|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **복부 CT 영상에서 다중 concatenation U-net 모델을 이용한 췌장 분류**   |  |  | | --- | --- | | 이 상 윤[[1]](#footnote-1) | 김 남 기[[2]](#footnote-2)1\* | | Sang-yun Lee | Namgi Kim |   **pancreas Segmentation using Multiple Concatenated U-net Model in Abdominal CT Images**  **요 약**  크기가 작고 다른 장기에 중첩되어 있는 pancreas은 CT image에서 그 region을 정확히 찾아 segmentation하기 어렵다. 본 논문에서는 machine learning을 통해 CTimage에서 pancreas region을 보다 정확히 segmentation할 수 있는 Multiple Concatenated U-Net (MCU-Net) model을 제안한다. MCU-Net model은 filter size가 다른 여러 개의 U-Net을 final layer에서 concatenated한 model이다. 실험 결과에 따르면 MCU-Net model의 accuracy는 91.6%로 기존 U-Net model에 비해 약 2% 높은 성능을 보였다.  **ABSTRACT**  The pancreas, which is small in size and superimposed on other organs, is difficult to accurately locate on the CT 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 CT images through machine learning. The multiple concatenated U-net model is a model that concatenated multiple U-net in the final layer with various sizes of filters, which perform better than the original U-net model. Experiments showed that the multiple concatenated U-net model had 91.6 percent accuracy, about 2 percent higher than the original U-net model. |

**1. 서론**

pancreas은 pancreatic cancer이나 pancreatitis과 같이 치사율과 위험성이 높은 병이 자주 발견되는 장기이기에 image analysis을 통한 병의 조기 발견이 매우 중요하다. pancreas에 발생한 병을 정확히 진단하기 위해서는 Radiology specialist가 많은 시간을 할애하여 abdominal CT image를 분석해야 한다. 그럼에도 불구하고 pancreas image에 대한 analysis accuracy는 다른 장기에 비해 낮다. 이를 개선하기 위해 최근에는 deep learning 기반의 automated medical image analysis system이 속속 등장하고 있다. CT image을 통해 pancreas에 발생한 병을 자동적으로 analysis하기 위해서는 먼저 CT image에서 pancreas region을 정확히 segment해야 한다. pancreas에 병이 있는지를 진단하기 위한 CT image는 최소 8channel 이상의 CT 장비에서 촬영한다. 4channel 또는 2channel의 CT 장비에서 촬영된 image은 해상도가 떨어져 pancreas에 대한 정확한 진단을 내리기 어렵기 때문이다. 그러나 8channel 이상의 CT 장비에서 촬영된 CT image에서도 크기가 작고 모양이 다양하며 위장의 뒤쪽 십이지장과 concatenated된 위치적 특성을 가지는 pancreas의 region은 정확히 segment하기 어렵다.

최근에는 CNN (Convolutional Neural Networks) 기반의 deep learning technology을 이용해 CT image에서 특정 장기의 region을 segment하거나 volume을 파악하는 computer-aided detection system이 활발히 대두되고 있다. CNN 기반 deep learning technology을 사용하면 lung 또는 retina blood vessel을 98%이상의 accuracy로 segment할 수 있다 [11][12]. pancreas 또한 이러한 CNN 기반 deep learning technology을 이용해 그 위치를 segment하는 방법들이 최근 제안되고 있다. [1][2][3]. 그러나 이 기술을 이용한다고 해도 pancreas의 region을 정확히 segment하는 것은 다른 장기의 region을 segment하는 것보다 훨씬 어렵다. TCIA (The Cancer Imaging Archive) CT-82 dataset[8]을 이용한 accuracy verification에서 pancreas segmentation는 가장 성능이 좋은 deep learning model을 적용할 때에도 DSC(Dice Similarity Coefficient)가 90%를 넘지 않는다 [13]. 따라서 본 논문에서는 pancreas region segmentation에서 이 보다 더 높은 accuracy를 가질 수 있는 deep learning model을 제안한다.

현재 pancreas region의 segmentation accuracy가 80% 이상인 deep learning model로는 U-Net model [1], Attention U-Net model [2], RSTN model[3] 등이 있다. 이 연구들은 모두 단일 patch 또는 단일 filter size를 사용한다. CNN 기반 deep learning model에서는 filter를 통해 image의 feature을 capture하는데 filter의 크기를 다르게 하면 feature도 달라진다. 이러한 사실을 바탕으로 본 논문에서는 filter size 가 다른 U-Net model들을 final layer에서 concatenated하여 pancreas region을 보다 정확하게 segmentation하는 MCU-Net (Multiple Concatenated U-Net) model을 제안한다.

본 논문의 구성은 다음과 같다. 2장에서는 기존 pancreas region을 segment하기 위해 사용된 deep learning model들에 대해 알아본다. 3장에서는 본 논문에서 제안하는 MCU-Net model의 구조 및 원리에 대해 설명하고. 4장에서는 실험 환경 및 실험 결과를 밝혀 다른 모델과의 성능을 비교한다. 마지막으로 5장에서는 결론 및 향후 연구 방향에 대해 논한다.

**2. 관련 연구**

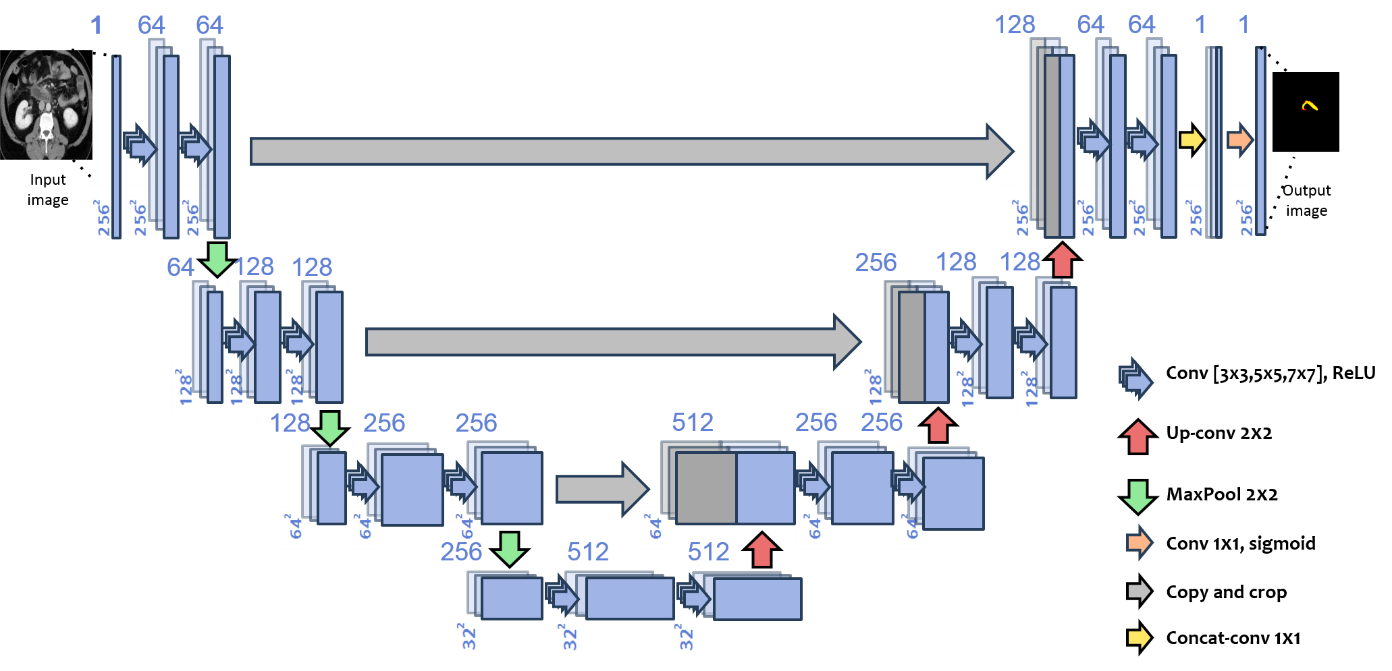
pancreas region을 segment하기 위해 초기에는 statistical model을 쓰거나 multi-atlas method을 사용했다. 그런데 statistical model은 전반적으로 segmentation accuracy가 낮았고, atlas method은 segmentation하고자 하는 pancreas가 training했던 atlas의 형태와 다른 경우 accuracy가 낮아지는 문제점이 있었다. atlas 기반 methods의 accuracy는 TCIA CT-82 dataset을 사용할 때 DSC가 평균 73.9%이다 [9][10]. 반면, CNN 기반 model인 U-Net [1], Attention U-Net [2], RSTN (Recurrent Saliency Transformation Network)  [3]은 같은 dataset을 사용한 결과 segmentation accuracy가 모두 80%이상이였다.

input image로부터 특정 기관 및 세포를 segment하는 U-Net model은 U자 형태의 네트워크로 구성된다. U-Net는 contracting path와 expanding path를 통해 image의 사소한 부분을 배제하고 feature가 되는 context information을 끝까지 전달한다. U-Net은 filter를 통과할 때 edge image의 context information을 loss없이 capture하기 위해 the values directly in the opposite direction of the edge 를 넣는 reflection padding을 사용한다. 따라서 edge value에 zero를 넣는 zero padding에 비해 information loss가 적다. U-Net model을 변형한 Attention U-Net[2] model은 feature에 집중할 수 있는 attention gate를 추가하였다. attention gate는 image의 feature information를 선택적으로 유지할 수 있게 도와 준다. Attention gate로 training된 model들은 target을 보다 잘 예측하기 위해 image의 두드러진 feature을 위주로 learning하게 됐고, 동시에 target과 관련되지 않은 region을 억제하는 방법을 implicit하게 learning한다. 이로써 model의 sensitivity와 accuracy를 개선하였다. Attention U-Net은 U-Net을 기반으로 만들어졌기 때문에 기존 CNN model처럼 image region을 겹쳐서 traverse하지 않고 patch 단위로 traverse해 training 속도가 빠르다. TCIA CT-82 dataset로 실험한 결과에 따르면 attention gate는 computing efficiency는 유지하면서 다양한 data learning을 통해 U-Net의 성능을 개선함을 알 수 있었다. RSTN [3]은 R-CNN과 같이 two-stage approach로 학습을 진행하여 신장, 비장, pancreas 등 몸의 작은 장기들을 segment하는 model이다. RSTN model은 첫 번째 stage에서 찾고자 하는 장기의 위치를 찾고, 두 번째 stage에서는 장기를 식별한다.

RSTN model은 training할 때 U-Net model에 비해 약 10배 정도의 시간을 더 필요로 한다. 그러나 data를 이용해 model을 training시킨 뒤에는 training된 model에 target image만 input하면 되는 deep learning 기반의 automated segmentation system에서는 training에 소요되는 시간이 크게 중요하지 않다. training 시간 보다는 model의 accuracy가 훨씬 더 중요하다. 따라서, 비교적 training 시간은 길더라도 accuracy가 높게 나오는 model이 더 가치가 있다고 볼 수 있다. 그러므로 본 논문에서는 여러 개의 U-Net을 concatenated하여 pancreas region을 보다 정확하게 segment하는 Multiple Concatenated U-Net (MCU-Net) model을 제안한다.

**3. MCU-Net model**

U-Net model은 CNN에서 filter와 비슷한 역할을 하는 patch 단위로 feature를 extract한다. U-Net을 확장한 MCU-Net model도 patch 단위로 input image를 traverse하여 feature를 extract한다. 같은 region을 여러 번 traverse하는 filter sliding 방식과 달리 지정된 strides만큼 건너뛰면서 input data를 traverse하는 patch 방식은 image region에 대한 중복된 feature searching이 없어 학습이 빠르다. 그러나 feature를 충분히 extract하지 못해 context information을 loss할 수 있다. 이를 방지하기 위해서는 learning speed가 다소 느리더라도 image를 다양한 관점에서 여러 번 traverse할 필요가 있다. 본 논문에서 제안하는 MCU-Net model은 context information을 보다 세밀하게 capture하기 위해, Faster-RCNN [15] model의 anchor box처럼, 한 image에 크기가 다른 patch를 여러 번 적용하여 다양한 관점에서 feature를 extract한다.



<그림 1> three patch sizes(3x3, 5x5, 7x7)를 사용한 Multiple Concatenated U-Net(MCU-Net) model

<그림 1>은 3x3, 5x5, 7x7 이렇게 3 개의 patch sizes를 사용하는 MCU-Net model의 전체 구조를 보여 주고 있다. 제안하는 MCU-Net model은 다양한 patch size를 가진 U-Net들을 final layer에서 concatenate한다. MCU-Net model에서 patch size가 (*ri*, *ci*) 인 U-Net *Li*에서 extract한 feature vector를 라 할 때, 서로 다른 크기의 patch를 사용하는 U-Net *L1*, *L2*, ..., *Ln*에서 extract된 features 은 출력 이전 단계에서 <식 1>과 같이 concatenated된 후 active function를 통해 예측 값 *U*로 출력된다.

<식1>

extract한 feature들이 concatenated된 final layer *U*를 통해 탐지한 pancreas region은 활성화 함수인 Softmax를 통해 output image *O*를 출력한다. MCU-Net model은 여러 patch size를 사용해 capture된 다양한 feature에 의한 결과 값을 동시에 고려했기 때문에 단일의 patch size를 사용하는 U-Net model에 비해 더 많은 context information를 얻을 수 있다. MCU-Net model은 extract한 feature 값들을 concatenate한 feature vector *U*를 통해 결과를 예측하기에, 여러 model을 따로 학습한 후 각 model의 결과 예측 값을 통합하는 ensemble method의 한 variation이라고 볼 수 있다. <그림 1>은 patch size가 각각 (3,3), (5,5), (7,7)인 3개의 U-Net을 concatenate한 MCU-Net model 구조를 보여 주고 있다. 그림에서 각 layer 위의 숫자는 channel 수를 나타내며, 왼쪽 아래 숫자는 x-y-size에 대한 feature map의 크기를 나타낸다. 그림에서 gray layer는 feature map을 copy하여 concatenate한 것을 의미한다. MCU-Net 구조는 multiple patches를 사용할 수 있도록 U-Net 구조를 확장한 것이다. 따라서 multiple patches를 사용하는 것 이외에는 U-Net 구조를 대부분 차용한다. MCU-Net 구조에서 아래로 내려가는 contracting path에서는 U-Net 구조와 마찬가지로 2x2 max pooling을 사용하여 한 단계 내려갈 때마다 1/2 down sampling을 수행하고 channel 수를 두 배씩 늘린다. 반대로 위로 올라 가는 expansive path에서는 contracting path에서 줄어든 feature map size를 다시 늘리기 위해 2x2 up-convolution을 수행하고 동시에 channel 수는 반으로 줄인다. input image가 들어오면 MCU-Net model은 first layer에서 256 x 256 size와 하나의 channel을 가진 feature map을 생성한다. second layer에서는 서로 다른 3개의 patch별로 convolution을 수행하여 3개의 multiple feature maps를 생성한다. 다음으로는 각 feature map마다 독립된 U-Net model에 따른 operation을 수행한다. 이렇게 수행된 3개의 U-Net model operation을 합치는 final layer에서는 3개의 U-Net 구조에서 출력된 feature maps을 concatenate하고 이를 하나의 256x256x1 feature map으로 1x1 convolution한다.

MCU-Net model을 training할 때 loss function는 DSC를 사용한다. DSC는 estimated region과 ground truth region이 얼마나 일치하는지 그 비율을 나타내는 값으로, <식 2>와 같다. DSC는 estimated region X와 ground truth region Y가 일치할수록 1에 가까운 값이 출력되고, estimated region이 ground truth region과 다를수록 0에 가까운 값이 출력된다.

< 식 2 >

**4. 구현 및 실험**

**4.1 data 집합**

실험에서는 pancreas segmentation에서 대표적으로 사용되는 pancreas-CT set[8]을 이용하여 model learning 및 evaluation을 진행하였다. pancreas-CT는 NIH(National Institutes of Health Clinical Center)에서 공개하여 제공하고 있으며, 총 82개의 3D 복부 pancreas CT scans image로 이루어져 있다. 피실험자는 53명의 남성과 27명의 여성으로 구성되어 있으며, 연령은 18 – 76세까지 다양하다. image의 해상도 및 사이즈는 512 x 512 크기에 1.5mm – 2.5mm 두께를 바탕으로 하고 있으며 총 10.2GB의 데이터 크기에 19,328개의 복부 CT image를 보유하고 있다. 본 논문에서는 집합을 4개의 부분 집합(fold)으로 나누어 cross validation하였다.

**4.2 실험 환경**

본 논문에서는 U-Net 기반의 model[4]에서 제공된 코드를 바탕으로 제안하는 MCU-Net model의 코드를 구현하였다. 실험에 사용된 Python 버전은 2.7.12이며, model은 Python 기반의 deep learning library인 Tensorflow(ver. 2.1.0), Keras(ver. 2.0.8)을 사용하여 구현했고, RAM 16GB와 그래픽카드 Geforce GTX 2080ti가 설치된 Ubuntu 16.04 LTS에서 실험을 수행하였다. 그 외 부가적인 library로는 numpy(ver. 1.14이상), pandas, matplotlib 등이 있다. 실험에 사용된 hyper paratmers로, Epoch는 10, learning rate은 1e-5, optimization function으로는 Adam을 사용하였다.

**4.3 실험**

본 논문에서는 MCU-Net의 성능을 입증하기 위해 비교 model을 구현하여 정량 실험을 진행했다. 성능 평가를 위한 척도로는 평균 DSC와 표준 편차 값을 이용하였다.

본 실험에서는 앞서 언급한대로 전체 dataset을 4개의 부분 집합으로 나누어 cross validation을 실행하였으며, 각 fold마다 model을 따로 training하고 evaluation하였다. <표 1>은 이 실험에 대한 정량적인 결과를 보여주고 있다. 표에서 제시된 결과 값들은 해당 fold을 테스트 결과로 사용했을 시의 평균 결과를 나타낸다.

<표 1>에서 M1 model은 기존에 설계되어 있던 U-Net 기반의 baseline model[1]을 pancreas 분류에 접합시킨 U-Net model [13]이고 M2 model은 Patch size가 (3,3)으로 같은 2개의 U-Net을 단순히 더하여 Softmax한 Ensemble model이다. M3 model은 patch size가 (3,3)인 3개의 U-Net을 ensemble한 model이다. M4 model은 2개의 U-Net으로 구성된 Multi U-Net model로써 각각의 patch size는 (3,3), (5,5)이다. M5 model의 경우 총 3개의 U-Net으로 구성된 Multi U-Net model로써 각각의 patch size는 (3,3), (5,5), (7,7)이다.

<표 1>의 실험 결과에서 알 수 있듯, fold에 따라서 학습의 accuracy의 차이가 발생할 수 있음을 보인다. ensemble을 적용한 model의 경우 큰 accuracy의 향상은 없었으며, 첫 번째 fold에서는 오히려 accuracy가 낮아졌다. 이에 반해 본 논문에서 제안하는 model(M5)은 모든 Fold에 대해서 accuracy가 높아진 것을 확인할 수 있다.

<표 1> Multiple Concatenated U-Net model 및 U-Net의 성능 실험

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Dataset (fold) | 1 | | 2 | | 3 | | 4 | | Total |
| AVG | DEV | AVG | DEV | AVG | DEV | AVG | DEV | AVG |
| 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 2>는 선행 연구들과 제안하는 model인 MCU-Net의 비교를 위한 성능 평가 실험 결과이다. 선행 연구의 실험 결과는 각 연구 논문과 관련된 코드를 직접 구현하여 결과 값을 출력했다. 모든 실험의 Loss function은 DSC이다. 실험 결과에서 알 수 있듯이 본 논문의 MCU-Net model이 이전의 연구들에 비해 다소 accuracy가 높다.

<표 2> 선행 연구 간의 성능 비교 실험

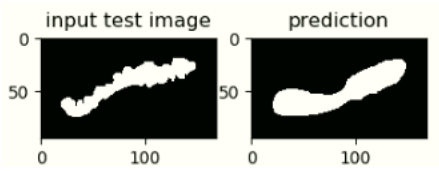
|  |  |
| --- | --- |
| **Model** | **DSC** |
| Recurrent Saliency Transformation Network Model [6][7] | 0.846 |
| U-Net Model [1] | 0.822 |
| Attention U-Net based Model | 0.832 |
| U-Net based Model | 0.898 |
| MCU-Net(Proposed) Model | **0.916** |

MCU-Net model의 정성적 평가는 <그림 2>와 <그림 3>을 비교해서 볼 수 있다. <그림 2>는 원본 model[4]의 정성적 평가 결과이며, <그림 3>은 MCU-Net model의 정성적 평가 결과이다. <그림 2>의 원본 model[4]의 경우 가운데에 해당하는 예측 segmentation image과 오른쪽의 정답 image과의 차이가 큼을 시각적으로 확인할 수 있으며, <그림 3>의 MCU-Net model의 경우 예측한 segmentation image과 정답 segmentation image이 <그림 2>에 비해 더욱 유사함을 확인할 수 있다.

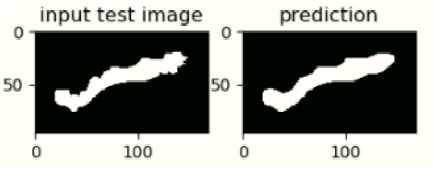
**5. 결론**

본 논문은 patch size가 다른 U-Net들을 concatenate하여 image에서 pancreas을 segmentation하는 MCU-Net deep learning model을 제안하였다. 제안하는 model의 성능을 평가하기 위해 NIH에서 제공하는 약 2만장의 복부 CT 집합을 이용하여 학습한 결과, 제안하는 model의 accuracy는 약 91.6%로 다른 학습 model들보다 높은 성능을 보였다. 이는 성능 비교를 위해 구현한 관련 연구 중 가장 높은 성능을 보였던 U-Net based Model[13]에 비해 1.8% 더 높은 accuracy이다. pancreas은 장기 중에서도 특히 탐지난이도가 높다. 따라서 MCU-Net은 pancreas뿐만 아니라 탐지 난이도가 비교적 낮은 다른 장기의 분류 문제에서도 좋은 결과를 보일 것이라 기대된다. MCU-Net은 기존 U-Net model에 비해 비교적 오랜 training 시간을 필요로 한다. U-Net의 개수가 늘어난 만큼 연산이 이루어지기 때문에 시간도 따라 늘어났다. 그러나 accuracy 면에서 기존의 model보다 나은 결과를 얻을 수 있기에 CAD와 같은 image 진단 system 및 소프트웨어에서 보다 나은 성능을 보일 수 것이다.

향후 연구로는 다른 CT 집합에서도 이번 실험과 같이 유의미한 결과가 나오는 지 확인해 볼 예정이다. 그 외에도 image를 활용한 CNN이 적용될 수 있는 분야에서 이러한 model이 유의미한 결과가 나타나는지 검증할 것이다. 그 후 U-Net이 아닌 다른 model에서도 multiple concatenate를 적용했을 때, 성능 향상이 나타나는지 검증할 예정이다.



<그림 2> U-Net model의 정성적 평가



<그림 3> MCU-Net model의 정성적 평가

**참고문헌(Reference)**

[1]Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015.

[2]Oktay, Ozan, et al. "Attention u-net: Learning where to look for the pancreas." arXiv preprint arXiv:1804.03999 (2018).

[3]Yu, Qihang, et al. "Recurrent saliency transformation network: Incorporating multi-stage visual cues for small organ segmentation." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018.

[4]https://github.com/snapfinger/pancreas-seg

[5]https://paperswithcode.com/sota/pancreas-segmentation-on-tcia-pancreas-ct

[6]YU, Qihang, et al. Recurrent saliency transformation network: Incorporating multi-stage visual cues for small organ segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018. p. 8280-8289.

[7]https://github.com/twni2016/OrganSegRSTN\_PyTorch

[8]Roth, H., Farag, A., Turkbey, E.B., Lu, L., Liu, J., Summers, R.M.: Data from pancreas-CT. The Cancer Imaging Archive(2016), http://doi.org/10.7937/K9/TCIA.2016.tNB1kqBU

[9]Oda, M., Shimizu, N., Roth, H.R., Karasawa, K., Kitasaka, T., Misawa, K., Fujiwara, M., Rueckert, D., Mori, K.: 3D FCN feature driven regression forest-based pancreas localization and segmentation. In: DLMI, pp. 222–230. Springer (2017)

[10]Wolz, R., Chu, C., Misawa, K., Fujiwara, M., Mori, K., Rueckert, D.: Automated abdominal multi-organ segmentation with subject-specific atlas generation. IEEE TMI 32(9) (2013)

[11]Md Zahangir Alom, Mahmudul Hasan, Chris Yakopcic, Tarek M. Taha, Vijayan K. Asari.:Recurrent Residual Convolutional Neural Network based on U-Net (R2U-Net) for Medical Image Segmentation (2018)

[12]Liangzhi Li, Manisha Verma, Yuta Nakashima, Hajime Nagahara, Ryo Kawasaki.:IterNet: Retinal Image Segmentation Utilizing Structural Redundancy in Vessel Networks (2019)

[13]Yijun Liu, Shuang Liu.:U-NETFOR PANCREAS SEGMENTATION IN ABDOMINAL CT SCANS (2018)

[14]Roth, H., Lu, L., Farag, A., Shin, H., Liu, J., Turkbey, E., Summers, R.: DeepOrgan: multi-level deep convolutional networks for automated pancreas segmentation. In: International Conference on Medical Image Computing and Computer- Assisted Intervention (2015)

[15]Shaoqing Ren, Kaiming He, Ross Girshick, Jian Sun. “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks”

1. 경기대학교, 경기도 수원시 영통구 이의동 산 94-6 [↑](#footnote-ref-1)
2. **1\*** 경기대학교, 경기도 수원시 영통구 이의동 산 94-6, 교신저자 [↑](#footnote-ref-2)