

Teeth Segmentation for Orthodontics based on Deep Learning

치열 교정을 위한 딥러닝 기반 치아 분할

Tae-Hoon Kim · Jong-Jin Park

김태훈* · 박종진†

Abstract

In this paper, we proposed a new UNet model to segment teeth from dental CBCT data for orthodontic treatment. The proposed model uses both the inter-connection and intra-connection proposed by UNet3+ and the nested convolution block proposed by UNet++ in order to utilize the skip connection structure designed to improve performance in the existing UNet series of models. Also, in order to reduce the number of parameters to be learned, the convolution operation is used once or twice in the convolution block. For performance improvement, deep supervision was used for learning on a total of 8 nodes. The proposed layered UNet model shows better segmentation results than the existing UNet3+ and has excellent accuracy even though a small number of image data is used for learning. As a result of the simulation, the proposed layered UNet model using two convolution operations was the best with loss function values of 0.92, 0.85, and 0.79 for training, validation, and test data.

Key Words

Teeth segmentation, Deep learning, UNets, Layered UNet, Orthodontics

1. 서론

최근 인공지능(Artificial Intelligence, AI)은 일상 생활에 영향을 미치는 모든 분야에 적용되어 많은 새로운 제품과 서비스가 개발되고 있다. AI 기술은 정확한 진단과 더 나은 환자 치료의 필요 때문에 의료 분야에 많은 영향을 미치고 있다[1]. 특히 AI를 이용한 의료영상 처리는 의료영상 기술 분야의 발전을 선도할 핵심 기술로 자리 잡고 있다. AI 의료영상 처리는 특징을 추출하여 처리하는 방법으로 최근 딥러닝(deep learning)에 기반을 둔 학습 모델이 주목을 받고 있다. 딥러닝은 학습 과정에서 데이터의 특징을 자체적으로 추출하는데, 영상 기기나 질환 부위에 상관없이 일관된 특징 추출 모델을 가진다. 최근 AI의 의료영상 처리 및 분석 기술은 일반촬영술을 시작으로 전산화단층촬영, 자기공명영상, 초음파촬영, 투시 촬영 등의 다양한 영상 장비에 적용되고 있으며, 다양한 질환을 대상으로 유의미한 판독 결과가 보고되고 있다[2].

치과에서도 AI는 고품질 환자 치료를 제공하도록 임상 의에게 도움이 되고, 예측 가능한 결과를 제공하여 복잡한 프로토콜을 단순화하기 때문에 점점 더 많은 관심 주제가 되고 있다. 임상 치과의 방사선과, 교정과, 치주과, 구강 내과 등에서 AI의 응용 사례 연구가 보고되었다[1]. AI는 치과 방사선과에서 이

미지 해석을 개선하고 충치 감지, 이미지 향상, 치료 계획 제안 등에 사용되고 있다. 교정과에서는 치열 교정 발치를 위한 진단, 치료 결과 분석 등에 사용되고 치주과는 급성과 만성 치주 질환의 구분, 치주 손상 치아의 진단 및 예측 등에 사용된다. 구강내과는 구강 암의 위험성 평가, 치조골의 약물 관련 괴사(MRONJ)의 예측 등에 사용된다.

또한 국외 벤처 회사에서는 2D(Dimension) 영상뿐만 아니라 3D CBCT(Cone-Beam Computed Tomography) 영상을 이용하는 다양한 소프트웨어를 개발하여 치아 분리를 통한 디지털 셋업의 자동화, 교정 치료 및 양악수술 진단 3D 분석, 정교한 악교정 및 임플란트 수술 계획 수립, 교정 치료 및 수술 전후 3D 결과 분석 등이 이루어지고 있다[1]. 또한 AI를 이용하여 해부학적 랜드마크의 주석과 뼈와 치아 구조의 분할(segmentation) 등이 자동적으로 수행되도록 많은 연구가 이루어지고 실용화되고 있다. 국내 기업들도 AI를 기반으로 다양한 CBCT 및 스캔 메쉬(mesh) 영상으로부터 개별 치아를 자동분할하고 치아 이동을 시뮬레이션하는 기술을 개발하여 치아 임플란트 및 보철 치료에서 크라운 배치와 크기 조정의 자동화, 치근과 치관의 인접 간섭과 충돌을 회피하는 AI 알고리즘을 탑재한 소프트웨어를 상용화하고 있다[3].

치과에서 교정은 치과의사의 경험에 많이 의존하는 분야로

† Corresponding Author : Dept. of Computer Engineering, Chungwoon University, Korea.
E-mail:jjpark@chungwoon.ac.kr
<https://orcid.org/0000-0002-8258-4008>

* Dept. of Computer Engineering, Chungwoon University, Korea.
<https://orcid.org/0000-0002-7882-7230>

Received: Jan. 20, 2023 Revised: Feb. 02, 2023 Accepted: Feb. 03, 2023

Copyright © The Korean Institute of Electrical Engineers

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

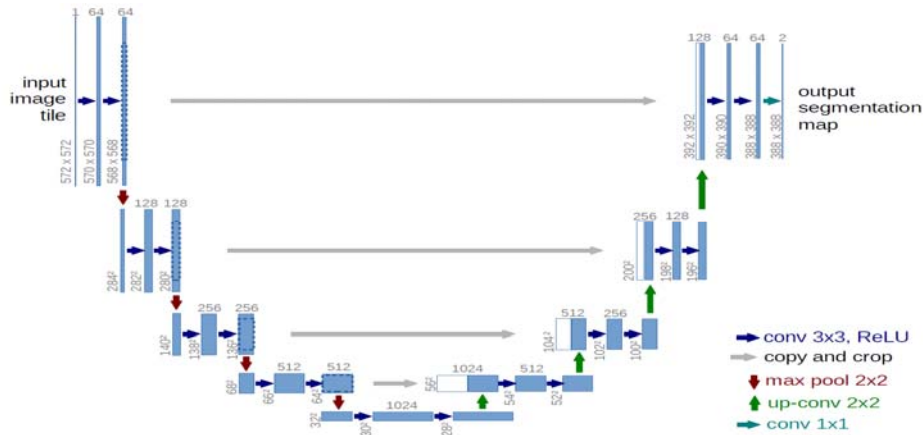


그림 1 UNet의 구조
Fig 1. Structure of the UNet

치료의 질이 의사에 의해 많은 영향을 받게 된다. 기존에는 X-ray 등 2D 진단 도구를 이용하여 각도와 길이 등을 측정하고 교정장치는 평균적인 장치를 사용하여 치료의 일관성이 부족하였다. 근래에는 CBCT 3D 모델에 의한 정밀한 진단과 각 환자에서 치아들이 들어갈 수 있는 골격의 공간 볼륨을 측정한 후 검증된 기준선을 근거하여 각 환자의 주어진 골격 안에 치아를 심미적이고 기능적으로 재배열한다. 이 배열을 근거로 개인에 최적화된 교정장치들을 제작하여 사용함으로써 치료의 질을 높이고 일관성 있는 치료 결과를 얻는 것에 도움을 줄 수 있다. CT 영상은 교정 치료 시 잇몸뼈의 두께와 밀도 확인, 치아 뿌리 길이와 모양 확인, 과잉치와 치아중 확인에 이용된다. 이러한 치료가 이루어지기 위해서는 뿌리를 포함한 치아의 분할이 매우 중요하다. 치아 CT에서 치아 뿌리까지 자동으로 분할하면 교정의 정확성을 높일 수 있다. 현재 상용화된 AI 기반 소프트웨어는 상단 크라운만 움직일 뿐 치근은 어디로 움직이는지 정확하게 알 수 없다.

AI의 한 분야인 딥러닝(Deep learning)은 의료 영상 분할에서 괄목할만한 성능을 내고 있다. 의료 영상처리의 한 분야인 의미론적 분할(Semantic Segmentation)에서 다양한 딥러닝 모델이 고안되었고[4]-[7], Ronneberger 등에 의해 제안된 UNet[5]이 우수한 성능으로 주목을 받고 있다. 초기의 딥러닝 모델인 CNN(Convolution Neural Network)은 기존의 이미지 분류 문제나 물체 검출 문제에서 우수한 성능을 보여주었다. 그러나 영상 분할 문제에서는 한계가 있었는데 이는 기존의 대부분의 영상 분류를 위한 네트워크들은 마지막 층에 완전 연결된(fully connected) 층으로 구성되어서 분류 문제나 물체 검출 문제에는 적합하지만, 픽셀 단위로 조밀한 분류를 해야 하는 문제에는 적합하지 않기 때문이다. 이러한 한계점을 해결하기 위해 FCN(Fully Convolutional Network)는 완전 연결층을 제거하고 모든 층을 합성곱 층(convolution layer)으로만 구성하여, 특징점 추출의 결과가 픽셀 별로 미리 정의된 몇 개의 클래스에 대한 확률 지도(probability map) 형태로 나오게 하였다. UNet은

이러한 FCN에 기반하고 인코더(encoder)와 디코더(decoder) 구조를 가짐으로써 의미론적 분할에서 매우 우수한 결과를 보여 준다[4], [5]. 이후 UNet을 개선한 UNet++, UNet3+ 등 다양한 모델이 제안되어 성능이 더욱 향상되었다[6], [7].

본 논문에서는 치과 교정 치료에 사용되는 3D CBCT 영상에서 UNet 계열의 딥러닝을 사용하여 치아를 분할하고 성능을 향상하기 위해 스킵 연결(skip connection)을 새로 설계한 레이어(Layered) UNet 모델을 제안한다. 제안된 Unet 모델은 Unet의 기본 구조에 기초하여 UNet++와 UNet3+에서 성능을 향상시키도록 설계된 스킵 연결 구조를 결합하였다. 제안된 레이어 UNet에 의해 분할된 치아는 특정 소프트웨어를 이용하여 골격 안에서 기준에 맞추어 재배열하는 셋업(setup) 모델을 통해 치료를 위한 진단 분석 및 치료 계획 설정과 3D 프린터용 stl 파일 추출 등에 사용될 수 있다.

2. 레이어(Layered) UNet

2.1 UNets

FCN은 이미지 분할에서 초기 CNN 모델들이 의미론적 분할에서 좋은 성능을 보여주지 못해서 이러한 한계를 극복하기 위해 제안된 모델이다. 그러나 FCN의 출력 특징맵은 너무 거칠어서(coarse) 원본 이미지 크기에 가까운 조밀한(dense) 특징맵으로 변환해줘야 한다. 이를 위해 업샘플링(Upsampling)을 수행하고 스킵 구조(Skip Architecture)를 도입하여 개선된 분할 결과를 얻었다[4]. 이러한 개선에도 불구하고 FCN은 원본 이미지와 같은 크기로 키우는 과정에서 위치와 경계 같은 정보에 대해서 손실이 발생하게 되고 이에 따라 부정확한 분할 결과를 얻게 된다.

UNet은 FCN의 단점을 보완하여 특정 이미지에 대한 단순한 레이블을 지정하는 것으로 이미지를 인식하지 않고 이미지 그 자체의 특정 영역을 레이블로 구현한다. 즉, 이미지 분할에 특화된 FCN을 활용해 이미지를 상세히 분할하고, 각 픽셀을 레

이블링해 특정 영역을 자동으로 검출한다. 그림 1은 UNet의 구조를 나타낸다. 그림과 같이 U자 형태를 가지므로 UNet이라는 이름이 붙여졌다[5].

UNet의 가운데를 중심으로 왼쪽 부분을 수축 경로(Contracting path) 또는 인코더(Encoder)라고 하고 오른쪽 부분을 팽창 경로(Expanding path) 또는 디코더(Decoder)라고 한다. 인코더는 이미지의 컨텍스트 정보를 얻고 디코더는 정확한 지역화(Localization) 정보를 추출한다. 인코더에서는 입력 이미지의 합성곱 연산이 연속해서 수행된 후 풀링 연산이 수행된다. 디코더는 인코더에서 얻어진 최종 특징맵으로부터 더 높은 해상도의 분할 결과를 얻기 위해 업샘플링(Upsampling)을 수행하며 수축 경로의 같은 수준의 특징맵을 확장 경로의 특징맵과 결합하여 합성곱(Convolution) 연산을 수행한다. 즉, 인코더에서 디코더로 정교한 정보를 넘겨주는 스킵(Skip) 연결을 통해 보다 더 선명한 분할 결과를 얻을 수 있다. 이를 통해 적은 수의 이미지로도 정확한 이미지 분할이 가능하게 되었다.

UNet++는 더 좋은 성능의 의료 이미지 분할을 위해 인코더와 디코더 하위 네트워크(sub-networks)가 일련의 중첩되고 조밀한 스킵 경로를 통해 연결되는 심층 감독된(deeply-supervised) 인코더-디코더 네트워크 구조를 가진다. 재설계된 스킵 경로는 인코더와 디코더 하위 네트워크의 특징맵 사이의 의미론적 간극(gap)을 줄이는 것을 목표로 하고 이를 통해 의료 이미지 분할에서 UNet보다 더 좋은 성능을 보인다.

UNet3+는 의료 영상 분할을 위해 전체 스케일(full-scale)로 연결된 구조이다. 정확한 분할을 위해 중요한 요소 중 하나는 다중 스케일(multi-scale) 특징을 결합하는 것이다. UNet++는 중첩되고 조밀한 스킵 연결을 가지는 아키텍처를 설계하여 변형된 UNet으로 개발되었다. 그러나 전체 스케일에서 충분한 정보를 탐색하지 못하고 여전히 개선의 여지가 있다. UNet3+는 전체 스케일 스킵 연결과 심층 감독(deep supervision)을 활용한다. 전체 스케일 스킵 연결은 다양한 스케일의 특징 맵에서 나온 상위 수준 의미(semantics)와 하위 수준 세부 정보를 통합한

다. 심층 감독은 전체 스케일에서 집계된 특징맵에서 계층적 표현을 학습한다. 그림 2는 Unet, Unet++ 그리고 UNet3+의 구조를 나타낸다[7].

UNet3+의 스킵 연결에서 특징맵 스택, X_{De}^i 는 다음 식(1)과 같이 계산되고 결과는 3x3 크기의 320개의 필터를 가지는 특징맵이다.

$$X_{De}^i = \begin{cases} X_{En}^i, & i = N \\ H\left(\left[\frac{\mathcal{C}\mathcal{D}(X_{En}^k)}{Scales: 1^{\#} \sim i^{\#}}\right]_{k=1}^{i-1}, \mathcal{C}X_{En}^i, \left[\frac{\mathcal{C}\mathcal{U}(X_{De}^k)}{Scales: (i+1)^{\#} \sim N^{\#}}\right]_{k=i+1}^N\right), & i = 1, \dots, N-1 \end{cases} \quad (1)$$

여기서, 함수 $\mathcal{C}(\bullet)$ 는 합성곱 연산을 나타내고, $\mathcal{H}(\bullet)$ 는 배치 정규화 및 ReLU 활성화함수가 이어지는 합성곱으로 특징을 집계하는 메커니즘을 구현한다. $\mathcal{D}(\bullet)$ 와 $\mathcal{U}(\bullet)$ 는 다운 및 업샘플링을 각각 나타내고, $[\bullet]$ 는 붙이기(concatenation)를 의미한다.

2.2 레이어(Layered) UNet

본 논문에서 치아 분할의 성능향상을 위해 제안된 레이어 UNet 모델은 그림 3과 같다. 이는 성능이 우수한 UNet3+를 다양한 깊이(depth)별로 구하여 이전 깊이에서 얻어낸 블록들 중 같은 레벨에 있는 블록들을 결합한(ensemble) 모양과 유사하다. UNet과 UNet++ 모두 전체 스케일에서 충분한 정보를 탐색하지 못하여 영상에서 장기(organ)의 위치와 경계를 명시적으로 학습하지 못한다. 이러한 문제를 해결하기 위해 UNet 3+의 각 디코더 층은 인코더의 작거나 같은 스케일(scale)의 특징 맵과 디코더의 더 큰 스케일의 특징 맵을 모두 통합하여 전체 스케일에서 미세한 세부 정보와 거친 의미 정보를 알아낸다[7]. 그림 2의 (c)에서 보는 것처럼 인코더와 디코더 간의 상호 연결(inter-connection)은 장기의 경계를 강조하는 풍부한 공간 정보를 가져오고, 디코더 간의 내부 연결(intra-connection)은 장기가 있는 위치 정보를 가져온다.

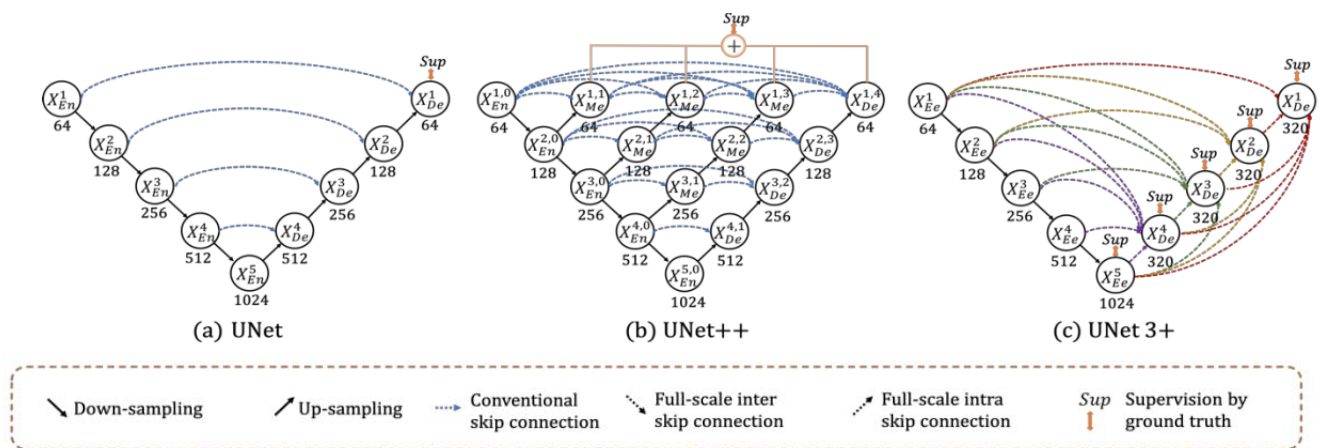


그림 2 UNet, UNet++, UNet3+의 구조
Fig. 2 Structures of UNet, UNet++, UNet3+

본 논문에서 제안된 UNet 모델은 UNet3+에서 제안한 상호 연결 및 내부 연결과 UNet++에서 제안한 중첩된 합성곱(convolution) 블록을 모두 사용한다. 다양한 깊이로 구현 레이어(Layer)별 UNet3+와 UNet++를 결합하여 구현하였으므로 레이어(Layered) UNet으로 명명하였다.

또한 총 8개 노드에 대해 심층 감독을 사용하여 학습하였다. 그림 3에서 회색으로 표시된 블록이 심층 학습에 사용된 블록이다. 학습되는 파라미터의 수를 줄이기 위해 합성곱 블록에서 합성곱이 1 또는 2회 연산된다. 이러한 구조를 통해 인코더와 디코더의 하위 네트워크의 특징맵 사이의 의미론적 간극을 줄이고 다양한 스케일의 특징맵에서 상위 수준의 의미와 하위 수준 세부 정보를 통합함으로써 의료 영상 분할의 성능을 향상할 수 있다. 그림 4는 레이어 UNet를 구축하는 단계를 깊이별로 보여준다.

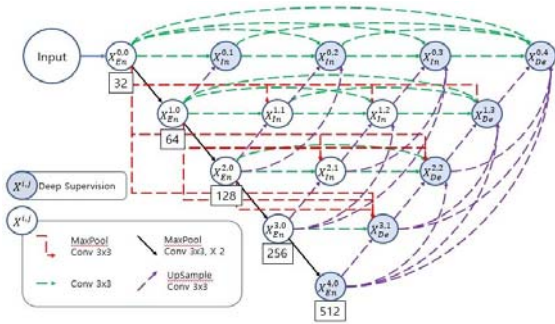
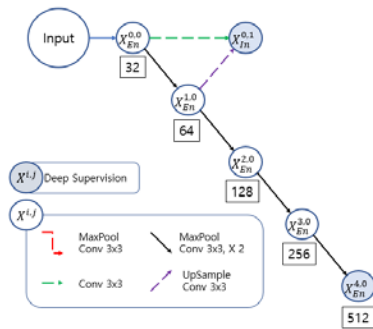
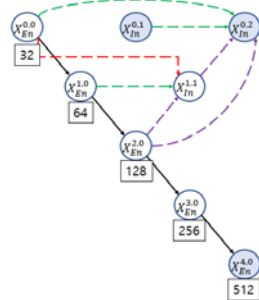


그림 3 제안된 레이어 UNet 모델
Fig. 3 Proposed Layered UNet Model



(a) 깊이 2(depth 2)



(b) 깊이 3(depth 3)

그림 4 깊이 2, 3 레이어 구축 과정
Fig. 4 Building process of layers of depth 2, 3

그림 3의 인코더 및 디코더의 각 블록 $x^{i,j}$ 에서 이루어지는 연산은 아래 식 (2)와 같다.

$$x^{i,j} = \begin{cases} H^2(Input) & i=0, j=0 \\ H^2(m(x_{i-1,j}, 1)) & i>0, j=0 \\ f^n \left(\begin{cases} H \left(\left[H(x_{i,k})_{k=0}^{k < i} \right] \right) \\ H(u(x_{i+k+1,j-k-1}, k+1))_{k=0}^{k < j} \\ H(m(x_{k,0}, i-k))_{k=0}^{k < i} \end{cases} \right) & j>0 \end{cases} \quad (2)$$

여기서, $H^n(x) = ReLU(Conv(x))$ with n times,
 $f^n(x) = ReLU(BatchNorm(Conv(x)))$ with n times,
 $u(x,n) = UpSample(x, 2^n, 'bilinear')$, $m(x,n) = MaxPool(x, 2^n)$,
 $[x_1, x_2, \dots, x_n] = Concatenate(x_1, x_2, \dots, x_n)$.

이때 각 블록의 f 연산 후의 채널수는 아래의 식 (3)과 같다.

$$\text{채널수} = (\text{Concatenate에 사용하는 블록들의 수}) \times (\text{초기 채널 수}) \quad (3)$$

그림 3과 식 (2), (3)에 의해 구현되는 블록의 한 예로서 그림 5는 $X^{1,3}$ 블록을 만드는 자세한 과정을 설명하고 있다. 먼저 $X^{0,0}$ 블록을 커널 크기를 2로 하여 풀링(maxpool) 연산을 진행한 후 합성곱(convolution) 연산을 진행한다(B1). $X^{1,0}$, $X^{1,1}$ 과 $X^{1,2}$ 블록은 각각 합성곱 연산을 한 후 채널수를 32로 만들어 주고, 세 개의 블록을 붙이기(concatenate) 연산을 한 후에 다시 합성곱 연산을 진행하여 32 채널로 만들어준다. 이때 이 결과를 하나의 블록으로 생각한다(B2). $X^{2,2}$, $X^{3,1}$, $X^{4,0}$ 블록은 각각 스케일(scale)을 2, 4, 8로 하여 업샘플링(upsampling) 연산을 한 후 합성곱 연산을 한다(B3, B4, B5). 이렇게 생성된 5개의 블록(B1, B2, B3, B4, B5)으로 합성곱 연산을 하여 32×5 , 즉 160개 채널의 $X^{1,3}$ 블록을 만든다.

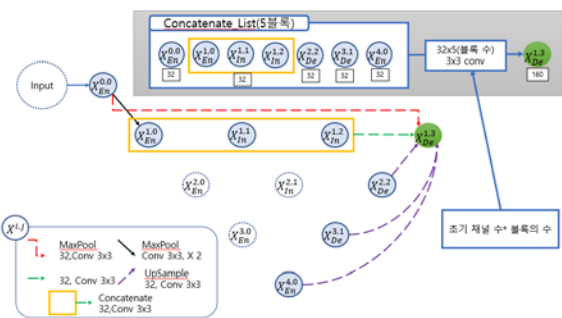


그림 5 $X^{1,3}$ 블록 구축 예
Fig. 5 An example of building $X^{1,3}$ block

3. 학습용 치아 이미지 데이터셋과 손실함수

본 논문에서는 환자 3명의 치아 CBCT 데이터를 사용하여

제안된 모델을 학습하고 평가하였다. 사용된 데이터는 모 치과 병원에서 제공되는 환자 데이터를 비식별화하여 사용하였다. 학습에 사용되는 치아 데이터셋은 환자 2명(환자 1, 2)의 CBCT 데이터이고 하나의 CBCT는 280개의 슬라이스 이미지로 이루어져 있다. 나머지 환자 1명(환자 3)의 CT 데이터는 학습된 모델의 평가를 위한 시험용(test)으로 사용되었다. 학습용 데이터 총 560개의 슬라이스 이미지에서 무작위로 448장을 학습용(train)으로 나머지 112장 검증용(validation)으로 사용하였다. 학습에 사용하지 않은 CBCT 데이터에 대한 모델의 성능 시험용으로 환자 3의 슬라이스 이미지 150장을 사용하였다.

본 논문에서 제안된 UNet을 이용하여 치아 분할을 위한 모델을 학습하기 위해 사용한 파라미터는 다음 표 1과 같다. UNet의 깊이는 그림 3에 보이는 것과 같이 5로 하였고 초기 특징맵의 수는 32로 다음 수준의 깊이로 가면서 2배씩 특징맵의 수가 늘어난다. 입력 이미지의 크기는 효율적으로 학습하기 위해 400×400 크기로 사용하였다. UNet 모델을 학습하기 위해 사용하는 평가지표인 손실함수(Loss function)는 Focal loss와 물체 검출 분야에서 평가지표로 많이 사용되는 IoU(Intersection over Union) Loss를 혼합하여 사용하였다. Focal loss는 부류 불균형이 심한 경우에 주로 사용한다. 영역 분할은 물체가 점유

한 영역보다 배경 영역이 훨씬 넓기 때문에 부류 불균형이 심한 문제이다. 물체 분할은 물체와 배경을 구분하는 작업이므로 부류가 2개인 경우에 해당한다. Focal loss의 계산식은 식 (4)와 같다[8].

$$FL = -(1 - p_t)^{\gamma} \log(p_t) \tag{4}$$

IoU란 두 영역의 교차영역의 넓이를 합영역의 값으로 나눈 값을 의미한다. 객체 검출에서 예측된 경계 상자의 정확도를 평가하는 지표 중 하나로 사용되며, 예측된 경계 상자와 실제 참값(ground truth) 경계 상자의 IoU를 해당 경계 상자의 정확

표 1 학습 파라미터

Table 1 Parameters for learning

Parameter	value
BATCH_SIZE	4
NUM_OF_EPOCHS	50
UNet_depth(UNet_level)	5
Initial_features	32
Image_height	400
Image_width	400

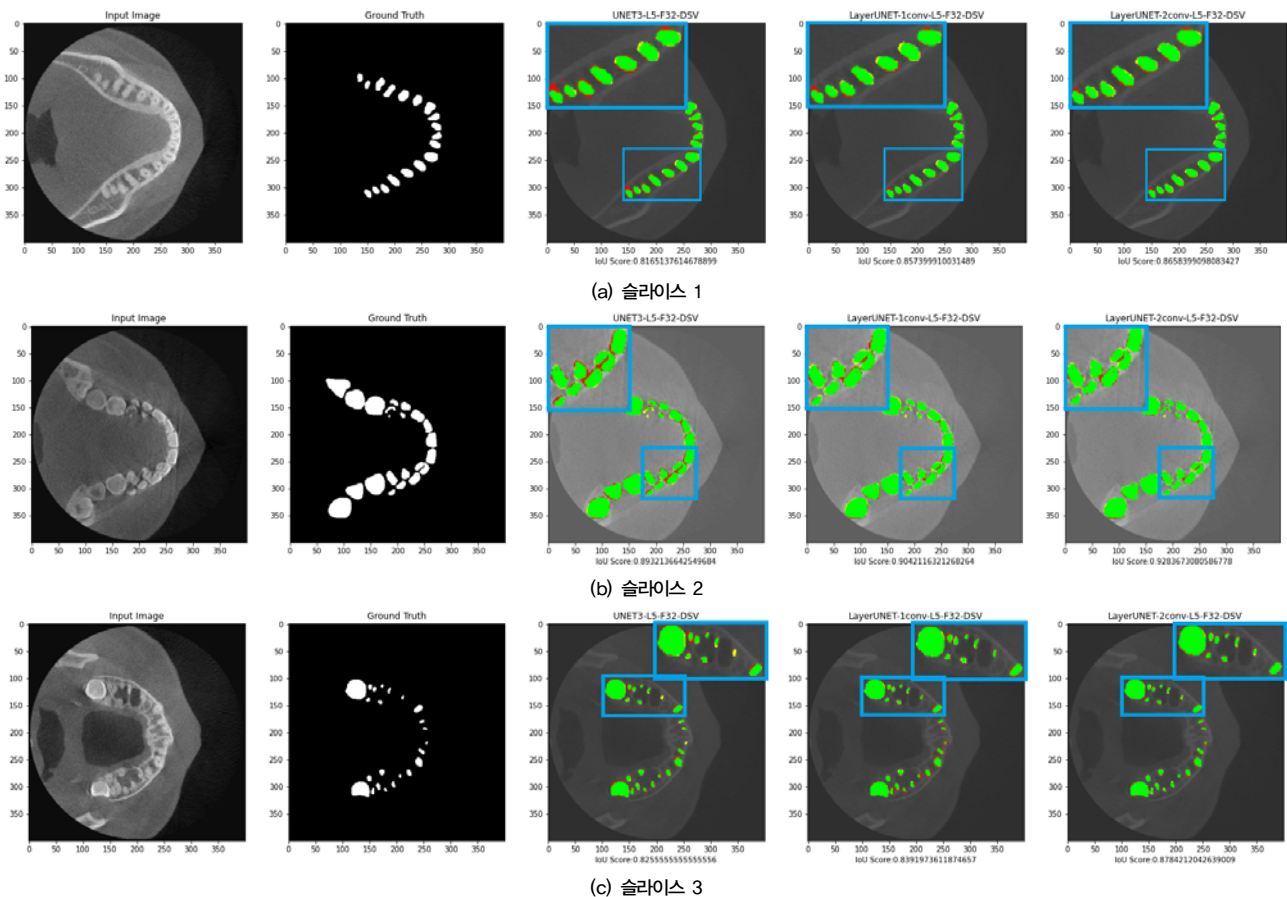


그림 6 각 UNet 모델의 환자 1과 2의 치아 분할 결과(학습용 데이터)
Fig. 6 Results of teeth segmentation of patient 1 & 2 by UNet Models(Training data)

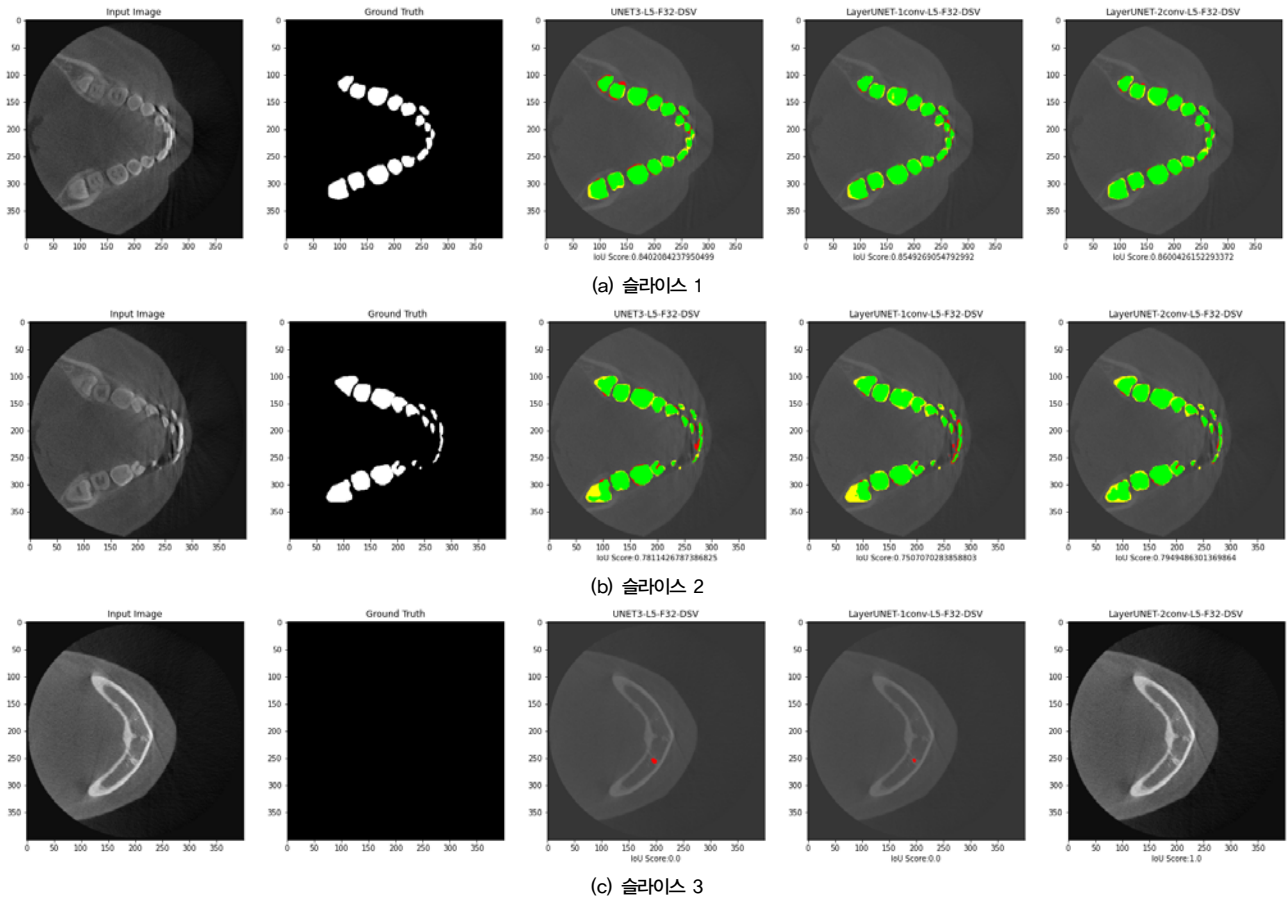


그림 7 각 UNet 모델의 환자 3의 치아 분할 결과(시험용 데이터)
 Fig. 7 Results of teeth segmentation of patient 3 by UNet Models(Test data)

도로 간주한다. IoU의 값이 클수록 물체가 잘 검출되고 있다고 할 수 있다.

4. 시뮬레이션 결과

제안된 UNet 모델을 이용하여 교정 치료를 받는 환자의 치아 CBCT 데이터셋에 대해 잇몸뼈를 구분하여 각 치아를 자동적으로 분할하는 시뮬레이션을 수행했다. 추출된 치아는 치아의 뿌리까지 포함하여 온전한 치아 모양을 나타낸다. 제안된 UNet 모델과 학습, 검증 및 시험 데이터셋을 이용한 시뮬레이션 결과는 표 2와 같다. 제안된 모델의 성능의 우수함을 보이기 위해 UNet3+의 결과와 비교하였다. 시뮬레이션 결과 제안된 레이어 UNet 모델의 합성곱 블록에서 2개의 컨벌루션 연산을 사용한 모델의 성능이 학습용, 검증용, 시험용 데이터셋 모두에 대해 가장 우수했다. UNet3+와 레이어 UNet 모두 심층 감독(DSV)을 사용한 모델이 심층 감독을 하지 않은 경우보다 더 우수했다. 학습하는 파라미터의 수는 레이어 UNet(2conv)이 가장 많다.

그림 6은 각 UNet 모델에 의한 환자 1과 2(학습용)의 치아 이미지 분할 결과의 질적 비교를 보여준다. 성능 비교를 위해

그림의 일부를 확대하였다. 그림 7은 각 UNet 모델에 의한 환자 3(시험용)의 치아 이미지 분할 결과의 질적 비교를 보여준다. 각 슬라이스별 그림에서 첫 번째 열의 그림은 원본 영상이고 두 번째 열의 그림은 참값(Ground Truth)이다. 세 번째 열은 UNet3+ DSV에 의한 결과이고 네 번째와 다섯 번째 열은 각각 제안된 레이어 UNet(1conv, 2conv)에 의한 결과이다. 제안된

표 2 시뮬레이션 결과
 Table 2 Results of Simulation

Result Model	Train	Validation	Test	Parameter (mil.)
UNET3+	0.8399	0.7973	0.6877	7.6
UNET3+ DSV	0.8848	0.8415	0.7406	7.6
Layer UNet	0.8765	0.8240	0.7222	8.4
Layer UNet DSV (1conv)	0.9091	0.8479	0.7744	8.4
Layer UNet DSV (2conv)	0.9177	0.8500	0.7888	10.0

UNet 모델은 여러 개의 치아로 이루어진 CBCT 영상에서 각 치아의 위치와 경계를 잘 추출하여 우수한 결과를 보여준다. 여기서 녹색으로 표시된 영역은 TP(True Positive), 빨강은 FP(False Positive), 그리고 노랑은 FN(False Negative)를 나타낸다. 그림에서 보이는 것과 같이 분할 결과 이미지의 질적 비교에서도 제안된 UNet 모델에 의한 결과가 더 우수한 것을 알 수 있다.

5. 결론

본 논문에서는 치과 교정 치료를 위해 치아 CBCT 데이터에서 치아를 분할하는 새로운 UNet 모델을 제안하였다. 제안된 모델은 기존의 UNet 계열의 모델에서 성능을 향상하기 위해 고안한 스킵 연결 구조를 활용하기 위해 UNet3+에서 제안한 상호 연결 및 내부 연결과 UNet++에서 제안한 중첩된 합성곱 블록을 모두 사용한다. 또한 학습되는 파라미터의 수를 줄이기 위해 합성곱 블록에서 합성곱연산을 1회 또는 2회 사용한다. 성능향상을 위해 총 8개 노드에 대해 심층 감독을 사용하여 학습하였다.

제안된 레이어 UNet 모델은 기존의 UNet3+ 보다 우수한 분할 결과를 보여주고 학습에 적은 수의 이미지 데이터를 사용하였음에도 우수한 정확도를 가진다. 시뮬레이션 결과 합성곱 2회 연산을 사용한 레이어 UNet 모델이 학습용, 검증용 그리고 시험용 데이터에 대해 손실함수 값, 0.92, 0.85 그리고 0.79로 가장 우수하였다. 합성곱 1회 연산을 사용한 레이어 모델은 학습용, 검증용 그리고 시험용 데이터에 대해 손실함수 값, 0.91, 0.85 그리고 0.77, 그리고 심층 감독을 사용한 UNet3+ 모델은 각 데이터에 대해 0.88, 0.84 그리고 0.74의 값을 가진다.

치과 교정 치료에 치아 분할 결과를 이용하면 기존의 방법보다 더 효율적인 치료를 할 수 있을 것으로 예상된다. 이를 위해 향후 다양한 교정 치료 환자 사례를 대상으로 학습하여 성능을 향상하고 개별 치아별로 분할하는 연구가 필요하다.

References

- [1] Yo-wei Chen, et al., "Artificial intelligence in dentistry: current applications and future perspectives," *QUINTESSENCE INTERNATIONAL*, vol. 51, no. 3, March 2020.
- [2] Hong Jun-Yong, Park Sang Hyun, Jung Young-Jin, "Artificial Intelligence Based Medical Imaging: An Overview," *Journal of Radiological Science and Technology*, vol. 43, no. 3, pp 195-208, 2020.
- [3] <https://www.denfoline.co.kr/news/articleView.html?idxno=21716>
- [4] J. Long, E. Shelhamer and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," *The IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3431-

3440, 2015.

- [5] 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*, pp. 234-241. Springer, 2015.
- [6] Z.W. Zhou, M.M.R. Siddiquee, N. Tajbakhsh and J.M. Liang, "UNet++: A Nested U-Net Architecture for Medical Image Segmentation," *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pp: 3-11, 2018.
- [7] Huimin Huang, et al., "Unet 3+: A full-scale connected unet for medical image segmentation," *2020 IEEE International Conference on Acoustics, Speech and Signal Processing(ICASSP)*, pp. 1055-1059. IEEE, 2020.
- [8] Tsung-Yi Lin, et al., "Focal loss for dense object detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 2, pp. 318-327, Feb. 2020.
- [9] Pesapane F, Codari M, Sardanelli F., "Artificial intelligence in medical imaging: Threat or opportunity? Radiologists again at the forefront of innovation in medicine," *Eur Radiol Exp*, Oct; 2(1):35, 2018.
- [10] Trinder JC, Wang Y, Sowmya A, Palhang M., "Artificial intelligence in 3-D feature extraction," *Automatic Extraction of Man-Made Objects from Aerial and Space Images*. 2nd ed., Basel: Birkhäuser; 1997.
- [11] Shen D, Wu G, Suk HL., "Deep learning in medical image analysis," *Annu Rev Biomed Eng*. 2017 Jun; 19:221-48.
- [12] RavR D, Wong C, Deligianni F, Berthelot M, Andreu-Perez J, Lo B, et al., "Deep learning for health informatics," *IEEE J Biomed Health Inform*, Jan;21(1):4-21, 2017.

저자소개

김태훈(Taehoon Kim)



He is currently a student of Dept. of Computer Engineering, Chungwoon University, Incheon, Korea. His research interests include AI, computer vision, medical AI.

박종진(Jongjin Park)



He received BS, MS and Ph.D degrees in Electrical Engineering from Yonsei University, Seoul, Korea, in 1989, 1991 and 1997 respectively. He was a Visiting Scholar at College of Engineering, University of Houston from 2010 to 2011. He is currently Professor of Computer Engineering, Chungwoon University, Incheon, Korea. His research interests include deep learning, computer vision, IOT and medical AI.