

MR 영상을 이용한 합성곱 신경망 기반의 상완골두 및 관절와 자동 분할

Automated Segmentation of Humeral head and Glenoid based on Convolution Neural Network using Magnetic Resonance Images

박예랑¹, 김영재², 김광기^{3*}, 이신우⁴

Ye Rang Park¹, Young Jae Kim², Kwang Gi Kim^{3*}, Sheen-Woo Lee⁴

요약

어깨 관련 질환에는 회전근개 손상, 관절와순손상, 어깨충돌증후군 및 습관성어깨탈구 등이 있다. 건강보험진료 통계에 따르면, 이러한 어깨병변으로 치료를 받은 환자수는 매년 증가하는 추세이다. 어깨 관련 질환을 진단하는 대표적인 방법으로 MRI(Magnetic Resonance Imaging, 자기공명영상)가 있다. 본 논문에서는 합성곱 신경망 알고리즘인 U-net을 이용해 MR 영상에서 상완골두와 관절와 부위를 자동 분할함으로써 전문의가 관련 질환을 더욱 빠르게 진단하는 데 도움을 주는 연구를 진행하고자 한다. 81명 환자의 MR 영상 총 2251장을 수집하여 이를 훈련셋과 테스트셋으로 사용하였다. 427장 데이터로 검증한 U-net으로 훈련된 어깨뼈 분할 모델의 평가 결과로 민감도 89.32%, 특이도 99.78%, 정확도 99.44% 다이스 계수 92.02%를 보였다. 추후 영상처리 기법을 통해 어깨뼈의 경계영역의 모호성을 해결하고 추가 데이터 수집을 통해 관절와 영역을 충분히 훈련시킨다면 보다 높은 정확도를 얻어 전문의의 진단에 큰 도움을 줄 수 있을 것이라 기대한다.

핵심어 : 인공지능, 딥러닝, 분할, 합성곱 신경망, 어깨 질환

Abstract

Shoulder-related disorders include joint and labial injuries, shoulder impingement syndrome, habitual shoulder dislocation and slab lesions. According to health insurance statistics, the number of patients treated

1 Department of Health Sciences and Technology, Gachon Advanced Institute for Health Sciences and Technology (GAIHST), Gachon University, Incheon, Korea [Graduate Student]

e-mail: rainpark823@gmail.com

2 Department of Biomedical Engineering, Gachon University College of Medicine, Incheon, Korea [Professor]

e-mail: kimyj10528@gachon.ac.kr

3 Department of Biomedical Engineering, Gachon University College of Medicine, Incheon, Korea [Professor]

e-mail: kimkg@gachon.ac.kr (Corresponding author)

4 Department of Radiology, Gachon University, Gil Medical Center, Incheon, Korea. [Professor]

e-mail: leesw1@gilhospital.com

* This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2020-2017-0-01630) supervised by the IITP(Institute for Information & communications Technology Promotion), the GRRC program of Gyeonggi province (GRRC Gachon 2017-B01) and the Gachon University Gil Medical Center (Grant number : FRD2019-11).

Received(February 19, 2020), Review Result(1st: February 29, 2020), Accepted(March 13, 2020), Published(March 31, 2020)



© 2020 The Authors. Published by NCISS.
This is an open access article licensed under the Creative Commons Attribution-NonCommercial 4.0 International License.
To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

with these shoulder lesions is increasing every year. MRI (Magnetic Resonance Imaging) is a representative method of diagnosing such shoulder-related diseases. In this paper, we will use U-net, a convolution neural network algorithm, to automatically segment humeral heads, joints in MR images to help doctors diagnose related diseases more quickly. A total of 2251 MR images of 81 patients were collected and used as training set and test set. As a result of evaluation of U-net trained scapular segmentation model validated by chapter 427 data, sensitivity of 89.32%, specificity of 99.78%, accuracy of 99.44%, and die coefficient of 92.02%. In the future, image processing techniques will solve the ambiguity of the scapular border and collect additional data to fully train the joints and regions, which will provide a higher level of accuracy and help the diagnosis for doctors.

Keyword : Artificial Intelligence, Deep learning, Segmentation, Convolution Neural Network, Shoulder disease

1. 서론

어깨 관련 질환에는 회전근개 손상, 관절와순손상, 어깨충돌증후군 및 습관성어깨탈구 등이 있다 [1][2]. 건강보험진료 통계에 따르면, 이러한 어깨병변으로 진료를 받은 환자수는 매년 증가하는 추세이다 [3]. 어깨 관련 질환을 진단하는 대표적인 방법으로 MRI(Magnetic Resonance Imaging, 자기공명영상)가 있다 [4]. MRI는 강한 자기장과 고주파의 상호작용을 이용해 인체 구조를 단면으로 재구성해내는 검사로, 인체의 해부학적 구조 변형 및 병리학적 정보를 파악할 수 있는 방법이다 [5]. MR 영상을 통해 어깨 부위의 국소적 연골 결손 여부를 확인하고 주변 연부 조직의 이상 유무를 관찰하여 수술 또는 치료 방향을 결정하게 된다 [6]. 하지만 MRI는 CT보다 피질골의 병변의 진단에 취약하며 검사 정보가 방대하여, 한 명의 전문의가 환자 한 명당 수십장에 달하는 영상을 확인하여 상완골두, 관절와 뼈의 결손이나 손상을 체크하는데 긴 시간이 필요하다. 대한민국의 임상 의사수는 1천명당 2.2명으로 OECD 평균인 3.3명보다 1.1명이 적다는 통계에 따르면 [7], 한 명의 전문의가 OECD 평균보다 더 많은 환자를 진료해야 하며 이에 따라 피로도도 증가할 것임을 유추할 수 있다.

의료 진단에서의 이러한 비효율성을 해결하기 위해 최근 의료 영상 분야에서 딥러닝을 활용한 사례가 급격히 증가하고 있다 [8]. 특히 이미지 분할 성능은 딥러닝(Deep learning)의 한 종류인 합성곱 신경망(Convolutional Neural Network, CNN)을 통해 기술발전의 고도화가 이루어지고 있다. 합성곱(Convolution)을 통해 이미지의 특징을 추출하고 서브샘플링(Subsampling)으로 그 크기를 효과적으로 제어하는 CNN 구조는 이미지 인식 분야에서 최고의 성능을 발휘하고 있으며, 특히 영상에서 원하는 영역을 분할하기 위한 활발한 연구가 진행 중이다 [9].

독일의 Heidelberg 대학교 연구팀에서는 U-net 모델을 사용하여 전신 CT 이미지에서 전체 뼈를 분할하는 연구를 진행해 0.92의 dice score를 얻은 바 있다 [10]. 데이터는 한 환자 당 최대 450장의 슬라이스로 이루어진 18명 환자의 PET/CT 데이터를 사용하였다. 캐나다의 한 AI 연구팀에서는 뇌의 MR 영상에서 다양한 CNN알고리즘을 이용하여 뇌종양을 분할하였고 최고 0.88의 dice score, 0.93의 specificity, 0.80의 sensitivity를 얻었다 [11]. 여기서는 55명 환자의 60여개의 고등급 및 저등

급 증양이 포함된 데이터를 사용하였다. 어깨 질환에 관련된 컴퓨터 자동 진단 연구로는, 미국 Michigan 대학교에서 머신러닝 기법 중 하나인 SVM을 이용한 회전근개 파열 환자의 예후 예측으로, AUC 0.87의 결과를 나타낸 사례가 있다 [12].

앞선 독일의 연구에서 사용한 데이터인 CT 영상은 뼈의 경계가 명확하여 뼈의 구조와 위치를 확인하기에 좋은 도구이지만 환자에게 직접 방사선이 노출된다는 문제점이 있다. 이에 본 논문에서는 방사선 노출의 위험성이 적은 MR 영상을 이용하여 합성곱 신경망 알고리즘인 U-net을 기반으로 상완골두와 관절와 부위를 자동 분할함으로써 전문의가 관련 질환을 더욱 빠르게 진단하는데 도움을 주는 연구를 진행하고자 한다.

2. 재료 및 방법

2.1. 데이터 셋(Data Set)

가천대길병원에 방문한 81명 환자의 MR 영상 총 2251장을 수집하였다. 훈련셋은 67명의 환자 데이터 1,824장을 사용하였고, 테스트셋은 검증에 신뢰성을 높이기 위해 훈련에 사용되지 않은 데이터인 14명 환자의 데이터 427장으로 구성하였다.

영상에서 분할하고자 하는 영역을 관심 영역(Region of Interest)이라고 하는데, 훈련의 정답으로 사용되는 이미지인 마스크 이미지를 만들기 위한 관심 영역 표시 작업이 필요하다. 전문의의 교육을 받은 전문가들이 관심 영역 표시 작업을 수행하였고, 해당 과정을 거친 모든 데이터는 전문의의 확인을 받았다. 관심 영역은 어깨 부위의 MR 영상에서 상완골두와 관절와에 해당하는 영역으로 지정하였다.

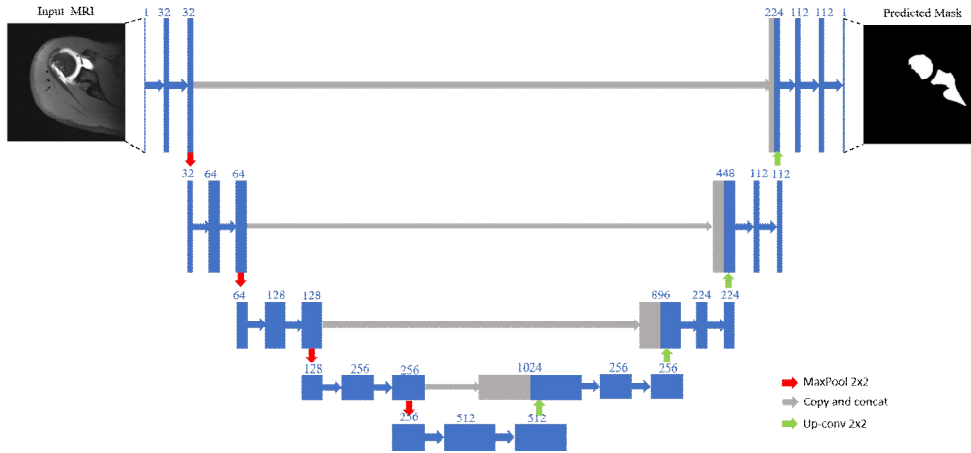
각 픽셀이 0~255사이의 값을 갖는 이진화이미지에서 관심 영역의 모든 픽셀은 255값으로, 나머지 영역은 0값을 부여하여 마스크 이미지를 생성하였고 이를 훈련 및 평가를 위한 레이블로 사용하였다.

2.2. 딥러닝 모델

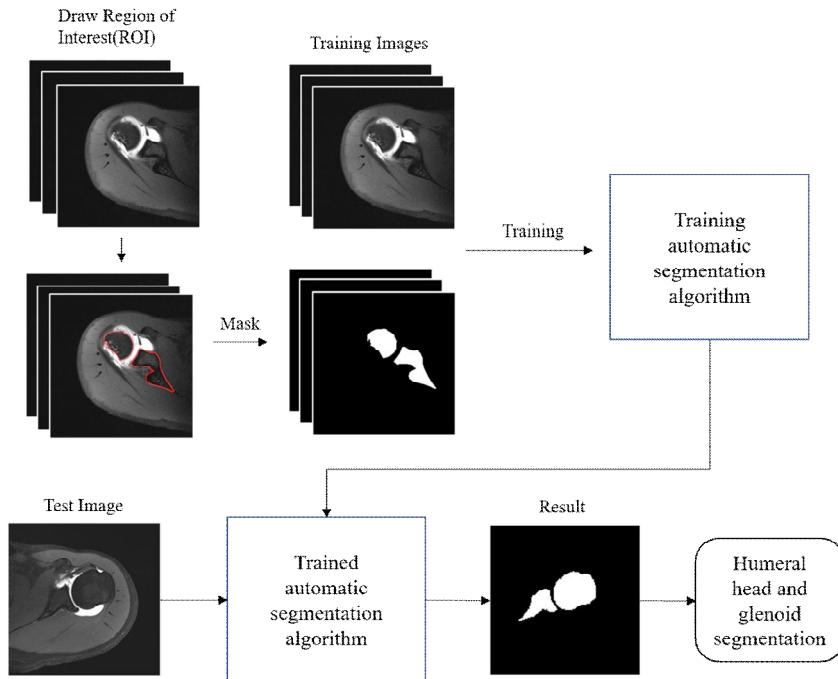
의료 영상 분할 분야에서 [그림 1]과 같이 가장 많이 사용되는 U-net 구조를 기반으로 어깨뼈를 분할하였다 [13]. U-net은 이미지 인식 단위(patch)에 대한 겹침(overlap) 비율이 적어 속도가 빠르고, 다층의 출력결과를 동시에 검증하여 기존 분할 모델의 단점인 인접 문맥(context) 파악과 지역화(localization)의 이율배반 관계를 극복하여 이 두 가지가 동시에 가능하다는 장점을 가진다 [14].

학습을 진행하는 데 사용된 파라미터는 컴퓨팅 파워에 맞추어 배치사이즈(batch size)는 2, 에폭(epoch) 300으로 설정해주었고, 컨볼루션 층을 구성하는 활성화 함수로는 연산 비용이 크지 않아

기존 활성화 함수보다 학습 속도를 빠르게 하는 Relu 함수를 적용했으며, 최적화 방법으로는 학습의 갱신강도를 스스로 조정함으로써 안정적으로 학습하게 하는 Adam 최적화를 사용하였다 [15]. 본 연구의 전체 흐름도는 [그림 2]와 같다.



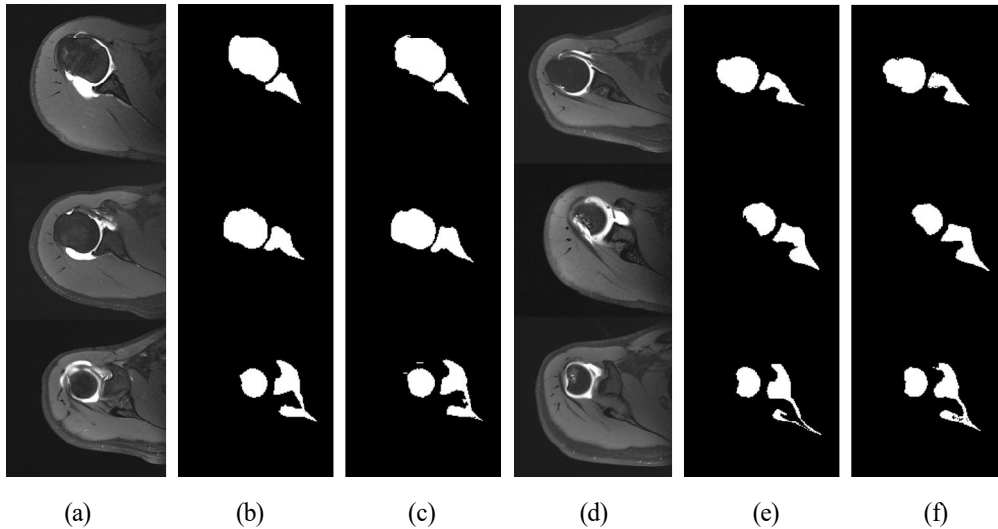
[그림 1] 상완골두와 관절와 분할 모델에 사용된 U-net 구조
 [Fig. 1] The Structure of U-net Used at Humeral head and Glenoid



[그림 2] 합성곱 신경망 기반의 상완골두와 관절와 분할 모델 플로우 차트
 [Fig. 2] The flowchart of humeral head and glenoid segmentation model based on convolution neural network

3. 결과

본 연구에서는 MR 영상을 이용하여 합성곱 신경망 기반의 U-net 상완골두와 관절와 분할 모델을 구축하였다. [그림 3]은 학습된 모델을 기반으로 테스트셋을 적용한 결과 이미지이다.



[그림 3] 자동 분할된 결과 이미지 예시 (a)(d) 원본 MR 영상 (b)(e) Ground truth 이미지 (c)(f) 딥러닝 모델로 예측된 분할 이미지

[Fig. 3] Examples of Images of automated segmentation result. (a)(d) Original MR images, (b)(e) Ground truth images, (c)(f) Predicted segmentation images by deep learning model

딥러닝 모델은 각 픽셀이 관심영역에 해당할 확률을 0~1 사이의 값으로 예측하고, 예측한 확률이 0.5를 초과하는 값은 1, 0.5 이하의 값은 0으로 나누게 된다. Ground truth 이미지는 관심영역 내부의 모든 픽셀이 1, 그밖의 값은 0으로 레이블링(labeling) 된다. 이렇게 모든 픽셀이 0과 1로 레이블링 된 예측 이미지와 ground truth 이미지의 각 픽셀을 비교하여 민감도와 특이도 및 정확도, 그리고 두 샘플의 유사성을 측정하는 다이스 계수로 평가하였다.

Ground truth 이미지에서 True는 값이 1인 픽셀, False는 값이 0인 픽셀이고, 예측 이미지에서 Positive는 값이 1인 픽셀, Negative는 값이 0인 픽셀을 의미한다. TP(True Positive)는 실제 True인 정답을 True라고 예측, FP(False Positive)는 실제 False인 정답을 True라고 예측, FN(False Negative)는 실제 True인 정답을 False라고 예측, TN(True Negative)는 실제 False인 정답을 False라고 예측하는 것을 말한다. 427장 데이터로 검증한 U-net으로 훈련된 어깨뼈 분할 모델의 평가 결과로 민감도 89.32%, 특이도 99.78%, 정확도 99.44% 다이스 계수 92.02%를 보였다.

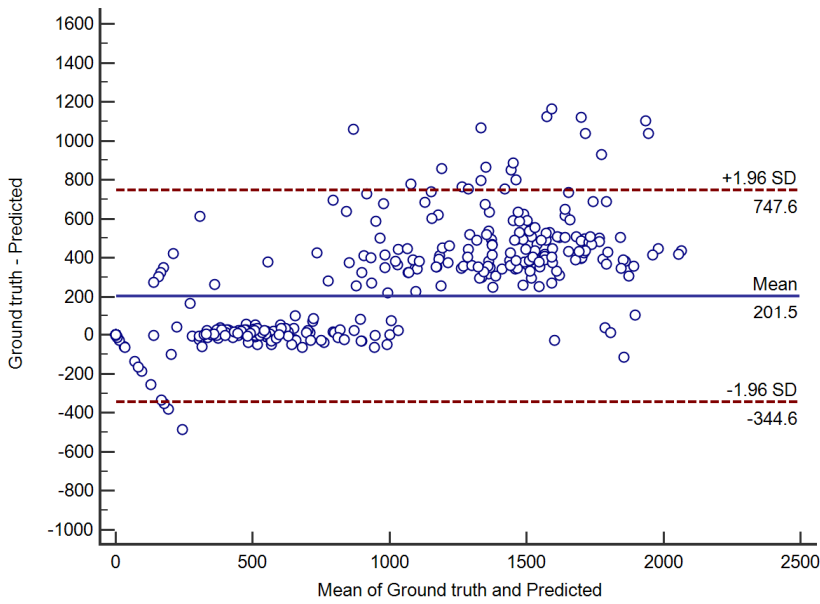
$$\text{민감도 (Sensitivity)} = \frac{TP}{TP+FN} \tag{1}$$

$$\text{특이도 (Specificity)} = \frac{TN}{FP+TN} \tag{2}$$

$$\text{정확도 (Accuracy)} = \frac{TP+TN}{T+P} \tag{3}$$

$$\text{다이스 계수 (Dice coefficient)} = \frac{TP*2}{T+P} \tag{4}$$

서로 다른 두 가지 데이터 간의 일치성을 분석하는 데 사용되는 방법인 Bland-altman 도표를 이용하여 결과를 확인하였다. 이는 ground truth 영역의 넓이와 예측 이미지의 넓이를 측정하여 그 차이의 정도를 확인하는 도표로서, 통계에서 실제 계측값과 추정값을 비교하는 데 많이 사용되는 방법이다. [그림 4]에서 x축은 ground truth 영역과 예측 이미지의 평균값이고, y축은 ground truth 영역의 넓이와 예측 이미지 영역 넓이의 차이값을 나타낸다. 아래 도표에서 y값의 평균은 201.5 이며, 일반적으로 표준편차(Standard Deviation) ±1.96 의 오차 범위를 허용한다.



[그림 4] Ground truth와 예측된 결과 이미지의 Bland-altman 도표

[Fig. 4] Bland-altman plot of Ground truth and predicted result image

4. 고찰

MRI는 어깨 질환 및 손상을 확인하기 위한 수단으로서 점차 방사능 노출 위험성이 있는 CT의 역할을 대체하고 있으며 진단 도구로 매우 유용하게 사용되고 있다. 한 환자당 얻어지는 수십장의 어깨 부위 MR 영상을 전문의가 일일이 확인하는 기존의 방법이 아닌 컴퓨터 보조 진단 기반의 자동 분할 연구에 대한 필요성이 있었다. 이에 본 논문에서는 어깨 질환 진단을 위한 합성곱 신경망 기반 어깨뼈 자동 분할 연구를 수행하였다.

앞서 언급한 독일 Heidelberg 연구팀의 CT 영상의 전체 뼈 분할 연구와 거의 비슷한 0.92의 다이슨 계수를 얻었지만, 본 연구에서는 골격의 구조가 CT 비교적 덜 뚜렷하게 보이지만 환자에게 방사선을 노출하지 않고도 검사가 가능한 MR 영상을 이용해 흡사한 결과를 냈다는 점을 고려하면 본 연구 결과가 임상에서 어깨 관련 질환 환자들에게 더 유용하다고 할 수 있다.

예측된 결과 이미지를 확인해보면 상완골두에서의 시작 지점에서 예측 정확도가 떨어짐을 확인할 수 있는데, 이는 상완골두의 시작지점이 연부 조직으로 덮여 있어 partial volume averaging에 의해 경계가 불분명하기 때문이다. 뿐만 아니라 상완골두에 비해 관절와 영역에서 분할 정확도가 낮았다. 상완골두의 경우 곤봉에 가까운 형태로, MR 축상 영상에서 경계가 비교적 뚜렷한 반면, 관절와는 상대적으로 복잡한 형태를 갖고 있기 때문에 다른 형태와 이어진 부분 중 관절와에 해당하는 부분만 분할하는 데 어려움이 있었다고 판단된다.

이러한 경계의 모호성을 해결하기 위한 방법으로 영상처리 기법을 적용하는 방안이 있다. 히스토그램 평활화(Histogram equalization)와 같은 기법을 적용하여 MR 영상에서의 대조도를 뚜렷하게 하거나, 주변 픽셀값들과 비교하여 레퍼런스 픽셀을 부각시키거나 약화시키는 샤프닝(Sharpening) 기법을 적용한다면 경계가 모호하여 분할이 잘 안되는 이미지에서의 정확도를 높일 수 있을 것이다.

또한 관절와 부위를 더욱 잘 분할하기 위해서는 더 많은 데이터를 수집하여 다양한 환자 케이스를 훈련한다면, 보다 높은 정확도를 얻을 수 있을 것이라 생각한다.

References

- [1] Y. S. Kim, J. H. Kim, "The Most Common Shoulder Disorders", *Journal of Rheumatic Diseases*, vol. 18, no. 1, May 2011, pp. 2-10, doi: 10.4078/jrd.2011.18.1.3.
- [2] A. J. P. Boeke, D. Windt, B. Koes, W. Deville, B. Jong, L. Bouter, "Shoulder disorders in general practice: prognostic indicators of outcome", *British Journal of General Practice*, vol. 46, no. 410, September 1995, pp. 519-523.

- [3] S. S. Leopold, R. Porcher, "Reporting Statistics in Abstracts in Clinical Orthopaedics and Related Research", *Clinical Orthopaedics and Related Research*[®], vol. 471, iss. 6, April 2013, pp. 1739-1740, doi: 10.1007/s11999-013-2984-9.
- [4] J. P. Iannotti, M. B. Zlatkin, J. L. Esterhai, H. Y. Kressel, M. K. Dalinka, "Magnetic resonance imaging of the shoulder", *J. Bone Jt. Surg.*, vol. 73, no. 1, January 1991, pp. 17-29.
- [5] V. S. Khooa, D. P. Dearnaley, D. J. Finniganb, A. Padhani, S. F. Tannerd, M. Leachd, "Magnetic resonance imaging (MRI): considerations and applications in radiotherapy treatment planning", *Radiotherapy and Oncology*, vol. 42, iss. 1, October 1997, pp. 1-15, doi: 10.1016/s0167-8140(96)01866-x.
- [6] M. B. Zlatkin, "MRI of the postoperative shoulder", *Skeletal Radiol*, vol. 31, January 2002, pp. 63-80, doi: 10.1007/s00256-001-0460-1.
- [7] H. Lee, Ko, Y. K. Kim, M. Kim, "The Effects of Medical Staffing Level on Length of Stay", *Journal of Korean Academy of Nursing Administration*, vol. 17, no. 3, September 2011, pp. 327-335.
- [8] J. G. Lee, S. H. Jun, Y. W. Cho, H. Lee, G. B. Kim, J. B. Seo, N. Kim, "Deep learning in medical imaging: General overview", *Korean Journal of Radiology*, vol. 18, no. 4, August 2017, pp. 570-584, doi: 10.3348/kjr.2017.18.4.570.
- [9] K. He, X. Zhang, S. Ren, J. Sun, "Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification", *IEEE International Conference on Computer Vision (ICCV)*, December 7-13, 2015, Santiago, Chile, doi: 10.1109/ICCV.2015.123.
- [10] A. Klein, J. Warszawski, and J. Hillengaß, "Automatic bone segmentation in whole-body CT images", *International Journal of Computer Assisted Radiology and Surgery*, vol. 14, no. 1, February 2019, pp. 21-29, doi: 10.1007/s11548-018-1883-7.
- [11] M. Havaei, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, C. Pal, P. M. Jodoin, H. Larochelle, "Brain tumor segmentation with Deep Neural Networks", *Medical Image Analysis*, vol. 35, January 2017, pp. 18-31, doi: 10.1016/j.media.2016.05.004.
- [12] A. E. Silver, M. P. Lungren, M. E. Johnson, S. W. O. Driscoll, K. An, R. E. Hughes, "Using support vector machines to optimally classify rotator cuff strength data and quantify post-operative strength in rotator cuff tear patients", vol. 39, October 2006, pp. 973-979, doi: 10.1016/j.jbiomech.2005.01.011.
- [13] O. Ronneberger, P. Fischer, T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation", *MICCAI*, October 5-9, 2015, Munich, Germany, pp. 1-8, doi: 10.1007/978-3-319-24574-4_28.
- [14] T. Falk, D. Mai, R. Bensch, Ö. Çiçek, A. Abdulkadir, Y. Marrakchi, A. Böhm, J. Deubner, Z. Jäckel, K. Seiwald, A. Dovzhenko, O. Tietz, C. D. Bosco, S. Walsh, D. Saltukoglu, T. L. Tay, M. Pr, "U-Net: deep learning for cell counting, detection, and morphometry", *Nature Methods*, vol. 16, January 2019. pp. 67-70, doi: 10.1038/s41592-018-0261-2.
- [15] D. P. Kingma, J. L. Ba, "Adam: A method for stochastic optimization," *International Conference for Learning Representations*, May 7-9, 2015, San Diego, CA, USA, pp. 1-15.