

딥러닝 기반의 대퇴골 영역 분할을 위한 훈련 데이터 증강 연구

최규진^O 신주연 경주현 경민호 이윤진*

아주대학교 미디어학과

{paganinist, wndus6165, kkjhh0207, kyung, yunjin}@ajou.ac.kr

Data Augmentation Method for Deep Learning based Medical Image Segmentation Model

Gyujin Choi^O Jooyeon Shin Kyung Joohyun Minho Kyung Yunjin Lee*

Department of Digital Media, Ajou University

요 약

본 연구에서는 CT 영상의 대퇴골 부위를 해부학적으로 의미 있게 변형하여 CT 영상의 대퇴골 영역을 분할하기 위한 컨벌루션 신경망(CNN)의 훈련 데이터를 증강하는 방법을 제안한다. 먼저 CT 영상으로부터 삼차원 삼각형 대퇴골 메쉬를 얻는다. 그 후 메쉬의 국소부위에 대한 기하학적 특성을 계산하고, 군집화하여 메쉬를 의미 있는 부분들로 분할한다. 마지막으로, 분할한 부분들을 적절한 알고리즘으로 변형한 뒤, 이를 바탕으로 CT 영상을 왜곡하여 새로운 CT영상을 생성하였다. 본 연구의 데이터 증강 방법을 이용하여 학습시킨 딥러닝 모델은 기하학적 변환이나 색상 변환 같이 일반적으로 사용되는 데이터 증강법과 비교하여 더 나은 영상분할 성능을 보인다.

Abstract

In this study, we modified CT images of femoral head in consideration of anatomically meaningful structure, proposing the method to augment the training data of convolution Neural network for segmentation of femur mesh model. First, the femur mesh model is obtained from the CT image. Then divide the mesh model into meaningful parts by using cluster analysis on geometric characteristic of mesh surface. Finally, transform the segments by using an appropriate mesh deformation algorithm, then create new CT images by warping CT images accordingly. Deep learning models using the data enhancement methods of this study show better image division performance compared to data augmentation methods which have been commonly used, such as geometric conversion or color conversion.

키워드: 데이터 증강, 딥러닝, 의료영상, 영역 분할

Keywords: Data augmentation, Deep learning, Medical image, Image segmentation

1. 서론

현재 다양한 분야에서 활용되고 있는 딥러닝 시스템을 학습시키기 위해서는 원본과 정답 영상의 쌍으로 이루어진 훈련 데이터가 필요하다. 특히, 의료 영상의 경우, 일반적인 영상에 비해 정답 영상이 포함된 훈련 데이터를 만들거나 구하는 것이 어렵다. 정답 영상을 만들기 위한 해부학적 전공지식이 수반되지 않으면 정확한 정답영상을 만들 수

없으며 의료 영상에 영상 편집 도구를 이용하여 장기의 종류를 표시하는 라벨링(labeling) 과정이 의료인들에게 노동집약적이므로 대량의 훈련 데이터를 쉽게 생성하기 힘들기 때문이다.

너무 적은 양의 훈련 데이터로 딥러닝 모델을 학습시키게 되면 과적합(overfitting) 문제로 인해 학습 후 테스트 데이터를 제대로 판별하지 못하게 될 수 있다 [1]. 소량의 데이터로 딥러닝 모델을 학습시키기 위해서 다양한 방법이 연구되고 있지만, 가장 널리 사용되는 방법 중의 하나가 바로 데이터 증강법(data

*corresponding author: Yunjin Lee/Ajou University(yunjin@ajou.ac.kr)

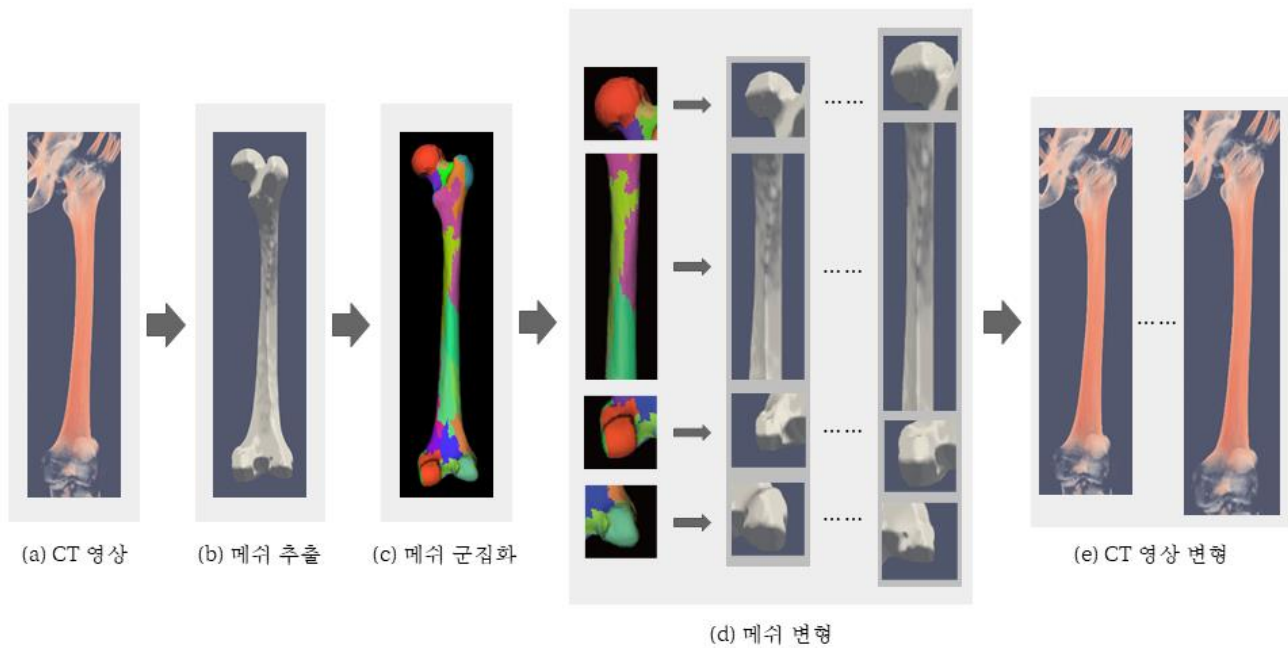


Figure 1: Overall process

augmentation)이다 [2,3,4]. 데이터 증강법이란, 원본 훈련 데이터를 기반으로 새로운 데이터를 만드는 과정을 말한다.

기존 연구에서는 영상 분할 모델의 데이터 증강을 위해서 아핀 변환(평행이동, 회전, 크기변환), 일정 영역 자르기(crop), 반전, 노이즈 추가 [2, 3] 혹은 영상 와핑(image warping) [5], GAN(Generative Adversarial Network) [4] 같은 방법들이 사용되어 왔다. 그러나 이와 같은 영상의 내용을 고려하지 않은 ‘임의적인 변환’ 등에 의한 데이터 증강법은 자칫 실제 의료 데이터에 내포된 해부학적, 생물학적 구조를 해칠 수 있는 위험이 있다. 그리고 이렇게 만들어진 데이터로 학습시킨 영상 분할 모델은 현실의 데이터에서 높은 성능을 보이지 못할 가능성이 있다. 따라서 본 연구에서는 해부학적으로 유의미한 CT 영상 변형을 통해 딥러닝 모델의 성능을 향상시키는 데이터 증강법을 제안하고, 기존 연구의 임의적인 변환들로 딥러닝 영상분할 모델을 학습시켰을 때의 성능을 각각 측정하였다. 그 결과, 본 연구의 데이터 증강 방법이 노이즈가 심한 의료 영상에서 영상분할 성능 향상을 보이는 것을 실험 데이터에 대해 확인하였다.

본 연구에서는 대퇴골을 이용하여 영상 분할 성능을 측정하려고 한다. CT 영상에서 대부분의 대퇴골 영역은 다른 신체 부위와 뚜렷한 경계를 이루고 있어서 기본적인 영상 분할 알고리즘으로도 높은 성능을 쉽게 얻을 수 있다. 그러나 연골로 덮인 부분 주위에서는 경계가 흐릿하여 기존의 방법으로는 매끄러운 영상분할 결과를 얻기 힘들기 때문이다.

본 논문의 구성은 다음과 같다. 2장에서 전체 과정에 대해서 살펴본 후, 3-6 장에서는 CT 영상으로부터 삼차원 삼각형 메쉬

추출, 계층적 면 군집화(hierarchical face clustering)를 이용한 메쉬 분할, 메쉬 변형, 그리고 변형된 메쉬를 이용한 CT 영상 와핑을 차례로 설명한다. 7 장에서는 실험 결과에 대해서 살펴보고 마지막으로 8 장에서 결론을 짓고 향후 연구 방향에 대해서 논의한다.

2. 전체 과정

데이터 증강 과정은 다음과 같이 크게 메쉬 추출, 계층적 면 군집 분석을 이용한 메쉬 분할, 메쉬 변형, CT 영상 변형, 그리고 마지막으로 딥러닝 기반의 영상분할 모델 학습으로 진행된다 (Figure 1 참고).

- 메쉬 추출: CT 영상으로부터 데이터 증강의 토대가 되는 삼차원 삼각형 메쉬 모델을 얻는다. 본 연구에서는 CT 영상에서 대퇴골 영역을 분할하는 것이 목적이므로, 대퇴골 부위의 메쉬를 추출하였다.
- 계층적 면 군집화를 이용한 메쉬 분할: 모델의 해부학적 특성을 반영하여 메쉬를 변형하기 위해 곡률 기반의 형태 분석과 면 군집화(face clustering)를 통해 메쉬를 의미 있는 부분들로 나눈다.
- 메쉬 변형: 대퇴골 각 부위별 크기의 통계 수치를 이용하여 각 부위에 적용할 변형 파라미터를 자동으로 생성하고, 이를 바탕으로 단일 메쉬로부터 다수의 변형된 메쉬를 자동으로 생성한다.
- CT 영상 와핑: 변형 전후 메쉬의 정점들 간의 위치 이동 벡터(displacement vector)를 CT 영상의 모든 영역에 대해

보간하고, 보간된 위치 이동 벡터를 이용하여 원본 CT 영상의 각 복셀(voxel)을 이동한다. 이 과정을 통해 다수의 변형된 메쉬들부터 다수의 와핑된 CT 영상을 얻을 수 있다.

딥러닝 기반의 영상분할 모델 학습 과정에서는 본 연구의 데이터 증강방법과 ‘임의적인 변환’ 들을 이용한 데이터 증강법으로 각각 훈련 데이터를 증강한 후, 각각의 훈련 데이터를 이용하여 영상분할 모델을 학습시키고, 성능을 비교하였다.

3. 메쉬 추출

CT 영상에 대퇴골 부위가 라벨링되어 있는 경우, 라벨링된 복셀 집합으로부터 마칭 큐브(marching cubes) 알고리즘 [6]을 적용하여 삼차원 삼각형 메쉬를 추출할 수 있다. 그러나 데이터 수집 과정에서 얻은 대부분의 CT 영상에는 대퇴골이 표시되어 있는 정답 영상이 포함되어 있지 않다. 원래대로라면 전문적인 해부학적 지식을 가진 의료인이 만든 정답 영상을 써야 하지만, 본 연구는 학습 성능의 향상을 실험적으로 증명하는 것이 목적이므로 연구자 본인이 직접 정답 영상을 라벨링하였다. 훈련 데이터의 모든 픽셀에 일일이 수작업으로 라벨링을 하는 것은 비효율적이므로 반자동적 알고리즘을 사용하였다. 먼저, MITK-GEM [7]을 사용하여 일부의 대퇴골 픽셀들을 라벨링한 후, 해당 프로그램의 삼차원 그래프 컷(3D graph cut) 알고리즘을 이용해서 대퇴골과 배경에 대한 정답 영상을 얻어 삼차원 삼각형 메쉬를 추출하였다.

4. 계층적 면 군집화를 통한 메쉬 분할

대퇴골 각 부위의 크기에 대한 통계 [8]를 보면 사람마다 대퇴골의 형태나 각 부위 간의 비율에 차이가 있다는 사실을 알 수 있다. 대퇴골 모델 전체에 대해서 동일한 변형을 적용하게 되면 이와 같은 해부학적 특성을 반영할 수가 없다. 대퇴골 각 부위별로 다르게 변형하기 위해서는 먼저 메쉬를 각 부위별로 나누어야 한다. 각 대퇴골의 각 부위를 살펴보면 골두는 구와, 그리고 대퇴골체는 원기둥과 같이 기본 도형과 유사한 형태를 띄고 있다는 것을 직관적으로 알 수 있다. 따라서 본 연구에서는 표면의 국소적 경향이 유사한 부위끼리 군집화하여 대퇴골의 각 부위를 나누었다. 이를 위해 본 논문에서는 Lee 가 제시한 방법 [9]을 사용하였다.

먼저, 메쉬 국소 표면의 기하학적 특성을 알아내기 위해서 메쉬 표면의 주곡률로부터 형태 인덱스(shape index)와 굴곡도(curvedness) [10]를 계산하였다. 주곡률로도 군집분석을 할 수는 있지만 주곡률은 값의 범위가 너무 크고 최대곡률과

최소곡률을 동시에 쓰지 않으면 국소 부위의 기하학적 특성을 바로 알아내기 어렵다. 형태 인덱스는 $(1, -1)$ 사이의 값으로 삼차원 표면이 어떤 형태에 가까운지를 판별하게 하며 굴곡도는 표면의 굴곡의 정도를 판별하기 위한 값으로 다음과 같이 정의된다 [1].

$$s = \frac{\pi}{2} \arctan \frac{\kappa_2 + \kappa_1}{\kappa_2 - \kappa_1} \quad (\kappa_1 \geq \kappa_2) \quad (1)$$

$$c = \sqrt{\frac{\kappa_1^2 + \kappa_2^2}{2}} \quad (2)$$

이때, κ_1 와 κ_2 는 각각 최소 곡률과 최대 곡률을 나타낸다. 군집 분석을 수행했을 때 울퉁불퉁한 메쉬의 표면이나 사소한 구조물이 군집분석 알고리즘에 미치는 영향을 줄이기 위해서 메쉬에 meanshift 알고리즘 [11]으로 형태 인덱스와 굴곡도의 잡음을 필터링하였다. (Figure 2 참조)

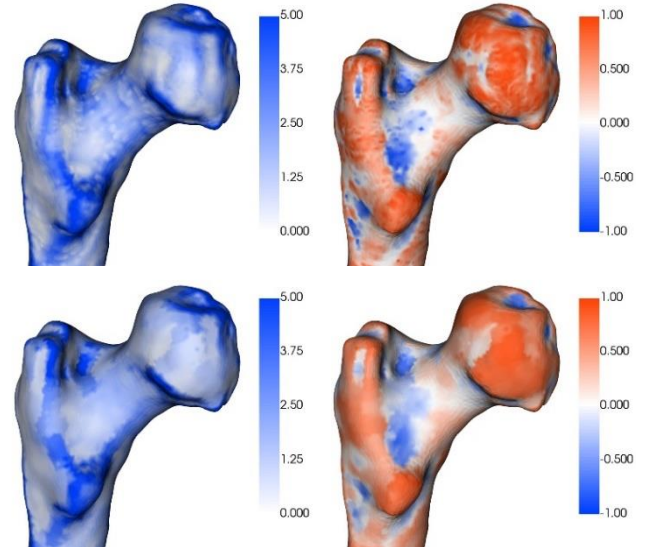


Figure 2: The results of curvedness (left column) and shape index (right column) calculated according to femur mesh model respectively. The results before filtering (upper row). Filtered by meanshift algorithm (lower row).

마지막으로, 형태 인덱스와 굴곡도가 비슷한 부분별로 메쉬를 나누기 위해 계층적 면 군집화 [12]를 이용하였다 (Figure 2 참고). 초기에 모든 면은 하나의 군집을 이루며 이웃한 두 군집 간에 수식 (3)과 같은 우선순위 값이 계산이 되고 우선순위 값이 작은 순으로 군집을 병합하였다. 이때, 면의 형태 인덱스와 굴곡도는 면에 속하는 정점에서 계산된 값의 평균을 사용하였다.

$$\omega_{A,B} = \|s_A - s_B\| + \alpha_1 \|c_A - c_B\| + \alpha_2 \|I_{A,B} - \max(I_A, I_B)\| \quad (3)$$

$$I_G = \frac{r_G^2}{4\pi a_G} \quad (4)$$

여기서, α_1 과 α_2 는 각 항에 적용되는 상대적인 가중치이며 s_i 와 c_i 는 군집 i 에 속하는 모든 면의 형태 인덱스와 굴곡도의 평균이다. 그리고, I_G 는 군집 G 의 모양 가중치를 나타내며, $I_{G,G'}$ 은 G 와 G' 을 병합한 이후의 모양 가중치를 나타낸다. r_G 는 군집 G 의 경계선의 길이를 나타내며, a_G 는 군집 G 의 넓이를 나타낸다. 형태 인덱스와 굴곡도만 이용해서 우선순위를 구성하는 경우 군집의 모양과 경계선이 불규칙해지는 단점이 있다. 군집의 경계선이 불규칙해지면 대퇴골을 의미 있는 단위로 잘 나누기 어렵고 각 부위별로 변형한 후 이들 사이를 연결할 때 연결이 부자연스러워진다. 따라서 우리는 군집을 병합하기 전과 후의 군집의 경계 길이의 차를 수식에 추가했다. 이렇게 하면 군집의 경계길이가 너무 길어지는 것을 방지할 수 있는데, α_2 값을 높일 수록 군집의 모양이 점점 원에 가까워진다고 할 수 있겠다. 본 논문의 결과에서는 α_1 과 α_2 의 값을 각각 0.7과 1.4로 하였다. Figure 3은 계층적 면 군집화를 사용하여 대퇴골 메쉬를 여러 부분으로 나눈 결과이다.

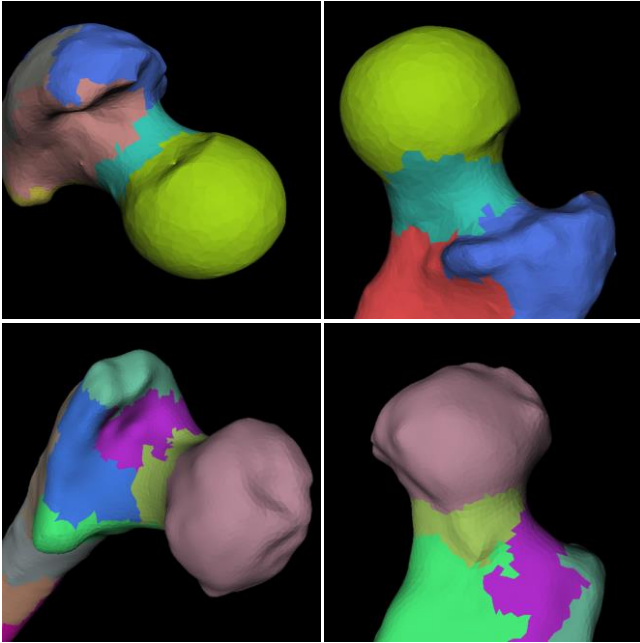


Figure 3: Mesh models of femur separated by specific parts. The result of the cluster analysis of smooth femur (upper row). The result of the clustering of femur with uneven surface (lower row).

5. 메쉬 변형

5 장에서는 4 장에서 여러 군집으로 분할된 대퇴골 메쉬를 해부학적으로 의미있게 변형하는 방법에 대하여 설명한다.

현실에 존재하는 대퇴골과 비슷한 데이터를 만들기 위해서 Wang 등 [8]이 소개한 대퇴골의 각 부위의 크기에 관한 확률변수를 사용했다. 각 확률변수들은 대퇴골 특정 부위의

물리적인 측정치와 연관되어 있으므로, 메쉬를 변형할 때 확률 변수를 참고하기 위해서는 메쉬의 물리적 크기를 알아야 한다. 본 논문에서는 원본 CT 영상 (DICOM 포맷)에 기록되어 있는 column spacing(mm), row spacing(mm), 그리고 slice thickness(mm) 속성들을 픽셀의 물리적인 크기로 사용하여 메쉬의 물리적인 크기를 계산하였다.

Wange 등[8]이 제시한 확률변수들 중, 4 장에서 얻은 군집으로 판별가능한 골두의 반지름 R_{fh} , 골두와 대퇴골체가 이루는 각 A_{ns} , 내측과 L_m , 외측과 L_l , 그리고 대퇴골체의 길이 H_f 를 사용하여 모델을 변형하였다(Figure 4 참고). 크기가 측정된 부위(R_{fh} , L_m , L_l , H_f)에서는 크기 변환(scaling)을, 각도가 측정된 부위(A_{ns})에서는 회전 변환(rotation)을 적용할 수 있다.

이제부터는 4 장에서 얻은 군집들을 이용하여 골두의 반지름 R_{fh} , 골두와 대퇴골체가 이루는 각 A_{ns} , 내측과 L_m , 외측과 L_l , 그리고 대퇴골체의 길이 H_f 를 구하는 방법을 서술할 것이다. 군집들을 구별하기 위한 절대적인 기준이 없으므로, 모양에 대한 휴리스틱 알고리즘으로 각 군집이 어떤 대퇴골 부위에 해당하는지 판별하였다. 대퇴골은 z 축 방향을 따라 길게 놓이므로 대퇴골에서 유의미한 군집을 찾을 때 z축의 좌표를 중요한 지표로 사용하였다. 우선 위쪽의 경계값 th_{upper} 와 아래쪽의 경계값 th_{lower} 두 가지를 정하였다. th_{upper} 와 th_{lower} 는 메쉬 위쪽과 아래쪽으로부터 전체 길이의 25% 만큼 내려가거나 올라온 부분의 z값으로 정하였다. 각수치는 다음과 같이 계산된다.

- R_{fh} : 전체 군집들 중 군집의 중점의 z좌표가 th_{upper} 이상인 군집들 중에서 이웃한 군집의 개수가 한 개인 군집(골두)을 찾는다. 그리고 군집의 중점과 정점들 간의 거리를 계산해서 평균값을 A_{ns} 로 사용한다.
- H_f : 군집의 중점의 z 좌표가 th_{upper} 이하, th_{lower} 이상인 군집들의 합집합 S 를 구하고, 이를 한 개의 군집으로 만든다. 그리고 군집 S 와 S 의 이웃 간의 병합 우선순위가 일정 이하인 동안 계속 병합을 한다. 최종 군집 S 의 (최대 S_z - 최소 S_z)를 H_f 로 사용한다.
- A_{ns} : R_{fh} 를 계산할 때 찾은 골두 군집의 중심(골두 군집에 속해 있는 정점들의 좌표의 평균)과 골두 군집에서 바로 인접해 있는 군집의 중심 쪽으로 향하는 벡터 x 와, H_f 에서 가장 처음에 구한 합집합 S 의 정점 집합에 대해서 주성분분석(principal component analysis) [13]을 수행하여 고유값의 크기가 가장 큰 고유벡터 y 를 찾으면, $\arccos(\frac{x \cdot y}{\|x\| \cdot \|y\|})$ 로 각도를 구할 수 있다.
- L_m , L_l : 외측과와 내측과가 군집분석에서 완벽하게 분리되지 않는 문제가 있다. 대신, 외측과와 내측과의

앞면은 분할이 잘 된다. 따라서 이 두 개의 군집들로부터 값을 추정할 수 있다. 먼저, th_{lower} 에 있는 군집들 중 형태 인덱스의 평균값이 가장 낮은 군집을 찾는다. 이 군집의 이웃중, 형태 인덱스의 평균값이 가장 큰 두 개의 군집 g_1 , g_2 를 찾는다. 그리고 두개의 군집의 중점을 잇는 방향의 벡터와 R_{fh} 에서 찾은 y 벡터의 외적을 구하여 벡터 z 를 구한다. 그 후, g_1 , g_2 의 중점으로부터 z 방향으로 향하는 직선과 메쉬와의 교차점을 구하고, 중점으로부터 교차점까지의 길이를 각각 L_m , L_l 로 사용한다.

Wang 등 [8]이 소개한 정규분포를 이용하여 5 개의 대퇴골 부위의 변수들(R_{fh} , L_m , L_l , H_f , A_{ns})의 분포로부터 값을 임의로 추출하여 메쉬 변형 시에 사용함으로써 대퇴골의 해부학적 특성을 유지한, 다양한 형태의 메쉬를 생성하였다.

이때, 각 부위의 형태적 특성에 차이가 있으므로 두 가지 메쉬 변형 방법을 사용했다. 먼저, 대퇴골체는 전체적인 길이만 변형하면 충분하므로 자유형상변형기법(free-form deformation, FFD)을 적용하였다 [4]. 대퇴골체를 변형하기 위해서 일정한 간격으로 나뉜 격자들의 기준점들 중, z 축 좌표가 H_f 의 th_{lower} 이상이고 th_{upper} 이하인 기준점들 사이의 z 축 간격을 정규분포에서 뽑힌 길이로 늘어나도록 늘렸다.

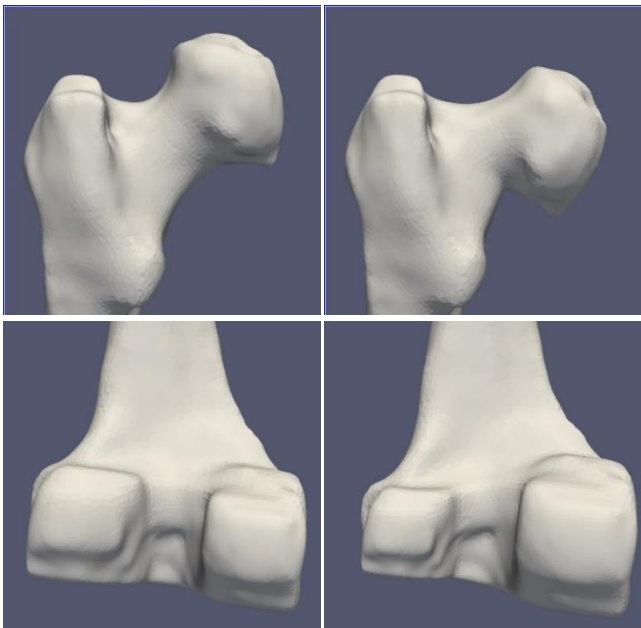


Figure 4: Original femur mesh model (left column), and modified mesh model using laplacian deformation on the outward surface (right column).

FFD는 구현이 쉽고 변형 방법이 직관적이거나 이를 이용하여 메쉬의 구조적 특징과 표면의 국소적 특징을 유지하면서 메쉬를 변형하기는 어렵다. 따라서 나머지 부위에서는 Laplacian deformation [15]을 적용하였다. Laplacian deformation을 사용하기 위해서는 변형이 일어날 부위와

그렇지 않은 부분을 나누는 경계 정점 집합이 필요하다. 이미 대퇴골 메쉬를 부위별로 나누었으므로 군집의 경계 정점들을 이용하여 경계집합을 계산할 수 있다. 경계는 군집의 경계 정점집합에 속한 각 정점의 3-ring neighborhood들의 합집합을 사용하였다. A_{ns} 를 변형할 때는 골두에 대해서 크기 변환만 적용했고, R_{fh} 를 변형하기 위해서 x 와 y 의 외적 z 를 구한 다음, z 를 축으로 하여 골두를 회전하였다. 마지막으로, L_m , L_l 에서 변형해야 할 내측과와 외측과를 정확히 분리하는 것이 불가능하기 때문에, 각각의 앞면의 크기만을 늘렸다.

6. CT 영상 와핑

6 장에서는 변형된 대퇴골 메쉬 모델을 기반으로 CT 영상을 와핑하는 방법에 대해서 서술할 것이다.

정점의 개수를 보존하면서 대퇴골 메쉬를 변형하였으므로 이동된 정점과 원본 정점을 일대일로 대응하여 원본 정점에 대응되는 CT 영상의 복셀이 변형 후에 얼마나 이동했는지를 나타내는 위치 이동 벡터를 모든 원본 정점에 대하여 계산할 수 있다. CT 영상의 모든 복셀에 대한 이동 벡터를 계산하기 위해 두 모델의 정점 간의 이동 벡터를 전체 CT 영상 영역에 대해 보간한다. 그런 다음, 보간된 벡터들을 이용하여 CT 영상 내의 복셀들을 이동시켜 변형된 모델과 대퇴골 영역이 일치하는 CT 영상을 생성한다(Figure 5).

이동 벡터 보간에는 multilevel B-splines 을 이용한 근사 방법 [16]을 사용하였다. B-splines 을 이용한 데이터 근사는 제어 격자의 해상도가 결과에 큰 영향을 준다. 해상도가 낮은 경우, 데이터의 전체적인 변화 혹은 형태는 잘 근사되나 근사 값과 보간하려는 값 간의 오차가 커지며 해상도가 높은 경우에는 오차는 작지만 데이터 부근에서만 정확한 근사가 이루어지므로 전체적인 변화는 잘 반영되지 않는다. multilevel B-splines 을 이용한 방법은 낮은 해상도의 제어 격자로부터 근사를 시작하여 보간하려는 값과 근사된 값 간의 오차를 다음 단계의 높은 해상도의 제어 격자로 근사한 후, 여기서 계산된 근사 값을 낮은 해상도에서 계산된 근사 값에 더 하는 과정을 여러 단계에 걸쳐 반복함으로써 낮은 해상도에서 얻게 되는 전체적인 경향성을 유지하면서 데이터 오차를 줄이는 방법이다.

제어 격자의 해상도가 $(l+3) \times (m+3) \times (n+3)$ 인 단계에서 계산되는 근사 함수 $f(s, t, u)$ 는 다음 식과 같이 정의된다.

$$f(s, t, u) = \sum_{p=0}^3 \sum_{q=0}^3 \sum_{r=0}^3 B_p(\hat{s}) B_q(\hat{t}) B_r(\hat{u}) \phi_{\alpha\beta\gamma}. \quad (3)$$

여기서 $\phi_{\alpha\beta\gamma}$ 는 (α, β, γ) 에 있는 제어점의 값이며 $\alpha = [sl] - 1 + p$, $\beta = [tm] - 1 + q$, $\gamma = [un] - 1 + r$, $\hat{s} = sl - [sl]$, $\hat{t} = tm - [tm]$, $\hat{r} = un - [un]$ 이다. B_p, B_q, B_r 은 균일 삼차 B-spline 기저 함수로 다음과 같이 정의된다.

$$B_0(\lambda) = (1 - \lambda)^3/6, B_1(\lambda) = (3\lambda^3 - 6\lambda^2 + 4)/6, \\ B_2(\lambda) = (-3\lambda^3 + 3\lambda^2 + 3\lambda + 1)/6, B_3(\lambda) = \lambda^3/6.$$

이때, $0 \leq \lambda \leq 1$ 이다.

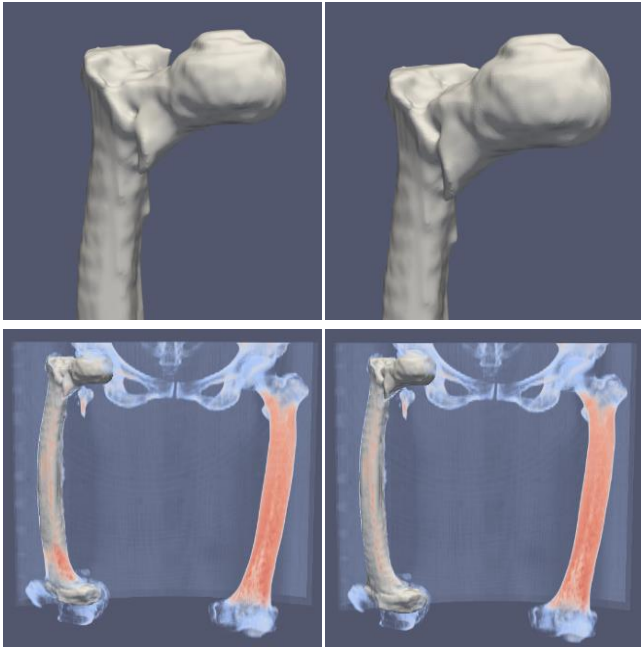


Figure 5: Original mesh model (upper left) deformed mesh model (upper right), Original CT image (lower left) CT image whose voxels are moved using interpolated vectors (lower right).

원본 메쉬와 변형된 메쉬의 정점 간의 이동 벡터를 보간할 때, (s, t, u) 는 CT 영상 내의 한 점이며 이동 벡터의 각 좌표값은 $f(s, t, u)$ 에 의해 계산되는 값이다. 이동 벡터는 삼차원 벡터이므로 각 축의 좌표값마다 근사 함수를 계산한다. 본 논문의 실험에 사용된 CT 영상의 해상도는 $512 \times 512 \times 419$ 이며 이동 벡터 보간 시에는 $4 \times 4 \times 4$ 해상도의 제어 격자로부터 $32 \times 32 \times 32$ 해상도의 제어 격자까지 총 4 단계로 계산하였다. 이동 벡터 보간을 통해 CT 영상 내의 각 복셀에 대한 이동 벡터를 계산할 수 있으며 이를 이용하여 backward mapping을 통해 CT 영상을 와핑했다.

7. 실험 결과

7.1 데이터 수집

아주대학교 병원에서 제공한 한 개의 하반신 CT 영상과 SMIR [17]에서 7 개의 전신 CT 영상을 실험에 사용하였다. 메쉬를 해부학적으로 유의미하게 변형할 때 대퇴골의 물리적인 크기가 필요하므로 DICOM 포맷의 파일 헤더에 픽셀 간격(pixel spacing)이나 단층 두께(slice thickness)와 같은 픽셀의 물리적인 크기를 계산할 수 있는 정보가 있는지 확인 후에 데이터를 수집하였다. 본 연구에서는 대퇴골 부위만을 사용할 것이므로, 전신 CT 영상에서 골반부터 무릎 뼈까지의 영상만을 잘라서 저장하였다.

7.2 딥러닝 모델 학습 및 결과 비교

학습 성능을 향상시키기 위해 몇 가지 영상 전처리를 수행하였다. CT의 HU(Hounsfield Unit) 값을 빼와 연골 부분만 나타나도록 0에서 2400까지로 절단(clipping)하였고, 연골부분의 대비를 효과적으로 높이기 위해 대비제한 적응 히스토그램 평활화(contrast limited adaptive histogram equalization) [18]를 적용하였다.

본 연구에서는 영상 분할을 위해 3D U-Net [2]를 사용하였다. U-Net은 크게 두 부분으로 구성되어 있다. 첫 번째는 영상의 특징을 추출하는 인코더(encoder) 부분이다. 그리고 두 번째는 인코더의 마지막 단계에서 추출된 특징 맵(feature map)에 상향 샘플링(up-sampling)을 하고, 이 결과를 인코더의 각 단계에서 추출된 특징 맵과 합치는 것을 반복하여 영상 전체에 대한 분할 결과를 생성해내는 디코더(decoder) 부분이다. 이때, 인코더와 디코더에서 특징 맵을 추출하기 위해 이차원 컨볼루션, 이차원 풀링(pooling), 그리고 이차원 상향 샘플링 연산이 사용된다. 3D U-Net은 U-Net에서 사용한 이차원 연산들을 삼차원으로 확장하였다. 그리고 딥러닝 모델 학습과정에서 발생하는 여러가지 문제를 완화하는 배치 정규화(batch normalization)를 적용하여 볼륨 데이터(volume data)에 대해서 영상분할 성능을 향상시킨 딥러닝 모델이다.

3D U-Net에서는 모든 연산이 삼차원으로 바뀐 탓에 학습에 걸리는 시간이 기하급수적으로 늘어났다는 단점이 있다. 하지만, 대퇴골의 해부학적 구조가 여러 장의 영상에 걸쳐서 나타나기 때문에 3D U-Net이 이차원 연산 기반의 영상분할 딥러닝 모델보다 본 연구에 더 적합한 방법이라고 생각하여 선택하였다. 그런데, 원본 CT 영상의 크기가 너무 커서 3D U-Net으로 학습시키기에는 계산 시간이 너무 오래 소요되므로, x, y, z 축 각각에 대하여 4, 4, 4 간격으로 하향 샘플링(down-sampling) 하였다.

8 개의 데이터 중, 5 개를 훈련 데이터(training data set)로, 나머지 세 개를 시험 데이터(test data set)으로 사용하였다. 훈련 데이터로 선택된 5 개의 데이터를 기존 방법과 본 논문에서 제안하는 방법을 이용하여 각각 50 개씩 증강된 데이터를 만든

뒤, 원본 데이터와 학습 데이터로 사용했다. 영상 분할 성능은 IoU(Intersection over Union) [19]로 측정하였으며 이는 아래와 같이 정의된다.

$$IoU = \frac{true\ positive}{true\ positive + false\ positive + false\ negative}$$

Table 1 에서 볼 수 있듯이, 1, 2 번에서는 큰 성능 향상이 없었지만, 영상의 노이즈가 심한 3 번 CT 에서 큰 성능 향상을 보였다. 또한, 1,2 번 CT 에서도 골두와 무릎뼈 부분의 영상분할 결과가 향상된 것을 알 수 있다 (Figure 6 참조).

Table 1: Segmentation experiment		
CT no.	Proposed (IoU)	Arbitrary (IoU)
1	0.98	0.96
2	0.98	0.98
3	0.97	0.92

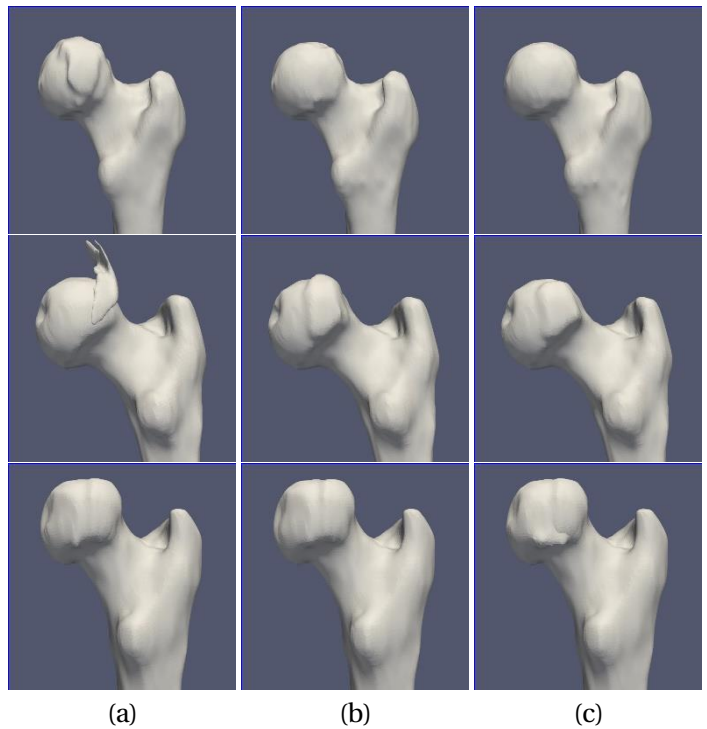


Figure 6: Comparison between the result in segmentation of the CT images of femoral head. Translating the binary label of the deep learning model to a 3D triangle mesh by marching cubes algorithm. Segmentation results of arbitrary augmentation method(a), proposed augmentation method(b) in turn. (c) is original mesh.

8. 결론 및 향후 연구 방향

본 논문에서는 CT 영상의 대퇴골 부위를 해부학적으로 의미 있게 변형하여 컨벌루션 신경망(CNN)의 훈련 데이터를 증강하는 방법을 제안하였다. 기존의 방법에 비해 극단적인 성능 향상을 보이지는 않지만, 노이즈가 심한 영상, 그리고 골두와 무릎뼈 주위 등, 경계가 뚜렷하지 않은 부위에서 비교적 좋은 성능을 보였다. 그런데 영상 분할 성능은 높게 측정되지만, 분할된 영역의 경계가 실용적으로 쓰이기에 충분히 깨끗하지 않으므로 향후, 영상 매팅(image matting)을 이용하여 이 문제를 개선해 볼 계획이다. 또한, 데이터 증강 시에 대퇴골 각 부위의 크기만 변형하기보다는, 메쉬 표면에 노이즈를 주거나 모양을 바꾸는 등, 다양한 변형 방법을 이용하여 딥러닝 모델의 일반화(generalization) 성능을 높일 수도 있을 것이다. 마지막으로, 4 장에서 대퇴골 메쉬를 의미 있는 부분들로 군집화 할 때 사용되는 우선순위의 각 항의 가중치에 따라 군집화 결과가 극단적으로 달라진다는 문제가 있다. 이 변수들의 조정을 좀더 효과적으로 할 수 있는 방법을 찾아볼 것이다.

감사의 글

이 성과는 2018 년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. NRF-2018R1A2B6002909).

References

[1] C. Rich, S. Lawrence, and C. -L. Giles, "Overfitting in neural nets: Backpropagation, conjugate gradient, and early stopping", *Advances in neural information processing systems*, 2001.

[2] O. Çiçek, et al, "3D U-Net: learning dense volumetric segmentation from sparse annotation", *International conference on medical image computing and computer-assisted intervention. Springer, Cham*, 2016.

[3] P. -F. Christ, et al, "Automatic liver and lesion segmentation in CT using cascaded fully convolutional neural networks and 3D conditional random fields", *International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham*, 2016.

[4] J. -Y. Zhu, et al, "Unpaired image-to-image translation using cycle-consistent adversarial networks", *IEEE Transactions on computer vision*, 2017.

[5] S. -C. Wong, et al, "Understanding data augmentation for classification: when to warp?", *IEEE Transactions on digital image*

computing: techniques and applications (DICTA), 2016.

[6] W. -E. Lorensen, and H. -E. Cline, "Marching cubes: A high resolution 3D surface construction algorithm", *ACM siggraph computer graphics. Vol. 21. No. 4*, 1987.

[7] P. Yves, et al, "Interactive graph-cut segmentation for fast creation of finite element models from clinical ct data for hip fracture prediction", *Computer methods in biomechanics and biomedical engineering 19.16*, pp. 1693-1703, 2016.

[8] L.Wang, K. He, and Z. Chen, "Statistical Analyses of Femur Parameters for Designing Anatomical Plates," *Computational and Mathematical Methods in Medicine*, pp. 12, 2016.

[9] Y. -J. Lee, "Mesh Scissoring: Contour-Based Mesh Segmentation. PhD dissertation", *Postech*, 2005.

[10] J. J-. Koenderink, and A. -J. Doorn, "Surface shape and curvature scales", *Image and Vision Computing 10*, pp. 557–565, 1992.

[11] D. Comaniciu, and P. Meer, "Mean shift: A robust approach toward feature space analysis", *IEEE Transactions on Pattern Analysis & Machine Intelligence 5*, pp. 603-619, 2002.

[12] M. Garland, A Willmott, and P. Heckbert, "Hierarchical face clustering on polygonal surfaces", *ACM Symposium on Interactive 3D Graphics 2001*, pp. 49–58, 2001.

[13] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis", *Chemometrics and intelligent laboratory systems 2.1-3*, pp. 37-52, 1987.

[14] T. -W. Sederberg, and, S. -R. Parry, "Free-form deformation of solid geometric models", *Annual conference on Computer graphics and interactive techniques 13th (SIGGRAPH '86)*, 1986.

[15] O. Sorkine, et al, "Laplacian surface editing", *Eurographics/ACM SIGGRAPH symposium on Geometry processing 2004 (SGP '04)*, pp. 175-184, 2004.

[16] S. Lee, G. Wolberg, and S. -Y. Shin, "Scattered data interpolation with multilevel B- splines", *IEEE Transactions on Visualization and Computer Graphics*, vol. 3, no. 3, pp. 228-244, 1997.

[17] M. Kistler, et al, "The virtual skeleton database: an open access repository for biomedical research and collaboration", *medical Internet*

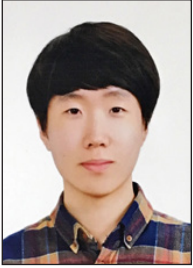
research, e. 245, 2013.

[18] K. Zuiderveld. "Contrast limited adaptive histogram equalization", *Graphics gems IV*. Academic Press Professional, 1994.

[19] P. Jaccard, "Étude comparative de la distribution florale dans une portion des Alpes et des Jura", *Bull Soc Vaudoise Sci Nat 37*, pp. 547-579, 1901.

[20] 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.

〈 저 자 소 개 〉



최 규 진

- 2015년 ~ 현재 아주대학교 미디어학과 재학
- 관심분야 : 컴퓨터그래픽스, 메쉬 프로세싱, GPU 가속
- <https://orcid.org/0000-0002-9222-3273>



경 주 현

- 2015년 ~ 현재 아주대학교 미디어학과 재학
- 관심분야 : 컴퓨터그래픽스, 메쉬 프로세싱, GPU 가속
- <https://orcid.org/0000-0002-7202-4285>



신 주 연

- 2016년 ~ 현재 아주대학교 미디어학과 재학
- 관심분야 : 컴퓨터그래픽스, 메쉬 프로세싱, GPU 가속
- <https://orcid.org/0000-0002-6256-0359>



경 민 호

- 1993년, 1995년 포항공과대학교 전자계산학과 졸업(학사, 석사)
- 2001년 Purdue University Dept. of Computer Science 졸업(박사)
- 2001년~2002년 Purdue University 방문 연구원
- 2002년~현재 아주대학교 미디어학과 교수
- 2008년~2009년 Purdue University 방문 교수
- 관심분야: 기하 모델링, 기하 알고리즘, GPU 가속
- <https://orcid.org/0000-0003-3037-4292>



이 윤 진

- 2005 포항공과대학교 컴퓨터공학과 박사
- 2006~2007 University of Michigan 박사 후 연구과정
- 2007~2008 서울대학교 BK21 연구교수
- 2008~2012 아주대학교 미디어학부 조교수
- 2013~2017 아주대학교 미디어학과 부교수
- 2017~현재 아주대학교 미디어학과 교수
- 관심분야: 컴퓨터그래픽스, 비사실적 렌더링, 메쉬 프로세싱, 표면 복원
- <https://orcid.org/0000-0002-5627-8314>