creative commons

COMMONS DEED

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

dCollection

치의과학박사 학위논문

파노라마방사선영상에서 딥러닝에 기반한 치근단 병변의 세그먼테이션

Deep learning-based periapical lesion segmentation from panoramic radiographs

2023년 2월

서울대학교 대학원

치의과학과 영상치의학 전공

송 일 석

파노라마방사선영상에서 딥러닝에 기반한 치근단 병변의 세그먼테이션

지도교수 허 민 석

이 논문을 치의과학 박사학위논문으로 제출함

2022년 12월

서울대학교 대학원

치의과학과 영상치의학 전공

송 일 석

송일석의 치의과학 박사학위논문을 인준함

2023년 1월

위 원 장 _____ (인)

부위원장 _____ (인)

위 원 _____ (인)

위 원 _____ (인)

위 원 _____ (인)

초 록

파노라마방사선영상에서 딥러닝에 기반한 치근단 병변의 세그먼테이션

송 일 석

서울대학교 대학원

치의과학과 영상치의학 전공

(지도교수 허 민 석)

1. 목 적

심층 신경망(deep neural network)은 의학 및 치의학 연구 분야에서 가장 유망한 인공지능 중 하나로 급부상하였으며, 이를 통해 질병을 조기에 발견할 수 있는 효과적인 진단 방법을 제공할 수 있다. 본 연구의 목적은 CNN 알고리즘을 이용하여 파노라마방사선영상에서 치근단 병변

을 세그먼테이션 하는 능력을 평가하는 것이다.

2. 방 법

치근단 병변을 가진 총 1000장의 파노라마방사선영상을 학습(n=800, 80%), 검증(n=100, 10%), 테스트(n=100, 10%) 데이터 세트로 각각 분류하였다. 파노라마방사선영상은 ResNet50 모델을 이용하여 일정한 크기로 절단된 영상을 학습에 사용하였으며, 학습은 시맨틱 세그먼테이션 방법의 대표적인 모델인 U-Net 모델을 이용하여 진행되었다. 치근단 병변을 확인하는 성능은 IoU 임계값에 따른 precision, recall, F1-score 값을 계산하여 평가하였다.

3. 결 과

IoU 임계값 0.3에서 180개의 병변 중 147개의 병변이 검출되었다. F1-score는 IoU 임계값 0.3, 0.4, 0.5에서 각각 0.828, 0.815, 0.742으로 확인되었다.

4. 결 론

이번 연구에서 치근단 병변의 감지 및 진단을 위한 딥러닝의 잠재적 유용성이 확인되었다. U-Net 알고리즘은 파노라마방사선영상에서 치근단 병변을 감지하는 데 상당히 좋은 성능을 제공하였다.

주요어 : 치근단 병변, 파노라마방사선영상, 딥러닝, 인공지능

학 번 : 2016-33728

목 차

요약(국문초록) i

목 차 iv

표 목차 v

그림 목차 vi

1. 서론 1

2. 문헌 고찰 7

3. 연구재료 및 방법 11

4. 결과 21

5. 총괄 및 고찰 31

6. 결론 35

참고문헌 36

Abstract (in English) 46

표 목차

Table 1. The number of segmented periapical lesions by U-Net. 22

Table 2. Measurements of precision (positive predictive value), recall (sensitivity), and F1-score for periapical lesion segmentation by U-Net. 23

Table 3. Precision, recall, and F1-score for periapical lesion segmentation by U-Net and FPN. 30

그림 목차

Figure 1. Venn diagram of artificial intelligence 2

Figure 2. Archtecture of neural network for cropping images using ResNet50. 12

Figure 3. Cropped panoramic images for training. 13

Figure 4. Training software called “Deep Stack.” 15

Figure 5. U-Net model architecture for the detection of apical lesions 17

Figure 6. Intensity histogram 20

Figure 7. Precision-recall curve 24

Figure 8. Training and validation curves for the pre-trained augmented model 26

Figure 9. An example of apical lesion segmentation from panoramic radiographs using U-Net 28

Figure 10. Feature pyramid network with ResNet50 encoder. 29

1. 서 론

치근단 병변은 치수가 괴사되어 근관이 감염되었을 때 방사선 영상을 통해 확인할 수 있으며, 치근단방사선영상, 파노라마방사선영상, 콘빔전산화단층영상에서 근단부에 방사선투과상으로 나타난다. 콘빔전산화단층영상이 일반적인 방사선영상에 비해 3차원적인 정보를 제공하여 상당히 높은 정확도의 진단을 가능하게 하지만[1], 촬영비용과 일반적인 방사선영상에 비해 상대적으로 높은 피폭량 때문에 일상적으로 촬영하기에는 제한이 있다. 그래서 치근단 병변을 진단하는 데 있어서 일반적으로 이용하는 방사선영상은 치근단방사선영상과 파노라마방사선영상이다. 특히, 파노라마방사선영상은 치아 우식, 치근단 병변, 치주 상태, 골내 결손과 이외 다른 치과 질환을 검출하기 위해 일상적으로 촬영하게 된다[2].

방사선검사는 다양한 개체를 구별할 수 있는 능력이 있음에도 불구하고 판독하는 사람에 따라 신뢰도에 차이가 발생할 수 있으며[3,4], 이러한 신뢰도는 검사자의 경험에 따라 영향을 받는다[5]. 따라서 치과질환의 예측 및 분석을 위한 경험적 접근 방식은 치과 임상가의 경험과 주관적인 평가기준에 상당히 의존하게 된다. 또한, 이러한 임상적 접근과 방법들은 치과질환을 조기에 검출하고 진단하는 과정에서 많은 시간과 노력이 요구되며, 신뢰할 수 있는 표준적인 임상 평가를 제공하는 것이 어렵다. 측방두부규격방사선영상에서 계측점 진단, 상악동염의 진단, 치아우식증과 치주염의 조기 진단 등은 많은 시간이

필요한 과정이며, 사람의 노동이 수반되는 과정이다.

최근에 등장한 머신러닝의 한 분야인 딥러닝에 의한 병변의 검출은 의료 및 치과 분야에서 상당히 높은 성능을 보이는 것으로 확인되었다 (Figure 1)[6,7]. 이와 같은 인공지능에 대한 연구가 활발히 진행되고, 학습에 의해 성능이 향상되면서 의료분야에서 딥러닝에 의한 학습모델의 진단능력이 점점 더 발전하고 있다.

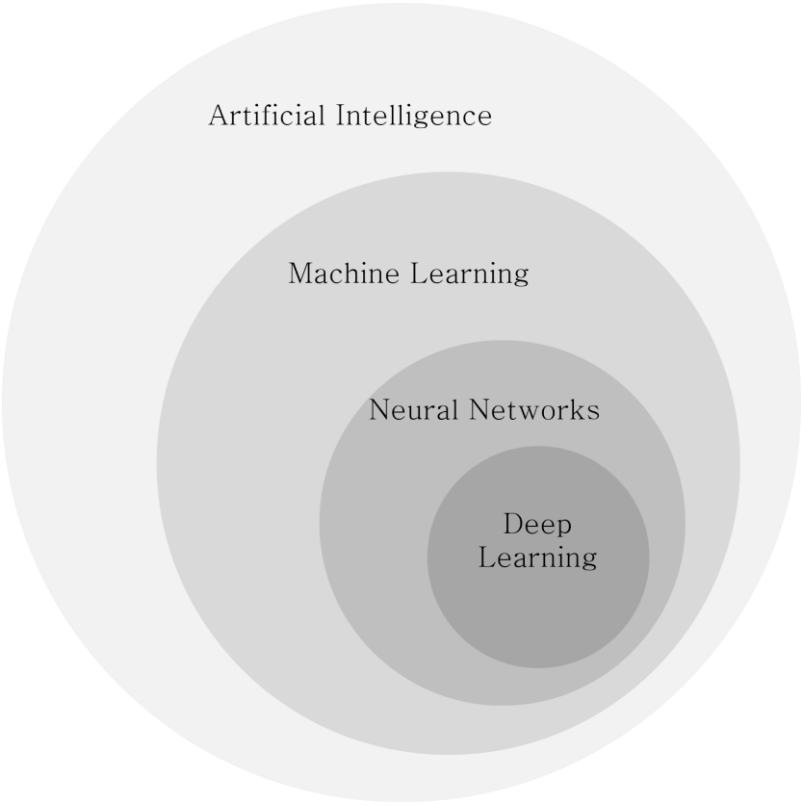

Figure 1. Venn diagram of artificial intelligence

인공지능은 인간의 인지 · 추론 · 판단 등의 능력을 고안된 알고리즘을 이용하여 컴퓨터로 구현하기 위한 기술을 포괄하는 의미의 용어로 사용된다[8]. 컴퓨터는 연산 능력이 뛰어나지만, 인간 공유의 능력인 인지 · 추론 · 판단 등의 능력은 가지고 있지 않다. 이러한 인간 고유의 능력을 컴퓨터에서 구현해 보고자 시작된 것이 인공지능이다.

1956년 Dartmouth 대학에서 열린 컨퍼런스에서 John McCarthy가 '인공지능(artificial intelligence)'이라는 용어를 처음 사용하기 시작했다[9]. 하지만 그 당시 연산처리 속도의 한계와 함께 최초의 인공신경망 알고리즘인 퍼셉트론(perceptron)의 근원적인 한계였던 XOR 선형분리 문제로 인해 연구가 빠르게 진행되지 못했다[10].

XOR 선형분리 문제란 0과 1의 XOR 진리표에 대한 선형분리가 불가능하다는 것을 의미하는데, 1980년대에 들어서 다층 퍼셉트론(multi-layer perceptron)을 학습시킬 수 있는 오차 역전파(error back-propagation) 기법이 도입되면서 이러한 문제가 해결되었다[11]. 오차 역전파 기법은 다층 신경망 학습에 사용되는 통계적 기법으로 출력층에서 발생한 결과를 입력층으로 전송하여 오차를 수정하면서 가중치를 재설정하는 기법이다. 하지만 오차 역전파 기법이 3층이 넘는 심층신경망에는 적용하기 어려운 문제를 가진 것으로 간주되면서 다시 한번 인공지능 발전의 정체기를 맞았다. 그리고 이러한 문제의 원인이 오차 역전파 기법에 있는 것이 아니고 퍼셉트론에서 사용하는 활성함수(activation function)의 문제라는 것이 밝혀졌다[12]. 2010년 정류 선형 유닛(rectified linear unit, ReLU)이라는 활성함수가

이용되면서 수백 수천 층의 심층신경망에서도 학습이 가능하게 되었고, 이것이 딥러닝 알고리즘 발전의 시초가 되었다.

딥러닝 분야에서는 여러 개의 단순 연산을 한 번에 처리할 수 있는 병렬 연산이 핵심 기술이다. 복잡하고 순차적인 연산을 하는 중앙처리장치(central processing unit: CPU)가 많은 수의 병렬 연산이 필요한 인공신경망에 적합하지 않음을 확인하고, 이를 해결하기 위해 다수의 병렬 연산을 한 번에 처리할 수 있는 그래픽처리장치(graphic processing unit: GPU)를 사용하고 있다[13]. 하지만 이것 역시 다중 레이어를 가진 신경망을 활용하게 되면 처리속도의 한계에 부딪히게 되기 때문에 CPU나 GPU가 아닌 신경망을 처리하는 인공지능 전용 하드웨어가 요구되고 있다.

머신러닝(machine learning)은 인공지능의 한 분야이며, 1959년 Samuel[14]은 머신러닝을 “기계가 일일이 코드로 명시하지 않은 동작을 데이터로부터 학습하여 실행할 수 있도록 하는 알고리즘을 개발하는 연구 분야”라고 정의하였다. 즉 사람이 학습하듯이 컴퓨터에도 데이터들을 입력하면 스스로 학습하게 함으로써 새로운 지식을 얻게 하는 분야이다[15].

딥러닝은 인공신경망에 기반을 둔 머신러닝 기술의 한 종류이다. 다층 구조로 설계하여 깊어진 인공신경망이 학습이 잘 이루어지지 않는 문제를, 학습을 위한 데이터들을 비지도 학습(Unsupervised Learning)을 통해 전처리하면 신경망이 깊어져도 학습이 잘 이루어진다는 것을 알아내고 이를 이용하였다[16]. 컴퓨터가 인간의

뇌에 있는 뉴런과 비슷한 입출력 계층을 활용하여 데이터를 학습하면 학습된 데이터를 사용하여 스스로 학습하여 인공지능 모델을 생성하고, 이렇게 학습된 특정 패턴과 정보는 계속적인 빅데이터의 학습을 통해 더욱 정교한 분류가 가능하게 된다. 신경망의 층이 많고 각 층마다 여러 개의 변수를 고려한다는 것이 딥러닝의 특징이다[17].

딥러닝을 활용한 컴퓨터 비전 분야에서 가장 많이 활용되는 것이 객체검출과 시맨틱 세그먼테이션(semantic segmentation)이다[17]. 객체검출 알고리즘에서는 주요 객체를 탐지하고 해당 객체를 중심으로 경계 박스를 표시하여 구분하며, 활용 분야로는 차량 감지와 추적, 감시 시스템 등이 있다[18]. 최근에는 안면인식, 영상 분류, 동영상 및 음성 인식 등과 같은 여러 다양한 분야에서도 많이 활용되고 있다[19-21]. 시맨틱 세그먼테이션은 디지털 영상을 여러 개의 화소 집합으로 나누는 과정으로, 영상 분할을 하여 영상의 표현을 해석하기 쉬운 것으로 단순화하여 변환하는 것이다. 이를 통해 영상 내 모든 화소에 대한 분류 결과를 보여주고, 영상의 각 부분이 어떤 의미를 가지고 있는지 구분할 수 있게 된다.

이러한 딥러닝 접근법을 이용하여 구강 및 악안면 영역의 방사선영상에서 경험에 기반한 진단의 문제를 효율적으로 극복할 수 있다[22]. 하지만 딥러닝 접근법을 이용하는 치과영역에서의 진단은 아직 연구단계에 머무르고 있으며, 그 성능에 대한 불확실성으로 인해 실제 진료에서 아직 널리 활용되지는 않고 있다.

본 연구에서는 파노라마방사선영상에서 치근단 병변을 스크리닝 하기

위해 CNN에 기반한 치근단 병변의 영상 분할을 수행하였다. 과거 연구에서는 학습 데이터 세트로 각각의 치아를 수작업으로 절단한 분할된 영상을 이용했던 반면[23], 이번 연구에서는 전체 파노라마방사선영상을 학습 데이터 세트로 사용하여 초기 단계의 치근단 병변을 효과적으로 검출할 수 있도록 하였다.

이전 연구들에서는 치과 질환의 검출을 위해 YOLO 신경망 모델을 사용했던 반면[24,25], 본 연구에서는 치근단 병변의 세그먼테이션을 위하여 U-Net 신경망 모델을 사용하였다. U-Net은 시맨틱 세그먼테이션의 대표적인 모델이며, 시맨틱 세그먼테이션은 화소 수준의 예측을 하는 딥러닝 방법이다. 그래서 적은 수의 화소를 가진 치근단 병변을 확인하는 연구에 U-Net이 더욱 적절할 것으로 판단하였다. 이번 연구의 목적은 U-Net 모델을 이용하여 파노라마방사선영상에서 치근단 병변을 분할하는 성능을 평가하는 것이다.

2. 문헌 고찰

치과 영역의 방사선영상에서 딥러닝을 활용한 진단에 관한 연구는 매우 활발하게 진행되고 있다. 구강악안면 영역에 발생하는 낭종과 종양의 검출[25], 치아 변화의 인식을 통한 개인 식별의 데이터베이스 자동화[26], 치아의 근단부에 발생하는 병변의 검출[23], 치아 우식의 탐지[27,28], 교정 진단[29], 치주치료[30] 등에서 많은 연구의 진전이 있었다.

Kwon 등[25]은 YOLOv3을 이용하여 1,282장의 파노라마방사선영상에서 구강악안면에 자주 나타나는 4가지 질환(함치성낭, 치근단낭, 치성각화낭, 법랑모세포종)을 자동으로 검출하고 진단하는 딥러닝 알고리즘을 개발하고 그 정확성을 평가하였다. 데이터 증강을 통해 학습 데이터를 12배로 증가시켜 학습시켰으며, 88.9% 민감도, 97.2% 특이도, 95.6% 정확도 및 0.94 AUC의 성능을 확인하였다. 함치성낭은 91.4% 민감도, 99.2% 특이도, 97.8% 정확도 및 0.96 AUC 값을 보였다. 치근단낭은 82.8% 민감도, 99.2% 특이도, 96.2% 정확도 및 0.92 AUC 값을 나타냈다. 치성각화낭은 98.4% 민감도, 92.3% 특이도, 94.0% 정확도 및 0.97 AUC 결과를 보였다. 법랑모세포종은 71.7% 민감도, 100% 특이도, 94.3% 정확도 및 0.86 AUC의 결과를 보였다.

Choi 등[26]은 파노라마방사선영상을 이용하여 딥러닝으로 사전

훈련된 객체인식 모델을 이용하여 치아의 변화를 인식 후 개인의 치열에 대한 데이터베이스를 자동으로 구축하고, 개인식별을 자동화하였다. 상위 20.0% 후보군을 추출하였을 때 84.0%의 성공률을 나타내었으며, 상위 10.0%, 5.0% 후보군을 추출하였을 경우에는 각각 72.7%, 59.4%의 성공률을 보였다. 남성의 경우 상위 20.0%, 10.0%, 5.0% 후보군 추출시 71.3%, 63.6%, 51.8%의 성공률을 보였고, 여성의 경우 같은 경우에 97.8%, 81.8%, 66.9%의 성공률을 보였다.

Ekert 등[23]은 85장의 파노라마방사선영상에서 절단한 249개의 각각의 치아 사진을 학습 데이터로 사용하여 치근단 병변을 검출하였다. 학습 데이터를 7배로 데이터 증강시켜 학습시켰으며, 0.85 AUC(area under the curve), 0.65 민감도, 0.87 특이도, 0.49 PPV(positive predictive value), 0.93 NPV(negative predictive value)의 성능을 확인하였다.

Lee 등[27]은 GoogLeNet Inception v3을 이용하여 3000장의 치근단 방사선 사진에서 치아우식을 검출하고 그 정확성을 평가하였다. 학습 데이터를 10배로 데이터 증강시켜 학습시켰으며, 소구치에서는 89.0% 정확도, 84.0% 민감도, 94.0% 특이도, 93.3% PPV, 85.5% NPV의 성능을 확인하였다. 대구치에서는 88.0% 정확도, 92.3% 민감도, 84.0% 특이도, 85.2% PPV, 91.3% NPV의 성능을 확인하였다.

Khan 등[28]은 U-Net, Xnet, SegNet, U-Net+Densenet121를 이용하여 206장의 파노라마방사선영상에서 치아우식, 치조골흡수, 치근간방사선투과상을 검출하고 정확성을 평가하였다. 학습 데이터를

다양한 변형을 적용하여 데이터 증강시켜 학습을 시행하고 성능을 확인하였다. U-Net, Xnet, SegNet, U-Net+Densenet121의 다이스 계수값은 각각 0.453, 0.405, 0.381, 0.434였다.

Setzer 등[29]은 U-Net을 이용하여 61개의 치근을 포함하는 20개의 콘빔전산화단층영상에서 치근단 병변을 검출하였다. 영상에서 각각의 복셀을 치근단 병변, 치아구조, 뼈, 수복재료, 배경으로 분할하는 성능을 평가하였다. 그래서 딥러닝에 의한 치근단 병변의 분할 성능은 0.93의 정확도, 0.88의 특이도, 0.87의 PPV, 0.93의 NPV로 나타났고, 다이스 계수값은 치근단 병변이 0.52, 치아구조가 0.74, 뼈가 0.78, 수복재료가 0.58, 배경이 0.95로 나타났다.

Bayrakdar 등[30]은 U-Net을 이용하여 470장의 파노라마방사선영상을 학습 데이터로 사용하여 치근단 병변을 검출하고 그 정확성을 평가하였다. 학습 데이터를 3배로 데이터 증강시켜 학습시켰으며, 70%의 IoU 임계값에서 0,92의 민감도, 0.84의 precision, 0.88의 F1-score 성능을 확인하였다.

Devito 등[31]은 인공지능을 이용하여 교익 방사선 사진에서 인접면 우식을 진단하는 연구를 하였다. 치아를 발치하여 광학현미경 사진을 촬영하고 이를 gold standard로 하여 전문가 집단과 인공지능 진단 모델의 성능을 비교하였을 때 인공지능 진단 모델이 더 높은 성능을 보여 진단 능력이 더 우수함을 보고하였다.

Park 등[32]은 YOLOv3 모델을 사용하여 1311장의 측방두부규격 방사선영상에서 자동으로 계측점을 측정하는 모델의 성능을 평가하였다.

학습 데이터를 통해 얻어진 모델을 이용하여 테스트 한 결과 80.4%~96.2%의 정확도를 나타내었다.

Lee 등[33]은 CNN을 기반으로 한 알고리즘을 사용하여 651명의 환자를 촬영한 1740장의 치근단 방사선 사진에서 치주적으로 손상된 치아를 검출하고 그 성능을 평가하였다. 소구치에서는 학습 모델이 82.8% 정확도, 82.6% AUC를 보였고, 치주과 의사는 79.7% 정확도, 79.3% AUC를 보였다. 대구치에서는 학습 모델이 73.4% 정확도, 73.4% AUC를 보였고, 치주과 의사는 76.6% 정확도, 76.4% AUC를 보였다.

이러한 구강악안면 영역에서 딥러닝을 활용한 진단에 관한 연구는 많이 진전되어 있는 단계이다. 더욱 흥미있는 사실은 데이터가 축적되고, 이러한 빅 데이터에 의해 학습된 모델은 더 정확하고 정교한 결과를 만들어 낼 수 있다는 사실이다. 연산 처리 장치의 발전은 이러한 속도를 더욱 가속화할 수 있다. 지금까지도 딥러닝에 의한 진단 능력이 상당한 수준에 이르렀음을 많은 연구들을 통해 확인할 수 있었고, 향후 연구에서의 결과가 더욱 기대된다.

3. 연구재료 및 방법

3.1. 데이터 준비

데이터 세트는 2018년부터 2019년까지 서울대 치과병원에 내원한 환자들의 1000장의 파노라마방사선영상으로 구성하였다. 파노라마방사선영상은 OP100 (Instrumentarium Imaging Inc., Tuusula, Finland)과 Rayscan alpha (Ray Co., Ltd., Hwaseong-si, Korea)를 사용하여 촬영되었으며, 모든 영상은 표준화된 절차와 방법으로 숙련된 방사선사가 촬영을 시행하였다. 연구에 사용된 1000장의 파노라마방사선영상에는 영상치의학전문의가 진단한 1691개의 치근단 병변들이 포함되어 있었다. 환자 1인당 1매의 파노라마방사선영상만을 사용하였으며, 혼합치열기의 영상은 제외하였다. 또한, 심한 노이즈나 흐림이 있는 방사선 영상들은 제외하였다. 이 연구는 서울대학교치과병원의 기관생명윤리위원회(IRB)에서 심의면제 승인(ERI19010)을 받고 진행되었으며, 모든 실험은 관련 지침과 규정에 따라 수행되었다.

3.2. 학습 데이터 영상의 절단 관련 영상 처리

파노라마방사선영상은 학습에 용이하도록 일정한 크기로 절단을 시행하였다. 이는 인공지능 모델을 이용하여 이루어졌으며, 사용한 신경망은 ResNet50을 이용한 모델이다[34]. 마지막 레이어에서는 GlobalAveragePooling2D를 이용하여 원하는 영상 크기의 좌측상단과 우측하단에 붉은 점의 X, Y 좌표를 얻게 된다 (Figure 2).

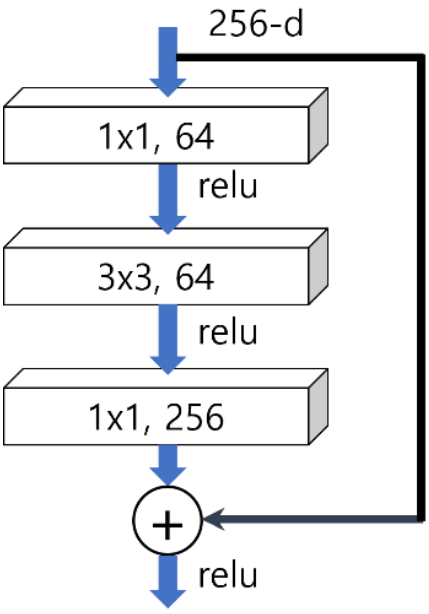

Figure 2. Architecture of neural network for cropping images using ResNet50. The network composed of three layers was used for cropping images. Panoramic images for training were automatically cropped by this model.

파노라마방사선영상의 좌측 상단과 우측하단의 점을 표시하여 인공지능 신경망에 학습시켜 구강의 크기나 영상의 치우침, 환자의 위치에 따른 영상의 왜곡 등에 관계없이 여러 상태의 파노라마방사선영상에서도 전체 치열을 포함하는 박스의 2개 점이 도출되도록 하였다. 그렇게 얻은 2개의 점을 기준으로 영상을 절단하여 최종적으로 학습하게 할 영상을 얻었다 (Figure 3).

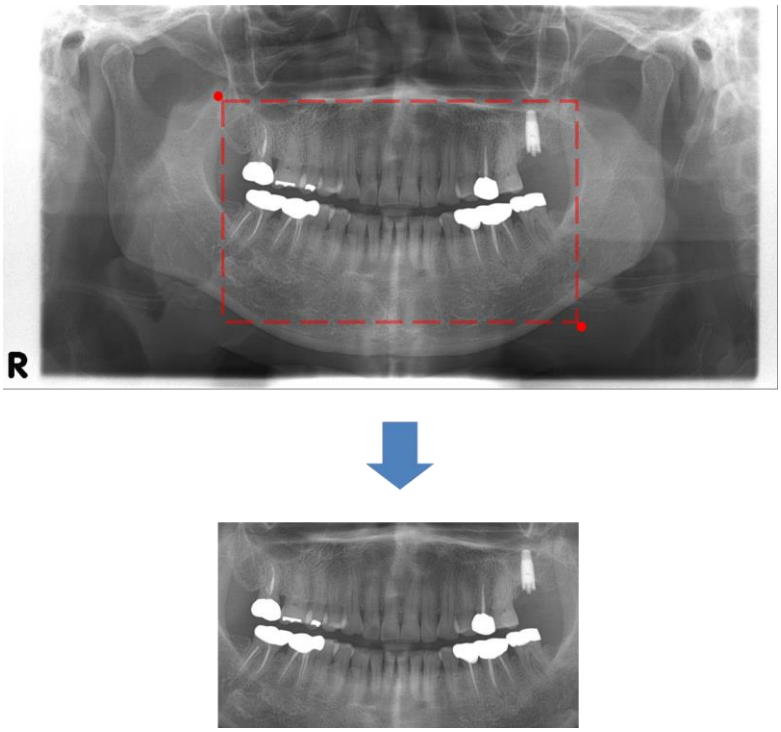

Figure 3. Cropped panoramic images for training. The model using ResNet50 marked two points. Based on these two points, panoramic images for training were cropped automatically.

3.3. 병변 영상의 잡음 제거 과정

영상 잡음을 제거하기 위해서 비국부 평균 잡음 제거 방식(non-local means denoising algorithm)을 사용하였다[35,36]. 국부 평균 잡음 제거 방식(local means denoising algorithm)은 주변의 화소 값을 평균값으로 변경하여 잡음을 제거하기 때문에 학습을 위한 병변 영상의 특이점 또한 사라지는 문제가 있다. 이에 반해 비국부 평균 잡음 제거 방식은 영상의 모든 화소의 평균값을 구하고, 특정 화소와 얼마나 유사한지 가중치가 부여되어 필터링 되는 방식이다. 국부 평균 잡음 제거 방식에 비해 선명도가 향상되고, 세부정보 손실이 줄어들게 된다.

3.4. 치근단 병변의 표기

치근단 병변들은 10년 이상의 경력을 가진 3명의 영상치의학전문의에 의해 진단되었으며, 2명 이상이 동의하는 것을 병변으로 인정하였다. 학습용 소프트웨어에 있는 폴리곤 표기 도구를 이용하여 파노라마방사선영상에서 수작업으로 병변을 표기하였으며, 병변의 외곽선을 빨간색 선으로 표기하였다 (Figure 4).

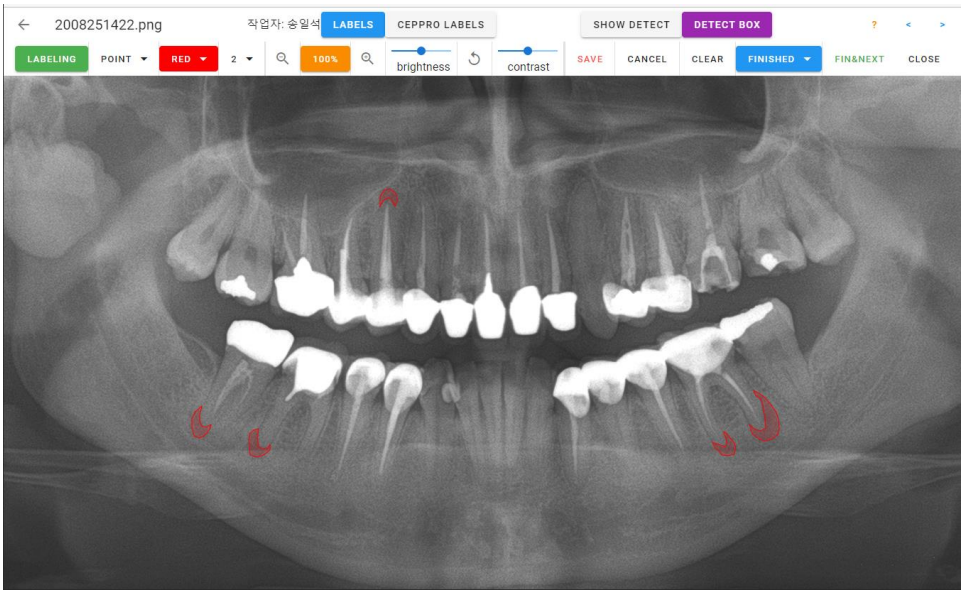

Figure 4. Training software called "Deep Stack." The lesions were labeled by outlining with polygon labeling tools.

3.5. 파노라마방사선영상의 전처리 및 증강

1976 x 976 화소의 해상도를 가진 총 1000장의 파노라마방사선영상을 960 x 480 화소의 해상도로 줄여서 PNG 파일 포맷으로 변환하였다. 데이터 세트는 데이터 증강을 하기 전에 훈련(800장[80%]), 검증(100장[10%]), 테스트(100장[10%])로 임의로 나누었다. 1511개의 치근단 병변이 학습을 위해 사용되었고, 180개의 치근단 병변이 테스트를 위해 사용되었다.

파노라마방사선영상의 수가 제한적이었기 때문에 학습 데이터를 늘리기 위해 데이터 증강이 시행되었다. 학습 데이터 세트는 온라인 증강 기법을 사용하여 각 단계에서 임의로 수평뒤집기, 흐림, 위치이동, 크기변화, 회전, 선예도, 양각처리, 대조도, 밝기, 그리드왜곡, 탄성변환 방법을 적용하여 증강되었다. 온라인 증강 기법은 최적화된 학습을 위해 각 단계에서 특정 증강 방법을 사용한다. 이 방법은 제한된 훈련 데이터 세트에서 전송 가능한, 더욱 효과적인 방법이다[37].

훈련 모델은 파이썬 오픈소스 프로그램 언어(v 3.6.12 ; Python Software Foundation, Wilmington, DE, USA; retrieved on August 17, 2020, from https://www.python.org/) 상에서 Pytorch (v 1.7.1) with CUDA (v 11.0)를 사용하여 최적화되었다. 학습은 4개의 V100 GPU core, Intel Xeon E5-2698 v4 CPU (Intel Corporation, Chandler, AZ, USA), 256 GB DDR4 RAM을 가진 NVIDIA DGX Station (Nvidia Corporation, Santa Clara, CA, USA)에서 진행되었다.

3.6. 심층 신경망 알고리즘의 구조

이 연구에서 미리 학습된 U-Net 심층 신경망이 전처리와 전이 학습을 위해 사용되었다. U-Net은 Olaf Ronneberger, Philipp Fishcer, Thomas Brox에 의해 2015년에 발표된 시맨틱 세그먼테이션 알고리즘 중 하나이며, 적은 데이터를 가지고 정확한 세그먼테이션을 할 수 있도록 Fully Convolutional Network (FCN)을 기반으로 고안된 알고리즘이다 (Figure 5) [38].

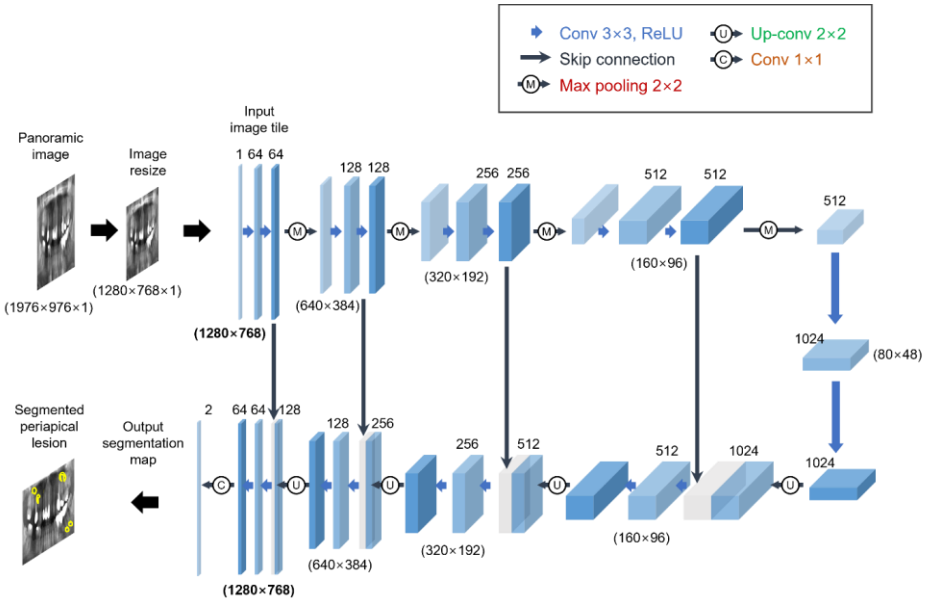

Figure 5. U-Net model architecture for the detection of apical lesions. Each blue box represents a multi-channel feature map. The number of channels is labeled on top of the box. The x-y-size is offered at the bottom of the box. White boxes represent duplicated feature maps. The arrows indicate the different operations.

U-Net 모델은 수축 경로(좌)와 확장 경로(우)의 두 가지 경로로 구성되어 있다[39]. 수축 경로에서 원본 방사선 영상들은 여러 레이어의 콘볼루션(3x3), 정류된 선형 유닛(ReLU) 활성화, 최대 풀링 연산(2x2) 및 잠재 공간으로의 압축을 하는 인코딩을 위한 입력으로 사용된다. 그리고 각 수축 과정의 단계에서 feature channel의 수는 2배가 된다. 확장 과정은 수축 과정이 역으로 진행되는 과정이다. CNN은 feature channel 수가 절반으로 감소한 feature map의 업콘볼루션 연산(2x2)을 통해 영상의 분할 마스크를 생성하고, 잠재 공간으로부터 압축된 정보를 해석하려고 시도하게 된다. 디코딩 연산의 나머지는 인코딩 과정(수축 경로에서 잘려진 feature map, 3x3 콘볼루션, ReLU)의 반대이다. 중요한 점은, 콘볼루션 과정 중에서 경계부 화소와 관련된 정보의 복원을 위해 잘려진 영상들이 반드시 필요하는 점이다. 64개 요소의 특징 벡터를 원하는 클래스 수에 매핑하기 위해 1x1 콘볼루션의 마지막 레이어가 최종적으로 사용된다. 그래서 네트워크는 23개의 콘볼루션 레이어를 가지게 된다.

3.7. 심층 신경망 모델의 검출과 분류 능력의 평가

개발된 CNN 모델의 성능은 모델을 학습할 때 사용하지 않은 테스트 데이터 세트만을 사용하여 평가되었다. 아래의 공식을 통해 IoU (intersection over union) 임계값에 따른 precision, recall, F1-score 값을 계산하였다. IoU는 영상 내에 있는 특정대상을 검출할 때 사용하는 평가지표로서 실제영역에 대해 인공지능 모델이 예측한 면적의 비율을 의미한다. 이 매개변수들은 CNN 모델의 성능을 평가하기 위해 일반적으로 사용된다.

$$Precision = \frac{TP}{FP + TP} \quad (1)$$

$$Recall = \frac{TP}{FN + TP} \quad (2)$$

$$F1 - score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad (3)$$

TP는 진양성, FP는 위양성, FN은 위음성을 나타낸다.

3.8. 히스토그램 분석

영상을 처리하는 과정에서 회색조수치를 기준으로 0~255까지의 값을 가진 화소 분포를 통해 히스토그램이 생성된다[40]. 낮은 값은 어두운 영상을 나타내고, 높은 값은 밝은 영상을 나타낸다. 이 히스토그램에서 치근단 병변은 81~91 사이의 회색조수치를 갖는 화소가 가장 많았다 (Figure 6).

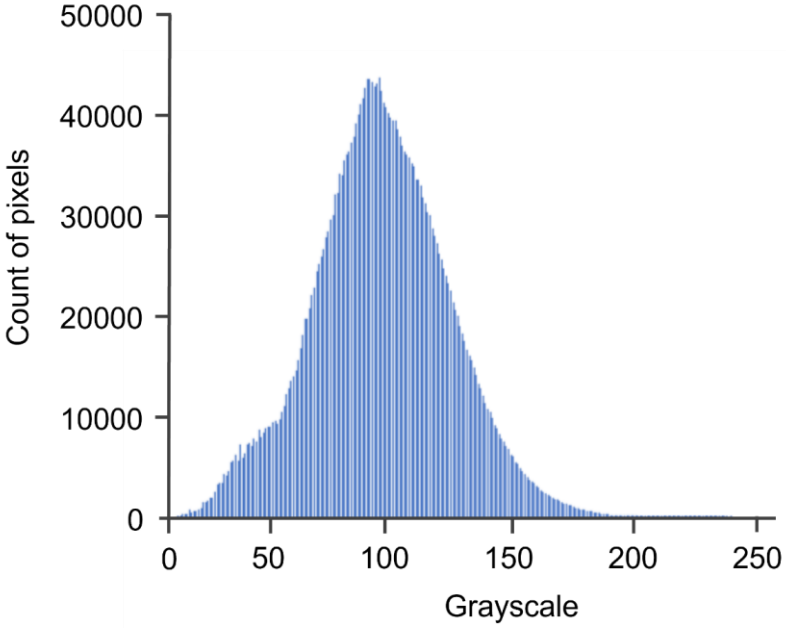

Figure 6. Intensity histogram. Overall consistency was identified without jagged edges for a specific labeled lesion.

4. 결 과

4.1. 딥러닝 세그먼테이션의 성능 평가

IoU (intersection over union) 임계값을 0.3으로 놓고 분석했을 때, 미리 학습된 CNN 모델은 180개 병변을 가진 테스트 그룹에서 147개의 병변을 검출하였다(진양성). 31개의 병변은 검출되지 못했고(위음성), 수동으로 표기되지 않았던 30개의 병변들이 검출되었다(위양성). IoU 임계값을 0.4으로 놓고 분석했을 때, 미리 학습된 CNN 모델은 180개 병변을 가진 테스트 그룹에서 143개의 병변을 검출하였다(진양성). 34개의 병변은 검출되지 못했고(위음성), 수동으로 표기되지 않았던 31개의 병변들이 검출되었다(위양성). IoU 임계값을 0.5으로 놓고 분석했을 때, 미리 학습된 CNN 모델은 180개 병변을 가진 테스트 그룹에서 125개의 병변을 검출하였다(진양성). 44개의 병변은 검출되지 못했고(위음성), 수동으로 표기되지 않았던 43개의 병변들이 검출되었다(위양성) (Table 1).

Table 1. The number of segmented periapical lesions by U-Net.

IoU 임계값	진양성	위음성	위양성
0.3	147	31	30
0.4	143	34	31
0.5	125	44	43

IoU: Intersection over Union

IoU 임계값을 0.3~0.5로 조정하면서 분석했을 때, IoU 임계값이 증가할수록 precision, recall, F1-score는 0.825-0.831 (IoU 임계값 ; 0.3)에서 0.740-0.744 (IoU 임계값 ; 0.5)로 점점 감소하였다. Precision과 recall 매개변수를 모두 반영하여 계산된 F1-score를 살펴보면, 더 낮은 IoU 값에서 더 높은 F1-score를 보였으며 이는 검출 능력의 향상을 의미한다(Table 2).

IoU 임계값을 0.3으로 설정했을 때, precision과 recall, F1-score는 0.831, 0.826, 0.828이었다. IoU 임계값을 0.4로 설정했을 때, precision과 recall, F1-score는 0.822, 0.808, 0.815이었고, IoU 임계값을 0.5로 설정했을 때, precision과 recall, F1-score는 0.744, 0.740, 0.742이었다(Table 2).

Table 2. Measurements of precision (positive predictive value), recall (sensitivity), and F1-score for periapical lesion segmentation by U-Net.

IoU 임계값	Precision	Recall	F1-score
0.3	0.831	0.826	0.828
0.4	0.822	0.808	0.815
0.5	0.744	0.740	0.742

IoU: Intersection over Union

Precision과 recall 값을 이용하여 precision-recall curve를 그릴 수 있으며, 학습시에 데이터 불균형이 있는 경우 이 precision-recall curve를 이용하여 모델의 성능 평가를 하게 된다 (Figure 7). 그래프 아래의 면적은 average precision을 의미하며, 그 수치는 0.746으로 측정되었다.

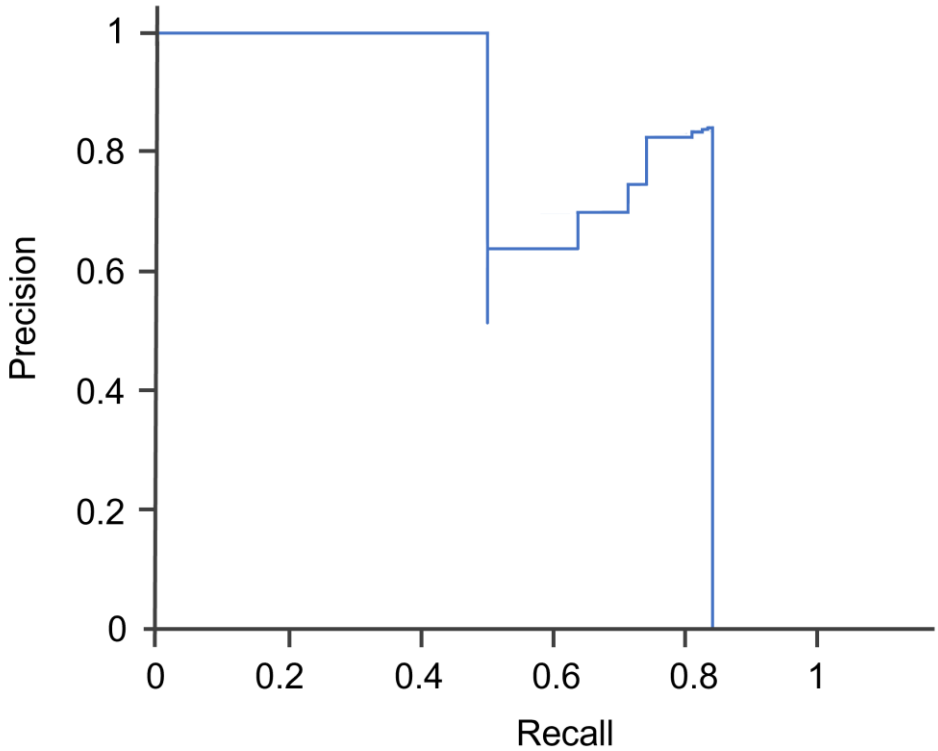

Figure 7. Precision-recall curve. The area under the curve means average precision.

4.2. 학습 epoch 증가에 따른 손실율 평가

학습과 검증 데이터 세트는 0~450 epoch 과정에서 평가되었고, 각 epoch는 전체 학습 데이터 세트를 통한 한번의 패스를 나타낸다 (Figure 8). 450 epoch를 반복하는 과정에서 학습과 검증 데이터 세트의 손실율은 감소하였다. 마지막 지점에서의 검증 데이터 세트의 손실율은 0.30, 학습 데이터 세트의 손실율은 0.18이었다. 학습과 검증 데이터 세트의 손실율은 400 epoch가 넘어가면서 값의 변화가 거의 없었으며, 이는 더 이상의 향상이 없음을 나타낸다.

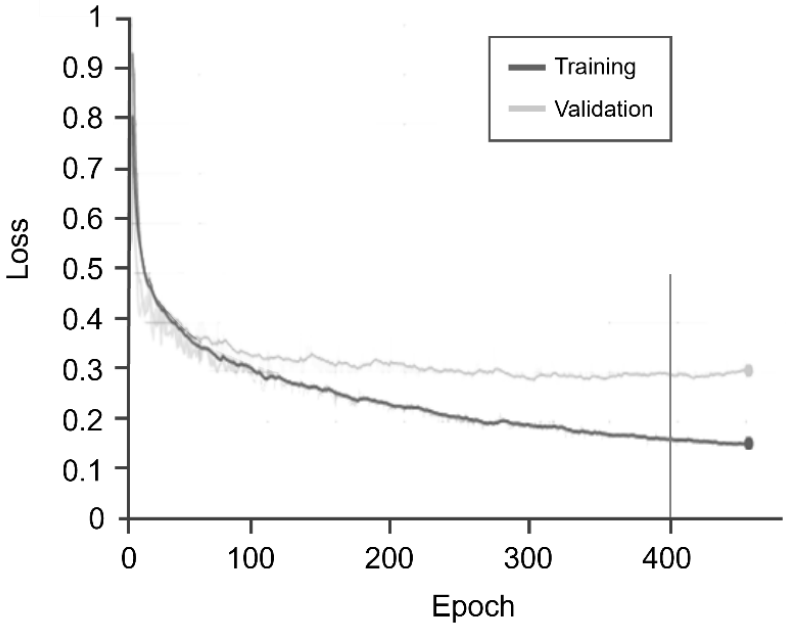

Figure 8. Training and validation curves for the pre-trained augmented model. Training and validation are performed for 450 epochs with each epoch representing 1 pass through the entire training and validation datasets. Overfitting is minimized under 400 epochs.

4.3. 학습된 딥러닝 모델의 치근단 병변 검출

U-Net을 이용해 미리 학습된 CNN 모델을 통해 치근단 병변이 검출되었다. 원본 영상이 테스트 데이터 세트로 준비되었다 (Figure 9a). 하얀색 선은 수동으로 표기한 것이고, 이것이 미리 학습된 CNN에 의해 검출되는 영역과 비교하기 위한 ground truth를 의미한다 (Figure 9b). 그리고 하얀색 영역은 CNN 모델이 치근단 병변의 영역을 표기한 것을 나타낸다 (Figure 9c). U-Net에 의해 검출된 영역은 수동으로 표기한 것과 유의미한 차이가 나지 않았다.

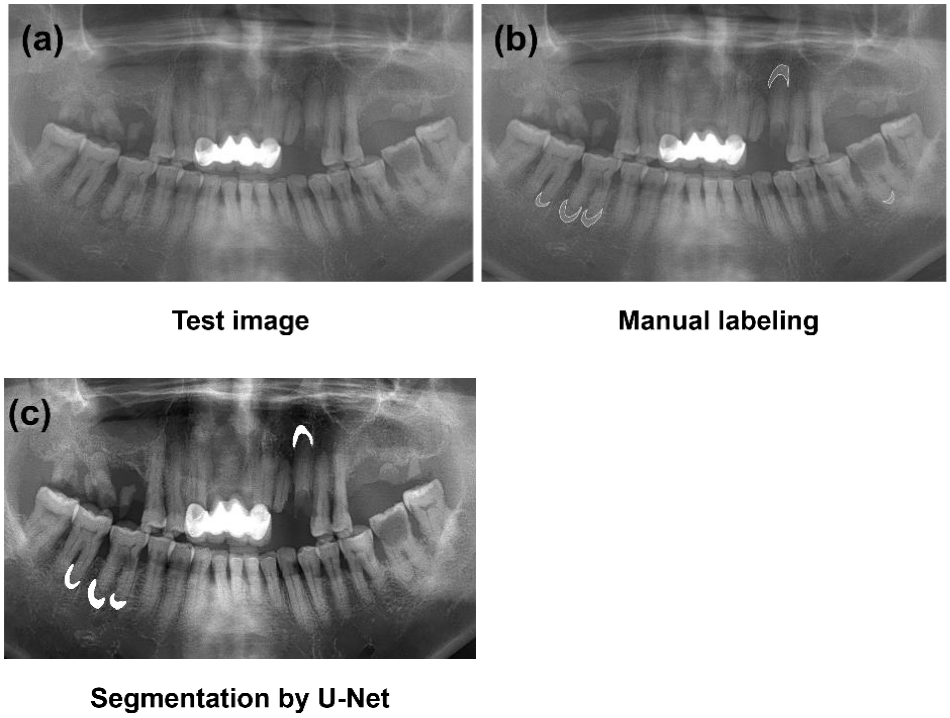

Figure 9. An example of apical lesion segmentation from panoramic radiographs using U-Net. Original panoramic radiograph is prepared for testing (a). White outlines denote the manual labeling (b) and white areas denote the CNN-generated areas by U-Net (c).

4.4. 다른 딥러닝 모델과의 성능 비교

시맨틱 세그먼테이션의 대표적인 모델은 U-Net과 FPN이다[41]. FPN 신경망 모델은 U-Net 구조와 유사하게 bottom-up 경로와 top-down 경로의 2가지 프로세스로 구성되어 있는 피라미드 네트워크가 특징이다 (Figure 10) [42].

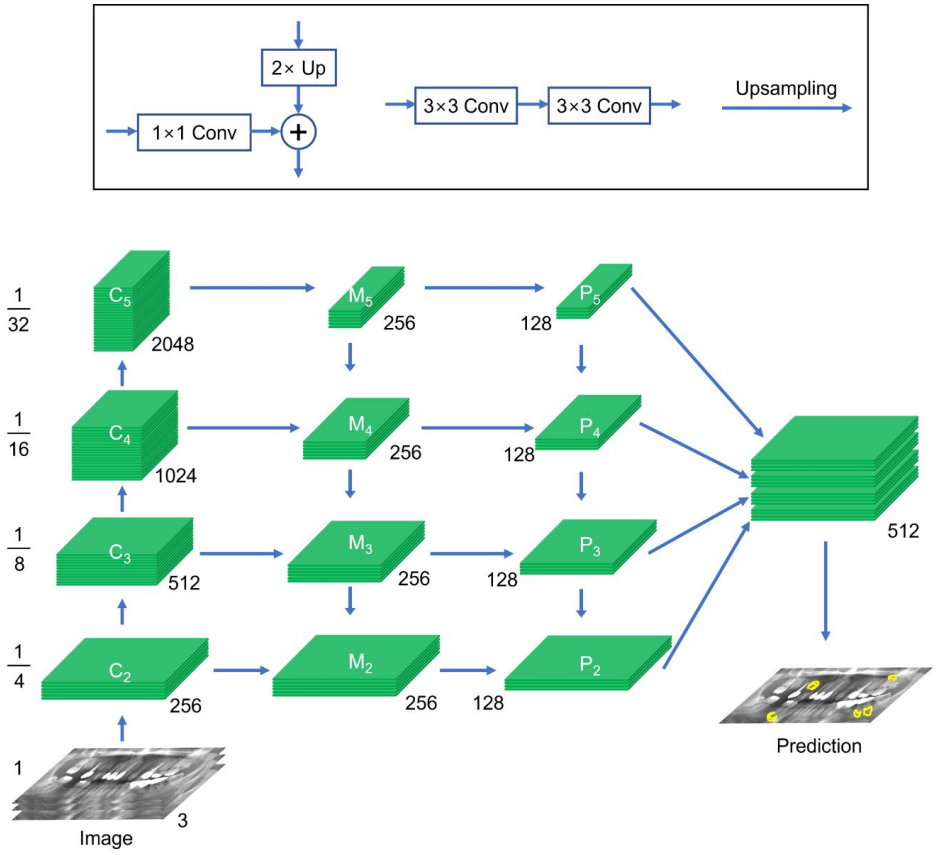

Figure 10. Feature pyramid network with ResNet50 encoder. In the left part of the architecture, the number of channels increases stepwise and the feature map decreases stepwise. In the last step, it is upsampled to the original image through feature mapping.

본 연구에서는 U-Net과 FPN 모델을 모두 활용하여 테스트를 진행하였다, IoU 임계값 0.3~0.5에서 precision과 recall, F1-score를 확인했을 때 근소한 차이로 U-Net에서 더 우수한 성능을 보여주었다(Table 3).

Table 3. Precision, recall, and F1-score for periapical lesion segmentation by U-Net and FPN.

IoU Threshold	U-Net			FPN		
	Precision	Recall	F1-score	Precision	Recall	F1-score
0.3	0.846	0.880	0.863	0.836	0.858	0.847
0.4	0.833	0.859	0.846	0.817	0.825	0.821
0.5	0.771	0.797	0.784	0.716	0.719	0.718

5. 총괄 및 고찰

치근단 병변을 진단하는 데 있어서 일반적으로 이용하는 방사선 영상은 치근단방사선영상과 파노라마방사선영상이다[43]. 특히 파노라마방사선영상은 환자의 초진 검사시에 일반적으로 촬영하게 되며, 한번의 촬영으로 치아와 악골 전체를 검사할 수 있기 때문에 진단이 어려운 초기 단계의 치근단 병변도 촬영 부위에 포함할 수 있다[44]. 하지만 이러한 방사선 검사는 영상을 판독하는 검사자의 경험에 따라 신뢰도에 차이가 발생할 수 있다.

딥러닝 기술의 발전을 통해 검사자의 능력에 영향을 받지 않고 인공지능 모델에 의해 치근단 병변을 진단하는 연구들이 있었다. 이러한 연구들에서는 주로 객체검출 알고리즘 모델인 YOLO 신경망을 사용하여 병변을 사각형 박스의 형태로 검출하는 방법을 사용하였다[45-47]. U-Net 신경망을 이용한 이전 연구가 있지만 신경망 학습을 위해 사용된 영상의 수가 적었다[30].

본 연구에서는 치근단 병변의 크기 자체가 작기 때문에 화소 단위의 접근이 필요하다고 판단하여 U-Net 신경망을 활용한 시맨틱 세그먼테이션 알고리즘 모델을 활용하였으며, IoU 임계값 0.3에서 F1-score가 0.828을 보여 높은 성능을 확인하였다.

딥러닝 기반의 치근단 병변의 검출은 검사자의 능력에 대한 의존도를 최소화하면서 치근단 병변의 자동 검출을 효과적이고 효율적으로 가능하게 하였다. 이번 연구에서는 치근단 병변을 검출하기 위해서

파노라마방사선영상을 사용하여 학습을 진행하였고, 인공지능을 이용한 초기 단계의 치근단 병변의 진단 가능성을 확인하였다.

데이터 증강은 CNN 모델을 학습하는데 일반적으로 사용된다[25,27,33]. 영상 분류, 객체검출 및 세그먼테이션을 위한 딥러닝 과정에서 데이터 증강은 꼭 필요한 과정이다[48]. 심층 신경망에는 제한된 학습데이터에 과적합되는 많은 매개변수가 있으며, 데이터 증강은 학습 데이터의 양과 다양성을 모두 증가시켜 과적합을 방지하고 일반화를 개선하는 데 사용된다[49].

이 연구는 효율적인 학습을 위해 적대적 훈련과 메타 학습을 모두 통합하는 온라인 데이터 증강을 채택하였으며, 온라인 방식으로 데이터 증강 및 대상 신경망의 학습을 최적화하였다. 온라인 증강의 장점은 오프라인 증강의 반대되는 특성을 가지고 있다. 온라인 증강 네트워크는 사전 학습이나 학습의 조기 중단 문제를 배제하면서 처음부터 끝까지 온라인 학습을 통해 대상 네트워크에 적용된다. 그리고 오프라인 학습 방법은 병렬 최적화 처리과정이 많이 때문에 일반적으로 분산 학습에 의존하지만 온라인 데이터 증강은 간단하고 학습하기 쉽다[50].

IoU는 두 가지 물체의 위치가 얼마나 일치하는지를 나타내는 수학적 지표이다[51]. 인공지능을 이용한 진단에서 실제 병변의 위치와 예측된 병변의 위치를 평가하는 방법으로 사용될 수 있다. IoU 임계값이 1에 가까울수록 실제 병변과 예측된 병변의 면적이 많이 일치해야 참으로 인식하고, 0에 가까울수록 실제 병변과 예측된 병변이 조금만 일치해도 참으로 인식하게 된다. 본 연구에서는 IoU 임계값을 0.3, 0.4, 0.5로

각각 설정하여 인공지능 모델의 성능을 평가하였다. 이 IoU 임계값은 치근단 병변이 아닌 다른 질환의 진단 연구에서 활용된 값보다는 다소 완화된 수치로 진단 기준을 설정한 점이 있다. 하지만 화소 단위의 학습을 하는 시맨틱 세그먼테이션 알고리즘을 활용한 치근단 병변의 검출에서는 작은 범위만 일치하더라도 치근단 병변의 예상 영역을 표시하는 것이라고 인정할 수 있으며, 유효한 결과를 나타낸다고 볼 수 있다.

모든 파노라마방사선영상은 학습 시간, 연산 장치의 비용과 저장 공간을 고려하여, 영상을 전체 치열을 포함하도록 자르고 960 x 480 화소의 해상도로 줄였다. 더 높은 해상도의 영상을 데이터 세트로 이용한다면 정확도와 구별 능력을 향상시킬 수 있을 것이며, 이는 이번 연구의 한계점으로 볼 수 있다[52,53]. 더 높은 성능의 연산 장치는 고해상도의 영상을 학습 과정에 사용할 수 있게 할 것이다.

이번 연구에서는 영구치를 가진 파노라마방사선영상만 사용되었고, 혼합치열기의 파노라마방사선영상은 제외하였다. 그 이유는 유치 아래에 영구치배가 중첩되어 보이기 때문에 학습이 어려울 것으로 판단하였고, 혼합치열기에 있는 영구치배의 근단부에 존재하는 치낭은 치근단 병변으로 인식될 가능성이 있기 때문이다. 혼합치열기의 파노라마방사선영상을 학습 데이터로 사용하기 위해서는 더 연구가 필요해 보인다.

인공지능 모델이 검출한 치근단 병변의 정확도를 평가하기 위해 10년 이상의 경력을 가진 3명의 영상치의학전문의가 진단하여 2명 이상이

동의하는 것을 ground truth로 설정하였다. 영상치의학전문의가 병변을 검출하고 직접 병변의 경계선을 표시하였기 때문에 이러한 과정에서 주관적인 관점이 들어갈 수밖에 없다[54]. 객관적인 ground truth를 설정할 수 있다면 더욱 정확한 학습이 가능할 수 있기 때문에 이에 대한 고민이 필요할 것으로 보인다.

이번 연구에서 사용된 파노라마방사선영상은 한 의료기관에서 촬영된 파노라마방사선영상만을 이용하여 학습 및 테스트가 진행되었다. 따라서 촬영장비가 달라지고 파노라마방사선영상의 질이나 영상의 특징이 바뀐다면 학습된 모델의 성능이 달라질 수 있다. 그래서 본 연구에서 개발된 인공지능 모델을 범용성 있게 사용하는 것은 추가적인 연구가 필요할 것으로 보인다. 다른 촬영장비를 이용하여 촬영한 파노라마방사선영상을 학습 데이터에 포함시킨다면 좀 더 일반적인 치근단 병변의 진단 도구로서 인공지능 모델을 활용할 수 있을 것이다.

6. 결 론

이번 연구에서 치근단 병변의 감지 및 진단을 위한 딥러닝의 잠재적 유용성이 확인되었다. U-Net 알고리즘은 파노라마방사선영상에서 치근단 병변을 감지하는 데 상당히 좋은 성능을 제공하였다. CNN에 기반한 병변의 검출은 치과 및 의료 질환을 조기에 발견할 수 있는 유능한 조력자가 될 수 있을 것이다.

참고문헌

1. Leonardi Dutra K, Haas L, Porporatti AL, Flores-Mir C, Nascimento Santos J, Mezzomo LA, et al. Diagnostic accuracy of cone-beam computed tomography and conventional radiography on apical periodontitis: a systematic review and meta-analysis. J Endod 2016; 42: 356-64.
2. Choi JW. Assessment of panoramic radiography as a national oral examination tool: review of the literature. Imaging Sci Dent 2011; 41: 1-6.
3. Nardi C, Calistri L, Grazzini G, Desideri I, Lorini C, Occhipinti M, et al. Is panoramic radiography an accurate imaging technique for the detection of endodontically treated asymptomatic apical periodontitis? J Endod 2018; 44: 1500-8.
4. Nardi C, Calistri L, Pradella S, Desideri I, Lorini C, Colagrande S, et al. Accuracy of orthopantomography for apical periodontitis without endodontic treatment. J Endod 2017; 43: 1640-6.
5. Parker JM, Mol A, Rivera EM, Tawil PZ. Cone-beam computed tomography uses in clinical endodontics: observer

variability in detecting periapical lesions. J Endod 2017; 43: 184-7.

6. Nagi R, Aravinda K, Rakesh N, Gupta R, Pal A, Mann AK. Clinical applications and performance of intelligent systems in dental and maxillofacial radiology: a review. Imaging Sci Dent 2020; 50: 81-92.

7. Hwang JJ, Jung YH, Cho BH, Heo MS. An overview of deep learning in the field of dentistry. Imaging Sci Dent 2019; 49: 1-7.

8. Russel S, Norvig P. Artificial intelligence: a modern approach: Pearson; 2021.

9. McCarthy, J. Artificial intelligence, logic and formalizing common sense. In: Thomason, R.H. (eds) philosophical logic and artificial intelligence. Dordrecht, The Netherlands ; Boston ; London : Kluwer academic publishers, c1989.

10. Olazaran M. A Sociological study of the official history of the perceptrons controversy. Social Studies of Science 1996; 26: 611–59.

11. Lee HM, Chen CM, Huang TC. Learning efficiency improvement of back-propagation algorithm by error saturation prevention method. Neurocomputing 2001; 41: 125-43.

12. Karlik B, Olgac AV. Performance analysis of various activation functions in generalized MLP architectures of neural networks. Int J Artif Intell Exp Sys 2011; 1: 111−22.

13. Pal S, Ebrahimi, E, Zulfiqar A, Fu Y, Zhang V, Migacz S, et al. Optimizing multi−GPU parallelization strategies for deep learning training. IEEE Micro 2019; 39: 91−101.

14. Samuel, Arthur L. Some studies in machine learning using the game of checkers. IBM J Res Develop 1959; 44: 206–26.

15. Jordan MI, Mitchell TM. Machine learning: trends, perspectives, and prospects. Science 2015; 349: 255−60.

16. Erhan D, Courville A, Bengio Y. Why does unsupervised pre−training help deep learning? Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics (AISTATS) 2010; 9: 201−8.

17. Feng D, Haase−Schutz C, Rosenbaum L, Hertlein H, Glaser C, Timm F, et al. Deep multi−modal object detection and semantic segmentation for autonomous driving: datasets, methods, and challenges. IEEE Trans Intell Trans Sys 2021; 22: 1341−60.

18. Wang H, Yu Y, Cai Y, Chen X, Chen L, Liu Q. A comparative study of state−of−the−art deep learning algorithms for

vehicle detection. IEEE Intell Trans Sys Mag 2019; 11: 82-95.

19. Wolf L, Hassner T, Maoz I. Face recognition in unconstrained videos with matched background similarity. Conference on Computer Vision and Pattern Recognition (CVPR) 2011; 529-34.

20. Connolly J, Granger E, Sabourin R. An adaptive classification system for video-based face recognition. Info Sci 2012; 192: 50-70.

21. Guo C, Yang Y. Implementation of a specified face recognition system based on video. IEEE 4th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC) 2019; 79-84.

22. Schmidhuber J. Deep learning in neural networks: an overview. Neural Netw 2015; 61: 85-117.

23. Ekert T, Krois J, Meinhold L, Elhennawy K, Emara R, Golla T, et al. Deep learning for the radiographic detection of apical lesions. J Endod 2019; 45: 917-22.

24. Yüksel AE, Gültekin S, Simsar E, Özdemir ŞD, Gündoğar M, Tokgöz SB, et al. Dental enumeration and multiple treatment detection on panoramic X-rays using deep learning. Sci Rep 2021; 11: 12342.

25. Kwon O, Yong TH, Kang SR, Kim JE, Huh KH, Heo MS, et al. Automatic diagnosis for cysts and tumors of both jaws on panoramic radiographs using a deep convolution neural network. Dentomaxillofac Radiol 2020; 49: 20200185.

26. Choi HR, Siadari TS, Kim JE, Huh KH, Yi WJ, Lee SS, et al. Automatic detection of teeth and dental treatment patterns on dental panoramic radiographs using deep neural networks. Forensic Sci Res 2022; 1-11.

27. Lee JH, Kim DH, Jeong SN, Choi SH. Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. J Dent 2018; 77: 106-11.

28. Khan HA, Haider MA, Ansari HA, Ishaq H, Kiyani A, Sohail K, et al. Automated feature detection in dental periapical radiographs by using deep learning. Oral Surg Oral Med Oral Pathol Oral Radiol 2021; 131: 711-20.

29. Setzer FC, Shi KJ, Zhang Z, Yan H, Yoon H, Mupparapu M, et al. Artificial intelligence for the computer-aided detection of periapical lesions in cone-beam computed tomographic images. J Endod 2020; 46: 987-93.

30. Bayrakdar IS, Orhan K, Çelik Ö, Bilgir E, Sağlam H, Kaplan FA, et al. A U-Net approach to apical lesion segmentation on panoramic radiographs. Biomed Res Int 2022; 15: 7035367.

31. Devito KL, de Souza Barbosa F, Filho WNF. An artificial multilayer perceptron neural network for diagnosis of proximal dental caries. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 2008; 106: 879-84.

32. Park JH, Hwang HW, Moon JH, Yu Y, Kim H, Her SB, et al. Automated identification of cephalometric landmarks: Part 1-comparisons between the latest deep-learning methods YOLOV3 and SSD. Angle Orthod 2019; 89: 903-9.

33. Lee JH, Kim DH, Jeong SN, Choi SH. Diagnosis and prediction of periodontally compromised teeth using a deep learning-based convolutional neural network algorithm. J Periodontal Implant Sci 2018; 48: 114-23.

34. Tian X, Chen C. Modulation pattern recognition based on Resnet50 neural network. IEEE 2nd International Conference on Information Communication and Signal Processing (ICICSP) 2019; 34-8.

35. Buades A, Coll B, Morel JM. Non-local means denoising. Image Processing On Line 2011; 1: 208–12.

36. Coupe P, Yger P, Prima S, Hellier P, Kervrann C, Barillot C. An optimized blockwise nonlocal means denoising filter for 3-D magnetic resonance images. IEEE Trans Med Imaging 2008; 27: 425-41.

37. Wei P, Ball JE, Anderson DT. Fusion of an ensemble of augmented image detectors for robust object detection. Sensors 2018; 18: 894.

38. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. Medical Image Computing and Computer-Assisted Intervention (MICCAI) 2015; 234-41.

39. Jean LA, Nick H, Daniel OB, Lena HL, Tobias B, Sonja H, et al. Deeply supervised UNet for semantic segmentation to assist dermatopathological assessment of basal cell carcinoma. J Imaging 2021; 7: 71.

40. Park GH, Cho HH, Choi MR. A contrast enhancement method using dynamic range separate histogram equalization. IEEE Trans Consumer Electronics 2008; 54: 1981-7.

41. Xu Z, Lu T. Joints Steering Gear Detection based on semantic segmentation. International Conference on Control, Robotics and Cybernetics (CRC) 2020; 196-201.

42. Seferbekov S, Iglovikov V, Buslaev A, Shvets A. Feature pyramid network for multi-class land segmentation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2018; 272-275.

43. Arslan ZB, Demir H, Berker Yıldız D, Yaşar F. Diagnostic accuracy of panoramic radiography and ultrasonography in detecting periapical lesions using periapical radiography as a gold standard. Dentomaxillofac Radiol 2020; 49: 20190290.

44. Rushton VE, Horner K. The use of panoramic radiology in dental practice. J Dent 1996; 24:185-201.

45. Tao J, Wang H, Zhang X, Li X, Yang H. An object detection system based on YOLO in traffic scene. 6th International Conference on Computer Science and Network Technology (ICCSNT) 2017; 315-9.

46. Liu C, Tao Y, Liang J, Li K, Chen Y. Object detection based on YOLO network. IEEE 4th Information Technology and Mechatronics Engineering Conference (ITOEC) 2018; 799-803.

47. Lee YH, Kim YS. Comparison of CNN and YOLO for object detection. J Semiconductor Display Technology 2020; 19: 85-92.

48. Chlap P, Min H, Vandenberg N, Dowling J, Holloway L, Haworth A. A review of medical image data augmentation techniques for deep learning applications. J Med Imaging and Rad Oncol 2021; 65: 545-563.

49. Ganesan P, Rajaraman S, Long R, Ghoraani B, Antani S. Assessment of data augmentation strategies toward performance improvement of abnormality classification in chest radiographs. Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) 2019; 841-844.

50. Liu G, Zhou W, Tian L, Liu W, Liu Y, Xu H. An efficient and accurate iris recognition algorithm based on a novel condensed 2-ch deep convolutional neural network. Sensors 2021; 21: 3721.

51. Rahman MA, Wang Y. Optimizing intersection-over-union in deep neural networks for image segmentation. Adv Visual Computing 2016; 234-44.

52. Sabottke CF, Spieler BM. The Effect of image resolution on deep learning in radiography. Radiol Artif Intell 2020; 2: e190015.

53. Lakhani P, Sundaram B. Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. Radiology 2017; 284: 574-82.

54. Lu SL, Xiao FR, Cheng JC, Yang WC, Cheng YH, Chang YC, et al. Randomized multi-reader evaluation of automated detection and segmentation of brain tumors in stereotactic

radiosurgery with deep neural networks. Neuro Oncol 2021; 23: 1560–8.

Abstract

Deep learning-based periapical lesion segmentation from panoramic radiographs

Il-Seok Song, D.D.S

Program in Oral and Maxillofacial Radiology

Department of Dental Science

Graduate School, Seoul National University

(Directed by Prof. Min-Suk Heo, D.D.S., M.S., Ph.D)

Objectives

Convolutional neural networks (CNNs) have rapidly emerged as one of the most promising artificial intelligence (AI) in the field of medical and dental researches, which can provide an effective diagnostic methodology allowing for detection of diseases at early

age. This study was, thus, aimed to evaluate performances for periapical lesion segmentation from panoramic radiographs using a deep CNN algorithm.

Material and Methods

A total of 1000 panoramic images showing periapical lesions were separated into training (n=800, 80%), validation (n=100, 10%), and test (n=100, 10%) dataset, respectively. All panoramic images were cropped using ResNet50 network. U-Net, representative network of semantic segmentation, was adopted for this study. The performance of identifying periapical lesions was evaluated after calculating precision, recall, and F1-score under specific IoU threshold.

Results

In the test group of 180 periapical lesions, 147 lesions were segmented from panoramic radiographs under IoU threshold of 0.3. The performance value of F1-score were 0.828, 0.815, and 0.742, respectively, under IoU threshold of 0.3, 0.4, and 0.5.

Conclusion

This study showed the potential utility of deep learning-guided approach for the segmentation of periapical lesions. The deep CNN algorithm using U-Net allows for a considerably high performance in detecting periapical lesions.

Keyword: Artificial Intelligence; Deep Learning; Periapical Periodontitis; Radiography, Panoramic

Student Number: 2016-33728