 대한 특징 탐색을 줄임으로써 학습 속도의 향상을 가져왔다. 그러나 패치 기반 탐색 방식은 스트라이드만큼 건너뛰며 탐색하기 때문에 맥락 정보 손실이 일어나게 되는데 이를 보완하기 위해서 본 논문에서는 Faster-RCNN [8] 모델의 anchor box와 같이, 다양

한 크기의 패치를 갖는 U-net 모델을 마지막 단계에서 결합함으로써 영상의 특징을 다양한 관점으로부터 추출하여 정보 손실을 줄인다.

MCU-net 모델은 서로 다른 크기의 패치를 가진 U-Net 들을 마지막 단계에서 서로 결합한 후 활성화 함수를 통해 예측된 값을 출력한다. 이 모델은 다양한 패치 크기에 의한 결괏값을 동시에 고려했기 때문에 단일 패치 모델에 비해 더 많은 context 정보를 얻는다. MCU-Net 모델은 입력 이미지 I 를 패치 크기가 서로 다른 n 개의 U-net 중 i 번째 U-net인 L_i 를 통해 특징 벡터 f_{L_i} 를 추출하고 이러한 특징 벡터들을 연결한 특징 벡터인 U 를 통해 전체 영상 중에서 채장 영역을 탐지하여 결과 이미지인 O 를 출력한다. 추출한 특징들이 결합된 U 는 <식 1>과 같이 나타낼 수 있으며, U 를 통해 탐지한 채장 영역은 활성화 함수인 Softmax를 통해 결과 이미지 O 를 출력한다.

$$U = \text{Concatenate}(f_{L_1}, f_{L_2}, \dots, f_{L_n}) \quad (1)$$

<그림 1>은 본 논문에서 제안하는 MCU-net 모델에서 패치 크기가 각각 (3,3), (5,5), (7,7)인 U-net 3개를 결합한 구조를 보여주고 있다. 각 Convolution에서는 ReLU 활성화 함수 사용했으며, 수축 경로에서는 2x2 max pooling을 사용했다. 매 계층을 내려갈 때마다 down sampling을 하여 채널 개수를 2배씩 늘렸다. 확장 경로에선 수축 경로 과정에서 줄어든 크기를 다시 키워야 하므로 2x2 up-convolution이 사용했고, 채널 개수는 반으로 줄었다. 회색 계층은 특징맵(feature map)의 복사본을 의미한다. 해당 구조에서 마지막 계층에선 각 U-net들을 256x256x1로 1x1 convolution하고 이를 연쇄 결합한다. 채널 개수는 진행된 U-net의 개수인데, 여기에서는 패치 크기가 다른 3개의 U-net을 연결했기 때문에 채널 개수는 3이 된다.

손실 함수로는 다이스 계수(dice coefficient)를 사용했다. 다이스 계수는 예측 영역과 실제 영역이 얼마나 일치하는지 그 비율을 나타내는 지표로써 예측 영역 X 와 실제 영역 Y 가 일치할수록 1에 가까운 값이 출력되고, 예측 영역이 실제 영역과 일치하지 않을수록 0에 가까운 값이 출력된다. 따라서 다이스 계수가 0이면 전혀 일치하지 않는다는 의미이고, 계수가 1이면 완전히 일치한다는 뜻이다.

4. Experiments

실험에서는 채장 분할에서 널리 사용되는 Pancreas-CT set[9]을 이용해 모델 학습 및 평가를 진행하였다. 이미지 해상도는 512 x 512 픽셀이고 총 19,328개의 복부 CT 이미지를 포함한다. 본 논문에서는 이 set을 4개의 부분 집합(fold)으로 나누어 교차 검증했다. [10]의 코드를 기반으로 학습 모델을 구현하였으며 표 1과 같은 환경에서 실험하였다.

<표 2>는 실험의 정량적인 결과를 보여주고 있다. 표에서 DSC는 dice coefficient를, Std. Dev.은 표준 편차를 의미한다. 표에 제시된 결괏값들은 해당 부분 집합을 테스트 결과로 사용했을 때의 평균 결과이다. M1 모델은 기존에 설계되어 있던 순수 U-net 기본 모델[3]을 채장 분류에 적용한 것 [13]이고, M2 모델은 패치 크기가 같은 2개의

U-net을 앙상블(ensemble)한 모델이다. M2 모델에서 패치 크기를 통일시키지 않을 경우 학습 케이스에서 약 83% 정도로 기존 모델에 비해 정확도가 3%정도 낮게 나왔으며, 테스트 케이스에선 측정이 불가능할 정도로 정확도가 낮아졌다. 따라서 M2 모델에서는 앙상블의 성능을 높이기 위해 패치 크기를 (3,3)로 통일시켰다. M3 모델은 3개의 U-net을 앙상블한 모델이고 패치 크기는 모두 (3,3)이다. M4 모델은 2개의 U-net으로 구성된 MCU-net 모델로써 각각의 패치 크기는 (3,3), (5,5)이다. M5 모델의 경우 총 3개의 U-net으로 구성된 MCU-net 모델로써 각각의 패치 크기는 (3,3), (5,5), (7,7)이다. 표에서 알 수 있듯이 앙상블을 적용한 모델의 경우 큰 정확도의 향상은 없었으며, 첫 번째 부분 집합에서는 오히려 정확도가 낮아졌다. 이에 반해 본 논문에서 제안하는 모델(M5)은 모든 부분 집합에 대해서 정확도가 높아진 것을 확인할 수 있다.

표 1 실험 환경

Table 1 Experimental environments

Parameters	Values
Python	ver. 2.7.12
Tensorflow	ver. 2.1.0
Keras	ver. 2.0.8
Ubuntu	ver. 16.04 LTS
Learning epoch	10
Learning rate	1e-5
Optimizer	Adam
Loss function	Dice coefficient

표 2 Multiple Concatenated U-net 모델 및 U-net의 성능 실험

Table 2 Performance of MCU-Net and U-Net models

Dataset (fold)	1		2		3		4		Total AVG
	AVG	DEV	AVG	DEV	AVG	DEV	AVG	DEV	
M1	0.864	0.038	0.906	0.013	0.907	0.010	0.915	0.014	0.896
M2	0.868	0.034	0.917	0.009	0.918	0.008	0.923	0.012	0.906
M3	0.860	0.034	0.918	0.009	0.917	0.008	0.923	0.011	0.904
M4	0.862	0.040	0.915	0.011	0.917	0.008	0.921	0.011	0.903
M5	0.876	0.033	0.928	0.008	0.929	0.008	0.931	0.008	0.916

<표 3>은 선행 연구들과 제안하는 MCU-net 모델의 성능을 비교한 실험 결과이다. 결과에서 알 수 있듯이 본 논문에서 제안하는 MCU-net 모델이 이전의 연구들에 비해 가장 나은 성능을 보여 주고 있다.

표 3 선행 연구 간 성능 비교 실험

Table 3 Performance comparison between models

Models	DSC
Recurrent Saliency Transformation Network Model [7][11]	0.846
pure U-Net Model [3]	0.822
Attention U-Net Model	0.832
U-Net Model	0.898
MCU-Net Model (Proposed)	0.916

<그림 2>는 U-net과 MCU-net 모델의 결과를 이미지를 통해 정성적으로 보여 주고 있다. 이를 통해 제안하는 MCU-net 모델이 기존 U-net 모델에 비해 채장 영역을 더 잘 분할함을 눈으로 확인할 수 있다.

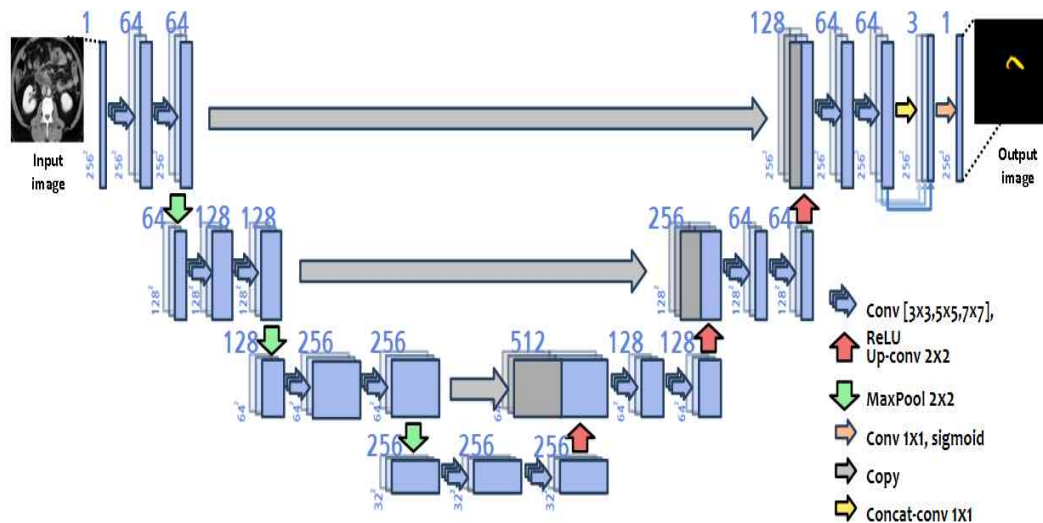


그림 1 MCU-Net 모델 구조
Fig. 1 The architecture of MCU-Net model

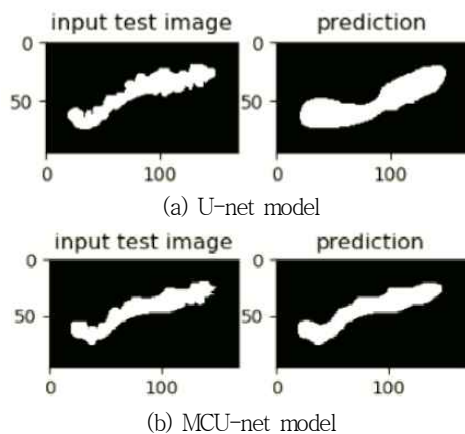


그림 2 U-Net과 MCU-Net 모델 결과 영상
Fig. 2 Result images of U-Net and MCU-Net models

5. Conclusions

본 논문은 영상에서 췌장을 분할하기 위해 패치 크기가 서로 다른 U-net 구조를 결합한 MCU-Net 모델을 제안하였다. 제안하는 모델의 정확도는 약 91.6%로 다른 학습 모델들보다 높은 성능을 보였다. 췌장은 장기 중에서도 특히 탐지난이도가 높다. 따라서 MCU-net은 췌장뿐만 아니라 다른 장기의 분류 문제에서도 좋은 결과를 보일 것이라 예상된다. 향후 연구로는 U-net이 아닌 다른 모델에 다중 결합을 적용했을 때 더욱 성능이 나아질 수 있는 분야가 무엇인지 찾고 그 결과를 정량적으로 분석해 볼 예정이다.

감사의 글

본 연구는 2022년도 경기대학교 학술연구비(일반연구과제) 지원에 의하여 수행되었음.

References

- [1] M. Z. Alom et al., "Recurrent Residual Convolutional Neural Network based on U-Net (R2U-Net) for Medical Image Segmentation", arXiv preprint arXiv:1802.06955, 2018.
- [2] Y. Liu and S. Liu, "U-net for Pancreas Segmentation in Abdominal CT Scans", IEEE Int. Sympo. on biomedical imaging, 2018.
- [3] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional Networks for Biomedical Image Segmentation," Int. Conf. on Medical image computing and computer-assisted intervention, 2015.
- [4] R. Wolz et al., "Automated Abdominal Multi-organ Segmentation with Subject-specific Atlas Generation", IEEE TMI, vol. 32, no. 9, 2013.
- [5] M. Oda et al., "3D FCN Feature Driven Regression Forest-based Pancreas Localization and Segmentation, DLMI, pp. 222-230, 2017.
- [6] O. Oktay et al., "Attention U-net: Learning Where to Look for the Pancreas," arXiv preprint arXiv:1804.03999, 2018.
- [7] Q. Yu et al., "Recurrent Saliency Transformation Network: Incorporating Multi-stage Visual Cues for Small Organ Segmentation," IEEE Conf. on Computer Vision and Pattern Recognition, 2018.
- [8] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks", Advances in neural information processing systems, vol. 28, 2015.
- [9] H. Roth et al., "Data from Pancreas-CT", The Cancer Imaging Archive, 2016
- [10] <https://github.com/snapfinger/pancreas-seg>
- [11] https://github.com/twni2016/OrganSegRSTN_PyTorch