

# 근감소증 진단을 위한 척추 분류 모델 개발 및 검증

이충섭<sup>1</sup>, 임동욱<sup>1</sup>, 노시형<sup>1</sup>, 박철<sup>2</sup>, 정창원<sup>1,3</sup>  
<sup>1</sup>원광대학교 의료융합연구센터  
<sup>2</sup>원광대학교병원 호흡기내과  
<sup>3</sup>원광대학교병원 의생명연구원, 중점연구센터  
 e-mail : {cslee99, dw1316, nosij123}@wku.ac.kr,  
 {cholssak21, mediblue}@wkuh.org

## Development and Validation of Spine Classification Model for Sarcopenia Diagnosis and Validation

Chung-sub Lee<sup>1</sup>, Dong-Wook Lim<sup>1</sup>, Si-Hyeong Noh<sup>1</sup>, Chul Park<sup>2</sup>,  
 Chang-Won Jeong<sup>1,3</sup>

<sup>1</sup>Medical Convergence Research Center, Wonkwang University  
<sup>2</sup>Division of Pulmonary Medicine, Department of Internal Medicine, Wonkwang University Hospital  
<sup>3</sup>Center Research Center of Biomedical Research Institute of Wonkwang University Hospital

### 요 약

컴퓨터 단층촬영(CT)을 활용한 골격근 단면적은 근감소증과 관련된 기능을 평가하는 데 사용된다. 일반적인 근감소증 연구는 요추 3번의 골격근량을 주로 보지만 암 또는 폐절제술과의 상관관계를 예측하기 위한 다양한 연구에서는 흉추 4번, 7번, 8번, 10번, 12번 다양한 수준의 골격근량으로 연구를 진행하고 있음을 알 수 있다. 본 논문에서는 흉부와 복부 CT 영상에서 근감소증 진단을 위해서 흉추와 요추의 영역별 슬라이스를 검출하기 위해서 CNN 구조의 EfficientNetV2를 전이학습하여 인공지능 모델을 개발하였다. 인공지능 모델은 전체 흉부 및 복부 CT 영상에서 Cervical, T1, T2, T3, T4, T5, T6, T7, T8, T9, T10, T11, T12, L1, L2, L3, L4, L5, Sacral 총 19 클래스를 검출하도록 하였다. Test 데이터셋을 사용하여 Confusion Matrix와 Grad-CAM으로 모델의 정확도를 시각화하여 보였으며 검증으로 인공지능 모델의 정확성을 측정하였다. 끝으로 우리가 개발한 다기관 공동연구 지원 플랫폼에 적용하여 시각화된 결과를 보였다.

### 1. 서론

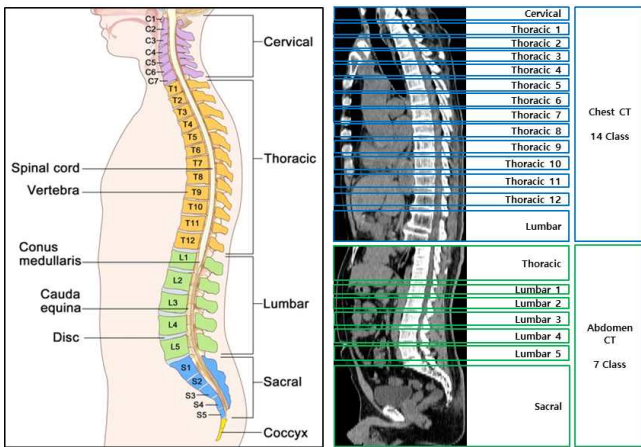
컴퓨터 단층촬영(CT) 스캔을 활용한 골격근 단면적은 근감소증, 골격근량 손실 및 노화와 관련된 기능을 평가하는 데 사용된다. 일반적으로 세 번째 요추의 3번 수준에서 골격근을 측정하여 컷오프 값을 기준으로 근감소증을 진단하는데 사용한다[1]. 흉부 CT 영상에서 근감소증을 진단하기 위해 흉추의 7번, 12번 수준에서 골격근량을 측정하여 분석하는 연구도 있다[2]. 또한 비소세포폐암(NSCLC)으로 폐절제술을 받은 환자의 사망 위험이 근감소증과 관련이 있는지 예측하기 위해서 흉추의 8번, 10번, 12번에서 골격근을 측정하는 연구도 있다[3]. 마지막으로 수술 전 직장암 환자의 흉추 4번, 12번, 요추 3번 수준에서 골격근의 변화와 재현성을 평가하는 연구 또한 진행되고 있다[4]. 일반적인 근감소증 연구는 요추 3번의 골격근량을 주로 분석하여 활용하지만 위의 다양한 연구에서는 흉추 4번, 7번, 8번, 10번, 12

번의 다양한 수준에서 골격근량으로 측정하여 근감소증 진단 연구를 진행하고 있음을 알 수 있다. 또한 요추와 흉추는 촬영하는 임상 CT 프로토콜이 다르기 때문에 기존에 개발된 요추 분류 모델[5]과 흉추 분류 모델도 개발되어야 영상분할과 정량분석을 자동으로 진행할 수 있다. 따라서 본 논문에서는 흉부와 복부 CT 영상을 입력하면 흉추의 T1~T12와 요추의 L1~L5 슬라이스를 자동 분류하는 인공지능 모델을 개발하고 검증을 통하여 흉추 분류 모델의 정확도를 향상시키고자 하였다.

### 2. CT 의료영상의 부위별 분류 및 학습데이터

흉부 CT 영상에서 흉추 T1~T12를 분류하기 위해서는 흉부 위쪽에 위치한 경추 1 클래스와 각 수준별로 흉부를 분류한 12 클래스 그리고 흉부 아래에 위치한 요추 1 클래스로 총 14 클래스로 구분하였다. 그리고 복부 CT 영상에서 위에 흉추 1 클레

스와 각 수준별로 요추를 분류한 5 클래스 그리고 척추 1 클래스 총 7 클래스로 구분하였다. 이미지의 분류는 디스크 부분은 제외하고 각 레벨에 해당하는 부위만 선택하여 4명의 레이블러가 레이블링하였고, 2명의 임상교수가 검증하였다. (그림 1)과 같이 (a) 척추의 해부학적 명칭에 따라 흉부와 복부 CT 영상의 세부 수준을 분류할 수 있도록 (b)와 같이 흉부 CT는 Cervical~Lumbar로 14 클래스로 복부 CT는 Thoracic~Sacral 7 클래스로 각 부위별로 수동 분류하여 학습 데이터를 구축하였다.



(a) 척추의 해부학적 명칭 (b) 의료영상의 분류  
(그림 1) 흉부와 복부 CT 의료영상의 분류 정의

척추 분류 모델을 개발하기 위해서 흉부와 복부 CT 데이터를 활용하였다. 학습, 테스트를 위한 데이터셋 현황은 <표 1>과 같다.

<표 1> 학습, 테스트 데이터셋 현황

Class	Train	Test	Class	Train	Test
Cervical	6045	1512	T10	2828	707
T1	1966	492	T11	3006	752
T2	1993	499	T12	3168	793
T3	2021	506	L1	1343	347
T4	2065	517	L2	1415	357
T5	2209	553	L3	1477	376
T6	2352	588	L4	1511	349
T7	2475	619	L5	1551	397
T8	2559	640	Sacral	9272	2254
T9	2668	668			

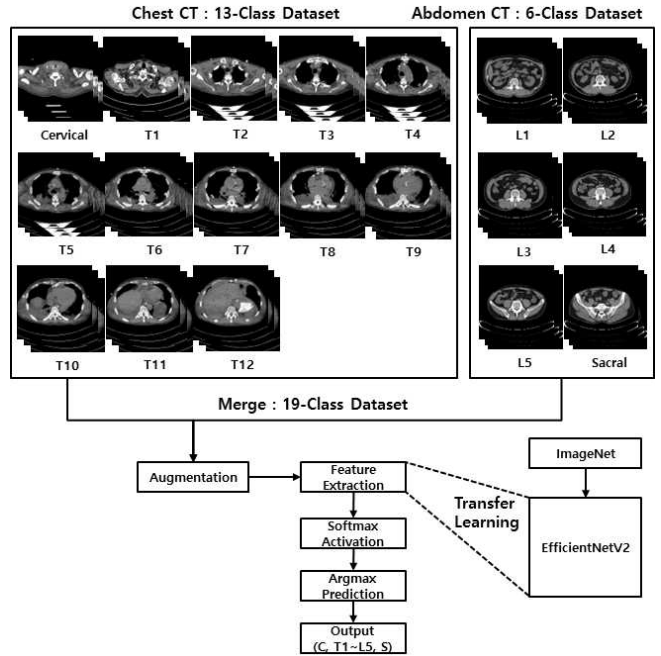
\* Unit : Image Count

### 3. 척추분류 모델 개발 및 시각화

#### 3-1. 척추 분류 모델

척추 분류 모델은 EfficientNetV2[6]로 개발하였

다. (그림 2)와 같이 학습을 위한 Input으로 총 19개의 클래스로 구성된 데이터셋으로 학습하여 모델을 생성하였다. 부족한 데이터셋의 개수를 보강하기 위해서 데이터 증강을 진행하였다. 각 클래스의 특징을 찾기 위해서 ImageNet의 EfficientNetV2를 전이 학습(Transfer Learning)하여 이미지를 분류하도록 학습하였다.



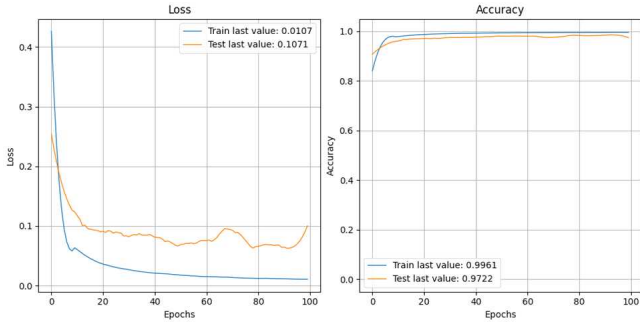
(그림 2) 척추 분류 모델의 전체 구조

학습을 위한 하이퍼파라미터는 아래 <표 2>와 같다.

<표 2> 학습을 위한 하이퍼파라미터

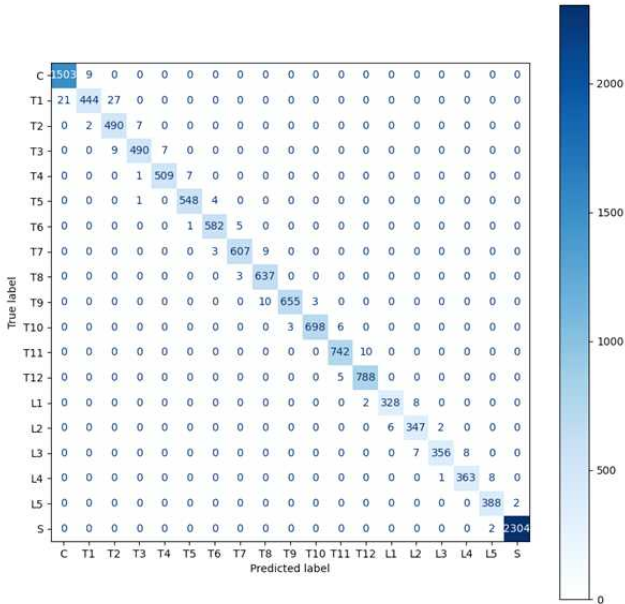
Name	Description
Input Size	512 x 512
Model	EfficientNetV2M
Activation	Softmax
Loss function	Categorical_CrossEntropy
Batch size	16
Learning rate	0.0001
Epochs	100
Optimizer	Adam

그리고 현재 모델의 학습이 CT 이미지의 원본 사이즈인 512x512 사이즈로 학습하였다. 개발된 분류 모델의 학습데이터 결과는 val\_loss: 0.0107 val\_accuracy: 0.9961이고 테스트데이터 결과는 val\_loss:0.1071 val\_accuracy: 0.9722로 History Plots은 (그림 3)과 같다. 베스트 모델로 92번째 Epoch의 모델이 저장되었다.



(그림 3) 학습 결과의 History Plots

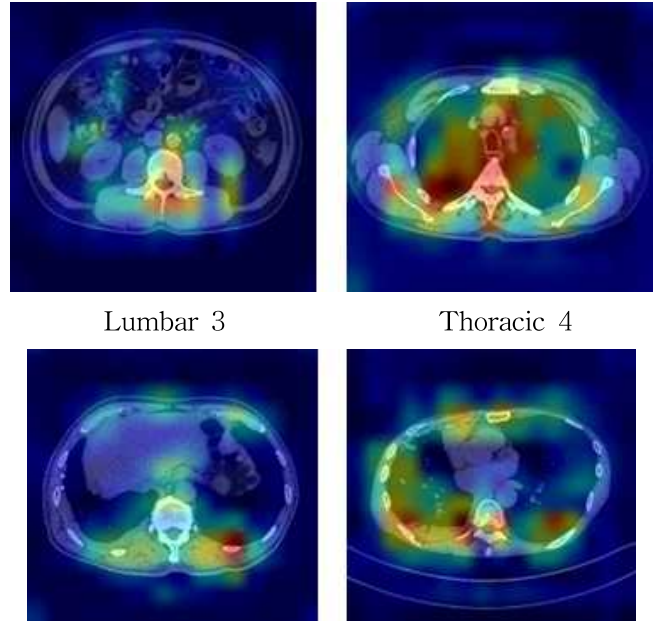
학습 후에 Test 데이터를 활용하여 분류 모델의 알고리즘 성능 평가한 결과는 (그림 4)와 같다.



(그림 4) Test 데이터의 Confusion Matrix

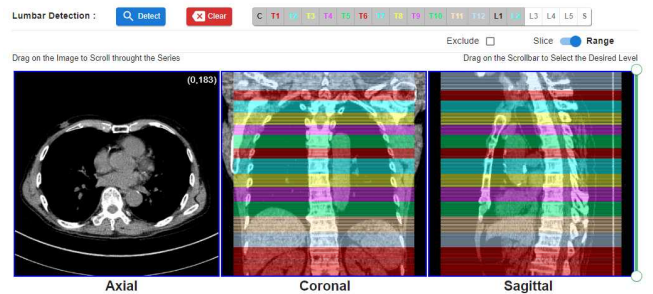
### 3-2. 분류 모델의 시각화

분류 모델의 정확도에 대한 성능평가가 이루어졌지만 시각화를 통해 확인할 필요가 있다. 인공지능이 어떤 특징을 추출해서 이런 결과를 예측했는지에 대한 요인을 분석하는 것도 중요하다. 이런 주요 요소를 찾는 방법으로 딥러닝 신경망이 어떤 근거로 분류 결정을 내렸는지 확인할 수 있는 XAI 기법[7]으로 Grad-CAM을 주로 사용한다. (그림 5)는 근감소증 진단에 중요한 요추 3번 슬라이스와 근감소증과 상관관계를 증명하는 다양한 연구에서 흉추 4번, 8번, 12번을 활용하는데 Grad-CAM을 통해 결과를 보여주고 있다. 보는 바와 같이 척추뼈를 중심으로 그 주변의 특징을 중심으로 영상을 분류하고 있음을 확인할 수 있다.



(그림 5) 요추, 흉추 슬라이스의 Grad-CAM

(그림 6)은 제안한 척추 분류 모델의 결과를 시각화하기 위해서 Sagittal 영상으로 변환하여 척추의 측면 모습으로 분류된 결과를 확인할 수 있도록 하였다.



(그림 6) 흉추 영역 선택 UI

기존에 개발한 다기관 공동연구 지원 플랫폼에 적용하였다. Axial 영상을 Sagittal 영상으로 재구성하여 각 척추 슬라이스를 수동으로 영역을 선택하여 라벨링하거나 분석 프로세스를 자동화 하였다. 특히, 척추 분류 모델을 적용하고, 시각화를 통해 사용자가 원하는 영역의 척추 슬라이스를 자동으로 분류하고 영상분석과 결과를 확인할 수 있도록 하였다.

### 4. 검증(Validation)

학습을 통하여 척추 분류 모델을 개발하고 검증과 테스트까지 완료하였다. 개발 완료된 모델을 검증하기 위해서는 모델의 예측한 값들을 별도의 데이터셋과 비교하는 검증을 수행한 결과는 <표 3>과 같다.

<표 3> 검증 결과

Class	I.V.	Acc.	Class	I.V.	Acc.
Cer.	231	231(0) 100%	T10	107	106(1) 99%
T1	75	70(5) 93%	T11	116	111(5) 96%
T2	77	68(9) 88%	T12	122	119(3) 98%
T3	80	74(6) 93%	L1	613	606(7) 99%
T4	81	76(5) 94%	L2	644	633(11) 98%
T5	85	80(5) 94%	L3	681	662(19) 97%
T6	88	75(13) 85%	L4	663	660(3) 99%
T7	91	80(11) 88%	L5	708	702(6) 99%
T8	100	91(9) 91%	Sac.	2254	2254(0) 100%
T9	106	91(15) 86%			

\* Unit : Image Count, I.V. : Internal Validation, Acc. : Accuracy, Cer. : Cervical, Sac. : Sacral

<표 3>은 학습에 사용되지 않은 16 Study 데이터를 가지고 검증을 진행하였다. 검증결과 모든 분류된 클래스에서 80% 이상으로 높은 검증결과를 보였다.

**4. 결론 및 향후연구**

본 논문에서는 CNN 구조의 EfficientNetV2를 전이학습하여 근감소증 진단을 위한 척추 분류 모델을 개발하였다. 그리고 모델을 라벨링 틀에 적용하여 전체 흉부 및 복부 CT 영상에서 19개 클래스의 슬라이스를 검출하도록 하고 근감소증 진단에 사용하고자 하는 영역을 선택하여 영상을 분할하고 정량분석을 진행할 수 있도록 하였다. 또한 학습 후 Test 데이터셋의 Confusion Matrix로 모델의 정확성을 검증하였다. 그리고 Grad-CAM이라는 XAI 기법을 통하여 인공지능 모델이 영상 분류에 대한 근거를 확인할 수 있어 예측 결과에 대한 신뢰성을 확인하였다. 향후연구로 다기관과의 공동연구 지원 및 2개 이상의 다기관 외부검증을 진행할 계획이다.

**사사표기**

본 연구는 보건복지부의 재원으로 한국보건산업진흥원의 보건의료기술 연구개발사업(HI18C1216) 그리고 과학정보통신부의 재원으로 한국연구재단(2021R1A5A8029876) 지원에 의하여 이루어진 것임.

**참고문헌**

[1] Derstine BA, Holcombe SA, Ross BE, Wang NC, Su GL, Wang SC. Skeletal muscle cutoff values for sarcopenia diagnosis using T10 to L5 measurements in a healthy US population. *Sci Rep.* 2018 Jul 27;8(1):11369. doi: 10.1038/s41598-018-29825-5. PMID: 30054580; PMCID: PMC6063941.

[2] Nemeč U, Heidinger B, Sokas C, Chu L, Eisenberg RL. Diagnosing Sarcopenia on Thoracic Computed Tomography: Quantitative Assessment of Skeletal Muscle Mass in Patients Undergoing Transcatheter Aortic Valve Replacement. *Acad Radiol.* 2017 Sep;24(9):1154-1161. doi: 10.1016/j.acra.2017.02.008. Epub 2017 Mar 30. PMID: 28365235.

[3] Troschel FM, Jin Q, Eichhorn F, Muley T, Best TD, Leppelmann KS, Yang CJ, Troschel AS, Winter H, Heußel CP, Gaissert HA, Fintelmann FJ. Sarcopenia on preoperative chest computed tomography predicts cancer-specific and all-cause mortality following pneumonectomy for lung cancer: A multicenter analysis. *Cancer Med.* 2021 Oct;10(19):6677-6686. doi: 10.1002/cam4.4207. Epub 2021 Aug 19. PMID: 34409756; PMCID: PMC8495285.

[4] Arayne, A.A., Gartrell, R., Qiao, J. et al. Comparison of CT derived body composition at the thoracic T4 and T12 with lumbar L3 vertebral levels and their utility in patients with rectal cancer. *BMC Cancer* 23, 56 (2023). <https://doi.org/10.1186/s12885-023-10522-0>

[5] Chung-sub Lee et al. "A Study on the Auto Lumbar Spine Classification Model Based on EfficientNetV2." *한국정보처리학회 학술대회논문집* 30, 1 (2023): 448-450.

[6] Tan, Mingxing and Quoc V. Le. "EfficientNetV2: Smaller Models and Faster Training." *ArXiv abs/2104.00298* (2021): n. pag.

[7] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh and D. Batra, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization," 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 2017, pp. 618-626, doi: 10.1109/ICCV.2017.74.