



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

의학박사 학위논문

인공지능 심층신경망을 이용한
코뼈 골절의 진단

Nasal bone fracture detection
utilizing a convolutional neural network

울산대학교 대학원
의 학 과
이 호 준

인공지능 심층신경망을 이용한
코뼈 골절의 진단

지도교수 장 용 주

이 논문을 의학박사 학위 논문으로 제출함

2022년 02월

울산대학교 대학원
의 학 과
이 호 준

이호준의 의학박사학위 논문을 인준함

심사위원장	최 승 호	인
심사위원	장 용 주	인
심사위원	정 유 삼	인
심사위원	김 지 희	인
심사위원	조 범 주	인

울 산 대 학 교 대 학 원

2022 년 02 월

<국문요약>

서론 : 코는 얼굴의 중심부에 위치하고 있으며 얼굴에서 외상의 빈도가 가장 높은 부위이다. 코뼈 골절의 진단은 전산화 단층 촬영 (CT)을 이용하는 것이 가장 정확도가 높지만, 모든 병의원 및 응급실에 CT 시설이 갖추어져 있지는 않기 때문에 단순 엑스선 촬영도 진단에 많이 이용되고 있다. 코뼈 주위에는 여러 구조물들이 겹쳐져 있고 혈관과 신경이 지나다니는 통로들이 존재하기 때문에 단순 엑스선 촬영만으로 경험이 부족한 비 전문과 의사들이 진단하기가 어려운 경우가 흔히 있다. 제한된 숫자의 해당과 전문의가 즉시 모든 결과를 판독하는 것도 현실적으로 어려운 상황이다. 이러한 경우 해당과 전문의의 수준으로 학습된 심층 신경망을 개발해 임상에 적용한다면 의사들의 업무량을 줄일 수 있고, 오진으로 인해 발생할 수 있는 문제들을 방지하여 환자들에게 적절한 시기에 치료를 받게할 수 있을 것으로 생각된다. 본 연구에서는 두부의 단순 엑스선 영상만으로 자동으로 1) 코뼈 골절 여부를 진단하고, 2) 골절 부위로 의심되는 곳을 표시해 줄 수 있는 복합적 딥러닝 모델을 개발하고자 하였다.

본론 : 총 2,312 장의 두부 측면 단순 엑스선 영상을 확보하였다 (한 사람 당 2장의 두부 엑스선 영상, 비골절 1,468장, 골절 844장). 두부 측면 엑스선에서 코뼈 부위만을 기계 학습에 이용하기 위해 코뼈의 위치를 라벨링 하였으며, 코뼈의 골절 여부를 라벨링 하였다. 데이터 증강 기법을 이용하여 데이터셋을 구축하였다. 본 연구의 목적을 달성하기 위하여 다음과 같은 4가지의 모델을 개발하였다. 1) 두부 엑스선 이미지로부터 코뼈의 영역을 자동 탐지하는 모델 (코뼈탐지모델, You Only Look Once version 5, YOLOv5), 2) 코뼈 골절 여부를 판단하는 모델 (Classifier, C 모델, EfficientNetB3), 3) 골절 엑스선 이미지로부터 비골절 엑스선 이미지를 생성하는 모델 (Generator, G모델, U-Net), 4) G 모델이 생성한 가상의 이미지가 자연스러운 엑스선 이미지 인지를 판별하는 모델 (Discriminator, D 모델, Convolutional Neural Network). 그리고 이 4가지의 모델을 복합적으로 사용하는 전체 딥러닝 모델을 구축하였다. 코뼈탐지모델의 결과로 두부 엑스선 이미지로부터 코뼈 영역만을 자동으로 추출할 수 있게 되었고, 추출된 영상은 일종의 전처리 과정으로서 다른 모델의 학습에 이용되었다. C 모델의 결과로 코뼈 단순 엑스선 영상으로부터 95%의 골절 여부 판단 정확도를 달성하였다. G-D-C 연결 모델의 결과로 코뼈 골절 의심 부위를 자동으로 예측하여 표시할 수 있는 모델을 개발하였다.

결론 : 상기 복합적 딥러닝 모델은 단순 엑스선 두부 영상에서 코뼈 부위를 탐지하여 코뼈 골절 여부를 비교적 정확하게 진단할 수 있었으며, 골절 부위를 탐지해 낼

수 있었다. 이를 이용하여 작은 규모의 병·의원, 군 야전 진료소, 원격의료, 수련의 교육 등에 활용될 수 있으며, 본 연구의 프레임워크를 확장함으로써 다양한 부위의 골절 진단에 활용이 가능할 것으로 생각된다.

<목차>

<국문요약>	i
I. 서론	1
II. 대상 및 방법	2
1. 데이터의 수집	2
2. 데이터의 분류 및 라벨링	2
3. 모델 훈련, 내부 검증, 평가 데이터셋 및 외부 검증 데이터셋	3
4. 모델	3
5. 딥러닝 모델 학습의 상세	4
5.1 모델 학습 단계의 요약	4
5.2 코뼈탐지 모델 학습 (두부 영상으로부터 코뼈 영역을 추출)	4
5.3 C모델 (골절여부 판정 모델) 학습	5
5.4 G-D모델 학습(코뼈 사진 생성)	5
5.5 G-D-C모델 학습 step 1 (비골 골절 영역 탐지)	6
5.6 G-D-C모델 학습step 2(골절 이미지로부터 비골절 이미지 생성)	7
5.7 실제 모델의 학습 환경 및 학습 소요 시간	8
III. 결과	8
IV. 고찰	10
1. 연구결과의 의의	10
2. 영역분할의 레이블링 없이 영역 탐지	10
3. 손실 함수로서의 모델	10
4. 생성된 영상에 대한 고찰	10
5. 향후 연구 방향 제언	11

V. 결론	11
<참고문헌>	12
<Abstract>	13

<그림 목차>

그림 1	수집된 엑스선 영상을 골절 (첫 행) 과 비골절 (둘째 행)로 라벨링	15
그림 2	코뼈의 위치를 라벨링	16
그림 3	골절 여부 판정 모델 (C 모델)의 기반인 EfficientNet의 구조	17
그림 4	비골 이미지 생성 모델 (G 모델)의 기반인 Skip-UNet의 구조	18
그림 5	생성 이미지 판별 모델 (D 모델)의 구조	19
그림 6	복합 딥러닝 모델의 전체 구조	20
그림 7	코뼈 탐지 모델의 학습	21
그림 8	C 모델의 학습	22
그림 9	G-D 모델의 학습	23
그림 10	G-D-C 모델의 학습	24
그림 11	Epoch에 따른 코뼈탐지모델의 손실 저하	25
그림 12	코뼈탐지 모델의 출력으로 코뼈 영역을 탐지한 예시	26
그림 13	학습 epoch에 따른 학습 정확도, AUC, 손실의 변화 그래프	27
그림 14	G-D 모델 학습에서 각 모델의 손실 변화 그래프	28
그림 15	G 모델의 입력 입력이미지 (첫 행)과 모델이 생성 이미지 (둘째 행)	29
그림 16	G-D-C step 1 모델의 학습에 따른 손실 변화 그래프	30
그림 17	골절 상태의 코뼈 이미지와 G-D-C 모델이 예측한 골절 의심 부위	31
그림 18	G-D-C step 2 모델 학습에 따른 손실 변화 그래프	32
그림 19	골절 상태의 엑스선과 (1열) G-D-C step 2 모델이 예측한 골절 부위	33
그림 20	인간이 눈으로 관측하는 것과 (판다) 모델이 높은 신뢰도로 판단하는 결과 (긴팔원숭이) 의 차이를 보이는 예제	34

I. 서론

코뼈는 얼굴의 중심부에 위치하고 있으며, 가장 앞으로 튀어나온 부위이기 때문에 얼굴에서 외상의 빈도가 가장 높은 부위이다. 외상으로 인해서 코뼈가 명백히 부러졌더라 하더라도, 수상 초기에는 부종 등으로 코 모양의 변형이 명확하지 않을 수 있다. 대부분의 코뼈 골절은 폐쇄 정복술을 이용하여 비교적 간단히 치료가 가능하지만, 적절한 시기에 치료를 받지 못한다면 미용적인 문제나 코막힘 등의 기능적 문제를 야기할 수 있다. 그러므로 수상 후 초기에 정확한 진단을 하고, 적절한 시기에 치료를 받아야 향후의 합병증을 예방할 수 있다.

현대 의학, 특히 외상 분야의 현대 의학은 진단 방사선 검사에 많은 부분을 의존하고 있으며, 이에 따라 방사선 촬영은 매우 방대하게 이루어지고 있다. 그러나 이를 제한된 숫자의 영상의학 전문의가 즉시 모두 판독하는 것은 현실적으로 어려운 상황이며, 진단이 제대로 이루어지지 않아 치료가 지연되는 경우가 드물지 않다¹. 코뼈 골절의 진단은 전산화 단층 촬영 (CT)을 이용하는 것이 정확도가 가장 높지만, 모든 병의원 및 응급실에 CT 시설이 갖추어져 있지는 않기 때문에 단순 엑스선 촬영으로 초기진단을 하는 경우가 흔하다. 코뼈는 주위에 여러 뼈와 봉합선으로 겹쳐져 있고, 신경 혈관들이 지나다니는 통로들이 있어서 골절선과 혼돈을 일으키는 경우가 있기 때문에 단순 엑스선 촬영만으로 골절을 진단하기 어려운 경우가 있으며 해당 전문과 및 영상의학 전문의들도 진단하기 애매한 경우가 흔하다¹.

역사적으로 오래전부터 컴퓨터의 보조를 이용해 이미지의 특정 영역을 발견하려는 노력이 있어왔다. 하지만 그동안 정확도가 떨어지고, 개선 속도가 느려서 이를 의학적으로 적용하기는 어려웠다. 최근 컴퓨팅 기술과 인공 지능 분야의 발전에 따라 디지털 영상을 정해진 기준에 따라 자동적으로 분류하는 것이 가능하게 되었다^{2,3,4}. 인공지능의 한 분야인 딥러닝이 출현하면서 컴퓨터 보조 탐지 (computer assisted detection, CAD)에 매우 높은 발전이 이루어졌고, 이를 의학에 적용하려는 시도가 일어나고 있다. 딥러닝을 이용한 이상영역 탐지는 기존의 CAD와는 완전히 다른 방식을 이용한다. 업무 특이적 (task specific) 이미지 모양을 지속적으로 학습하여 해석하는 작업이다. 이런 학습을 위해서 기존의 연구들은 대부분 사람이 일일이 라벨링을 한 자료를 바탕으로 지도학습을 이용한 것이었다. 특정 업무의 전문가가 대량의 이미지 데이터셋을 세심하게 라벨링하여 정상과 비정상을 나누어 학습시키는 과정이 필요했다. 전문가적 라벨링을 하여 분류한 충분한 양의 이미지를 컴퓨터가 학습하고 나면 컴퓨터는 그 라벨링을 한 전문가의 정확도로 판단을 내릴 수 있게 된다. 단순 엑스선 영상을 이용한 골절의 진단^{3,4,5}, 피부 병변과 망막 병변의 분류 등에 대한 연구 결과가 발표되고 있다.

국내외에서 코뼈 골절 진단에 딥러닝을 적용한 연구는 아직 없으며, 기존 연구들에서는 골절부위를 표시하기 위해서 해당과 전문의가 일일이 골절 의심부위에 표시를 해두는 '라벨링'을 하여야 했다. 그러나 이러한 방법은 학습을 위한 데이터셋의 양

이 엄청나게 많을 경우, 전문 인력의 많은 노동량이 필요하며, 해당과 전문의조차도 골절 의심부위가 어디인지 표시를 하기 모호한 상황일 때는 유용한 방법이 아닐 수 있다. 본 연구에서는 두부 엑스선 영상으로부터 코뼈의 이미지만을 추출하고, 코뼈의 골절 여부를 진단하며, 이상 탐지 기술을 이용한 비지도 학습 방법으로 골절 부위 표시를 할 수 있는 복합 모델을 개발하고 이 모델의 성능을 평가하고자 하였다. 본 연구는 적대적 생성 학습 기법 (Generative Adversarial network, GAN)의 판별자 (discriminator)를 사용하여 생성자 (generator)가 진짜와 차이 나지 않는 가짜 영상을 생성할 수 있는 점에 착안한 연구 설계이다. GAN과 유사하지만, 본 연구에서는 별개의 골절 여부 분류자 (classifier)를 함께 사용하여 비골절 영상과 차이 나지 않는 영상을 생성하도록 하였다. 판별자를 사용하여 실제 엑스선 영상과 유사한 영상인 동시에 분류자가 비골절로 판단할 수 있는 영상을 생성할 수 있다면, 입력에 사용된 골절 영상에 대하여 비골절인 영상을 생성할 수 있지 않을까 확인하는 것이 본 연구의 핵심 아이디어이다. 이런 방법으로 골절 영상에 대한 비골절 영상을 구할 수 있다면 두 영상을 비교하여 골절 부위 탐지에도 사용할 수 있을 것으로 기대하였다.

II. 대상 및 방법

1. 데이터의 수집

한림대학교 의료원 소속의 두 병원 (한림대학교성심병원, 춘천성심병원)에서 단순 두부 엑스선 영상을 수집하였다. 2007년 1월 1일부터 2019년 12월 31일 까지 코의 외상으로 상기 병원들에 내원하여 단순 두부 측면 엑스선 영상 및 안면골 CT를 모두 촬영한 환자들이 대상이었고, 후향적으로 단순 두부 측면 엑스선 영상, CT 영상 및 의무기록을 검토하였다. 단순 두부 측면 엑스선 영상은 연구에 연루된 두 병원의 picture archiving and communication system (PACS)에서 512x512 픽셀의 회색 스케일 이미지로 수집되었다. 수집된 엑스선 영상은 다양한 엑스선 장비에서 촬영된 것이었다. 환자 관련 개인 정보는 엑스선 이미지 추출시에 포함되지 않았다. 대상의 제외 기준은 다음과 같다. (1) 코뼈 골절 이외에도 다발성 안면골절이 존재하는 환자, (2) 기존에 코뼈 골절의 병력이 있는 환자 (3) 코뼈 골절로 정복술을 시행한 이후에 엑스선 촬영을 시행한 환자.

최종적으로 1156 명의 환자로부터 2312장의 두부 엑스선 이미지 (비골절 : 1468장, 골절 : 844장) 를 수집하여 모델 훈련, 검증, 평가에 사용하였다. 본 연구는 한림대학교 춘천성심병원 임상연구윤리위원회의 승인을 받은 연구이며 (승인번호 : CHUNCHEON 2021-05-010-001) 헬싱키 선언에 따라 수행되었다.

2. 데이터의 분류 및 라벨링

복합적 딥러닝 모델을 개발하기 위하여 수집된 두부 엑스선 영상에 두가지의 라벨

링을 시행하였다. 첫번째, 수집된 두부 단순 엑스선 영상을 코뼈 골절 이미지와 비골절 이미지로 분류하였다. 이 라벨링은 1명의 영상의학과 전문의와 1명의 이비인후과 전문의가 의무기록 상의 신체검사 소견, 두부 단순 엑스선 영상, 안면골 CT를 후향적으로 모두 검토하여 시행하였으며 두 의견이 일치하는 이미지만을 골절 혹은 비골절 데이터셋으로 분류하였다 (그림 1).

두번째, 코뼈의 영역을 표시하기 위하여 코뼈의 시작점 (비근점)과 코뼈가 끝나는 지점 (비골첨)을 두 점으로 표시하였다 (그림 2).

3. 모델 훈련, 검증, 평가 데이터셋

모델 훈련, 검증, 평가 데이터셋은 8:1:1 로 하였다. 학습을 위하여 좌우 반전, 상하 좌우 이동 (최대 5%), 회전 (최대 20도), 축소 확대 (5%)를 이용한 데이터 증강 기법을 사용하였다. 학습의 손실함수는 각각의 모델들마다 다르게 설정하였으며 아래에 기술하였다. 모델들의 손실값들을 모니터링하여 더 이상 내려가지 않고 평탄선이 지속될 경우 수렴에 이르렀다고 판단하였다.

4. 모델

연구 목표 달성을 위하여 다음의 4가지의 모델들을 구축하였다. Python 3.6 프로그래밍 언어를 사용하였다.

(1) 두부 엑스선 이미지로부터 코뼈의 영역을 자동 탐지하는 모델을 구축하였다 (코뼈탐지모델). 이미지에서 특정 물체의 위치를 검출하기 위해 사용되는 딥러닝 모델인 YOLO (You Only Look Once)의 최신 버전인 YOLO version 5 (YOLOv5) 모델을 코뼈 탐지에 이용하였다. 이 모델은 현재 논문으로 발표되지는 않은 상태이며 프로그램 코드로만 공개되어 있다⁶. PyTorch 1.7를 사용하였다.

(2) 코뼈 골절 여부를 판단하는 모델 (classifier, C 모델)을 구축하였다. EfficientNetB3 모델을 사용하였다⁷ (그림 3). TensorFlow 2.5.0을 사용하였다.

(3) 골절 엑스선 이미지로부터 비골절 엑스선 이미지를 생성하는 모델 (Generator, G모델)을 구축하였다. U-Net 모델을 사용하였다⁸ (그림 4). TensorFlow 2.5.0을 사용하였다.

(4) G모델이 생성한 가상의 이미지가 자연스러운 엑스선 이미지인지를 판별하는 모델 (Discriminator, D 모델)을 구축하였다. CNN (Convolutional Neural Network)을 사용하였다 (그림 5). TensorFlow 2.5.0을 사용하였다.

상기의 4가지 모델들을 이용하여 복합적 딥러닝 모델을 구성하였다 (그림 6).

5. 딥러닝 모델 학습의 상세

5.1 모델 학습 단계의 요약

- 1) 코뼈탐지 모델(object detection)모델을 학습하였다. 두부 영상에서 코뼈 부분만을 추출, 일종의 전처리로 활용하였다.
- 2) 골절 여부 판정 (Classification)모델 학습하였다. 주어진 코뼈 이미지로부터 골절 여부를 판정하였다.
- 3) 코뼈 이미지 생성(Generation)모델과 코뼈 이미지 적합도 판정(Discriminative)모델을 동시 학습하였다. 골절 및 비골절 데이터를 기반으로 딥러닝 모델이 코뼈 이미지를 생성하고(G모델), G모델이 생성한 이미지가 자연스러운 엑스선 영상인지를 판단(D모델) 하도록 하였다. 적대적 생성 학습 기법(Generative Adversarial learning, GAN)을 활용하였다.
- 4) G-D-C모델 학습 1단계를 시행하였다. 앞서 구축한 C모델을 기반으로 G모델을 학습시킴으로써 골절 영역을 탐지하도록 유도하였다.
- 5) G-D-C모델 학습 2단계를 시행하였다. 앞선 단계의 골절 영역 탐지 결과를 활용하여 새로운 G모델을 구축하여 비골절 이미지를 생성하였다.

5.2. 코뼈 탐지모델 학습 (두부 영상으로부터 코뼈 영역을 추출)

코뼈탐지모델은 그림 7을 타겟으로 모델 학습을 진행하였다. 이 학습 과정을 통해서 모델은 다양한 상황에서 촬영된 두부 방사선 이미지를 입력으로 받아 코뼈의 부위를 특정할 수 있다. 본 연구에서 코뼈 탐지 모델은 이후의 모델들과는 별개의 모델로서 학습되었으며 코드상으로 분리되어 있다. 이 모델의 출력은 이후 모델들에 있어 일종의 이미지 전처리 과정으로써 사용되었다.

모델 학습과 관련된 상세 파라미터는 다음과 같다.

- 입력 : 512x512 사이즈의 흑백 영상
- 각 픽셀은 0~255의 값
- 총 1715개 (골절 900개, 비골절 815개) 중 데이터 비율;
학습 : 검증 : 평가 = 1371 : 172 : 172 의 비율
- 출력: 코뼈 부분의 상자 좌표
- 데이터 증강 기법: 이미지당 밝기 +-25%, salt and pepper 노이즈 5%,
위치이동 최대 25%
- Optimizer: Adam
- 학습률: 0.01
- 학습 epoch: 100 epoch

- 배치 사이즈: 16

5.3. C모델 (골절여부 판정 모델) 학습

골절 여부 판정 모델은 그림 8을 타겟으로 모델 학습을 진행하였다. 이 학습 과정을 통해서 코뼈 엑스선 영상을 보고 골절 여부를 판정할 수 있게 된다. 앞선 코뼈 탐지모델과 결합하면 일반적인 두부 엑스선 이미지로부터 코뼈 골절 여부를 판정하는 것이 가능하다.

모델 학습과 관련된 상세 파라미터는 다음과 같다.

- 입력: 256x256x3 사이즈의 비골 부위 흑백 영상. 비골 탐지 모델에서 출력된 비골 부위를 리사이징하였으며, RGB 3채널의 분류자 모델을 사용하기 위해 256x256x1의 1채널 영상을 RGB 3채널로 변환하여 입력하였다.
- 0~255의 값
- 총 1400개 (골절 700개, 비골절 700개) 중 데이터 비율;
학습 : 검증 : 평가 = 1120 : 140 : 140 의 비율
- 출력: 골절 여부 (비골절 = 0, 골절 = 1)
- 데이터 증강 기법: 좌우 반전, 상하좌우 이동(최대 5%), 회전(최대 54도), 대비값 증감(최대 5%), 축소확대(최대 5%)
- 학습 손실 함수: Binary Cross Entropy (BCE)
- Optimizer: Adadelta optimizer
- 학습률: 0.001
- 학습 epoch: 2000 epoch
- 배치 사이즈: 64

5.4. G-D모델 학습(코뼈 사진 생성)

코뼈 사진 생성 학습은 그림 9를 타겟으로 모델 학습을 진행하였다. 이 학습 과정을 통해서 코뼈 이미지를 생성하고 (G모델), 생성된 이미지가 코뼈 엑스선 이미지로 적합한지 판단 (D모델)할 수 있게 되었다. 적대적 생성 학습 기법(Generative Adversarial learning)을 활용하였다.

모델 학습과 관련된 상세 파라미터는 다음과 같다.

- 입력 : 256x256x1 사이즈의 코뼈 부위 흑백 영상
- 코뼈 탐지 모델에서 출력된 코뼈 부위를 리사이징, 0~1로 normalized
- 총 2000개 (골절 1300개, 비골절 700개) 중 데이터 비율;
학습 : 검증 : 평가 = 1600 : 200 : 200 의 비율
- 출력: 코뼈 이미지(G모델), 생성된 코뼈 이미지 적합도 판정치(D모델)
- 데이터 증강 기법: 좌우 반전, 상하좌우 이동(최대 5%), 회전(최대 20도),

- 축소확대(최대 5%)
- 학습 손실 함수. lat(latency) Dfeat : 출력 전 마지막 값
 - G모델 loss: $L_G = L_{adv,G} + 40L_{con} + L_{lat}$
 - D모델 loss: $L_D = L_{adv,D} + L_{lat}$
 - $L_{adv,G} = BCE(y, D(G(X)))$: 입력 영상과 생성 영상의 D출력 (진위 여부)이 유사하도록 함.
 - $L_{con} = |X - G(X)|$: 입력과 생성된 이미지의 차이를 최소로 해야 함
 - $L_{lat} = MSE(D_{feat}(X), D_{feat}(G(X)))$: D출력을 같게 하기 위해 최종 레이어 값을 유사하게 함.
 - $L_{adv,D} = BCE(y_{natural}, D(X)) + BCE(y_{unnatural}, D(G(X)))$
 - X : 입력 이미지, Y : 출력 레이블, D_{feat} : D모델의 FC layer 출력치
 - BCE : Binary Cross Entropy, MSE : Mean Squared Error
 - $D(X)$: D모델의 출력, $G(X)$: G모델의 출력
- 학습률: 0.0001
- 학습 epoch: 80 epoch
- 배치 사이즈: 32

5.5. G-D-C모델 학습 step 1 (코뼈 골절 영역 탐지)

G-D-C모델의 첫 번째 학습은 코뼈 골절 영역을 탐지하기 위해 진행되었다. 그림 10을 타겟으로 모델 학습을 진행하며, 본 학습 과정을 통해 모델은 골절 이미지로부터 골절 의심 영역을 추정할 수 있게 되었다.

모델 학습과 관련된 상세 파라미터는 다음과 같다.

- D, C모델을 고정하고 G모델을 학습한다. C를 이용하여 masking 함수를 구성하였다. G의 입출력을 비교하여 mask를 생성하였다.
- 입력: 256x256x1 사이즈의 코뼈 부위 흑백 영상
- 코뼈 탐지 모델에서 출력된 코뼈 부위를 리사이징
- 0~1로 normalize
- 데이터 구성: 골절 데이터 700개
- 출력: 비골 이미지(G모델), 골절 예측 영역의 이미지가 변형되어 출력됨.
- 데이터 증강 기법: 좌우 반전, 상하좌우 이동(최대 5%), 회전(최대 20도), 축소확대(최대 5%)
- 학습 손실 함수
 - $L = 80L_G + L_D + 4L_C$
 - $L_G = |M(X)X - M(X)G(X)|$: 원 이미지와 생성된 이미지의 차이를 최소로 해야 함
 - $L_D = BCE(y_{real}, D(G(Adjust(X))))$: D모델이 실제 비골 이미지로 판단한 정

도

- $L_C = BCE(y_{normal}, C(Adjust(X)))$: C모델이 골절 여부를 정확히 판단한 정도
 - $Adjust(X) = XM(X) + (1-M(X))G(X)$: 실제 골절 영역과 관계 없는 곳이 변경되어 L_C 가 작아지는 것을 방지. 골절 영역만 변경이 되도록 유도한다.
 - $M(X)$: C모델 마지막 conv layer의 출력치. 값이 큰 부분이 골절 영역, 값이 작은 부분이 배경 영역.
- Optimizer: Adam
 - 학습률: 0.0002
 - 학습 epoch: 200 epoch
 - 배치 사이즈: 32

5.6. G-D-C모델 학습 step 2(골절 이미지로부터 비골절 이미지 생성)

G-D-C모델의 두 번째 학습은 골절 이미지로부터 비골절 이미지를 생성하기 위해 진행되었다. 그림 10을 타겟으로 모델이 학습되며, 첫 번째 학습과 사용하는 목적 함수를 달리하였다. 이 학습 과정을 통해 비골절 이미지를 출력할 수 있었다.

모델 학습과 관련된 상세 파라미터는 다음과 같다.

- D, C모델을 고정하고 G모델을 학습한다. G-D-C step 1의 G 모델로부터 mask 함수를 생성한다.
- 입력 : 256x256x1 사이즈의 비골 부위 흑백 영상
- 비골 탐지 모델에서 출력된 비골 부위를 리사이징
- 0~1로 normalized
- 데이터 구성: 골절 데이터 700개
- 학습 손실 함수 (목적은 (4)와 유사)
 - $L = 80L_G + L_D + 4L_C$
 - $L_G = | XM(X) - M(X)G(X) |$
 - $L_D = BCE(y_{real}, D(G(Adjust(X))))$
 - $L_C = BCE(y_{normal}, C(Adjust(X)))$
 - $Adjust(X) = XM(X) + (1-M(X))G(X)$
 - $M(X)$: (4)의 G모델에서 유도. $| G(X)-X | > 0.06$ 인 위치를 1로, 이외를 0으로 설정. 골절 영역은 입력이 많이 변형되어 1이고, 배경 부분은 그대로라 0이 된다.
- Optimizer: Adam
- 학습률: 0.0002
 - 학습 epoch: 400 epoch
 - 배치 사이즈: 32

5.7. 실제 모델의 학습 환경 및 학습 소요 시간

- 학습 환경 : Nvidia Tesla T4 GPU, RAM 64GB
- C모델 학습 소요 시간: 8시간
- G-D모델 학습 소요 시간: 1시간
- G-D-C모델 학습 1단계 소요 시간: 1시간 45분
- G-D-C모델 학습 2단계 소요 시간: 3시간 35분

Ⅲ. 결과

1. 코뼈 탐지모델의 결과

두부 엑스선 영상으로부터 코뼈 영역을 탐지하는 모델을 구축하였다. 학습과정 그래프를 살펴보면 학습과정에 따라 과적합 없이 손실값이 적절하게 낮아져 적절한 모델학습이 이루어졌음을 확인하였다 (그림 11). 학습된 코뼈영역 탐지모델에 기존 두부 방사선 이미지를 넣으면 코뼈 영역을 정확히 탐지함을 확인하였다 (그림 12). 해당 결과는 이후 모델들에 있어 일종의 이미지 전처리 과정으로써 사용되었다.

2. 골절 여부 판정 모델의 결과

비골의 골절 여부를 판단하는 C 모델의 골절 예측 정확도는 95% (SD 0.02%), AUC는 0.99 (SD 0.003)의 우수한 성능을 달성하였다 (그림 13). 모델 평가 결과, 골절 증례에 대해서는 84례 중 75례를 정확히 판단하였으며, 비골절 증례에 대해서는 84례 중 81례를 정확히 판단하였다.

3. 골절 이미지로부터 비골절 이미지 생성 모델의 결과

코뼈 이미지를 생성하는 모델을 구축하였다. G-D모델의 적대적 학습 과정에서 두 모델 모두 적절한 loss로 수렴하여 모델의 학습이 올바르게 이루어졌음을 확인하였다 (그림 14). 그러나 학습 과정의 그래프와는 별개로 인간의 육안으로 관찰되는 코뼈 생성 이미지는 코뼈의 비골절 상태를 잘 복원해 내지 못한 것으로 보였다 (그림 15).

4. 코뼈 골절 영역 탐지 모델의 결과

코뼈 골절시 골절 의심부위 표시 모델 (G-D-C step 1)은 C모델, D모델을 고정한 채로 G모델을 새로 학습하며 골절 의심 부위를 예측하게끔 학습시켰다. G, D, C모델 모두 학습에 따라 loss가 적절히 감소하며 수렴하여 올바른 학습이 진행됨을 확인하였다 (그림 16). 모델 출력 결과를 보면 골절 상태의 비골 이미지를 기반으로 G모델이 골절 의심 부위를 적절하게 예측함을 확인하였다 (그림 17).

5. 코뼈 비골절 이미지 생성 모델

코빼 골절 시 비골절 이미지 생성 모델 (G-D-C step 2)은 G-D-C 모델의 새로운 학습을 통해 골절 상태의 이미지로부터 비골절 상태의 이미지를 추측하도록 학습을 진행하였다. G, D, C 모델 모두 학습에 따라 손실이 감소하고 수렴함을 확인하였다 (그림 18). 모델 출력 결과를 보면 골절 상태의 코빼 이미지를 기반으로 비골절 이미지를 적절히 추측하고 골절 부위를 예측함을 확인하였다. 그러나 골절 부위를 자세히 보았을 때 비골절 이미지 생성 부위가 노이즈의 형태로 나타남을 확인할 수 있다 (그림 19).

IV. 고찰

1. 연구결과의 의의

딥러닝 모델들을 기반으로 하여 입력된 두부 단순 방사선 영상으로부터 (1) 코빼 부위를 탐지하고, (2) 골절 여부를 판단하며, (3) 의심 골절 부위를 복구하는 방식으로 골절부를 표시하는 AI 시스템을 구축하였다. 딥러닝 모델들을 이용한 전체 모델을 통해 1) 두부 엑스선 이미지로부터 코빼골절 여부를 판정할 수 있고, 2) 골절/비골절에 대한 레이블 데이터만으로 골절 부위를 탐지해낼 수 있었다. 모델 학습 시 서로 다른 개별 목적으로 학습시킨 모델을 손실 함수로서 사용하였다. 본 연구에 사용된 방법은 기존의 보고를 찾기 어려운 새로운 시도로, 상기 기술한 방법을 골절 탐지에 사용할 수 있는 가능성을 제시했다는 의의도 있다고 생각된다.

2. 영역분할의 레이블링 없이 영역 탐지

영역 분할(segmentation)을 딥러닝으로 구현하기 위해서는 영역에 대한 레이블링 데이터가 필요하다. 이는 단순한 분류작업의 레이블링에 비해 많은 시간을 필요로 하고 전문가급의 인력이 필요하게 된다. 본 연구에서는 골절 부위를 탐지해내는 일종의 영역 분할 결과를 얻으면서도, 이에 대한 레이블링 작업은 영역의 레이블링이 아닌 단순 골절 여부를 나타내는 레이블링만을 사용하였다. 즉, 본 연구의 방법론을 적용하면 영역 레이블링이 아닌 골절 여부에 대한 레이블링만으로도 영역 분할의 효과를 얻을 수 있다.

3. 손실 함수로서의 모델

일반적인 딥러닝 모델 학습은 목적으로 하는 바를 수학적 함수로 정의한 손실 함수를 사용한다. 그러나 추상적인 목적에 대해서는 이를 수학적으로 정의하기가 어렵다. 예를 들어, 현실에서 보이는 것과 같은 자연스러운 영상의 생성이 목적인 경우 이를 수학적으로 정의하는 것은 거의 불가능하다. 딥러닝에서는 구체적인 계산 관계로 정의된 손실함수를 최소화하는 방향으로 학습이 진행되는데, 이 경우에는 수학적으로 정의된 함수 이외에도 별도로 구현된 입출력 함수이기만 하면 된다. 이를 활용하여 본 연구에서는 GAN의 discriminator가 수집된 데이터의 패턴을 학습시

키고 다양한 모델에 손실 함수로서 이용되었다. 이 D모델이 학습한 정보는 '얼마나 엑스선 영상스러운지'에 대한 정보이며, 이를 수학적으로 정의하기는 어렵지만 GAN을 활용함으로써 이를 손쉽게 정의할 수 있었다.

4. 생성된 영상에 대한 고찰

본 연구를 시작할 당시의 초기 아이디어는 골절/비골절 여부를 판정하는 C모델과 자연스러움을 판정하는 D모델을 활용하여 골절 입력 영상에 대한 비골절 영상을 생성하는 것이 목표였다. 이러한 모델링이 가능하다면 골절 영상이 회복된 상태의 비골절 영상을 생성해내고, 원래의 골절의 이미지와 가상으로 회복된 이미지의 차이를 겹쳐 놓음으로서 자연스럽게 골절 영역을 탐지할 수 있을 것이라고 생각하였다. 그러나 연구 결과 생성된 영상은 GAN 특유의 학습 어려움으로 인하여 실제 x-ray 영상에 비해 크게 자연스럽지는 못하였다. GAN을 사용하여 손상된 이미지의 특정영역을 복원 생성하는 인페인팅 (inpainting) 작업처럼 분류자의 결과에 영향을 크게 끼치는 부위에 대한 mask를 구하고 이를 가지고 생성자가 집중적으로 생성할 수 있도록 범위를 한정하여 성능을 개선하도록 하였다. G-D-C 스텝 1에서 mask를 생성하고, 스텝2에서 이렇게 생성된 mask를 가지고 실제 비골절 영상을 생성하도록 하였다. 스텝1과 스텝2의 두개의 단계로 진행된 이유는 생성자가 영상 생성시 GAN의 학습 어려움의 문제를 해결하기 위하여 가이드를 두어 학습이 좀 더 원활히 되게 하기 위함이었다. 만약 좀 더 쉽게 학습되는 GAN의 연구결과가 공개되어 이를 사용할 수 있는 경우 두 단계로 분리되지 않는 결과적으로 더 간결한 방법으로 학습이 가능할 것으로 기대된다.

즉, 결과적으로 골절 부위 탐지라는 목적에 부합한 모델을 개발하였으나, 골절 영상으로부터 매우 자연스러운 비골절 영상 생성은 하지 못하였다. 딥러닝의 생성모델에 의한 성공적인 결과들이 보고되고 있지만, 실제 학습을 진행해 보면 인간의 눈으로 보기에 깔끔하게 생성하는 것이 매우 어려운 것으로 알려져 있다. 본 연구의 결과로 생성된 비골절 영상들에 대한 C모델의 출력은 큰 신뢰점수의 비골절이며, 자연스러움을 판정하는 D모델의 출력도 역시 큰 신뢰점수의 자연스러움이었다. 그러나 인간의 눈으로 보았을 때에는 그다지 자연스럽지 못하였다. 이러한 점이 딥러닝의 취약점으로 보고되고 있으며, 그 예시를 그림 20에서 볼 수 있다¹⁰.

이러한 결과는 학습 데이터의 부족, 모델의 확장성 부족 등 기타 사항에 기인하며, 이러한 이유로 생성모델의 학습이 쉽지 않은 것으로 알려져 있다. 본 연구에 사용된 두부 측면 엑스선 영상은 다양한 환경에서 다양한 기기와 세팅으로 촬영된 것으로, 코뼈 탐지 모델을 통해 크롭된 코뼈 부분의 사이즈와 비율이 다 달랐다. 이 크롭된 이미지를 통일된 입력으로 사용하기 위해 256x256 해상도로 리사이즈 하면서 정보 왜곡이 있을 수 있고, 이것 또한 연구 결과에 악영향을 주었을 가능성도 있을 것으로 생각된다.

5. 향후 연구 방향 제언

데이터의 보완, 모델 구조의 개선 등으로 자연스러운 비골절 영상의 생성을 개선하는 것이 필요하다. 성공적인 폐쇄 정복술로 골절 이후 치료된 환자들의 데이터를 추가적으로 수집하여 전후 변화를 예측하는 모델을 구축하는 것이 개선에 도움이 될 것으로 생각된다.

V. 결론

본 연구에서는 두부 엑스선 이미지로부터 비골 영역을 추출하고, 골절 여부를 판단하며, 골절 의심 부위를 추정하며 비골절 상태의 이미지를 예측하는 모델을 개발하고 성능을 검증하였다. 이는 코뼈 골절 진단에 있어 해당과 전문의 수준으로 학습된 심층신경망 모델이며, 이를 임상에 적용한다면 인력 부족과, 오진으로 인해 발생할 수 있는 문제들을 사전에 방지하고, 환자들에게 적절한 시기에 치료를 받게 할 수 있을 것으로 기대된다. CT 시설을 갖추지 못한 작은 규모의 병의원들, 영상 장비와 전문 인력이 부족하고 병원 방문이 어려운 군 야전 진료소 등에서 유용하게 활용될 것으로 기대되며, 원격의료에 관한 활발한 논의가 이루어지는 요즘, 본 연구에서 개발된 결과물을 적극적으로 활용한다면 지방 소도시의 작은 병원에서 상급 병원으로의 전원 여부를 결정하는 데에 큰 도움이 될 것으로 생각된다. 본 연구에서 개발한 프레임워크¹¹를 확장함으로써, 보다 다양한 부위의 골절 진단에 활용될 수 있는 기법을 개발할 수 있을 것으로 기대된다. 이를 통해 임상과의 진단을 보조하고, 수련의의 진단 능력을 키울 수 있는 교육적 도구로서도 활용이 가능할 것으로 기대된다.

<참고문헌>

1. Hallas P, Ellingsen T. Errors in fracture diagnoses in the emergency department-characteristics of patients and diurnal variation. BMC Emerg Med 2006; 6:4.
2. Gulshan V, Peng L, Coram M et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. JAMA 2016; 316:2402-2410.
3. Kim DH, MacKinnon T. Artificial intelligence in fracture detection: transfer learning from deep convolutional neural networks. Clin Radiol 2018; 73:439-445.
4. Chung SW, Han SS, Lee JW et al. Automated detection and classification of the proximal humerus fracture by using deep learning algorithm. Acta orthopaedica 2018 89; (4): 468-473
5. Lindsey R, Daluiski A, Chopra Set al. Deep neural network improves fracture detection by clinicians. Proc Natl Acad Sci U S A 2018; 115:11591-11596.
6. <https://github.com/ultralytics/yolov5>
7. <https://ai.googleblog.com/2019/05/efficientnet-improving-accuracy-and.html>
8. Olaf Ronneberger, Philipp Fischer, Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. International Conference on Medical image computing and computer-assisted intervention . Springer, Cham, 2015.
9. Michal Drozdal, Eugene Vorontsov, Gabriel Chartrand et al. The importance of skip connections in biomedical image segmentation. Deep learning and data labeling for medical applications. Springer, Cham, 2016. 179-187.
10. Ian J. Goodfellow, Jonathon Shlens, Christian Szegedy, Explaining and Harnessing Adversarial Examples. ICLR (Poster) 2015
11. https://github.com/entleehj/nose_fracture_detector

<Abstract>

Nasal bone fracture detection utilizing a convolutional neural network

Ho Jun Lee
The Graduate School
University of Ulsan, College of Medicine

Introduction : Nasal bone fractures is one of the most common types of facial bone fractures in the world, because the nasal bone is the most protruding part of the face, and its bony thickness is thin. Currently, especially in the trauma field, fracture diagnosis depends largely on the diagnostic imaging. Fine facial bone computed tomography (CT) scans can make a final diagnosis for the bony framework injury. However, not all the hospitals or clinics are equipped with CT. So, the simple nasal bone X-ray has been used for screening the nasal bone fractures. Since nasal bone is penetrated by several neurovascular bundles, and has suture lines with other facial bones, doctors who are not familiar with this area can make miss diagnosis with simple x-ray. Recently, artificial intelligence (AI) has shown potential to produce automated diagnosis on medical images in various fields of medicine such as clinical skin photographs or retinal fundus photos. Some achieved good performance in the field of diagnostic radiology including simple x-ray. However, there have been paucity of reports which used AI technology on nasal bone fracture detection using simple x-ray. Therefore, present study aimed to develop and assess convolutional neural network (CNN) models (1) which detect nasal bone fracture on simple x-ray images automatically and (2) which can highlight the fracture site using unsupervised anomaly detection technique.

Materials and Methods : 2312 images (2 image per person) of simple facial bone lateral x-ray were prepared. It included 1468 normal images and 844 fractured images. We classified the x-ray images into normal and fractured, and the position of the nose was labelled on the x-ray. We constructed 4 different model to achieve our goal. (1) Model that detect nasal bone area from facial

bone simple x-ray(nasal bone detecting modal, YOLO5), (2) Model that classify the x-rays into normal or fractured images (Classifier, C model, EfficientNetB3), (3) Model that generate non-fractured nasal bone images from fractured x-ray images (Generator, G model, U-net), (4) Model that discriminate the generated x-ray image natural or not (Discriminator, D model, Convolutional neural network). We connected these four different models to construct complex deep learning model, and it was trained after augmentation of a training dataset. The ability of this complex deep learning model on diagnosing the nasal bone fracture, and the fracture site detection were evaluated.

Results : Nasal bone detecting model was well trained, and it detected nasal bone area from facial bone x-ray. It was used as a preprocessing for the complex deep learning. C model produced the diagnostic accuracy of 95% (SD 0.02%), the area under the receiver operator characteristic curve (AUC) was 0.99 (SD 0.003). As a result of G-D-C complex model, fracture site was automatically detected and highlighted on the fractured x-ray image.

Conclusion : Authors developed a complex deep learning model which can automatically diagnose nasal bone fracture on the simple facial bone x-ray. This model also could automatically highlight the fracture site on the simple facial bone x-ray through unsupervised learning method. The idea of this model and its framework can be modified and improved in the future, and it can be applied to another part of fracture detection. It may play a role in the small sized clinics, military field clinics, rural hospital. And it also could be used for the medical student and novice doctor training.

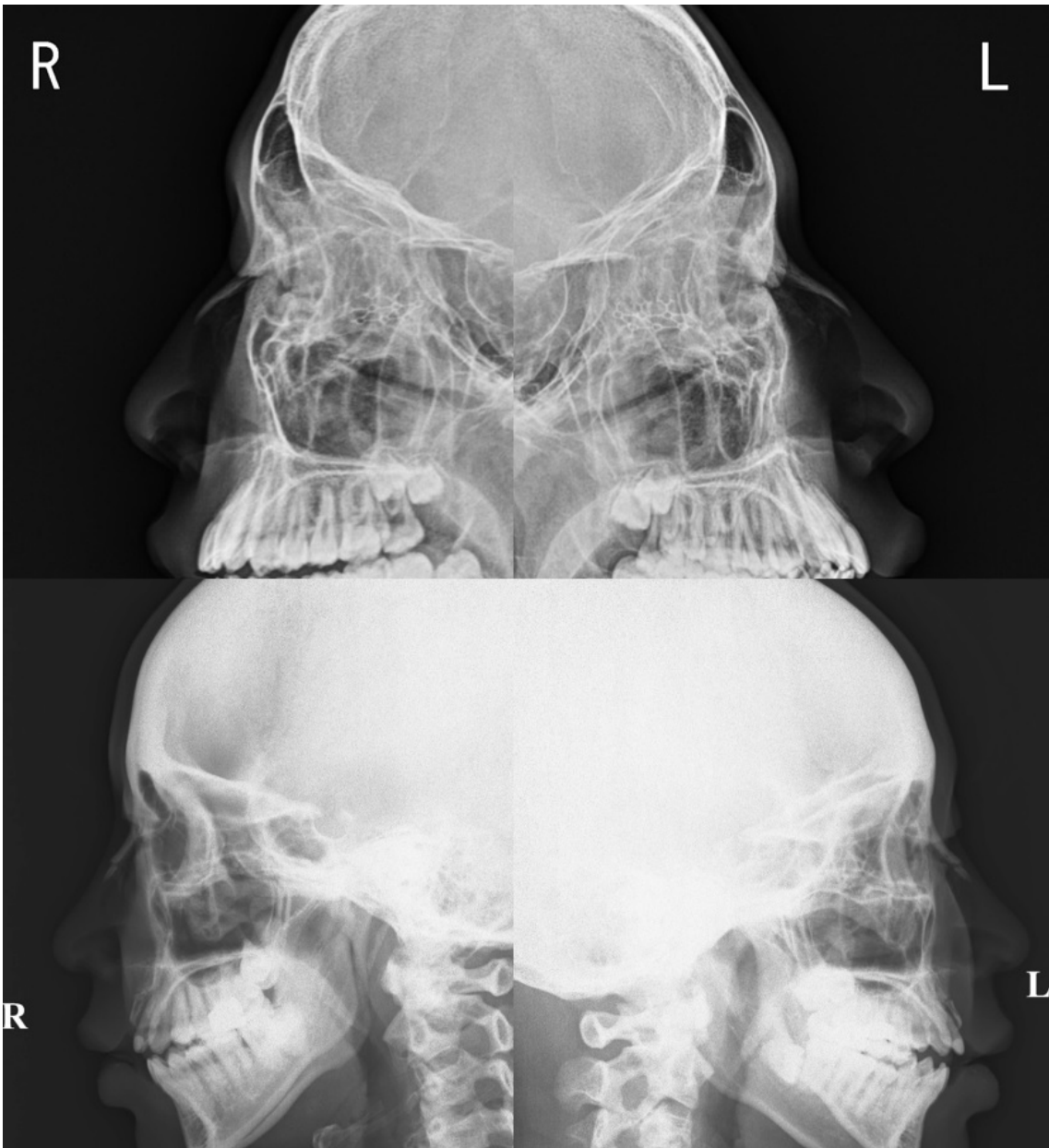


그림 1. 수집된 엑스선 영상을 비골절 (첫번째 행) 및 골절 (두번째 행) 이미지로 라벨링을 시행한 예시

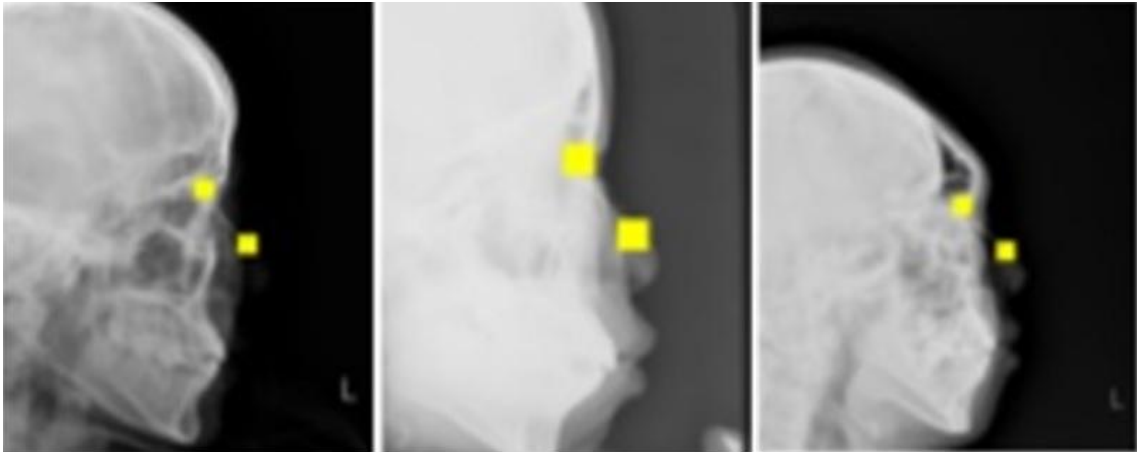


그림 2. 코뼈의 위치가 라벨링 되어 있는 예시



그림 3. 곱절 여부 판정 모델의 기반이 되는 EfficientNet의 구조

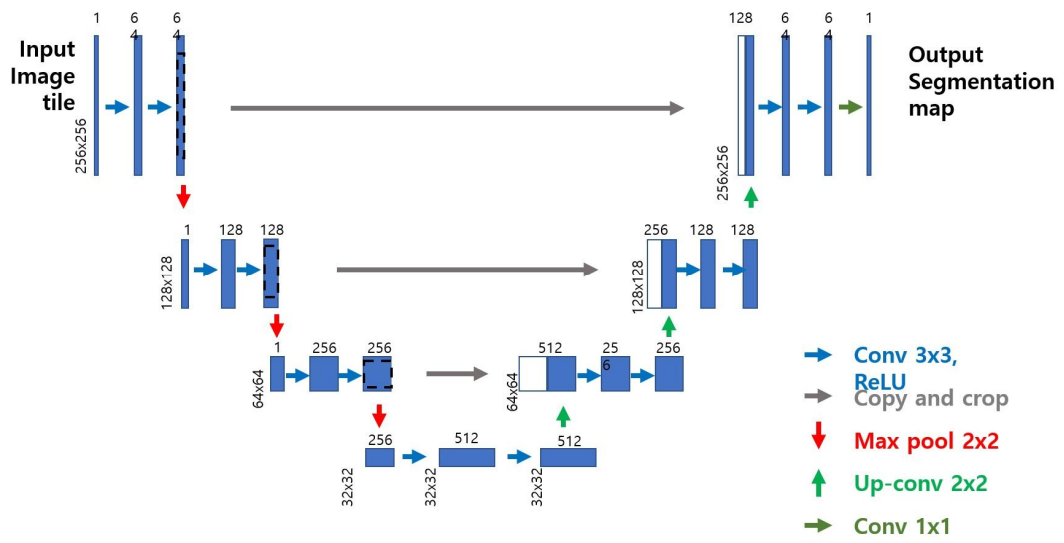


그림 4. 비골 이미지 생성모델 (G모델)인 Skip-UNet의 구조

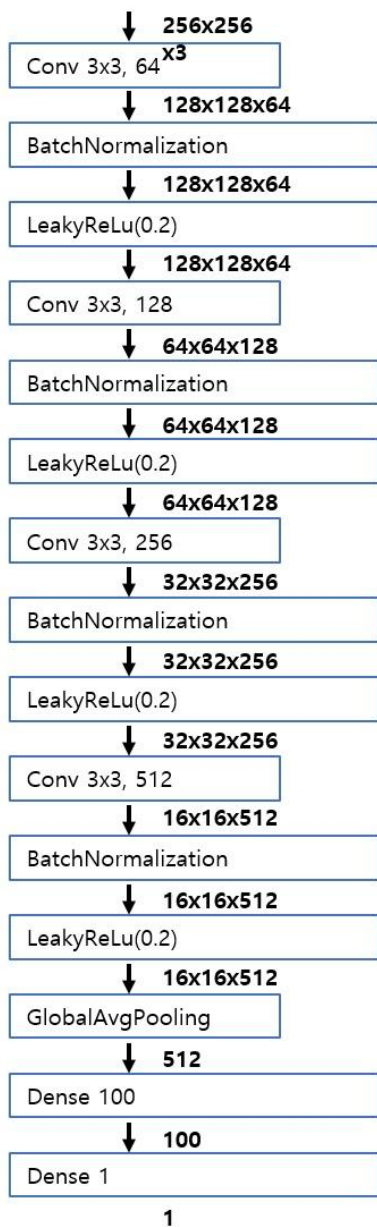


그림 5. 생성 이미지의 판별 모델 (D모델)의 구조

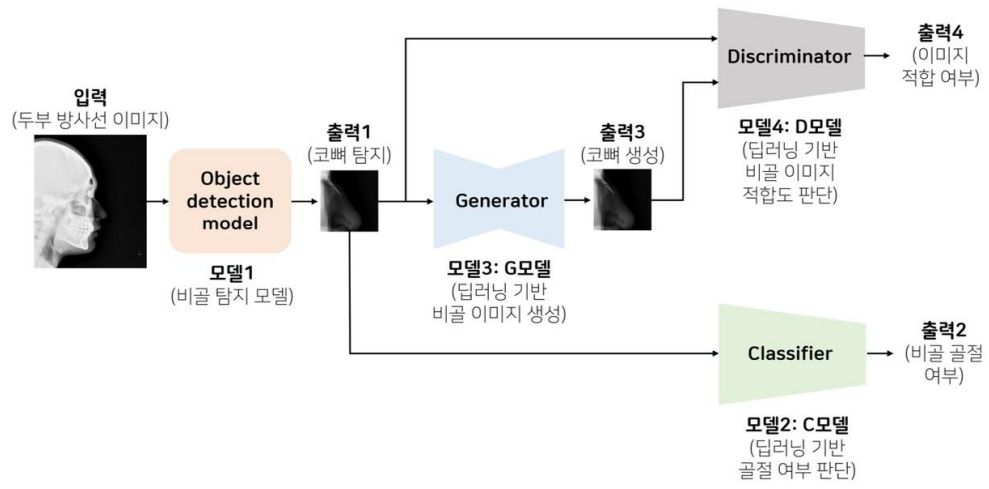


그림 6. 딥러닝 모델의 전체 구조

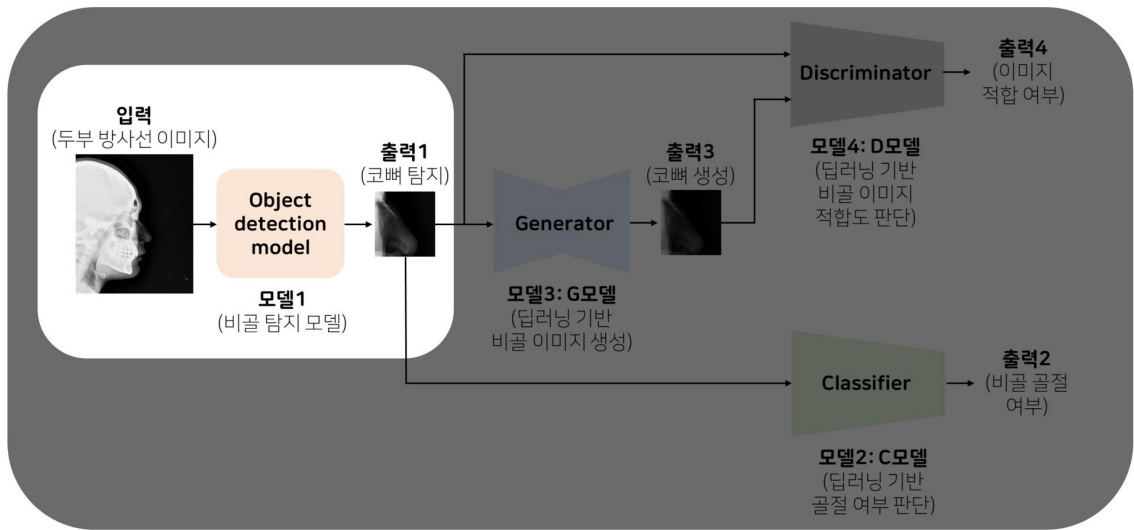


그림 7. 코뼈 탐지 모델의 학습

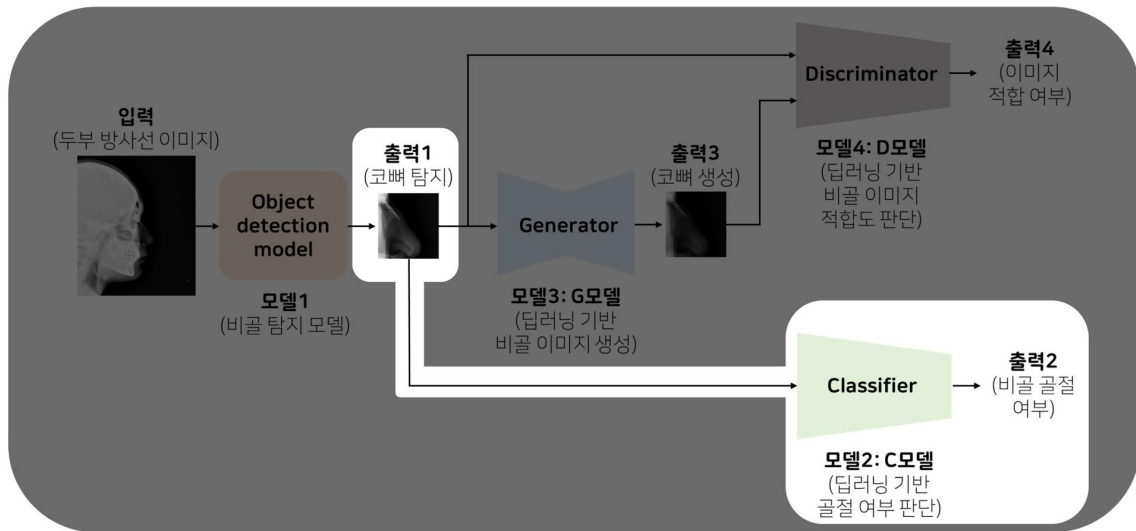


그림 8. C 모델의 학습

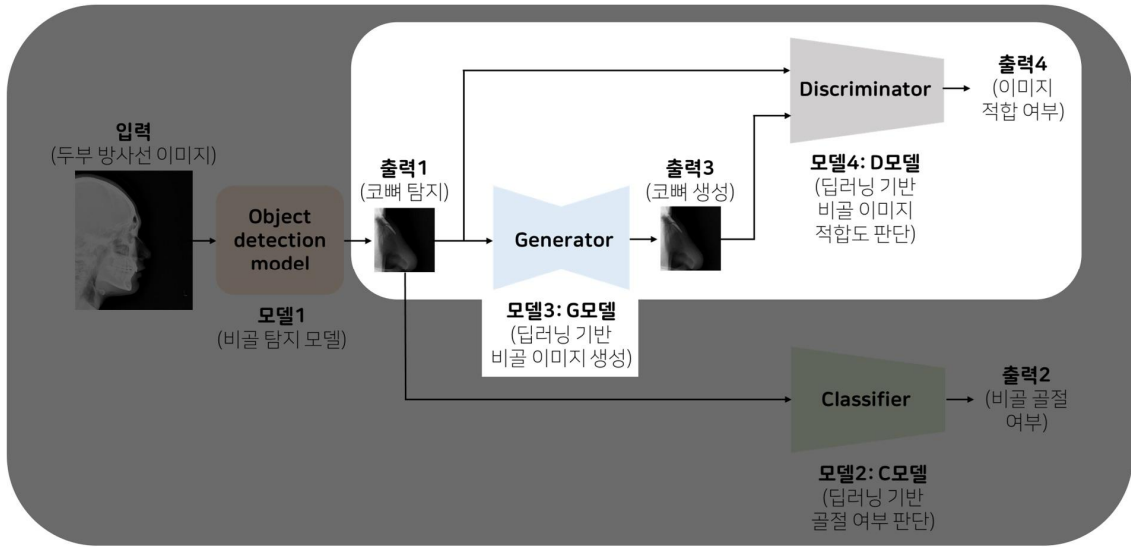


그림 9. G-D 모델의 학습

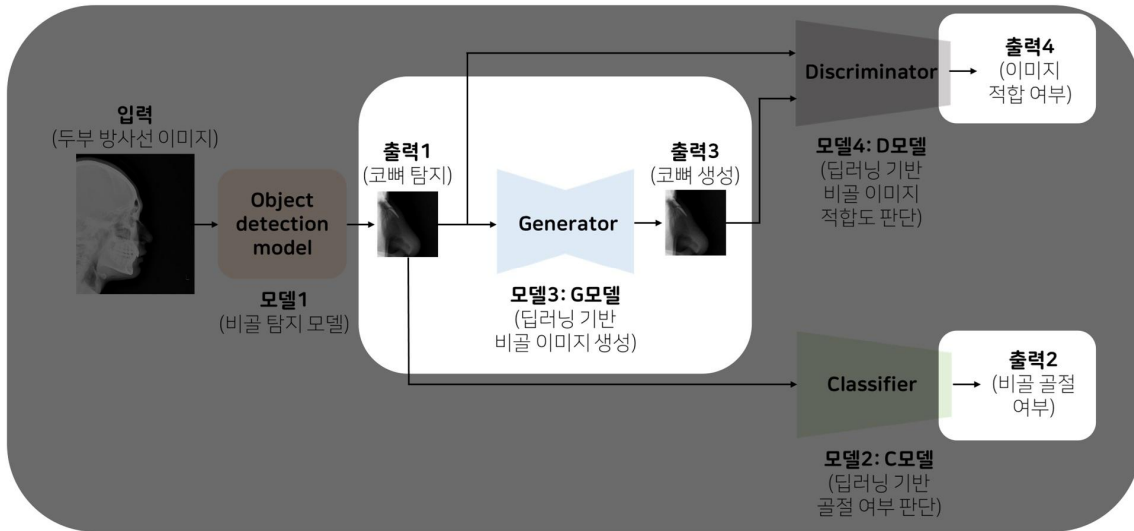


그림 10. G-D-C 모델의 학습

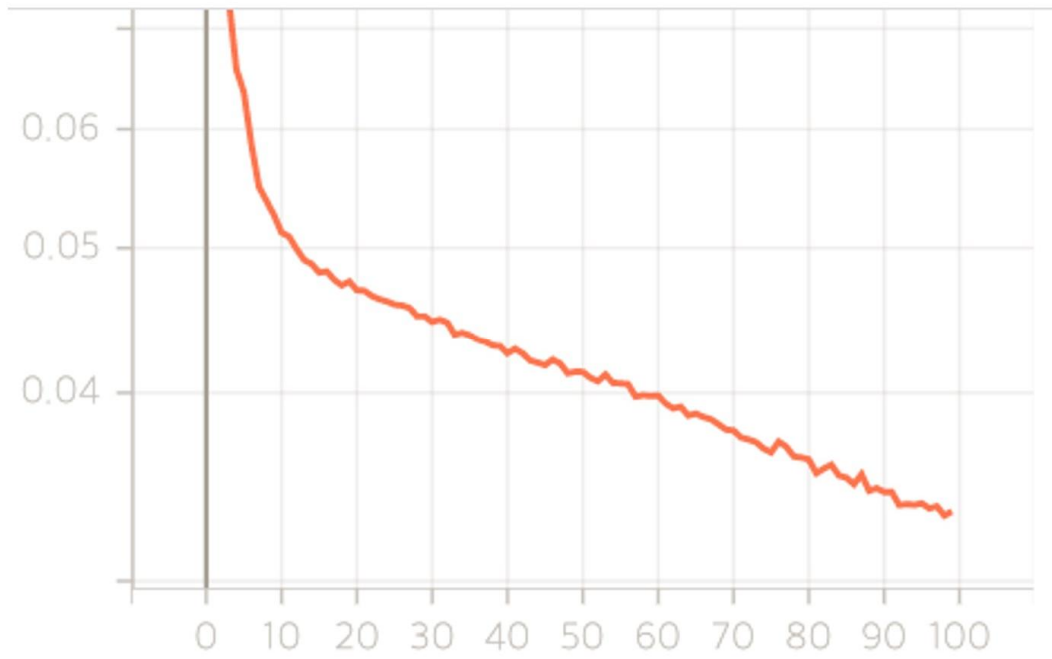


그림 11. epoch에 따른 코뼈탐지 모델의 손실 저하 (x: epoch, y: model loss)



그림 12. 코뼈탐지 모델의 출력으로 두부 엑스선 이미지 (왼쪽)로부터 코뼈영역 (오른쪽)을 탐지한 예시

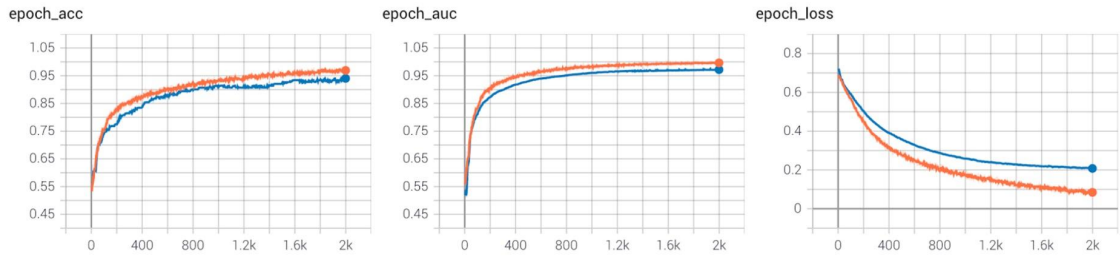
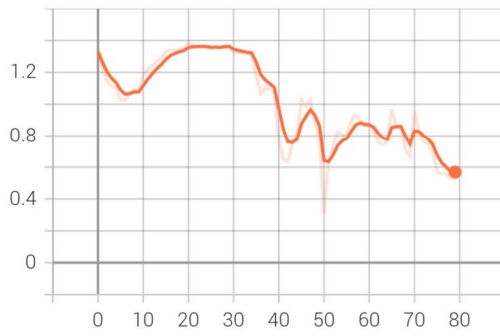


그림 13. 학습 epoch에 따른 학습 정확도, AUC, 손실의 변화 그래프

d_loss
tag: d_loss



g_loss
tag: g_loss

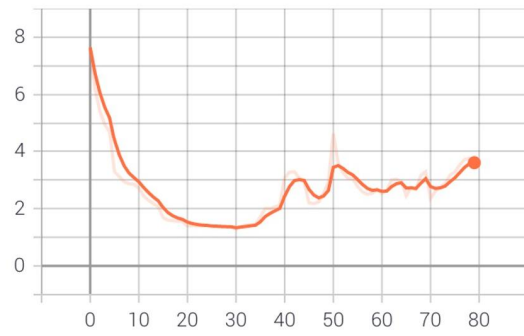


그림 14. G-D 모델 학습에서 각 모델의 학습 과정에 따른 손실 변화 그래프 (x: epoch, y: loss)

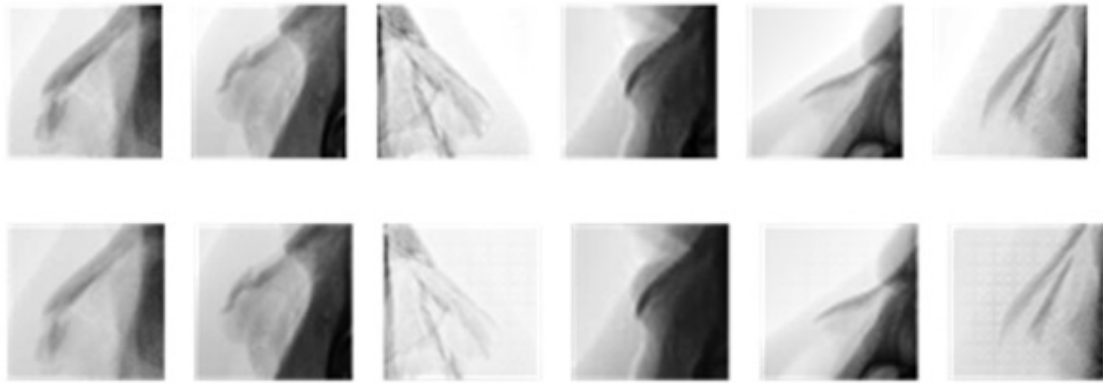


그림 15. G 모델의 입력 이미지 (첫번째 행) 과 모델의 생성 이미지 (두번째 행)

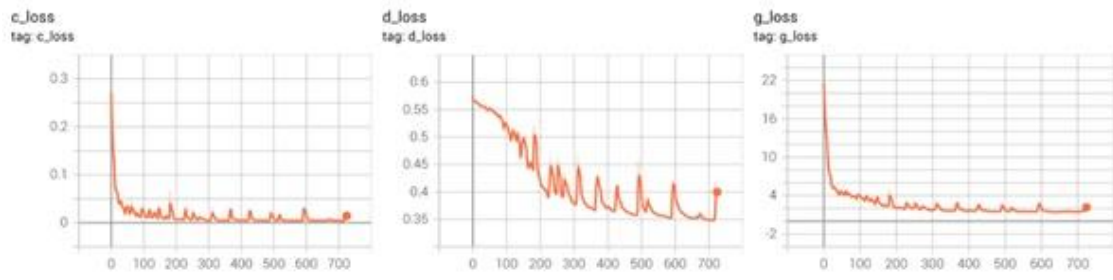


그림 16. G-D-C step 1 모델의 학습에 따른 손실 변화 그래프 (x: epoch, y: loss)

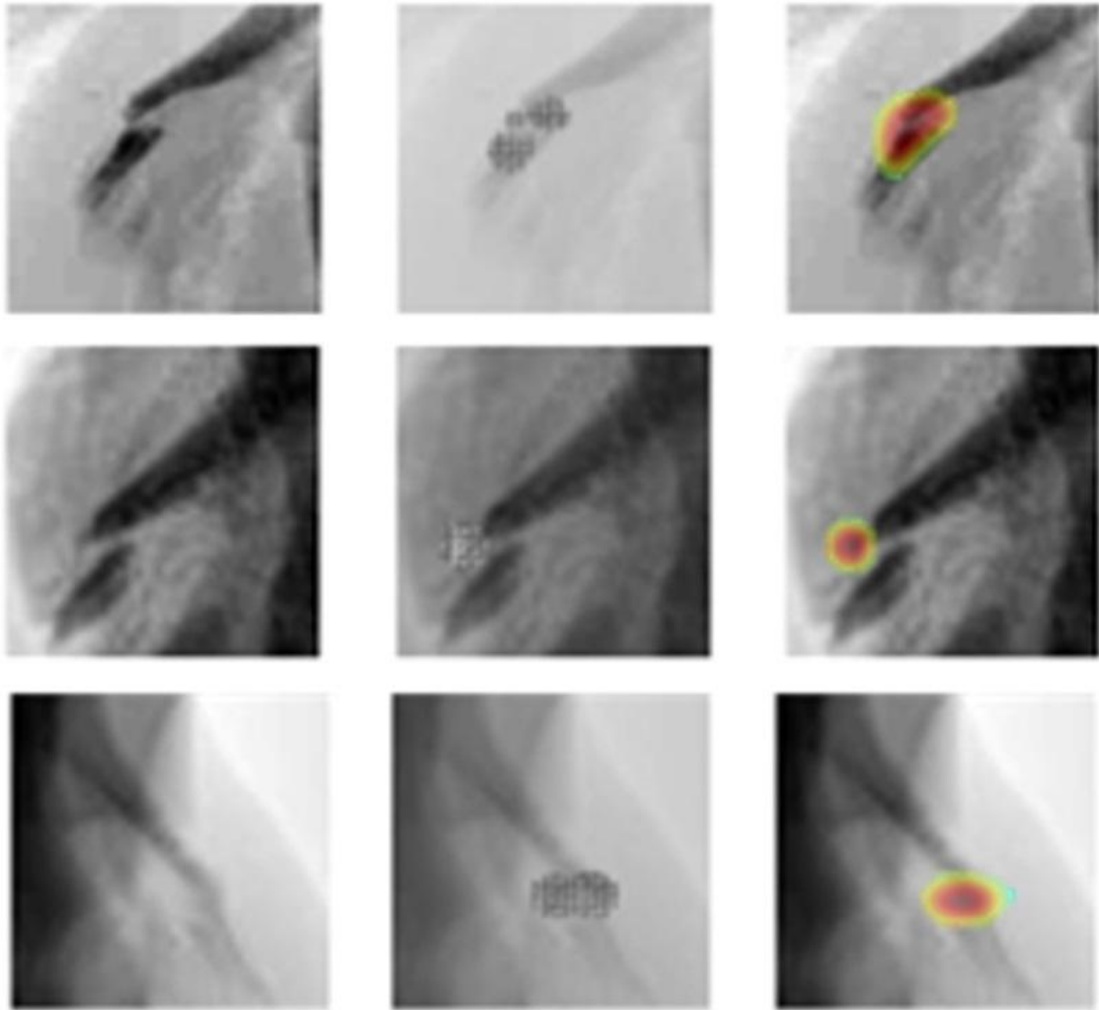


그림 17. 곰질 상태의 코빠 이미지와 G-D-C 모델이 예측한 곰질 의심 부위

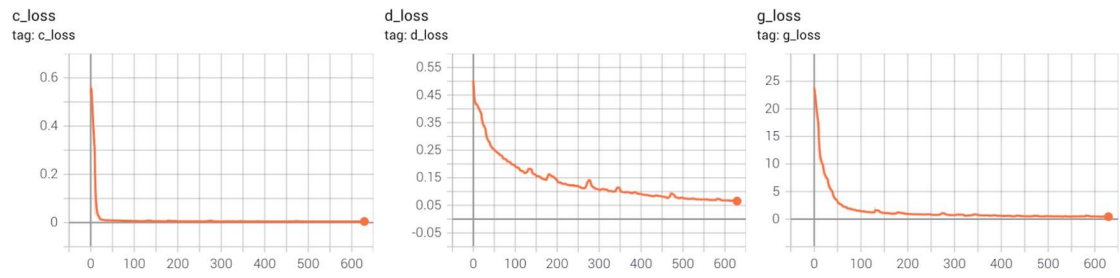


그림 18. G-D-C step 2 모델 학습에 따른 손실 변화 그래프 (x: epoch, y: loss)

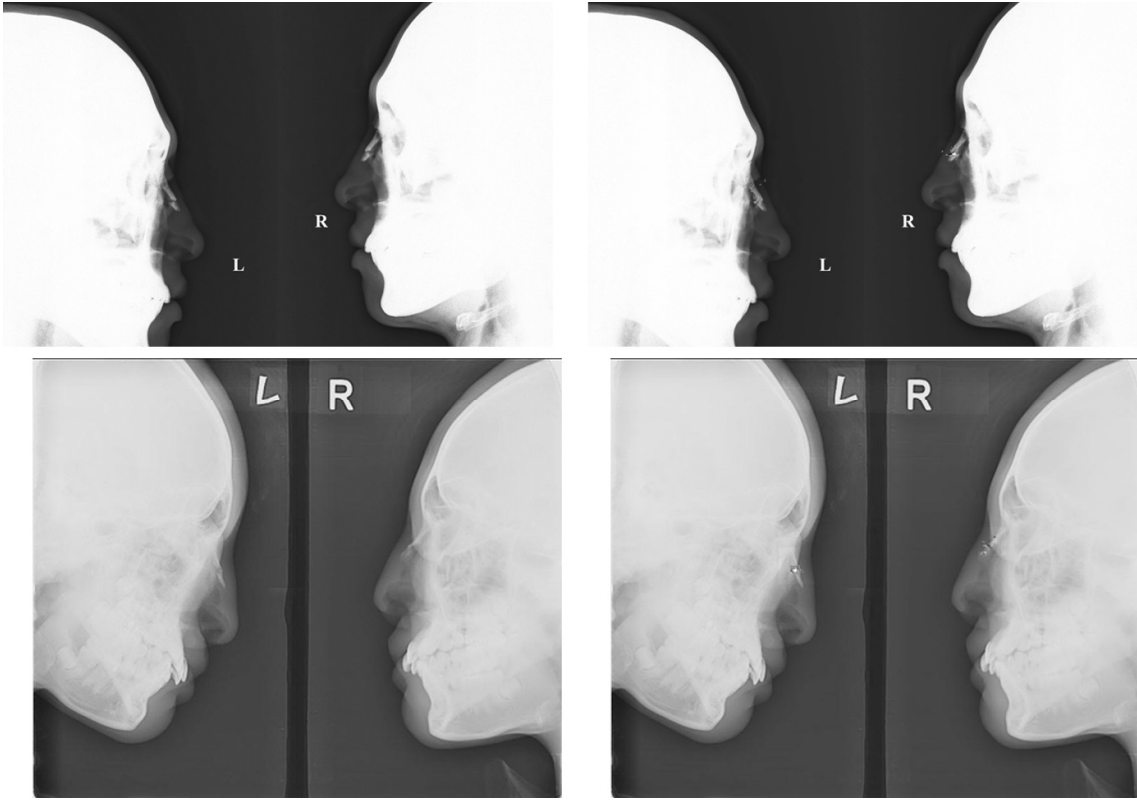
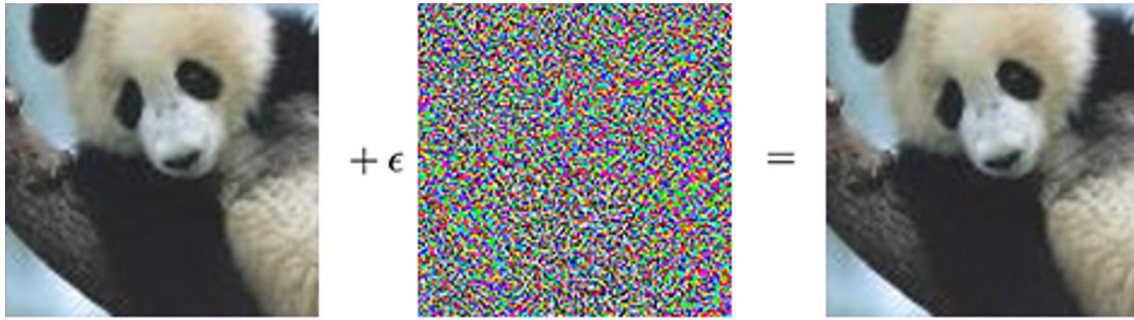


그림 19. 골절상태의 두부 엑스선 이미지와 (1열), G-D-C step 2 모델이 예측한 골절부위 (2열)



"panda"

57.7% confidence

"gibbon"

99.3% confidence

그림 20. 인간이 눈으로 관측하는 것과 (판다) 모델이 높은 신뢰도로 판단하는 결과 (긴팔원숭이)의 차이를 보이는 예제