

# Fully Automatic Heart Segmentation Model Analysis Using Residual Multi-Dilated Recurrent Convolutional U-Net

Sang Heon Lim<sup>†</sup> · Myung Suk Lee<sup>††</sup>

## ABSTRACT

In this paper, we proposed that a fully automatic multi-class whole heart segmentation algorithm using deep learning. The proposed method is based on U-Net architecture which consist of recurrent convolutional block, residual multi-dilated convolutional block. The evaluation was accomplished by comparing automated analysis results of the test dataset to the manual assessment. We obtained the average DSC of 96.88%, precision of 95.60%, and recall of 97.00% with CT images. We were able to observe and analyze after visualizing segmented images using three-dimensional volume rendering method. Our experiment results show that proposed method effectively performed to segment in various heart structures. We expected that our method can help doctors and radiologist to make image reading and clinical decision.

Keywords : Deep Learning, Artificial Intelligence, Heart Segmentation, Algorithm, ANN, CNN

# Residual Multi-Dilated Recurrent Convolutional U-Net을 이용한 전자동 심장 분할 모델 분석

임 상 현<sup>†</sup> · 이 명 숙<sup>††</sup>

## 요 약

본 논문에서는 딥 러닝 기반의 전-자동 심장 분할 알고리즘을 제안한다. 본 논문에서 제안하는 딥 러닝 모델은 기존 U-Net에 residual recurrent convolutional block과 residual multi-dilated convolutional block을 삽입하여 성능을 개선한 모델이다. 모델의 성능은 테스트 데이터 세트를 전-자동 분할한 결과와 영상의학 전문가의 수동 분할 결과를 비교하여 분석하였다. CT 영상에서 평균 96.88%의 DSC, 95.60%의 precision과 97.00%의 recall 결과를 얻었다. 분할된 영상은 3차원 볼륨 렌더링 기법을 적용하여 시각화한 후 관찰하여 분석할 수 있었다. 실험 결과를 통해 제안된 알고리즘이 다양한 심장 하부 구조를 분할하기에 효과적인 것을 알 수 있었다. 본 논문에서 제안하는 알고리즘이 전문의 또는 방사선사의 임상적 보조역할을 수행할 수 있을 것으로 기대한다.

키워드 : 딥러닝, 인공지능, 심장분할, 알고리즘, 인공신경망, 합성곱신경망

## 1. 서 론

2017년 통계청 자료에 따르면, 순환계통 질환 중 심장 질환으로 인한 사망률은 60.2%로 악성신생물 다음으로 가장 높은 2위를 차지하고 있으며, 매년 증가하고 있는 추세이다 [1]. 심장 질환의 진단은 주로 컴퓨터 단층촬영(Computed Tomography, CT) 또는 자기공명영상(Magnetic Resonance

Imaging, MRI)과 같은 비침습적 투시 검사를 통해 이루어진다. 이와 같은 방법은 침습적 진단방법에 비해 심장 질환을 악화시킬 염려가 적을 뿐만 아니라 심장의 단면을 정밀하게 관찰할 수 있으며, CT 혈관 조영술과 같은 방법으로 심장 혈관을 살펴보는 등 다양한 방면에서 심장 질환을 평가하기 용이하다.

CT는 심장을 고화질의 영상으로 관찰할 수 있으며, 심장의 해부학적인 이상을 검출하기에 적합하다. CT에서 발견할 수 있는 치명적인 질환은 상행대동맥 박리증, 대동맥류, 좌심실류, 좌심방 부속기의 혈전, 그리고 중심부 폐동맥의 혈전 등이 있다. MRI영상에서는 조직 간의 색조 대비가 뚜렷하여 정상 조직과 비정상 조직을 구별하기 용이하며, 심장 마비를

<sup>†</sup> 준 회 원 : 계명대학교 의용공학과 학사과정  
<sup>††</sup> 정 회 원 : 계명대학교 타볼라라사칼리지(과학과기술) 조교수  
Manuscript Received : September 18, 2019  
Accepted : November 9, 2019  
\* Corresponding Author : Myung Suk Lee(mslee@kmu.ac.kr)

겪은 환자의 심장 손상 부위를 관찰할 수 있는 장점이 있다. 또한 MRI는 전반적이거나 국소적인 심장기능의 평가에 사용되거나 심장의 형태진단, 선천성 심장병, 대동맥 질환, 판막성 심질환, 심근증, 심장종양 등을 진단하는데 용이한 검사법이다. 이에 따라 CT와 MRI는 환자의 임상적 진단, 치료, 심질환의 발견에 필수적인 영상 진단 장치이며, CT 및 MRI 영상을 분석하는 것은 심장 질환 진단에 있어서 매우 중요한 부분 중 하나이다.

CT와 MRI가 심질환의 해부학적 이상을 판독하기 위한 영상진단 장치로 꾸준히 이용됨에 따라, 심장의 하부 구조들을 정확히 분할하는 것은 심장의 기능을 정량적으로 평가하고 분석하는 것에 있어서 중요한 과정 중 하나가 되었다. 영상 분석의 첫 번째 단계는 영상을 분할하는 것으로 영상을 유사한 특징 및 속성을 가지는 부분으로 나누는 것이 주된 목적이다. 이 과정을 통하여 의료영상을 진단함에 있어서 의미 있는 부분을 찾아내고 분석에 용이한 영상으로 표현할 수 있다. 의료영상에서의 의미적 분할(semantic segmentation) [2, 3]은 서로 다른 장기 또는 신체부위를 분할하는 용도로 사용한다. 이 때, 픽셀 및 복셀이 포함하고 있는 영상의 강도, 색상, 질감 등과 같은 특징을 이용하여 영상을 분할한다. 기존의 전통적인 영상 분할 방법으로 임계 값 방법, 군집 방법, 경계선 방법 등의 기술이 있으나, 최근에는 인공신경망(Artificial Neural Networks, ANN)을 깊게 쌓아 성능을 높인 심층신경망(Deep Neural Networks, DNN) 기법이 다양한 컴퓨터 과학적으로 접근하는 문제해결에서 높은 성능을 내고 있다[4, 5].

기존의 전통적인 영상처리 방법 및 패턴인식 기반의 기계학습 방법은 특징 추출의 모든 과정에서 연구자가 직접 개입하여 조정해야 하므로, 모든 경우의 영상을 정확하게 분할하는 것에 한계가 있다. 특히, 의료영상 분할 문제의 경우 의료영상 특성상 노이즈가 많으며, 개인에 따라 영상에서 관찰되는 특징의 편차가 크기 때문에 기존의 영상처리 방법으로 분할하기에는 어려움이 따른다. 이에 따라, 많은 임상적 연구는 여전히 심장의 하부 구조들을 분할하는데 있어서 영상의학 전문가에게 의존하고 있다. 그러나 이 과정은 심장의 단면 영상을 하나씩 분석해야 하므로 과도한 노동과 시간이 소요되는 단점이 있다. 또한 심장의 하부구조들은 물리적인 조직으로 구분되어 있지 않기 때문에 정확한 분할을 위해서는 관찰자의 전문적인 해부학적 지식이 필요하며, 전문가에 따라서 분할 결과가 달라지는 경향이 있다[6].

따라서 본 논문에서는 인공신경망을 이용한 딥 러닝을 기반으로 심장의 구조를 전-자동으로 분할하는 인공지능 모델을 제안한다. 제안하는 모델을 이용해 심장의 3차원 볼륨 렌더링 또는 정량적 분석에 이용하는 컴퓨터보조진단(Computer Aided Diagnosis, CAD)시스템을 마련한다면, 객관적이고 일관성 있는 결과를 전문의에게 제공하여 심장질환의 임상적 진단에 보조할 수 있을 것으로 기대한다.

## 2. 배경 연구

### 2.1 의료영상처리

병원에서 다루는 의료영상은 방사선, 핵의학, 자기장, 초음파 4가지 촬영기법으로 분류된다. 이러한 의료영상들은 비침습적 투시 검사로써, 이와 같은 방법은 침습적 진단방법에 비해 질환을 악화시킬 염려가 적을 뿐만 아니라, 병변의 단면을 정밀하게 관찰할 수 있으며, CT 혈관 조영술과 같은 방법으로 혈관을 살펴보는 등 다양한 방면에서 병변을 분석하기 용이하다. 과거 아날로그 의료영상과 다르게 최근 대부분의 의료영상들이 디지털화됨에 따라, 픽셀 또는 복셀로 이루어진 의료영상을 컴퓨터로 처리하는 의료영상처리 기술이 발전되기 시작하였다. 의료영상처리는 노이즈 필터링, 감마 조정, 영상 분할 등과 같은 방법으로 의료영상의 가시성을 개선하여 병변의 판독을 도울 수 있는 장점을 가지고 있다. Fig. 1에서 간 초음파 영상에서 영상처리를 이용한 가시성 개선의 예를 보여주고 있다. 원본 영상보다 더 명확한 영상이 제공되므로 병변을 판독하는데 도움이 될 수 있다.

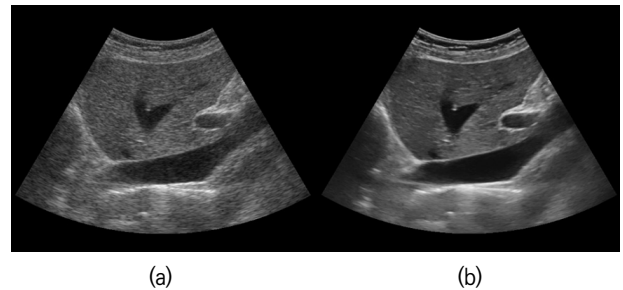


Fig. 1. Example of Improving Visibility using Image Processing in Liver Ultrasound(US) Image, (a) Original Liver US Image, (b) Liver US Image after Image Processing

### 2.2 패턴인식 및 머신 러닝

영상처리를 통해 의료영상의 가시성을 개선함에 그치지 않고, S. Sivakumar 등(2013)[7]은 특성 공학 기술을 기반으로 의료영상에서의 특성 추출을 통해 머신러닝 분류기(classifier)를 학습시키는 연구를 진행 하였다. 의료영상에서의 패턴인식 및 머신러닝 기법은 다양한 의료영상에 가지고 있는 특성을 일반화하여 분석하거나 분류하는 것을 목적으로 많은 연구가 진행되었다. 이 기법은 특성 추출 과정에서 엔지니어가 개입하여 머신 러닝 모델이 학습하기 적합한 특성 선택을 해주어야 한다. 그러나 인체의 특성상 개인에 따라 촬영되는 의료영상의 다양함과 인체의 복잡한 매커니즘 속에서 필요한 병변 및 장기 영역을 일반화하여 분할하거나 분류가 어려워 현재까지 의료영상에서의 머신 러닝 기법은 몇 가지의 병변 및 장기에만 좋은 성능을 내고 있다.

Fig. 2는 의료영상 분류를 위한 기계학습의 흐름도를 나타낸다. 기계학습을 위한 과정을 크게 3가지로 나누었다. 먼저, 영상처리를 통해 영상의 가시성을 개선한 뒤 특징 선택과 특징

추출을 통해 의사결정에 영향을 끼치는 특징을 지정한다. 이후, 분류기의 자동 학습과 수동 검증을 통해 기계학습을 수행한다.

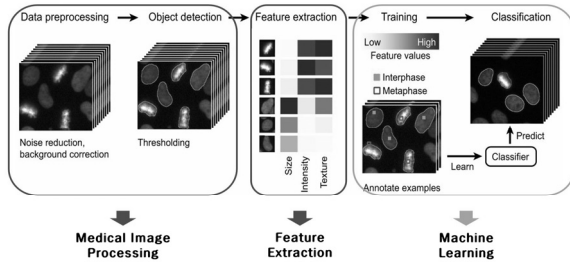


Fig. 2. Flowchart of Machine Learning for Medical Image Classification

### 2.3 심층신경망(DNN)

최근 딥 러닝이라는 기계학습 기법이 핵심 기술로 대두되면서 관련 기술 및 응용 분야에 대한 관심 또한 높아지고 있다. 딥 러닝은 머신 러닝의 여러 범주 중 한 분야로, 생물의 신경계를 모방한 인공신경망을 기반으로 기존의 인공신경망 모델이 얇은 층의 뉴런 모델들의 연결로 구성되어 있다면, 딥 러닝은 뉴런 모델의 층을 깊게 쌓아올림으로써 신경망의 학습 능력을 높이는 모델이다. 상기한 바와 같이 머신 러닝을 이용한 컴퓨터보조진단 시스템에서는 특징 선택 과정이 엔지니어에 의해 수동으로 이루어진다. 반면, 딥 러닝 기법은 전처리 과정, 학습 모델의 일부 초-매개변수를 조정하고 모델의 구조를 수정하는 과정들만 수동으로 이루어진다. 이는 기존의 머신 러닝 기법과 비교하여 모델을 학습하기 위해 필요한 상당한 노력과 시간을 단축할 수 있을 뿐만 아니라, 많은 양의 데이터를 제공해준다면 더욱 좋은 성능을 보여주는 장점을 가진다. 이러한 장점에 따라 딥 러닝은 의료영상을 분석하고 시각화하기에 가장 적합한 기법으로 자리 잡고 있다.

Fig. 3은 딥 러닝 기반의 영상 분류 예시를 보여주고 있다. 크게 특징 추출부분과 분류부분으로 나눌 수 있다. 분류과정은 기존의 기계학습 방법과 같이 분류기의 학습과 검증을 통해 성능을 평가할 수 있다. 그러나 순전파 및 역전파 기법을 통해 특징 추출 과정 또한 전-자동으로 진행된다는 점이 다르다.

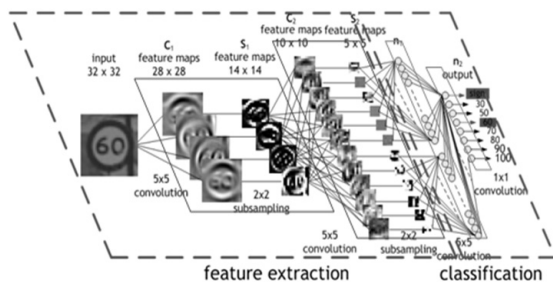


Fig. 3. Deep Learning Based Image Classification Example

### 2.4 심층신경망 기반의 영상 분할 모델

O.Ronneberger 등(2015)[8]이 제안한 U-Net은 'U' 모양의 구조를 가지고 있으며, 의료영상에서의 의미적 분할을 위한 convolution - up convolution 네트워크이다. 이 연구에 따르면, 일반적인 합성곱 신경망의 용도는 하나의 이미지에 대한 출력이 단일 클래스 레이블인 분류 작업이나, 의료 영상 내 병변 영역의 분할은 처리과정에서 지역 정보를 포함해야 하므로 각 픽셀에 대한 class label이 필요하다. U-Net 학습에 사용된 softmax 함수는 아래의 Equation (1)과 같으며, 입력 영상에 대한 정답의 추정 값을 구하는 역할을 한다.

$$p_k(X) = \frac{\exp(a_k(X))}{\sum_{k=1}^K \exp(a_k(X))} \quad (1)$$

또한, 각 정답 픽셀에 대한 교차 엔트로피 손실(cross entropy loss) 함수는 Equation (2)와 같으며 기존의 교차 엔트로피 손실 함수에  $w(X)$ 라는 가중치 값이 추가되어 있다.

$$E = \sum_{X \in ohm} w(X) \log(p_{l(X)}(X)) \quad (2)$$

$w(X)$ 의 값은 영상 내 분할하고자 하는 객체가 여러 개일 경우 객체 사이의 X위치에 있는 픽셀에 가중치를 부여하는 때 개변수이다.  $w(X)$  값을 구하는 방법은 Equation (3)과 같다.

$$w(X) = w_c(X) + w_0 \times \exp\left(-\frac{(d_1(X) + d_2(x))^2}{2\sigma^2}\right) \quad (3)$$

$w_c(X)$  값은 X의 위치에 해당하는 빈도수에 따라 값이 결정된다. 즉, 주변에 존재하는 픽셀들이 같은 클래스 레이블을 가지는 경우 값이 높게 부여된다. exp 식은 X위치에서 가장 가까운 관심영역(Region of Interest, ROI)의 거리를 나타내는 함수  $d_1$ 과 두 번째로 가까운 거리를 나타내는 함수  $d_2$ 를 포함하여 병변 영역에 해당하는 픽셀의 사이가 좁을수록 더 높은 가중치를 부여하는 식이다.

Equation (2)에서  $p_{l(x)}$ 의  $l(x)$ 의 값은 softmax 함수에서 정답의 레이블에 해당하는 k 값을 반환하는 함수이다. 따라서 교차 엔트로피 손실 함수에서는 log함수 안에 k 값을 반환하므로 각 픽셀에 해당하는 정답 확률을 가져와  $w_c(X)$ 의 가중치를 부여받아 Fig. 4의 웨이트 맵을 출력할 수 있었다. Equation (1), (2), (3)에 따라 결정되는 웨이트 맵을 통하여 U-Net이 학습할 때 얼마나 가중치를 부여하여 학습하는지 관찰할 수 있다.

Fig. 4의 U-Net 구조에서 보는 바와 같이, U-Net의 구조는 일반적인 분할 모델인 합성곱 신경망(Fully Convolutional Neural Network, FCNN)[9]과 다르다. U-Net은 convolutional

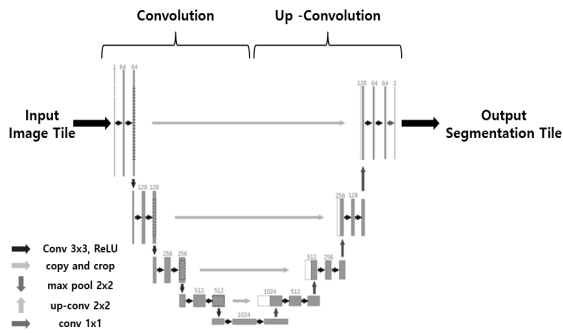


Fig. 4. U-Net Structure

filter를 통과한 특징을 이용하여 각 픽셀에 대한 분류를 한 뒤, up-convolutional filter를 통과하여 컨볼루션 기법을 적용한다. 이 과정은 일반적인 분할모델과 다르지 않으나, skip connection 기법을 도입함으로써 기술적, 구조적으로 개선한 모델이다. skip connection 기법은 분류기를 통과하기 이전의 특징 맵을 분류기를 통과한 특징맵과 병합(concatenate)하여 목표 개체의 위치를 원본 이미지와 정확하게 매칭하기 위한 기법이다. 이를 통해 U-Net은 특히 의료영상에서 좋은 성능을 내는 모델로 알려져 있다.

### 3. RMDRC U-Net(Residual Multi-Dilated Recurrent Convolutional U-Net) 모델

#### 3.1 연구방법

본 연구에서 수행한 모든 과정은 Google Colaboratory [10](이하 Colab)에서 클라우드 컴퓨팅을 이용하여 진행하였다. Colab은 구글에서 제공하는 컴퓨팅 서버로 접속하여 사용하는 것으로, 공동작업 문서관리, 데이터 분석, 파이썬, 딥러닝, GPU 컴퓨팅 등 파이썬 뿐만 아니라 딥러닝에 특화된 작업을 수행할 수 있다. 또한 딥러닝에 필요한 라이브러리, NVIDIA CUDA Toolkit, Cudnn 등이 모두 지원되며, 인터넷에만 연결되어 있으면 별도의 라이브러리 없이 모바일을 포함한 모든 기기에서 접속하여 소스 코드를 수정하거나 실행할 수 있다. 그러나 Colab을 이용한 컴퓨팅은 제한된 메모

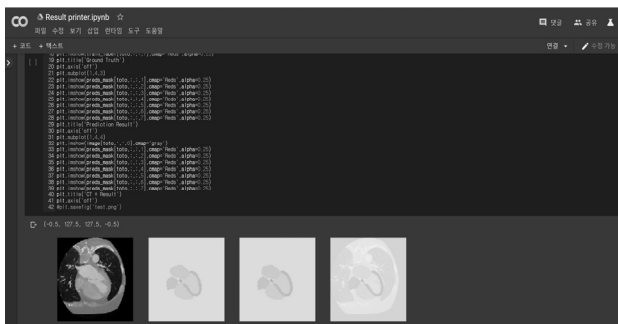


Fig. 5. Cloud Computing Session on Google Colaboratory

리(RAM)내에서 학습해야 하는 한계점이 있다. 학습에 필요한 데이터는 구글 드라이브에 업로드하고 Colab에서 불러와 사용할 수 있으며 Fig. 5와 같다.

#### 3.2 RMDRC U-Net 모델 구조

본 논문에서는 의료영상 분할에 좋은 성능을 내는 것으로 알려진 기존 U-Net을 개선한 모델을 이용하여 Residual Multi-Dilated Recurrent Convolutional U-Net(RMDRC U-Net)의 모델을 제안 하였다. 제안하는 인공지능망 모델의 구조는 다음 Fig. 6과 같다.

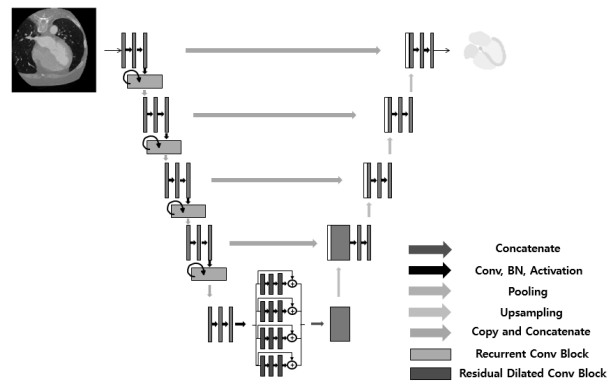


Fig. 6. RMDRC U-Net Model Structure  
Conv, Convolutional; BN, Batch Normalization

Fig. 6에서 보는 바와 같이 RMDRC U-Net은 기존 U-Net 모델에 Recurrent Convolutional block(RC block)[11]과 Residual Multi-Dilated Convolutional block(RMDC block)[12]을 삽입하여 개선하였다. RC block은 Fig. 7에서 나타내며, 각 convolution layer 간의 weight를 공유함으로써 convolution layer가 물체 특징의 통계적 정규성을 포착할 수 있는 능력을 향상시켜준다. 이는 CT 영상진단장비에서 나온 DICOM(Digital Imaging and Communications in Medicine) modality[13]의 특징을 고려한 것이다. CT DICOM은 인체를 연속적인 단층영상으로 표현하는 의료영

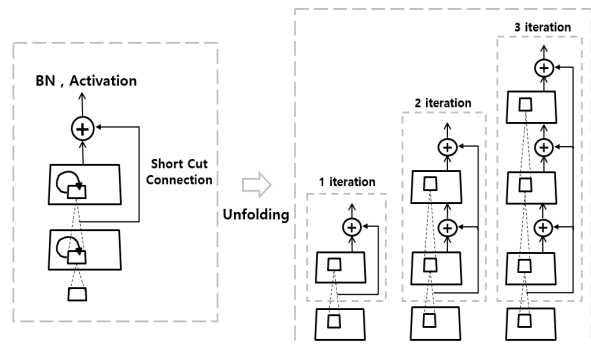


Fig. 7. Recurrent Convolutional Block

상이므로, 연속적인 영상을 다루는 것에 뛰어난 성능을 보이는 recurrent convolution 기법을 적용하였다.

RMDC block에 사용된 dilated convolution 기법은 convolution layer에 dilated rate 파라미터를 추가한 것이다. 이 방법은 풀링을 수행하지 않고도 수용 영역(receptive field)을 늘릴 수 있기 때문에 공간차원 손실이 적으며, dilated rate의 크기에 따라 필터링을 하는 영역이 커지므로 영상에서 특징을 추출하기 용이한 방법이다. 풀링 기법은 공간차원에 대한 다운샘플링을 위해 사용된다. 가장 일반적인 세팅은  $2 \times 2$ 의 receptive field를 가진 max pooling이다.  $2 \times 2$  크기의 max pooling의 경우 가로, 세로에 대해 각각 절반으로 다운샘플링 하므로 입력 중 75%의 activation 값이 손실된다. 또 다른 세팅은  $3 \times 3$  receptive field에 stride를 2로 설정하는 것이다. Max pooling에 3보다 큰 receptive field를 가지는 경우는 너무 많은 정보를 손실하게 되므로 잘 사용되지 않는다. 많은 정보 손실은 모델의 성능 하락을 일으킬 가능성이 있기 때문이다. Fig. 8에서는  $3 \times 3$  dilated convolution을 이용한 다운샘플링 방법(Fig. 8(a))와  $3 \times 3$  풀링 방법(Fig. 8(b))을 비교하였다.

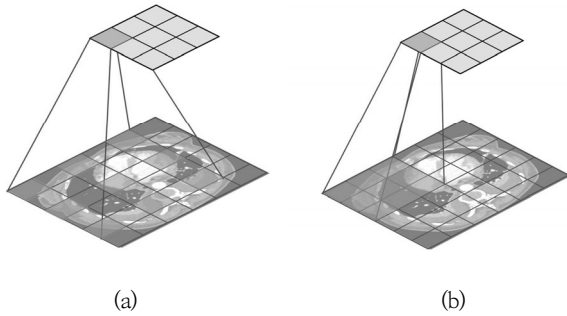


Fig. 8. Comparison of Down Sampling Methods  
(a) Dilated Convolution Methods (b) Pooling Methods

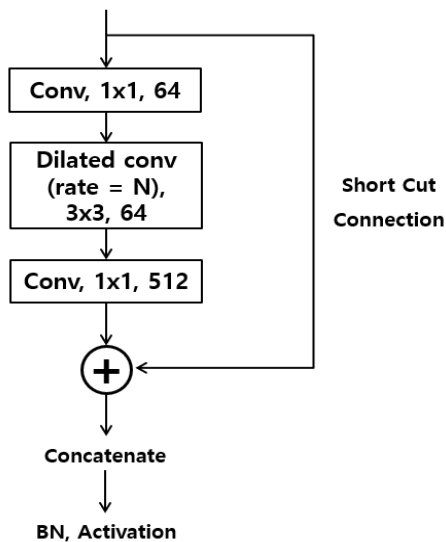


Fig. 9. Residual Multi-dilated Convolutional Block

본 논문에서는 Fig. 9와 같은 dilated block의 dilated rate값 2, 4, 8, 16을 가지는 4개의 convolution을 병합하여 과대적합을 방지하고 다양한 receptive field에서 추출된 특징을 모두 고려하도록 모델을 설계하였다.

본 논문에서 제안하는 모델에 삽입 한 RC block과 RMDC block 모두 short cut connection을 사용하고 각 convolutional block에서 입력과 출력 간의 잔차를 학습함으로써 gradient vanishing 문제를 완화하고 안전성을 강화하였다[14].

#### 4. 실험 결과

본 연구에서는 실험데이터로 Multi-Modality Whole Heart Segmentation Challenge(2017)[15]에서 제공받은 환자 20명의 CT 데이터를 사용하였다. 제공받은 CT데이터는 전처리 과정을 거치지 않은 raw-data를 학습하였으며, 전문가의 수동분할 label은 Fig. 10과 같이 multi-class로 분할하여 학습하였다(one-hot-encoding). label 데이터의 분할은 class에 따라 pixel value가 고정되어 있으므로, 이 값을 기반으로 분할하였다. 모든 환자의 영상 데이터는 각각의 2차원 영상으로 slice하여 training, validation, test set (6:2:2)의 비율로 나누어 실험하였다. 환자 20명의 DICOM 영상을 증식(augmentation)하였다. 데이터 증식 기법을 적용한 학습 데이터의 수는 기존 데이터의 2배이며, Fig. 11과 같이 영상을 shifting, rotating하거나 zoom하여 training set을 증식하였다.

이 방법은 기존 training set으로부터 더 많은 training 데이터를 생성하여 새로운 데이터에 일반화하고 과대적합을 방지할 수 있다. 학습은 batch size를 64로 진행하였으며, Adam optimizer[16]를 사용하였다. Adam의 초기 학습율은 0.001로 시작하여 validation 정확도가 5번 saturation 되면 현재 학습율의 20%만큼 줄여 학습율을 업데이트하였다. 모델이 학습하는 중 15번의 epochs 동안 학습의 진전이 없을 경우 학습을 중단하는 early stopping 기법을 이용하였으며, 실험 시 일반적으로 100~120 epochs 사이에서 학습이 중단되었다. 모델의 손실 함수는 generalized dice loss function[17]을 사용하였다. 인공지능망 모델의 성능을 평가하기 위해 심장구조 분할결과를 수동 분할한 결과와 DSC(Dice Similarity Coefficient), precision, recall을 계

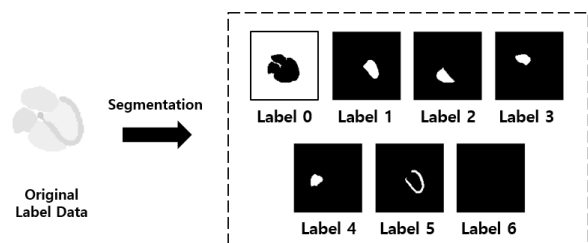


Fig. 10. One-hot-encoding of Label Data for Training Model

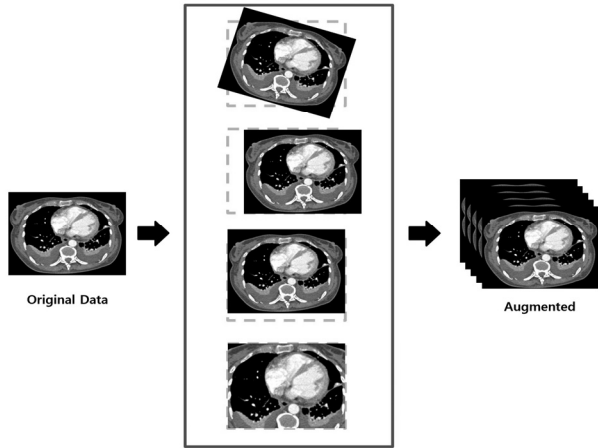


Fig. 11. Data Augmentation using Shifting, Rotating, and Zoom Methods

산하여 비교하였다. Fig. 12과 같이 결과를 3차원 볼륨 재구성을 통해 시각화 하였다.

$$DSC = \frac{2TP}{2TP + FP + FN} \quad (4)$$

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

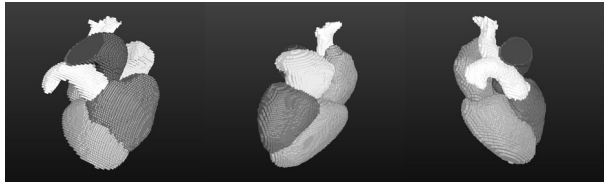


Fig. 12. 3D Volume Rendering of the CT Heart Infrastructure

본 연구에서는 기존의 U-Net 모델, U-Net에 RC block을 삽입한 모델, RMDC block을 삽입한 모델, 두 block을 모두 삽입한 모델의 결과를 비교 분석하였다. DSC는 두 영상이 얼마나 같은 정보를 가지고 있는지 판단할 때 사용되는 척도로, 값이 100%면 두 영상이 같은 영상이라고 판독할 수 있다. precision은 정확도라고 불리기도 하며, 모델이 예측한 정답이 실제 정답인 경우가 어느 정도인지 판단하는 척도이다. recall은 재현율이라고 불리기도하며, 실제 정답 중 모델이 정답이라고 판독한 비율을 나타내는 통계 방법이다.

실험에 의하면 Table 1과 같이 RC block과 RMDC block을 모두 삽입한 모델의 DSC가 96.88%로 가장 좋은 성능을 보였다. 특히 상행 대동맥의 분할 결과가 높은 정확도를 보였으며, 폐동맥의 분할 결과는 낮은 정확도인 93.36%를 Table 2에서 보였다. 폐동맥의 경우, 육안으로 확인하려

라도 CT 영상에서 차지하는 비중이 크지 않기 때문에 다른 해부학적 구조물과 비교해 상대적으로 정확도가 높지 않은 것으로 판단할 수 있다.

Table 1. Performance Comparison Experiment Result

Model	DSC	Precision	Recall
Simple U-Net	95.70%	94.90%	96.53%
U-Net+RC	96.44%	95.20%	96.63%
U-Net+RMDC	96.53%	94.80%	96.83%
U-Net+RC+RMDC	96.88%	95.60%	97.00%

\*RC, recurrent convolutional block;

\*RMDC, residual multi-dilated convolutional block

Table 2. Segmentation Results of the Heart Substructure

Class	DSC	Precision	Recall
Left Ventricle	97.30%	94.20%	97.50%
Right Ventricle	94.90%	95.07%	94.73%
Left Atrium	97.50%	96.34%	97.46%
Right Atrium	95.10%	90.80%	96.90%
Left Myocardium	96.04%	93.85%	96.14%
Ascending Aorta	98.40%	96.73%	98.80%
Pulmonary Artery	93.36%	91.70%	93.36%

본 논문에서 제안한 RMDRC U-Net 모델의 검증 및 시험 과정을 거친 후, 제안한 모델을 이용하여 CT 영상에서의 심장을 3차원으로 가시화하고 각 클래스의 직경을 구하는 프로그램을 제작하였다. 본 프로그램은 로컬 컴퓨터의 CT DICOM 파일을 전-자동으로 분할한 뒤, 3차원 가시화 및 직경을 분석하고 제공하여 짧은 시간 내에 시각적인 정보와 임상적인 정보를 Fig. 13에서 동시에 볼 수 있다.



Fig. 13. 3D Visualization and Analysis Program of CT Cardiac Images Based on Pre-trained RMDRC U-Net Model

## 5. 결 론

본 연구는 딥 러닝을 이용하여 심장 구조를 자동화된 다중 클래스로 분할하고 3차원으로 가시화하여 의사의 진단을 보조하는 인공지능 모델을 제안하였다. 심장 구조를 분할하는 방법으로 RC block과 RMDC block을 삽입하여 기존 U-Net을 개선한 인공지능망 모델을 사용하였다. RC block은 각 convolution layer 간의 weight를 공유함으로써 convolution layer가 물체 특징의 통계적 정규성을 포착할 수 있는 능력을 향상시켜주므로, CT 영상에서 모델의 정확도를 개선하는데 도움이 되었음을 확인하였다. RMDC block은 dilated convolutional filter를 사용하여 더 넓은 receptive field를 고려하였다. dilation rate은 각각 2, 4, 8, 16으로 설계하여 다중 연결하여 과대 적합에 대비하였다. RC block과 RMDC block 모두 short cut connection 기법을 추가하여 gradient vanishing 문제를 완화하고 안전성을 강화하였다. 실험 결과, 평균 DSC는 제안한 모델이 기존 U-Net보다 1.18% 향상되었으며, recurrent convolution block만 삽입한 U-Net보다 0.44% 높았으며, residual multi-dilated block만 삽입한 U-Net보다 0.25% 높았다.

Precision과 recall 값 분석에 따르면, 평균적으로 precision 값이 recall 값보다 낮았으며, 우 심방을 제외한 각 class에 대한 결과 또한 precision 값이 더 낮은 것을 알 수 있었다. 이는 모델이 예측한 결과의 true positive의 비율이 높았으나, false positive의 비율 또한 높다는 것을 의미한다. 예측 결과가 전체적으로 오검출의 비율이 높지 않았으나, 본 연구에서 사용된 20명의 환자 CT DICOM보다 더 많은 케이스와 분량의 데이터가 필요할 것으로 판단된다.

결론적으로, 기존 U-Net에 RC block과 RMDC block을 추가함으로써 CT 영상에서 심장의 하부구조들을 분할하는 성능이 향상된 것을 확인하였다. 또한, 의료영상 분할 및 3차원 가시화 이후 각 클래스에 해당하는 부위의 직경을 계산하여 임상적 분석을 보조할 수 있었다. 본 연구에서 제안한 딥 러닝 모델을 통해 실시간으로 심장의 구조를 자동으로 분할하고 3차원으로 가시화 된 심장모델과 각 클래스의 최대직경을 제공함으로써 의료영상 전문의 또는 방사선사의 영상 판독과 의사결정을 보조할 수 있을 것으로 기대한다. 향후 CT 뿐만 아니라 MRI를 포함하는 다양한 modality의 영상을 3차원으로 가시화 할 수 있는 다중 모델을 연구하고자 한다.

## References

- [1] 2017 Cause of Death Statistics, Seoul: National Statistical Office, pp. 10, 2018.
- [2] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation," *Proceeding of The IEEE Conference on Computer Vision and Pattern Recognition*, pp.580-587, 2014.
- [3] H. W. Noh, S. H. Hong, and B. H. Han, "Learning Deconvolution Network for Segmentation," *Proceeding of The IEEE International Conference on Computer Vision*, pp. 1520-1528, 2015.
- [4] D. Cireşan, U. Meier, and J. Schmidhuber, "Multi-column Deep Neural Networks for Image Classification," *Proceeding of The IEEE Conference on Computer Vision a (32nd Pattern Recognition)*, pp.3642-3649, 2012.
- [5] M. Havaei, A. Davy, D. W. Farley, A. Biard, A. Courville, Y. Bengio, C. Pal, P. M. Jodoin, and H. Larochelle, "Brain Tumor Segmentation with Deep Neural Networks," *Medical Image Analysis*, Vol. 35, pp. 18-31, 2017.
- [6] A. F. Frangi, W. J. Niessen, and M. A. Viergever, "Three-Dimensional Modeling for Functional Analysis of Cardiac Images: A Review," *IEEE Transactions on Medical Imaging*, Vol.20, pp.2-5, 2001.
- [7] S. Sivakumar and C. Chandrasekar, "Lung Nodule Detection using Fuzzy Clustering and Support Vector Machines," *International Journal of Engineering and Technology*, Vol. 5, No.1, pp.179-185, 2013.
- [8] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *Proceeding of International Conference on Medical Image Computer-Assisted Intervention*, Vol.9351, pp.234-241, 2015.
- [9] J. Long, E. Shelhamer, and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," *Proceeding of The IEEE Conference on Computer Vision and Pattern Recognition*, pp.3431-3440, 2015.
- [10] Google Colaboratory, <https://colab.research.google.com>, (Accessed Aug, 6, 2019)
- [11] M. Liang, and X. Hu, "Recurrent Convolutional Neural Network for Object Recognition," *Proceeding of IEEE Conference on Computer Vision and Pattern Recognition*, pp.3367-3375, 2015.
- [12] F. Yu, V. Koltun, and T. Funkhouser, "Dilated Residual Networks," *Proceeding of The IEEE Conference on Computer Vision and Pattern Recognition*, pp. 472-480, 2017.
- [13] P. Mildnerberger, M. Eichelberg, and E. Martin, "Introduction to the DICOM standard," *European Radiology*, Vol.12, pp. 920-927, 2002.
- [14] S. H. Lim, H. S. Choi, H. J. Bae, S. K. Jung, J. K. Jung, and M. S. Lee, "Multi-Class Whole Heart Segmentation using Residual Multi-dilated Convolution U-net," *The KIPS Spring Conference 2019*, Vol.26, No.1. pp.508-510, 2019.

- [15] X. Zhuang and J. Shen, "Multi-scale patch and multi-modality atlases for whole heart segmentation of MRI," *Medical Image Analysis*, Vol.31, pp.77-87, 2016.
- [16] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," *Proceeding of International Conference on Learning Representations*, 2015.
- [17] C. H. Sudre, W. Li, T. Vercauteren, S. Ourselin, and M. J. Cardoso, "Generalized Dice Overlap as a Deep Learning Loss Function for Highly Unbalanced Segmentations," *Proceeding of Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pp.240-248, 2017.



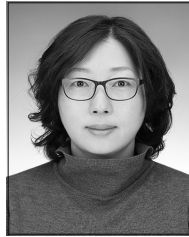
### 임 상 현

<https://orcid.org/0000-0002-1298-5180>

e-mail : smion123@naver.com

2014년 ~ 현 재 계명대학교 의용공학과  
학사과정

관심분야 : 인공지능, 기계학습,  
인공신경망, 의료영상



### 이 명 속

<https://orcid.org/0000-0003-0916-4071>

e-mail : mslee@kmu.ac.kr

2001년 계명대학교 컴퓨터공학부(공학사)

2003년 계명대학교 컴퓨터공학과(공학석사)

2009년 계명대학교 컴퓨터공학과(공학박사)

2013년 ~ 현 재 계명대학교 타블라라사칼리지(과학과기술) 조교수  
관심분야 : 컴퓨터 네트워크, 컴퓨터 교육, 인터넷 윤리, 고등교육,  
OER, 학습 분석, 인공지능