



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](#)

치의과학석사 학위논문

치과용 임플란트의
조직학적 분석 보조 수단으로서
인공지능 딥러닝 기법의
활용 가능성에 대한 연구
: 주변 조직과의 경계면 분할에 대한 고찰

2022년 8월

서울대학교 대학원
치의학과 구강병리학전공
이 민 선

국문초록

목 적

이 연구의 목적은 치과용 임플란트 조직 이미지의 뼈 조직형태계측에서 인공지능 딥러닝 기법이 효율적이고 일관적인 보조수단으로 활용이 가능한 지 고찰하는 것이다. 이를 위해 염색된 색상 기반의 선행 개발 방식과 딥러닝 기법을 비교하여 임플란트와 주변 조직으로의 이미지 분할 정확도를 평가하였다.

방 법

본 연구에서는 토끼의 경골에 식립된 치과용 임플란트 표본 16장에서 200배율로 촬영된 이미지를 사용하였다. Ground truth로 사용하기 위한 이미지 주석(annotation) 처리를 통해 하나의 이미지 전체를 임플란트 영역과 조직 영역의 두가지 클래스로 구분하였다. 관심 영역은 두가지 클래스의 경계선을 중심으로 설정하였다. 연구에서 사용한 인공지능 방법론은 kNN과 딥러닝 U-Net 모델 두가지로, kNN은 색상을 나타내는 각 클래스의 대표 픽셀값(instance)에 기반하여 분할이 이루어진다. 딥러닝 U-Net 모델은 인공지능 기반 의료 이미지 분할 연구에서 활발하게 사용 중인 모델로, 본 연구는 가장 기본적인 구조인 Vanilla U-Net을 사용하였다. 두가지 방법론 간의 결과를 비교하기 위해 딥러닝 훈련에 사용하지 않은 테스트 이미지를 사용하였고, Ground truth 이미지와 예측된 이미지를 픽셀 단위로 비교하여 정확도와 민감도, Area under ROC curve(AUROC)로 평가하였다. 그리고 Intersection Over Union(IOUS), Dice similarity coefficient(DSC)의 두가지 지표로 나타내어 분할 정확도를 비교하였다.

결 과

kNN에서 임의로 한 장의 이미지를 선택하여 instance를 지정하였고, 모든 이미지 데이터에서 예측 결과를 추출하여 Ground truth와의 정확도를 IOU 지표로서 비교하였다. 관심 영역의 정확도는 65~96% 범위에서 평균 90.39%로 나타났다. 이중 65.83%로 가장 낮은 정확도를 보인 이미지에 맞추어 다시 instance를 지정하였을 때 해당 사진은 91.39%까지 정확도가 상승하였고, 예측 결과 이미지에서 각 클래스의 서로 영역으로 침범 없이 경계가 명확해지는 양상을 확인하였다. 이를 통해 평가하고자 하는 조직 이미지가 변경될 때마다 그 이미지에 맞추어 기준값을 조정해야 하는 관찰자의 수동 개입이 매번 필요하다는 한계점을 확인하였고, 이는 선행 개발된 염색된 색상 기반 방식에서 나타나는 한계점과 동일하였다.

관심 영역에서 임플란트를 기준으로 sensitivity는 테스트 이미지 평균 kNN은 90.68%, 데이터 증강을 진행한 U-Net 모델에서는 98.32%, accuracy는 kNN에서 94.89%, U-Net 모델에서 97.86%, AUROC는 kNN에서 93.69%, U-Net 모델에서 99.79%로 나타났다. IOU는 kNN에서 84.19%, U-Net 모델에서 93.4%였으며, DSC는 kNN에서 91.03%, U-Net 모델에서 96.58%를 보이며 딥러닝 기법을 사용하였을 때 기존의 색상 기반 분할 방식과 비교하여 우수한 성능이 확인되었다. 더불어 분할 결과 이미지를 확인하였을 때 kNN에 비해 두가지 클래스로의 분할이 서로의 영역으로 침범 없이 더욱 명확한 양상을 확인하였다.

결 론

인공지능 딥러닝 기법은 치과용 임플란트의 조직학적 분석에서 중요한 지표인 Bone to implant contact(BIC)와 Bone area(BA)를 측정할 때 나타날 수 있는 관찰자 간의 결과 차이를 보완하며 분석의 일관성과 효율성을 향상시키는 보조 도구로의 발전 가능성이 확인되었다.

다양한 양상을 포함한 조직 이미지를 확보하고, 전문가의 주석을 토대로 한 훈련이 동반된다면 조직 영역에서 뼈와 뼈가 아닌 영역으로의 분할 정확도까지 달성하여 병리학 실험실에서 뼈조직 분석의 보조 수단으로 활용할 수 있을 것으로 판단하였다.

주요어 : 뼈 조직형태계측(bone histomorphometry), 치과용 임플란트 분할 (dental implant segmentation), 딥러닝(deep learning), 의료 이미지 분석 (medical image analysis)

학 번 : 2020-24988

목 차

제 1 장 서론	1
제 1 절 연구의 배경 및 목적	1
제 2 장 연구 방법	5
제 1 절 데이터베이스 구축	5
1. 데이터 수집	5
2. 이미지 주석 처리 및 ground truth 이미지 생성	7
3. 관심영역(Region of interest, ROI) 설정 방법	7
제 2 절 인공지능 기계학습 방법론	8
1. kNN	8
2. U-Net	9
제 3 절 인공지능 딥러닝 실험 방법	11
1. 데이터셋 준비	11
2. 데이터 증강	13
3. 모델 매개변수(parameter) 선정	13
4. 역치(threshold) 설정을 통한 결과 이미지 이분화	13
5. 최종 예측 이미지 생성	13
제 4 절 기계학습 실험 환경	14
제 5 절 평가 방법	14
제 3 장 연구 결과	15
제 1 절 딥러닝 데이터 구성	15
제 2 절 임플란트와 조직으로의 경계 분할 결과	16
1. kNN 결과	16
2. 딥러닝 모델 파라미터 선정	34

3. 딥러닝 결과 비교	36
제 4 장 고찰	49
참고문헌	54
Abstract	62

표 목 차

[표 1] 하나의 이미지에 맞춰 instance를 지정한 kNN 결과 - 전체이미지	18
[표 2] 하나의 이미지에 맞춰 instance를 지정한 kNN 결과 - 관심영역	25
[표 3] 'No.13'에 맞춰 instance를 지정한 kNN 결과	32
[표 4] 테스트셋 4장의 관심영역에서 딥러닝과 kNN 결과 비교(1) - Accuracy, Sensitivity, Specificity, AUROC	37
[표 5] 테스트셋 4장의 관심영역에서 딥러닝과 kNN 결과 비교(2) - IOU, DSC	41

그 림 목 차

[그림 1] 토끼의 경골에 식립된 임플란트조직 슬라이드 이미지	5
[그림 2] 주석 처리를 통한 Ground truth 설정	7
[그림 3] 관심영역 설정 과정	8
[그림 4] U-Net 모델 구조도 및 입력/출력 영상 시각화 ..	10
[그림 5] 딥러닝 전체 흐름도	11
[그림 6] 패치 생성 및 mirroring 과정	12
[그림 7] 데이터셋 구성	15
[그림 8] 16장 전체 이미지에서 kNN 결과	19
[그림 9] 16장 관심영역에서 kNN 결과	26
[그림 10] 정성적인 부분에서 한계를 보인 kNN 결과 제시	31
[그림 11] 평가 이미지에 맞추어 매번 instance를 조정해야 정확도와 향상된 경계 분할이 확보되는 한계점 제시	33

[그림 12] 학습 수에 따른 정확도 산출 그래프	35
[그림 13] 평가지표 IOU 산출 도표	38
[그림 14] 평가지표 DSC 산출 도표	39
[그림 15] 방법론 간 결과 이미지 비교	42
[그림 16] 딥러닝 Data augmentation 전/후 결과에서 정성적으로 향상된 케이스 제시	47

제 1 장 서론

제 1 절 연구의 배경 및 목적

치과용 임플란트(Dental implant)는 21세기 초반부터 상실된 치아를 회복하기 위한 최우선 순위의 치료법으로 선택되었으며[1], 국내에서 가장 많이 생산되는 의료기기이다[2]. 장기적인 임상적 성공을 위하여 다양한 생체재료와 표면처리 기술을 이용한 연구 개발이 지속되고 있으며, 이 과정에는 동물 실험을 포함한 실험실 단계에서 임플란트 안정성 확인을 위해 골 유착(osseointegration)을 평가하는 것이 필수적이다[3]. 이를 평가하는 방법 중 가장 신뢰도가 높은 것은 ‘조직형태계측학적 분석(Histomorphometric analysis)’으로[4], 골유착 정도를 나타내는 중요한 지표에는 뼈와 임플란트의 직접적인 접촉 길이를 백분율로 나타내는 Bone to implant contact(BIC)가 있다[5].

뼈 조직형태계측(Bone histomorphometry)은 뼈의 형성, 재형성과 관련된 정량적인 평가가 가능하다[6]. 전통적으로 임플란트와 주변의 뼈조직은 박절, 연마 및 염색된 후 현미경을 이용하여 촬영된 이미지에서 숙련된 전문가에 의해 정량적, 정성적으로 평가된다[3]. 현재 사용하는 이 방법은 소수의 전문가에 제한되어 수행되기에 시간이 많이 소모되고, 시각적 특징을 식별하여 수동으로 표시하는 작업이 포함되기에 각자의 판단 기준에 따라 결과 차이가 발생할 수 있다[7], [8]. 따라서 이를 보완하기 위해 관찰자의 시간을 포함하여 효율적이고 편향을 최소화 할 수 있는 일관적인 방법에 대한 연구들이 진행되어 왔으나 아직 확립되지 않았다[6], [9].

뼈 조직형태계측을 위해서는 이미지 데이터에 관심 영역(Region of

interest, ROI)을 정의하는 과정이 필요하다. 일부 영역을 분할(segmentation)하는 과정으로[10], 적절하고 정확한 분할이 선행되어야 신뢰할 수 있는 분석이 가능하기에 매우 중요한 단계이다[3]. 이는 관찰자가 수동으로 측정 영역에 선을 그리는 방식으로 수행되어 왔다. 최근에는 컴퓨터 기술의 발전과 함께 임플란트 영역이나 골소주 영역(trabecular area), 유골 영역(osteoid area) 등 필요한 매개변수를 반자동으로 측정할 수 있는 알고리즘이 다양하게 개발되었다[6]. 지금까지 보고된 소프트웨어들은 관심 영역을 수동으로 지정하거나, 불필요한 영역을 조정하는 것과 같은 관찰자 개입이 매 평가마다 필요하였다[11].

Sarve 등의 연구[3]에서는 뼈 임플란트 조직 표본의 현미경 이미지에서 자동화된 정량화 도구 개발을 위해 새로운 알고리즘을 구축하였으며, 염색된 색상의 강도에 기반하여 분할이 이루어진다. 토끼의 tibiae에 식립된 임플란트 표본 현미경 이미지에서 ‘자동 방식’과 ‘반자동 방식’을 비교하기 위해 뼈와 뼈가 아닌 영역, 임플란트라는 세가지 클래스로의 분할을 확인하고, 매개변수(Bone area, Bone to implant contact) 정확도를 비교하였다. BIC를 측정하였을 때 관찰자의 수동적 개입이 진행된 ‘반자동 방식’과 ‘자동 방식’에서 유의한 차이가 나타나지 않았으나, 기존의 ‘반자동 방식’에서 정확도가 더 높은 양상으로 나타났다. 해당 연구에서는 그럼에도 ‘반자동 방식’은 시간 소모가 증가하고 객관성이 낮아진다는 점을 언급하였다.

Arganda-Carreras 등의 연구[10]에서는 분할 정확도를 향상시키는 방법론으로서 인공지능 기계학습(machine learning) 기법이 도입된 소프트웨어를 소개하였다. 이 소프트웨어는 상업적 소프트웨어와는 다르게 무료로 배포되어 연구자들이 자유롭게 사용할 수 있는 오픈소스로서 많이 사용하는 ‘ImageJ’의 일부 기능이다. 이 도구 역시 염색에 따른 색상

을 기반으로 분할한다. 결과를 보았을 때 유사하게 염색된 두 색상을 다른 클래스로 구분하지 못하는 점이 확인되었고, 필요한 영역만을 측정하기 위해서는 관찰자의 수동 개입이 수반되어야 했다[11].

반자동, 혹은 자동화로 보고된 뼈 조직형태계측 소프트웨어는 모두 염색된 색상만을 기반으로 분할하는 알고리즘이다. 하지만 염색 과정에 따라 샘플 간 나타날 수 있는 염색의 정도 차이, 임플란트가 혈거워진 경향이 있을 경우 일부 경계면에서 매우 어둡게 염색되는 양상 등 변수가 발생할 수 있다[3]. 또 Hong 등의 연구[9]에서는 현미경 카메라의 노출 시간이나 불균일한 조명으로 이미지에 음영이 생긴 경우 관찰자의 수동 분할에 영향을 끼치고 이미지 분석 결과에 차이가 발생되었음이 보고되었다. 색상을 기반으로 한 알고리즘의 분석에서는 혼란을 야기할 수 있는 경우의 수가 많이 존재하기에[3] 이를 보완할 수 있는 방법에 대한 연구가 필요하다.

조직 슬라이드를 디지털 이미지로 변환하여 사용하고, 컴퓨터 모니터를 통해 분석하고 평가하는 것과 같이 병리학 분야에서 디지털 시스템은 필수적인 구성요소가 되었다[12]. 현재 “디지털 병리(Digital pathology)” 시대에서는 영상 분석 소프트웨어를 이용하여 계측 분석을 하는 ‘컴퓨터 보조 진단(computer-aided diagnosis, CAD)’을 넘어[13], 자동으로 관심 영역(Region of interest, ROI)을 탐지(detection), 분할(segmentation)하고 진단(diagnosis), 분석(analysis)하는 ‘인공지능(Artificial intelligence, AI)’ 기법을 활용한 시도가 활발히 보고되고 있다 [12], [14].

2016년 이후 의료 이미지 분석에서 가장 많이 사용하는 인공지능 기계학습의 방법은 “딥러닝(Deep learning)” 알고리즘이다[15]. 인공지능망

(artificial neural network, ANN)’의 한 종류로서[16] 딥러닝 중에서도 “합성곱 신경망(convolutional neural network, CNN)”은 주어진 이미지에서 인간의 개입 없이 직접 특징을 추출하고 최적의 결과를 위해 스스로 학습하며, 깊은 신경망 구조를 통해 점차 높은 수준의 특징까지 계층적으로 학습한다[17]. 하나의 픽셀이 아닌 주변 픽셀들과 통합된 국소적 특징을 추출하고, 추출된 특징을 전역에 적용할 수 있는 CNN은 뛰어난 성능으로 의료계에서 주로 사용되는 방법론이다. 기계학습을 이용한 이미지 분석에서 고전적인 기술로 많이 사용되는 방법론인 ‘k-최근접 이웃(k-nearest neighbor, kNN)’과 알고리즘과 비교하여 이미지 분할, 분류 정확도에 훨씬 우수한 성능을 보였다[18]. kNN은 반자동, 자동 방식으로 개발된 소프트웨어와 유사하게 픽셀의 강도(pixel intensity)를 기반으로 분할이 정해지는 알고리즘이다.

따라서 본 연구에서는 임플란트 조직 슬라이드 이미지의 뼈 조직형태계측에 인공지능 딥러닝 기법을 적용하여 보고, 관찰자의 수동 개입에 대한 의존도를 낮춘 효율적이고 일관적인 보조 도구로서 활용 가능성이 있는지 확인해보고자 하였다. 이를 위해서는 가장 먼저 임플란트와 주변 조직의 명확한 분할이 확보되어야 하기에, 경계면에 대한 분할 정확도를 평가하였다. 그리고 색상을 기반으로 하는 k-최근접 이웃(kNN) 알고리즘을 적용하여 보고, 색상 외의 특징까지 학습하여 결과를 도출하는 딥러닝 기법의 정확도를 비교하였다.

제 2 장 연구 방법

딥러닝 기법이 신뢰성 있는 임플란트 뼈 조직형태계측 보조 수단으로서 역할을 해내기 위해 가장 필수적으로 평가되어야 하는 것은 임플란트와 주변 조직으로의 명확한 경계 분할이다. 이를 위해 임플란트가 식립된 뼈 조직 슬라이드 이미지를 수집하고 실험에 필요한 데이터를 준비하여 데이터베이스를 구축한다.

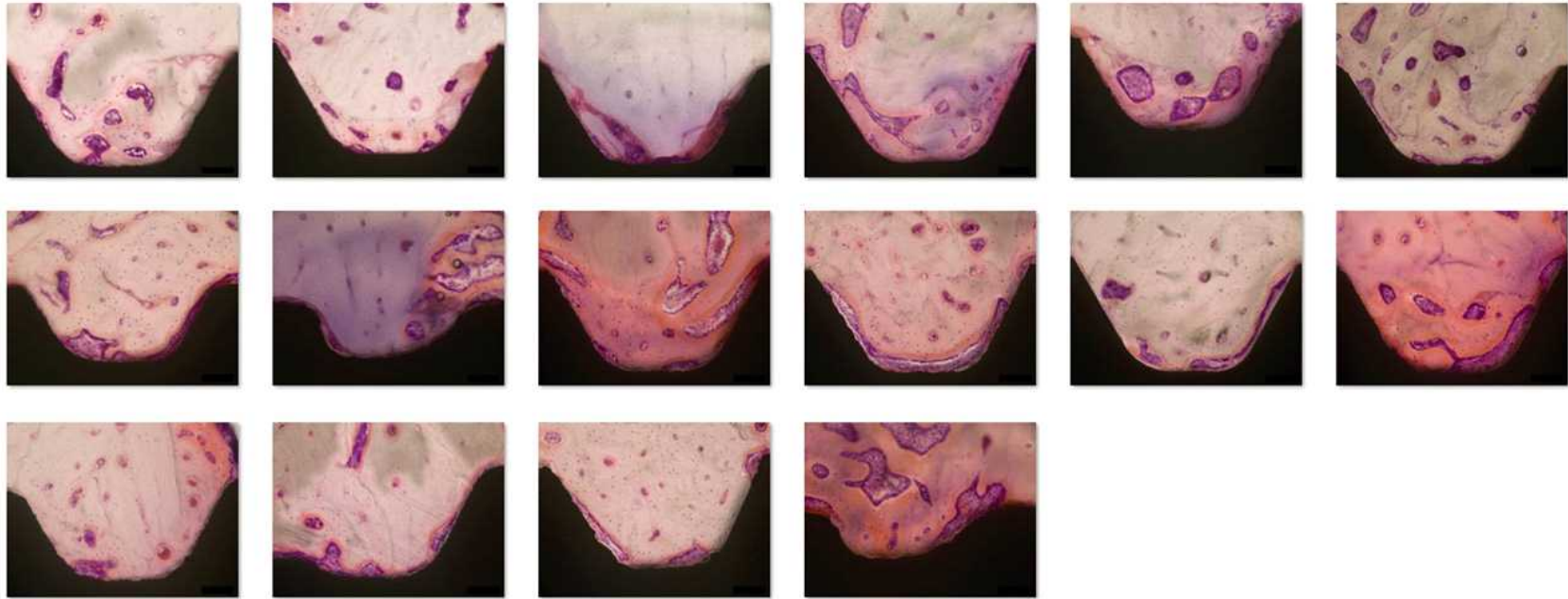
먼저 색상을 나타내는 픽셀값에 의해 분할이 이루어지는 kNN에 이미지를 적용하여 결과와 한계점을 확인한다. 이를 극복할 수 있는 대안으로서 딥러닝 기법에 적용하여 결과를 확인하고 kNN과 비교한다. 이후 분할 정확도를 더욱 향상시키고 한계점을 보완할 수 있는 방안에 대해 고찰하여 실험을 조정한 후 결과를 확인한다.

제 1 절 데이터베이스 구축

1. 데이터 수집

데이터는 이미 수행된 Kim 등의 연구 동물실험에서 수집되었다[19]. 토끼의 경골에 식립된 임플란트 표본은 ‘Olympus bx51’ 광학현미경과 ‘Spot RT3’ 디지털 카메라를 이용하여 200배율에서 촬영되었고, 총 16장이 데이터로 활용되었다.

그림1. 토끼의 경골에 식립된 임플란트 조직 슬라이드 이미지

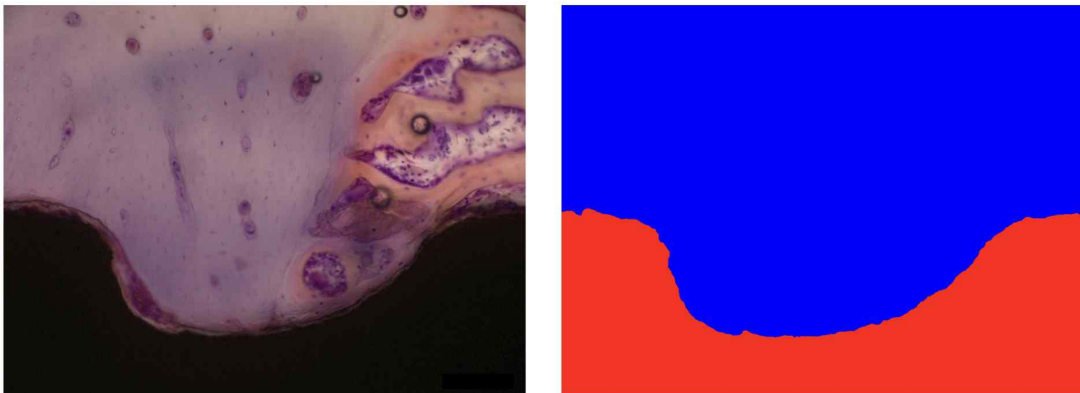


2. 이미지 주석(annotation) 처리 및 Ground truth 이미지 생성

임플란트 조직 슬라이드 이미지에서 분할 클래스는 (1) 임플란트 영역, (2) 임플란트가 아닌 조직 영역의 두가지로 나누었다. 주석(annotation)은 Adobe Photoshop (Adobe Systems, San Jose, CA)의 ‘올가미 도구’를 이용하였고, 두가지 클래스로 영역 분할 후 ‘페인트 통’ 도구를 이용하여 각 클래스를 RGB 값 ‘255, 0, 0’, ‘0, 0, 255’으로 표시하여 사용하였다[그림2].

그림2. 주석 처리를 통한 Ground truth 설정

(a) 원본 (b) Ground truth

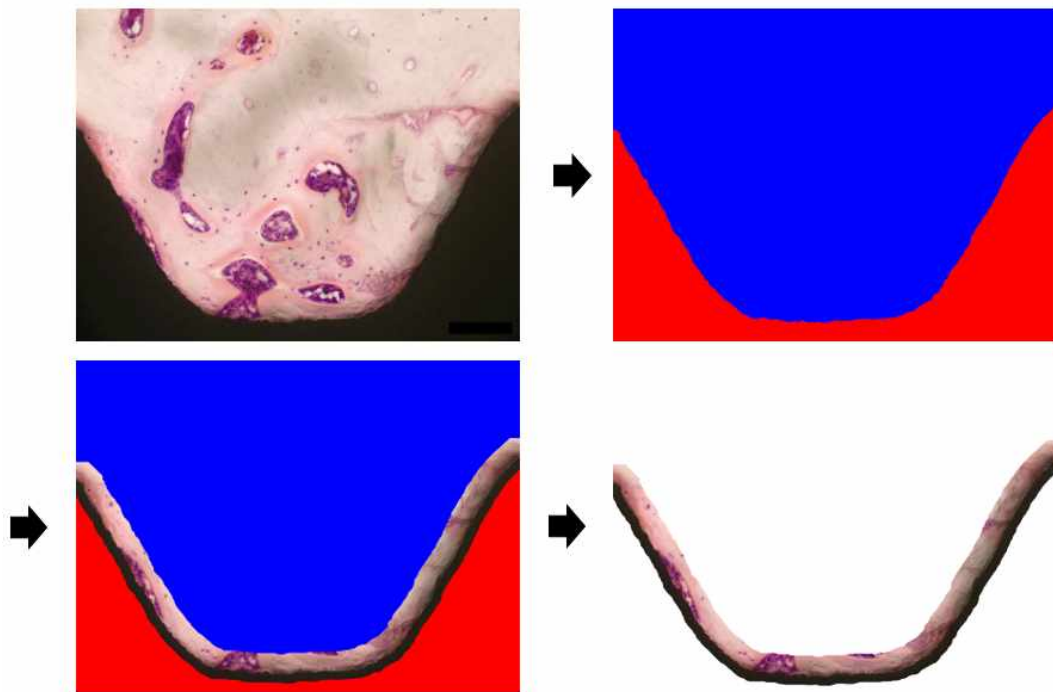


3. 관심영역(Region of interest, ROI) 설정 방법

본 연구의 평가 대상은 두가지 클래스의 명확한 경계 분할이기에 관심영역을 두가지 클래스의 경계선 중심으로 설정하였다. 관심영역 설정을 위해 ‘자동선택영역 도구’를 이용하여 각 클래스 영역을 선택한 후 ‘선택-수정-축소’ 기능에서 (1) 임플란트 영역의 경우 30픽셀을 적용하

고, (2) 조직 영역의 경우 70픽셀을 적용하였다. 그 후 축소된 두 클래스 영역에 해당하지 않는 영역을 관심영역으로 지정하여 평가에 사용하였다.

그림3. 관심영역 설정 과정



제 2 절 인공지능 기계학습 방법론

1. kNN

이미지 분할에서 kNN은 분할하고자 하는 각 클래스의 대표 픽셀값 (pixel intensity)을 관찰자가 지정해주며, 픽셀값은 RGB 값에 의한 것으로 색상을 나타내는 수치이다. 이 대표 픽셀값을 'instance'라고 부르며,

하나의 픽셀은 instance와의 유사도를 k개의 인접 픽셀과 비교하여 분할 결과를 결정 짓는다. 뼈 조직형태계측에서 선행 연구개발된 방법은 모두 색상을 기반으로 영역을 분할하였기에 해당 방법론을 채택하였다.

본 연구에서는 한 장의 조직 슬라이드 이미지에 맞추어 instance를 지정한 후 16장에서의 결과를 추출하였다. 이후 주석(annotation)과 가장 일치 정확도가 낮은 이미지에 맞추어 다시 결과를 추출하였다.

2. U-Net

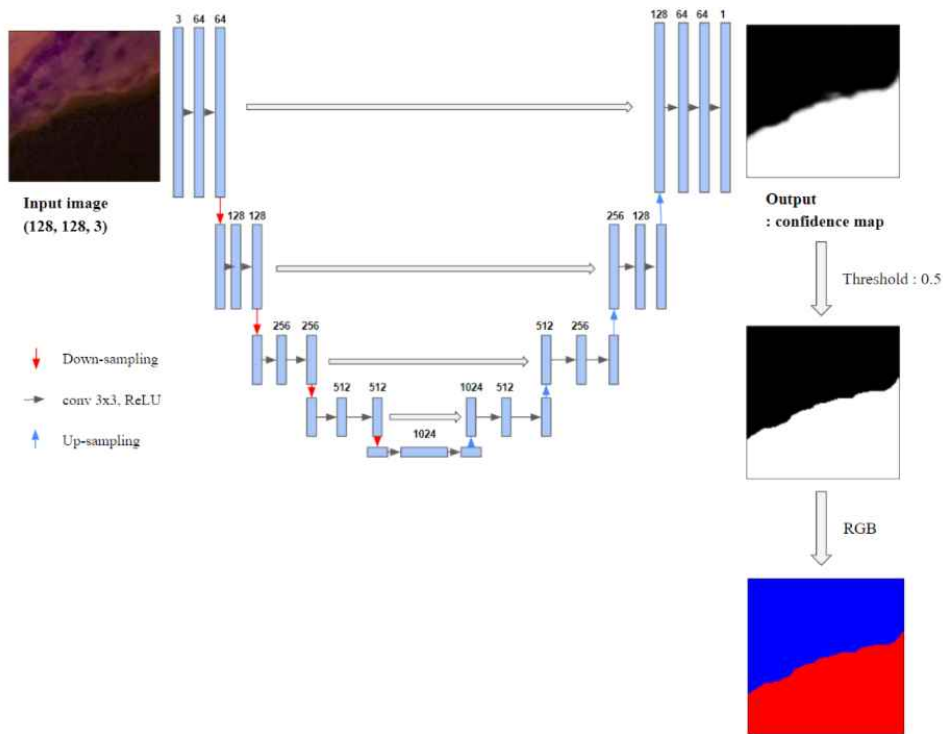
이미지 분할(Image segmentation) 연구에 대표적으로 사용되며 의료 이미지 분할에서도 활발하게 사용 중인 이 모델은 ‘인코더-디코더(Encoder-Decoder)’ 구조로서, 대칭적인 U자 모양을 이룬다[그림4]. 인코더에서 이미지 전반의 픽셀 간 정보를 수집하여 특징 지도(feature map)를 추출하고, 디코더에서 추출된 특징을 기반으로 정확한 위치에 각 픽셀이 해당하는 클래스를 표시하여 분할 결과를 나타낸다. 이 과정에서 U-Net의 중요한 특징 중 하나인 ‘skip connection’을 이용하는데, 입력 이미지에서 특징을 추출하며 위치 정보가 손실되는 것을 방지함으로써 결과적으로 분할 결과를 향상시키는 역할을 한다.

최근까지 이미지 분할 성능의 개선을 위해 기본적인 U-Net 골격을 기반으로 한 다양한 모델이 제안되어왔다. 현재 임플란트 조직 슬라이드 이미지의 뼈 조직형태계측에 있어 딥러닝 기법을 적용한 선행 연구가 없으므로, 본 연구에서는 가장 기본적인 구조의 U-Net 모델인 Vanilla U-Net을 훈련 모델로 사용하였다.

U-Net은 원본 이미지의 크기가 지나치게 큰 경우 학습을 위한 CPU

의 메모리 점유율이 커지는 문제를 해결하기 위해 이미지를 작은 크기의 패치로 나누어 학습을 진행한다[20]. 입력 영상에 따른 출력 영상은 ‘confidence map’으로서 각 픽셀이 0~1 범위의 분할되는 클래스에 할당하는 확률로 나타나며, 이 확률에 역치(threshold)를 설정하여 최종 분할이 결정된다. 최종 분할이 결정된 이미지는 gray scale의 이진 이미지(binary image)로 나타나며, 후처리를 통해 각 클래스를 RGB 값 ‘255, 0, 0’, ‘0, 0, 255’로 나타내었다.

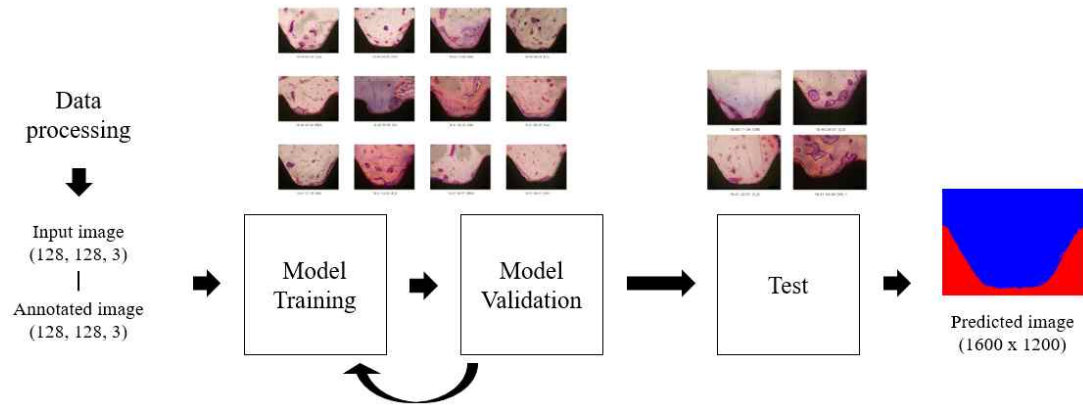
그림4. U-Net 모델 구조도 및 입력/출력 영상 시각화



제 3 절 인공지능 딥러닝 실험 방법

딥러닝을 위한 학습의 전체적인 흐름도는 다음과 같다.

그림 5. 딥러닝 전체 흐름도



1. 데이터셋 준비

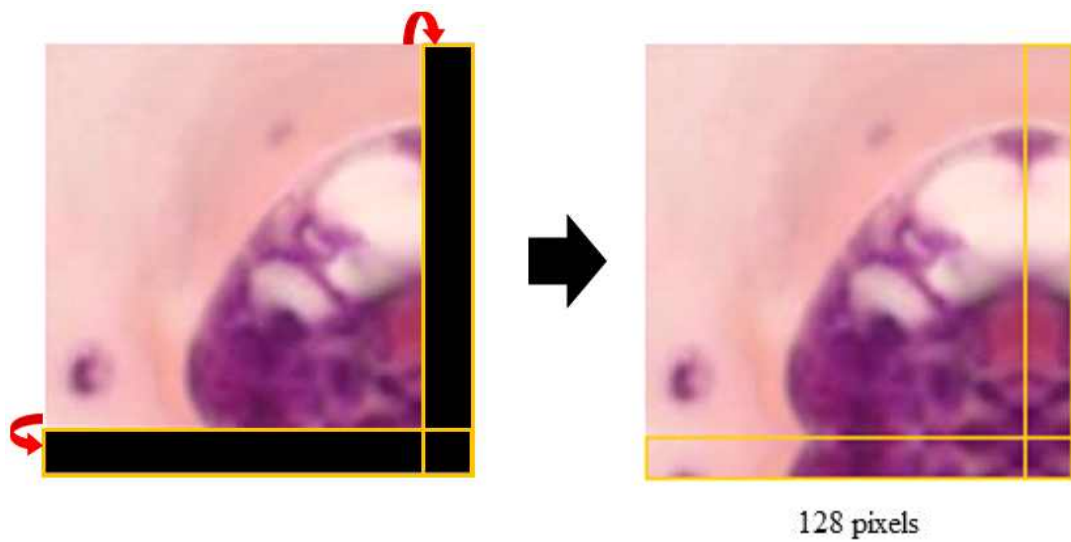
수집된 총 16장의 데이터는 3개의 RGB 채널로 이루어진 (1600x1200) 크기로, 훈련과 검증, 그리고 테스트에 각 8장, 4장, 4장을 사용하였다. 훈련 이미지 8장은 (128x128) 크기의 패치로 잘라서 훈련 세트를 구성하였으며, 패치는 전체 이미지(1600x1200)의 왼쪽 위에서부터 오른쪽 아래까지 순차적으로 잘라낸다. 이때 전체 이미지에서 패치의 크기를 채우지 못하는 경우 mirroring 기법을 사용하였다[그림7].

그림 6. 패치 생성 및 mirroring 과정

(A) 패치 생성



(B) mirroring 과정



2. 데이터 증강

데이터 증강 기법에는 회전(Rotation), 이동(Translation), 반전(Flip)의 세가지 방법을 사용하였다. 하나의 훈련 이미지 당 회전의 경우 -10° ~ $+10^{\circ}$ 도의 범위, 이동의 경우 상/하/좌/우 200픽셀 내의 범위에서 임의(random)로 적용되도록 하였으며, 반전의 경우 좌/우 반전이 적용되었다. 따라서 12장의 훈련 이미지는 각 이미지마다 세가지 기법이 각 범위 내에서 무작위로 적용되는 과정을 6번씩 수행되었다.

3. 모델 매개변수(parameter) 선정

배치 사이즈(Batch size)는 24, 학습률(Learning rate)는 0.001로 지정한 후 훈련 epoch를 250으로 하여, model checkpoint 함수를 통해 전체 훈련 epoch 중 validation IOU(intersection over union)가 가장 높은 모델의 epoch 수를 선택하였고, 이때의 가중치(weight)를 함께 저장하여 모델을 채택하였다.

4. 역치(threshold) 설정을 통한 결과 이미지 이분화

출력 영상은 confidence map으로서 각 픽셀이 0~1 범위의 확률로 나타난다. 이때 역치를 0.5로 설정하였고, 0.5 이상인 경우 임플란트, 0.5 미만인 경우 조직 영역으로 분할되어 이진 이미지로 나타난다. 후처리를 통해 임플란트는 RGB 값 (255, 0, 0), 조직은 (0, 0, 255)로서 변환한다.

5. 최종 예측 이미지 생성

패치를 적용하여 훈련과 예측이 끝난 분할 이미지는 패치의 원래 배

치에 맞추어 재배열한 뒤, 본래 전체이미지 크기인 (1600x1200) 크기 만큼을 왼쪽 위부터 잘라내어 최종 결과 이미지를 생성한다.

제 4 절 기계학습 실험 환경

본 연구에서 U-Net 모델의 구현과 기계학습을 위해 사용한 환경은 다음과 같다. 기반 프레임워크로는 Python 3.8 기반 Jupyter notebook과 Ubuntu 18.04.3 환경을 사용하였다. 데이터 전처리 및 추론(inference)을 위해 OpenCV와 Matplotlib 라이브러리를 사용하였으며, 훈련 과정의 고속화를 위해 Tensorflow-gpu 2.5.0, cudnn 8.2.1, cudatoolkit-11.3.1, GeForce RTX 3090의 환경을 사용하였다.

학습 과정에서 Loss의 경우 Binary cross entropy를 사용하였으며, Activation function은 Sigmoid(binary), optimizer는 adam 함수를 사용하였다.

제 5 절 평가 방법

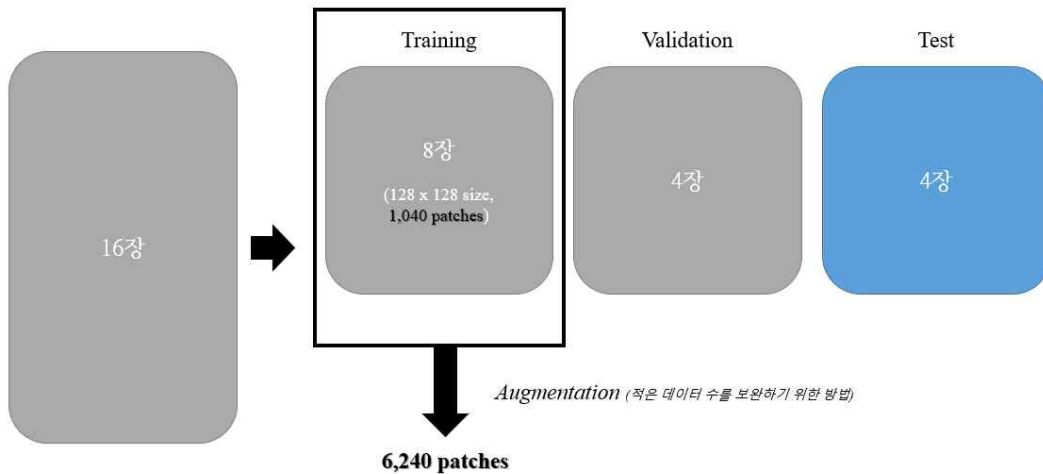
기계학습 방법론 간의 결과는 ground truth 이미지와 픽셀 단위로 비교하여 정확도(Accuracy)와 민감도(Sensitivity), AUROC(Area under ROC curve)로 평가하였다. 그리고 이미지 분할에서 대표적으로 사용되는 Intersection Over Union(IOUS, Jaccard similarity)와 Dice similarity coefficient(DSC)의 두가지 지표를 사용하여 비교하였다. 테스트 이미지 4장은 훈련 과정에서 사용하지 않은 데이터로 구성되었다.

제 3 장 연구 결과

제 1 절 딥러닝 데이터 구성

본 연구에서 합성곱신경망(CNN)의 학습을 위해 구축한 치과용임플란트 조직 슬라이드 이미지는 8마리의 토끼의 tibiae에 식립된 16개의 임플란트 조직 슬라이드로, 데이터셋 구성은 [그림8]과 같다. 전체 이미지를 (128x128) 사이즈로 패치화 하였을 때 총 훈련 데이터 수는 augmentation이 진행되지 않은 경우 1,040 패치, augmentation이 진행된 경우 6,240 패치가 훈련에 사용되었다. 이후 4장을 통해 검증 단계를 거쳐 모델을 선정하고, 테스트 데이터 4장으로 결과를 비교하였다.

그림 7. 데이터셋 구성



제 2 절 임플란트와 조직으로의 경계 분할 결과

임플란트가 식립된 뼈 조직 슬라이드 이미지에서 임플란트와 조직 영역으로의 분할 성능 비교를 위해 kNN과 U-Net 모델의 딥러닝 방법과 비교하였다.

1. kNN 결과

kNN의 경우 색상을 나타내는 픽셀값에 기반하여 각 클래스 영역으로의 분할이 이루어지는 방식이다. 먼저 총 16장의 이미지 중 한 장을 무작위로 선정하여 해당 이미지에서 관찰자가 지정한 instance(“① 임플란트 영역, ② 조직 영역”을 대표하는 픽셀값)를 기준으로 경계 분할 결과 이미지와 IOU 지표를 이용한 정확도를 추출하였다. 결과에서 임플란트 영역은 RGB 값 (255, 0, 0)으로 빨간색, 조직 영역은 (0, 0, 255)으로 파란색으로 나타내었다.

전체 이미지에서 분할 결과를 확인하였고[그림9, 표1], 16장의 정확도 평균이 98.61%로 나타났다. 관심 영역을 지정하여 비교하였다[그림10, 표2]. 이때 전체 이미지에서는 임플란트와 조직 영역으로의 경계 외에도 명확한 임플란트, 혹은 조직 부위까지 평가하는 픽셀의 대상이 되기에 정확도가 90% 이상 높게 나타난 이미지라 하더라도, 경계 부위에서 서로의 영역을 침범하며 분할이 명확하게 나타나지 않는 양상을 보였다(No.,8, 13 참고). 16장에서 전반적으로 정확도가 높게 나타나기 때문에 경계 부위에 대한 방법론 간 성능 평가를 하기에 적합하지 않다는 것을 확인하였다. 따라서 임플란트와 조직의 경계에 대한 일정 영역을 관심영

역으로 설정하였고, 16장의 정확도는 65~96%의 범위에서 평균 90.39%로 나타났다. 가장 낮은 정확도를 보인 사진은 No.13으로 65.83%를 보였으며, 해당 이미지에 맞추어 관찰자는 instance를 새로 지정하여 16장의 분할 결과를 추출하였다. 두 번째 kNN 실험 결과는 [표3]과 같다.

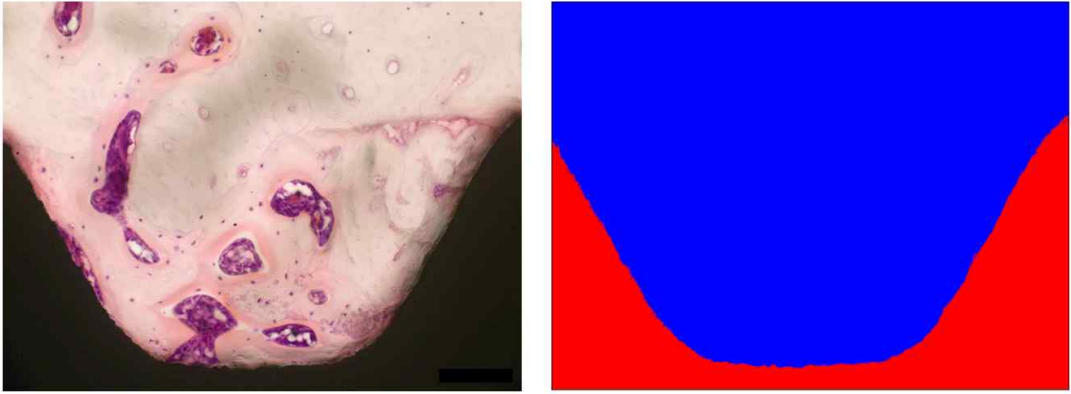
No.13 이미지의 정확도는 65.83%에서 91.39%까지 상승하였고, 다른 이미지에 맞춰 instance를 적용하였을 때보다 평가 대상 이미지에 맞춰 instance를 지정하였을 때 임플란트와 조직의 경계가 서로 영역으로의 침범 없이 명확해지는 양상을 확인할 수 있다[그림11]. 이를 통해 평가하고자 하는 이미지가 변경될 때마다 그 이미지에 맞춘 기준값을 조정해야 하기에 관찰자의 수동 개입이 매번 필요하다는 한계점을 확인하였다.

표 1. 하나의 이미지에 맞춰 instance를 지정한 kNN 결과 - 전체 이미지

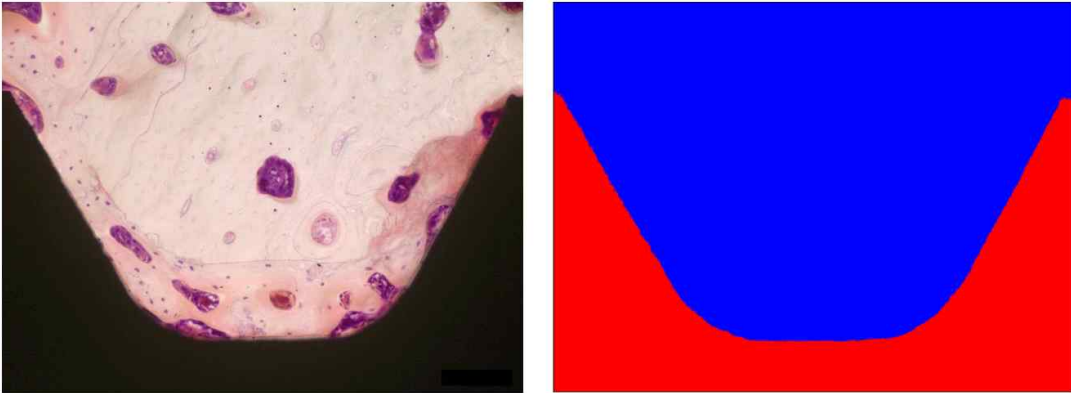
No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Mean
정확도	98.82	98	98.26	99.29	99.16	99	99.58	98.01	98.65	99.49	99.47	97.66	94.26	99.38	99.61	99.04	98.61

그림 8. 16장 전체 이미지에서 kNN 결과

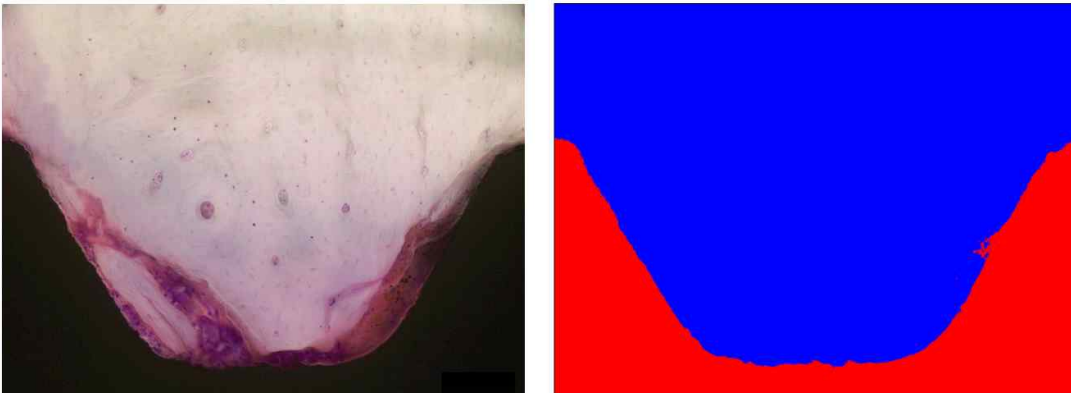
No.1



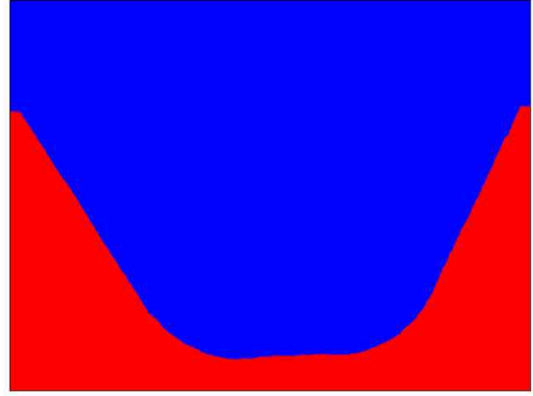
No.2



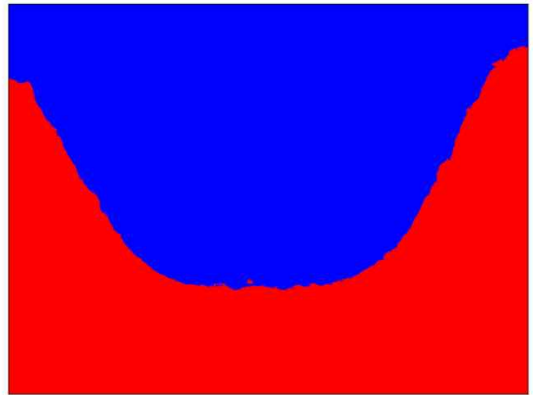
No.3



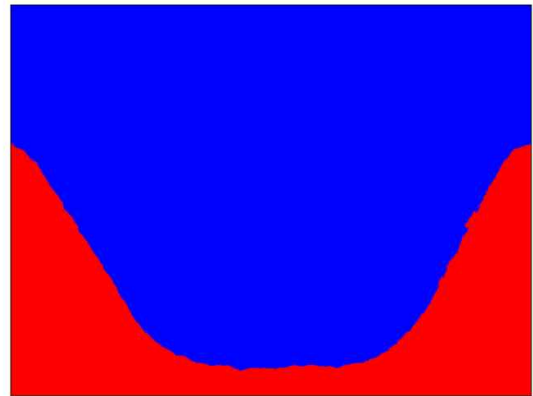
No.4



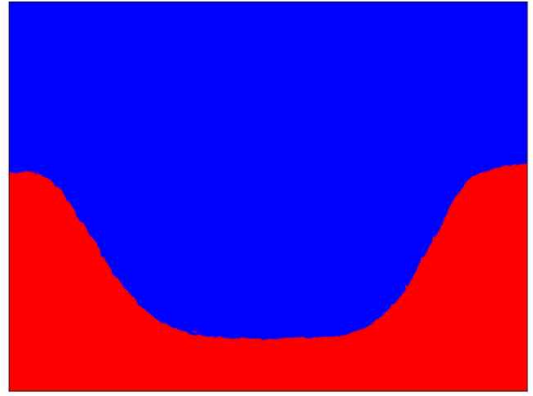
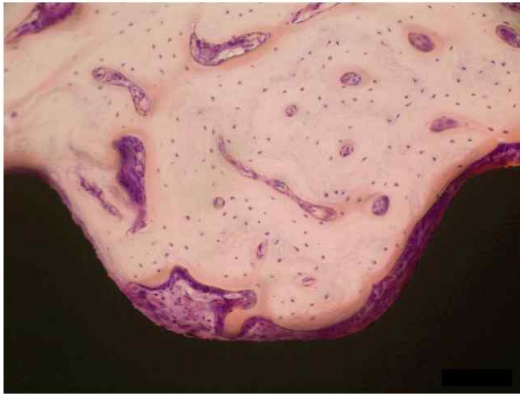
No.5



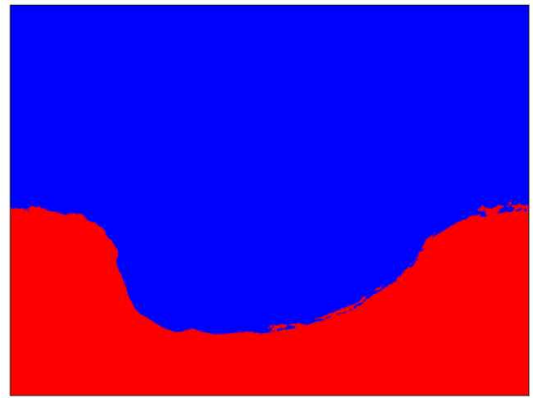
No.6



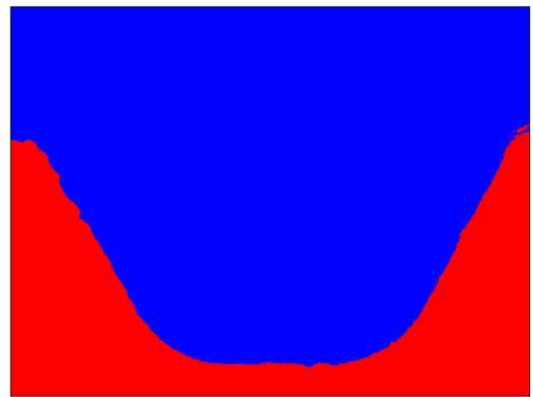
No.7



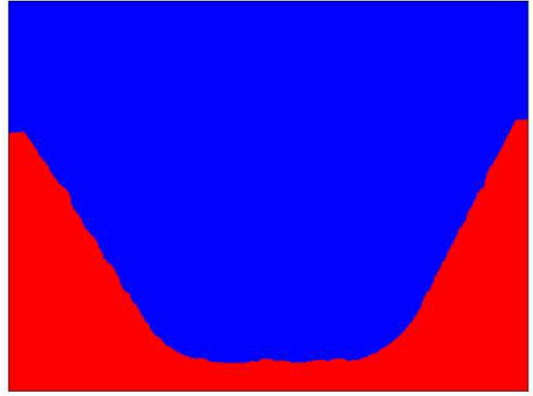
No.8



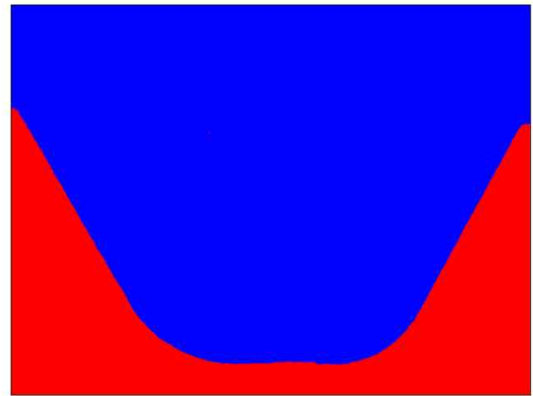
No.9



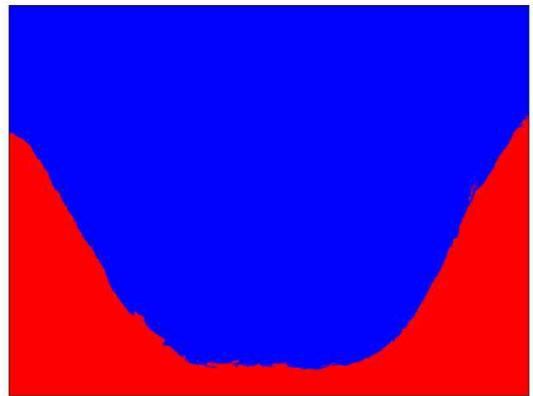
No.10



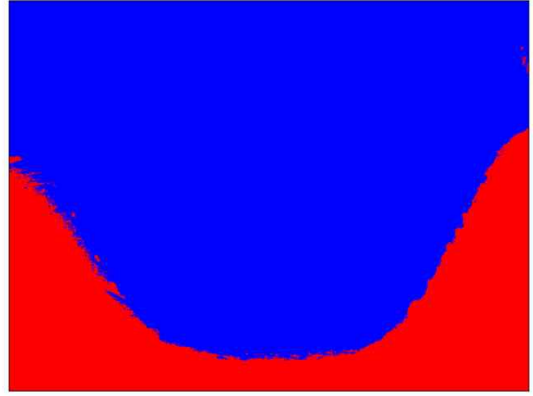
No.11



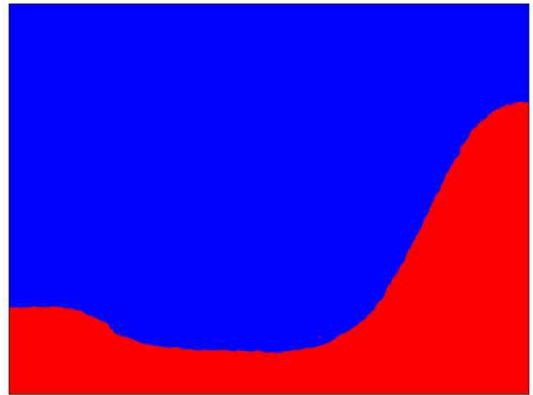
No.12



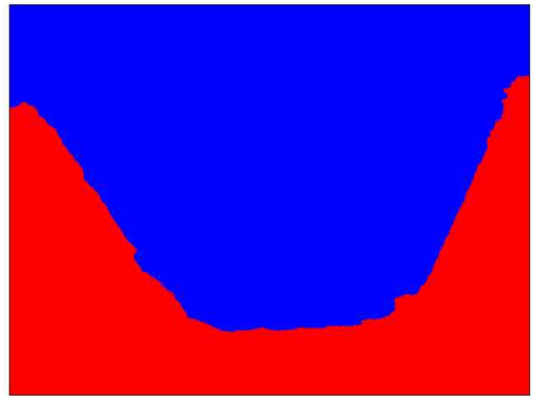
No.13



No.14



No.15



No.16

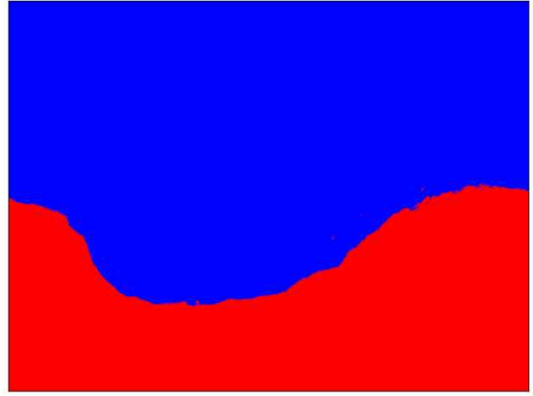


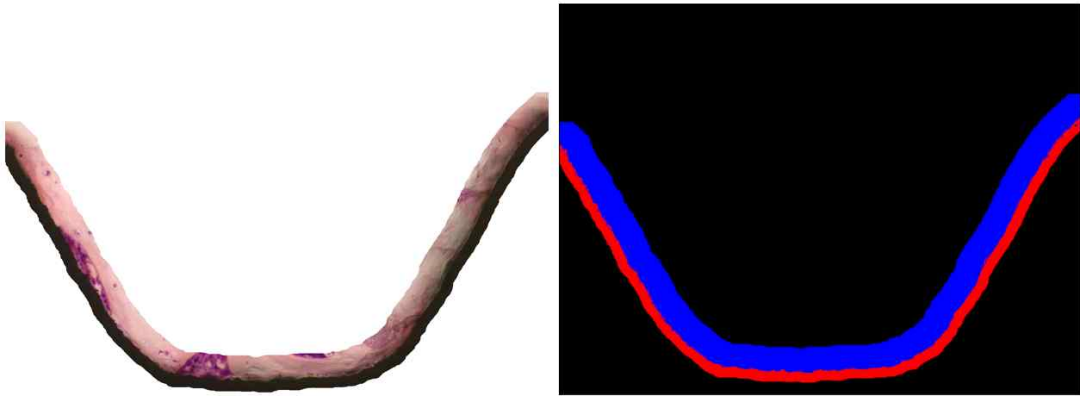
표 2. 하나의 이미지에 맞춰 instance를 지정한 kNN 결과 - 관심영역

가장 분할 정확도가 낮게 나타난 이미지와 정확도 수치는 회색 음영으로 나타내었다.

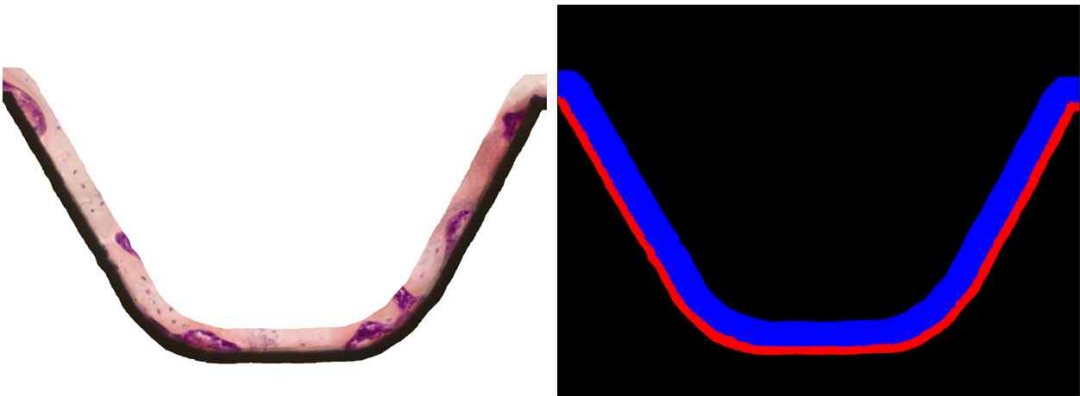
No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Mean
정확도	92.84	83.52	89.64	95.18	91.13	93.89	96.51	84.41	91.5	96.77	96.54	86.45	65.83	95.2	96.63	90.17	90.39

그림 9. 16장 관심영역에서 kNN 결과

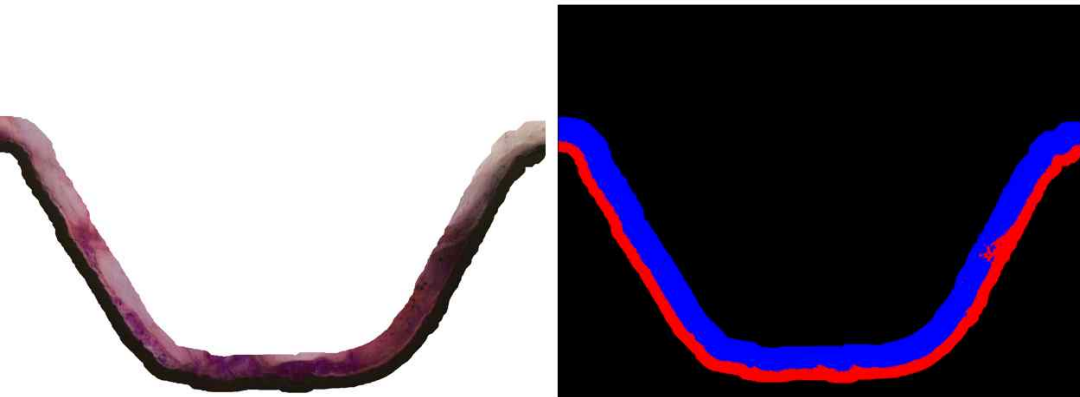
No.1



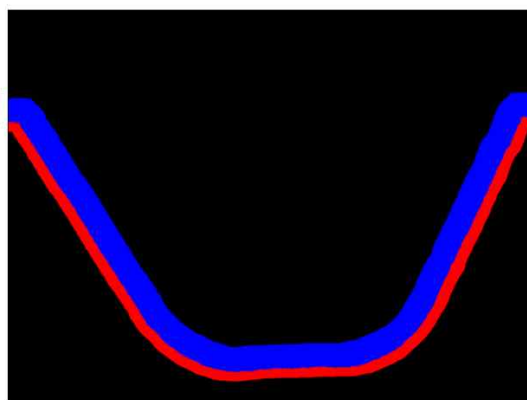
No.2



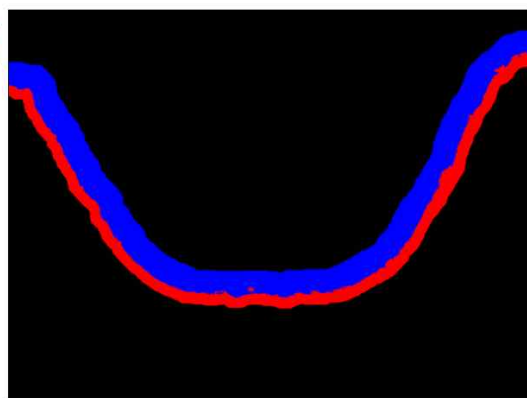
No.3



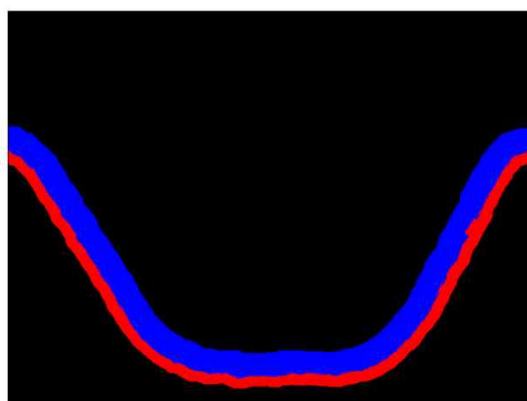
No.4



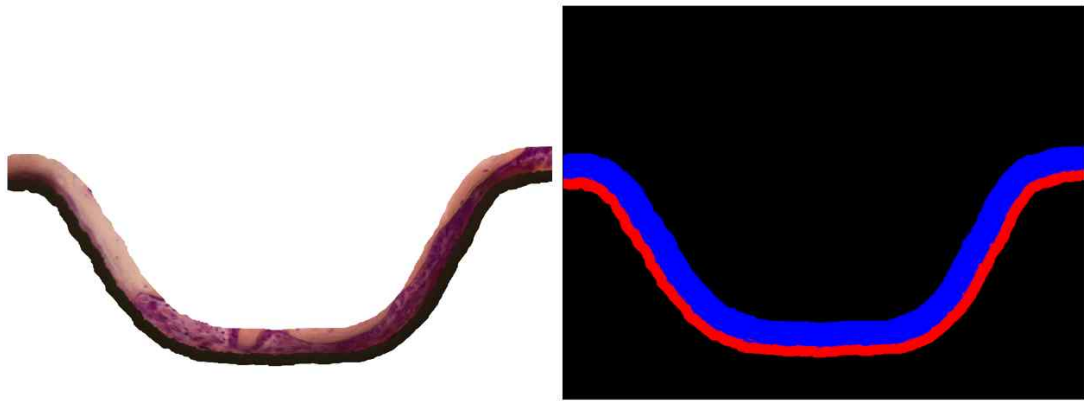
No.5



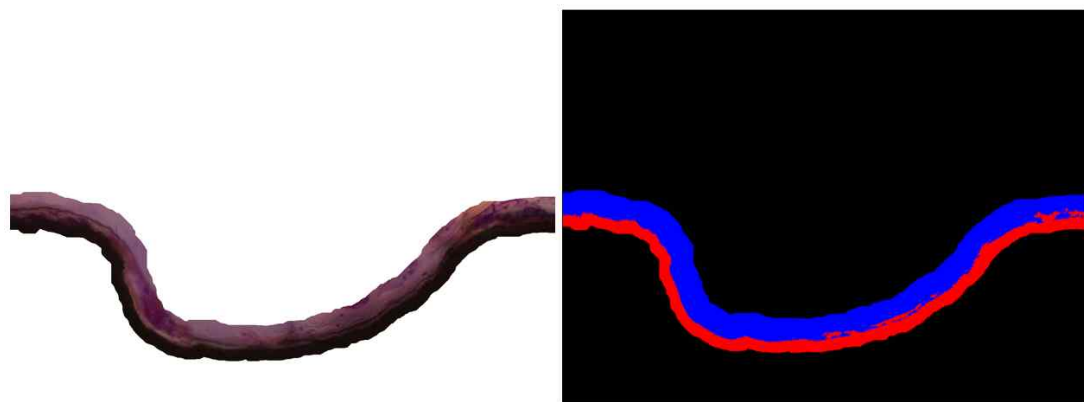
No.6



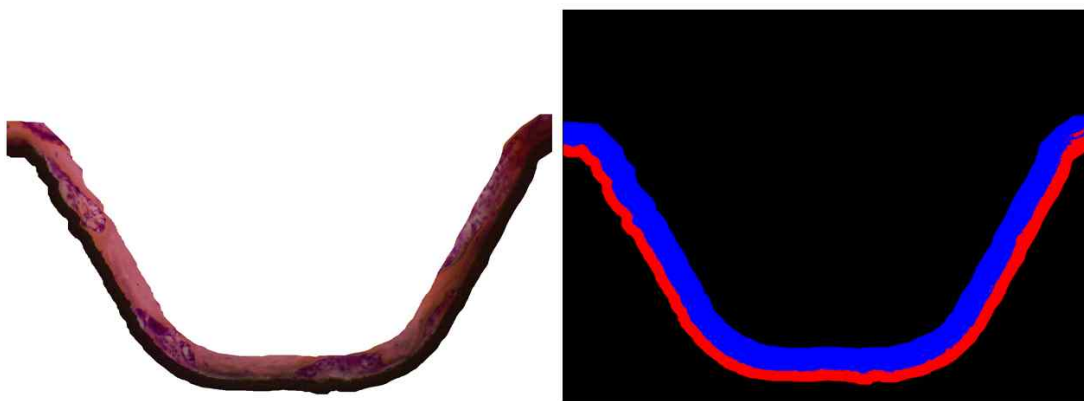
No.7



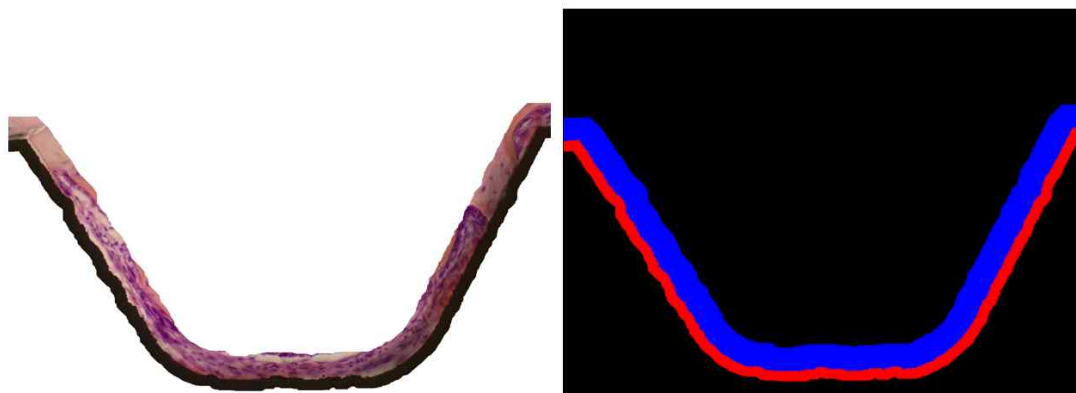
No.8



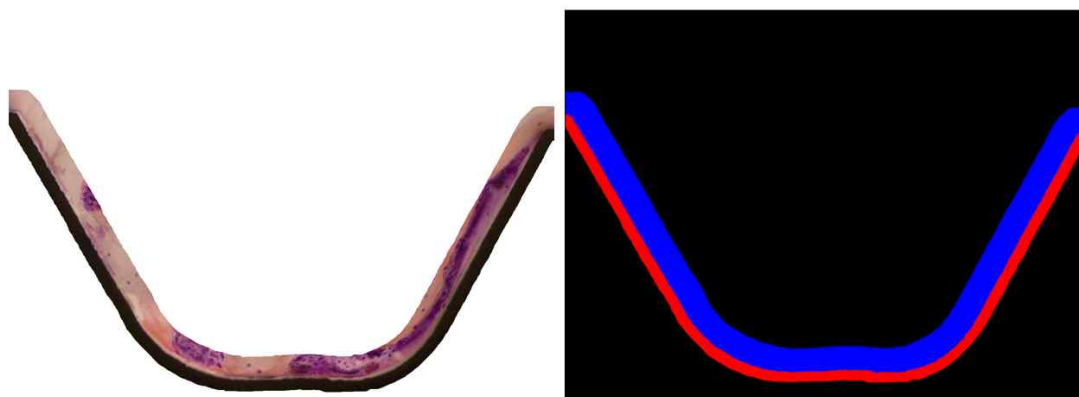
No.9



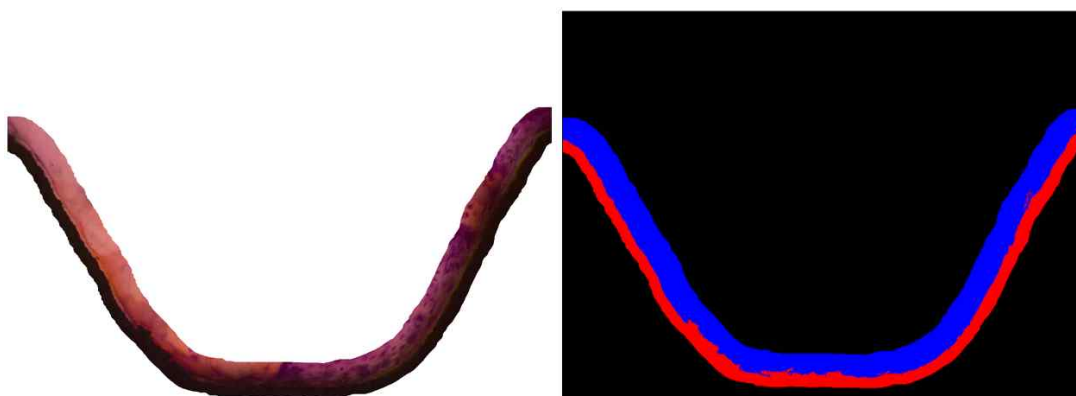
No.10



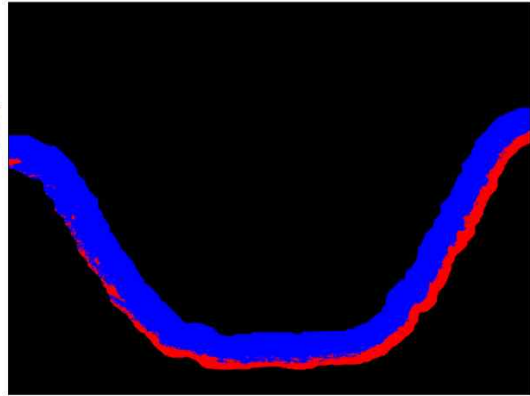
No.11



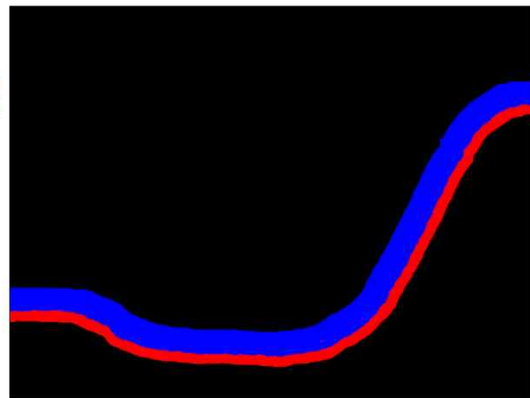
No.12



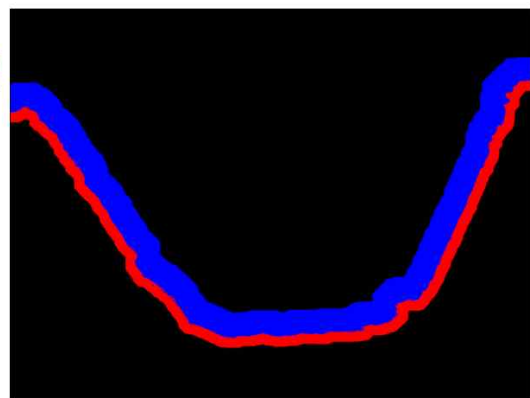
No.13



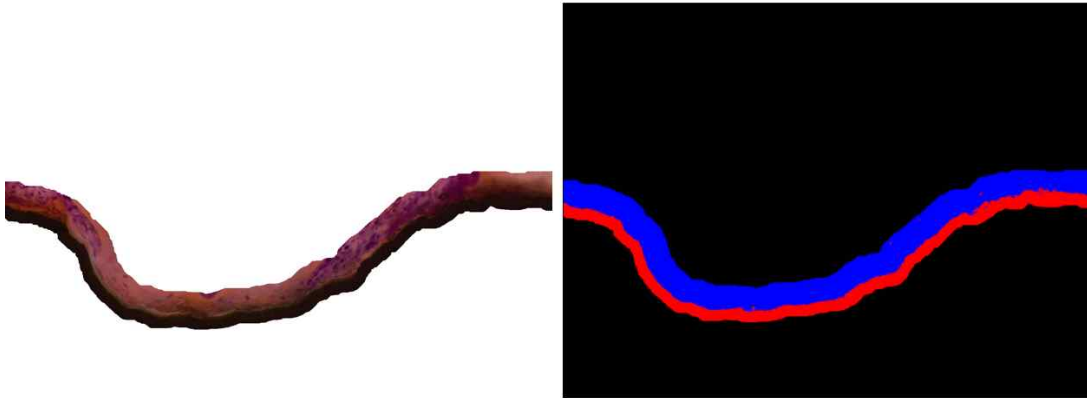
No.14



No.15



No.16



No.1의 이미지를 보면[그림10] 분할 정확도가 관심영역에서 90% 이상을 나타내며 정량적인 성능을 확보했음에도 정성적으로 ‘임플란트, 조직’ 서로의 영역으로 침범이 나타나며 불명확한 경계로 분할이 된 것을 확인하였다.

그림 10. 정성적인 부분에서 한계를 보인 kNN 결과 제시



표3. 'No.13'에 맞춰 instance를 지정한 kNN 결과

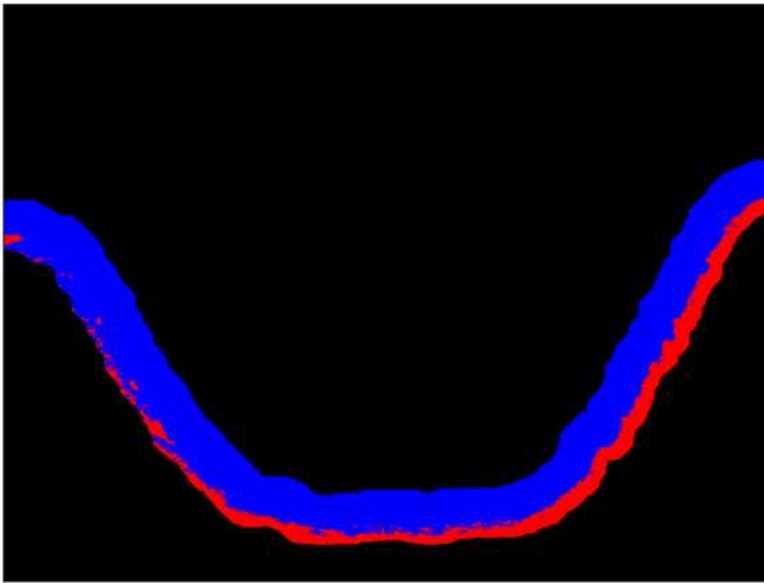
[표2]에서 65.83%로 가장 정확도가 낮게 나타난 No.13의 정확도가 91.39%까지 향상되었음을 확인하였다.

No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Mean
정확도	95.42	95.46	73.41	88.93	78.62	86.48	81.28	67.76	68.42	92.85	90.28	61.75	91.39	85.8	95.98	68.96	82.67

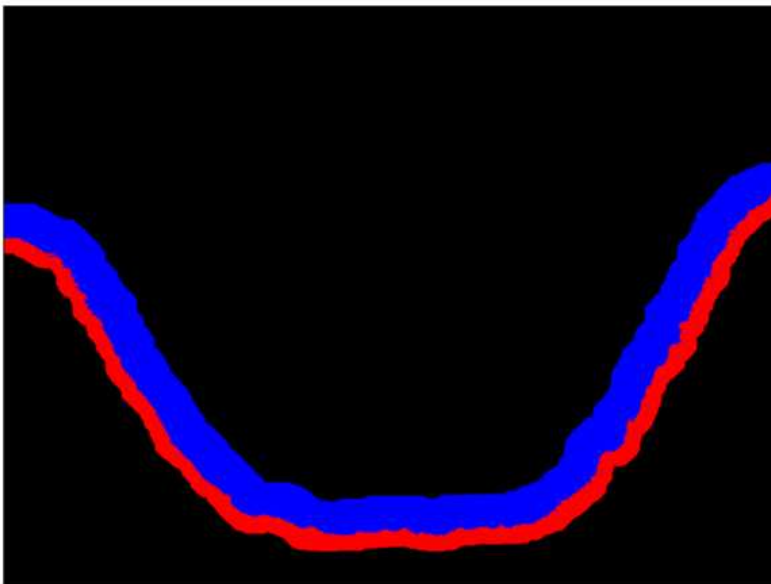
그림 11. 평가 이미지에 맞추어 매번 instance를 조정해야 정확도와 향상된 경계 분할이 확보되는 한계점 제시

다른 이미지의 instance를 적용하였을 때 정확도가 65.83%로 나타났으나, 해당 이미지에 맞춰 조정한 후 91.39%로 상승하였고 조직과 임플란트의 경계가 명확해졌다.

(A) 다른 이미지의 instance 적용한 kNN 결과



(B) 해당 이미지의 instance 적용한 kNN 결과



2. 딥러닝 모델 파라미터 선정

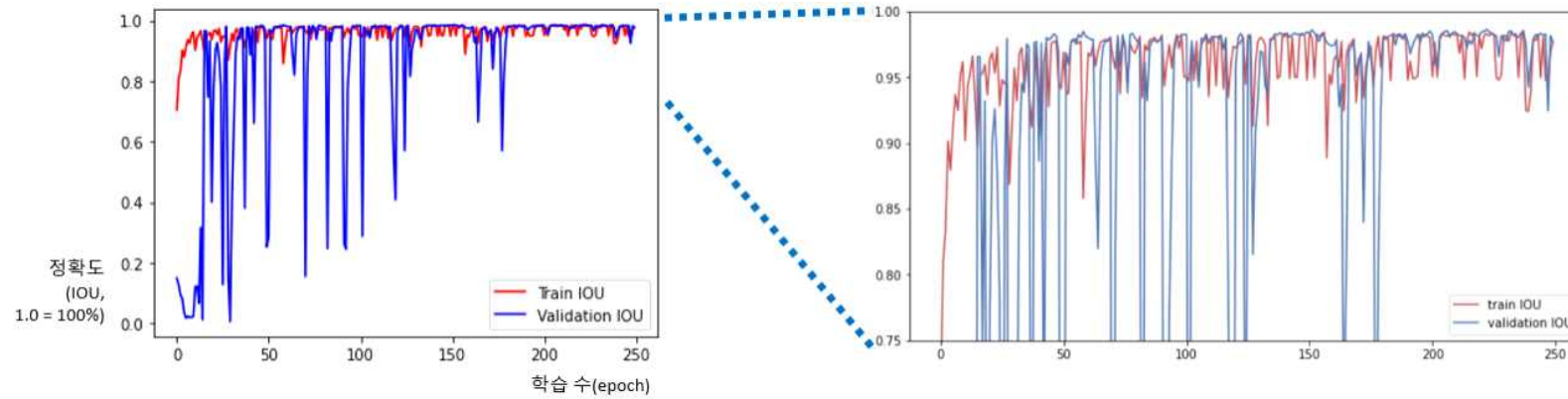
데이터 증강 전과 후의 모델에서 epoch 수를 250로 하여 각각 훈련 후 4장의 데이터로 검증하였고, 그 중 Validation IOU가 가장 높은 모델의 epoch 수를 선택하였다.

데이터 증강 전의 모델에서는 222 epoch가 선정되었고, 데이터 증강 후 모델에서는 247 epoch이 선정되어 해당 epoch 모델의 가중치 (weight)를 저장하여 최종 모델로서 사용하였다.

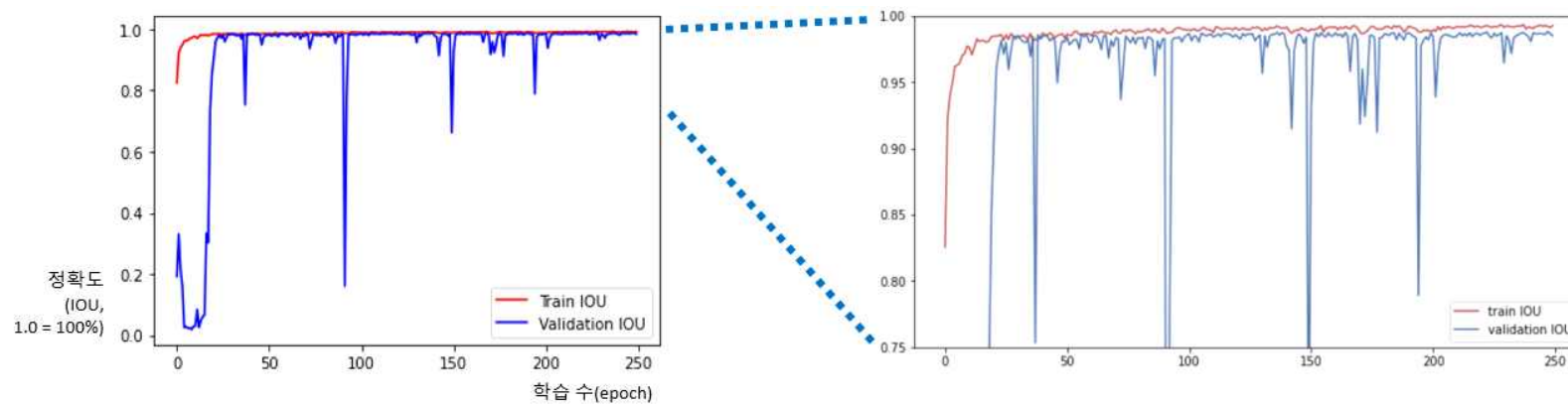
그림 12. 학습 수(epoch)에 따른 정확도(IOU) 산출 그래프

검증 데이터에서 IOU가 가장 높은 epoch을 선정하였으며, IOU ‘75~100%’의 범위를 확대하여 나타내었다.

(A) Data augmentation 하지 않은 모델



(B) Data augmentation 진행한 모델



3. 딥러닝 결과 비교

훈련과 검증이 끝난 딥러닝 U-Net 모델에 4장의 테스트 이미지를 적용하였으며, 모델 성능은 ground truth 이미지와 기계학습 결과 예측된 이미지를 픽셀 단위로 각각 비교하여 혼동 행렬을 통한 진양성(True positive, TP), 위양성(false positive, FP), 위음성(false negative, FN), 진음성(true negative, TN)으로 나타내었다. Ground truth를 기준으로 임플란트 영역을 임플란트로 예측한 경우 TP, 임플란트 영역을 조직으로 예측한 경우 FN, 조직 영역을 임플란트로 예측한 경우 FP, 조직 영역을 조직으로 예측한 경우 TN으로 표현된다. [식1, 2, 3]에 따라 정확도 (accuracy)와 민감도(sensitivity), 특이도(specificity), 그리고 AUROC(area under receiver operating characteristic curve)를 측정하여 데이터 증강 전/후 딥러닝 모델을 kNN과 비교하였다[표4]. 이때 AUROC는 X축을 False Positive Rate, Y축을 True Positive Rate으로 하여 수치가 1에 가까울수록 FP 대비 TP 비율이 높은 것으로 예측을 잘 하는 모델이라고 판단할 수 있다.

식 1.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

식 2.

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

식 3.

$$\text{Specificity} = \frac{\text{TN}}{\text{FP} + \text{TN}}$$

표 4. 테스트셋 4장의 관심 영역에서 딥러닝과 kNN 결과 비교(1)

전체 이미지			
	kNN	Augmentation 하지 않은 모델	Augmentation 진행한 모델
Accuracy	0.9931	0.9951	0.9972
Sensitivity	0.9882	0.9991	0.9982
Specificity	0.9956	0.9930	0.9967
AUROC	0.9919	0.9999	0.9999

관심 영역			
	kNN	Augmentation 하지 않은 모델	Augmentation 진행한 모델
Accuracy	0.9489	0.9625	0.9786
Sensitivity	0.9068	0.9931	0.9832
Specificity	0.9672	0.9492	0.9766
AUROC	0.9369	0.9954	0.9979

그리고 최종 결과 이미지와 ground truth의 이미지 분할을 직관적으로 비교하기 위해 테스트셋 4장의 관심영역에서 가장 많이 사용되는 지표인 Intersection over the union(IOU), Dice similarity coefficient(DSC)를 이용하였고, 결과 이미지와 정확도는 [그림13], [표5]과 같다.

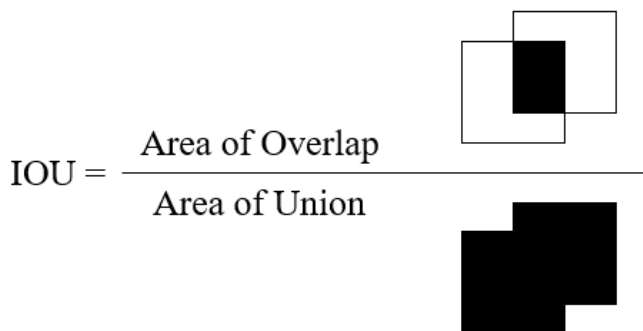
1) IOU(Intersection Over Union, Jaccard similarity)

두 집합의 교집합을 합집합으로 나누어 주는 [식4]을 사용한다[21]. 이때 GT는 Ground truth, SR은 딥러닝 결과로서 추출된 Segmentation Result이다. 이미지로서 나타내면 [그림13]과 같다.

식 4.

$$IOU = \frac{|GT \cap SR|}{|GT \cup SR|}$$

그림 13. 평가지표 IOU 산출 도표



2) Dice similarity coefficient(DSC)

Dice similarity coefficient는 F1 score와 같은 식을 가진다.

IOU와 달리, 두 집합의 교집합 크기를 각 집합의 넓이의 합으로 나누어 주는 [식5, 그림14]을 사용한다[22]. F1 score로서 [식6]으로 나타낼 수 있다.

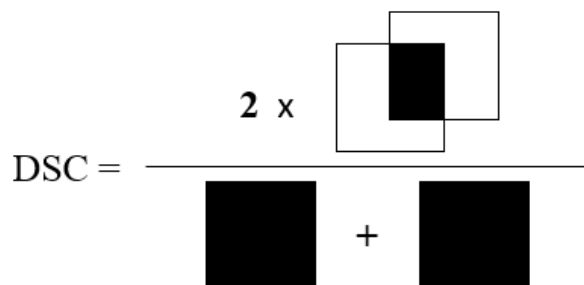
식 5.

$$DSC = \frac{2 | GT \cap SR |}{| GT | + | SR |}$$

식 6.

$$DSC = \frac{2 \times TP}{(TP+FP) + (TP+FN)}$$

그림 14. 평가지표 DSC 산출 도표



딥러닝 U-Net 모델에서 데이터 증강을 진행했을 경우, 4장의 테스트 데이터에서 모두 kNN에 비해 정확도가 상승하였고, 임플란트와 조직 간의 영역 구분이 서로의 침범 없이 더욱 명확해지는 양상을 확인할 수 있다[그림15]. 관심영역 외의 영역은 검정색으로 마스킹 처리를 진행하여 시각화하였다.

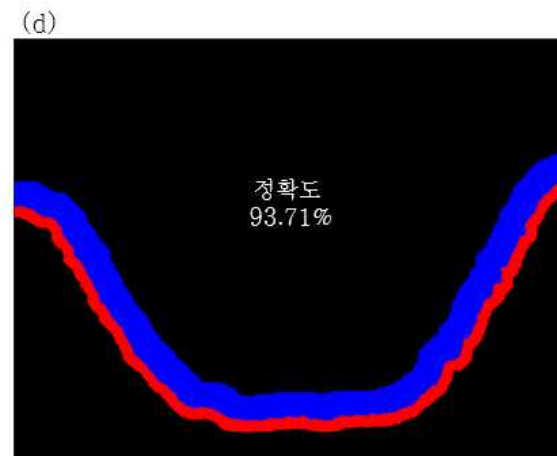
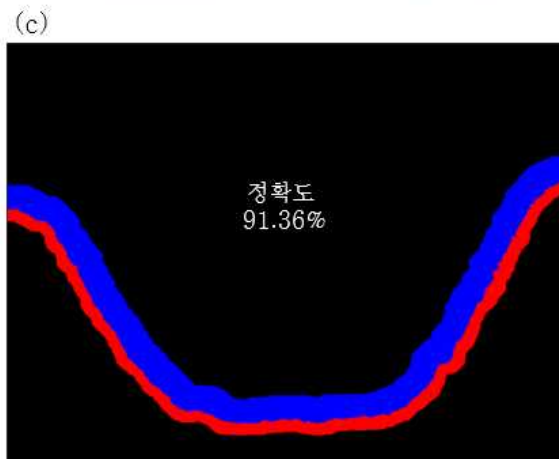
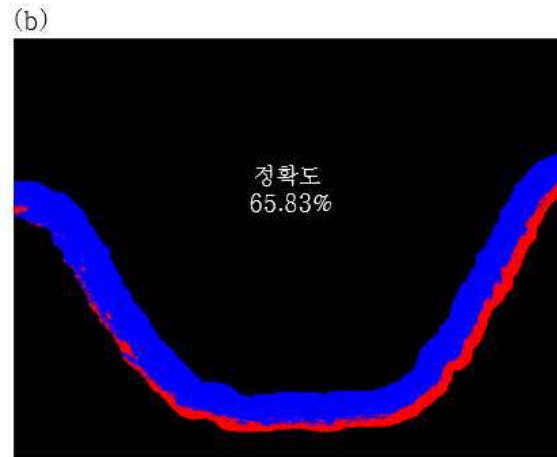
표 5. 테스트셋 4장의 관심 영역에서 딥러닝과 kNN 결과 비교(2)

IOU	Test No.				평균
	1	2	3	4	
kNN	89.64	91.13	65.83	90.17	84.19
Augmentation 하지 않은 모델	87.59	87.17	91.36	89.84	88.99
Augmentation 진행한 모델	92.15	92.82	93.71	94.91	93.40

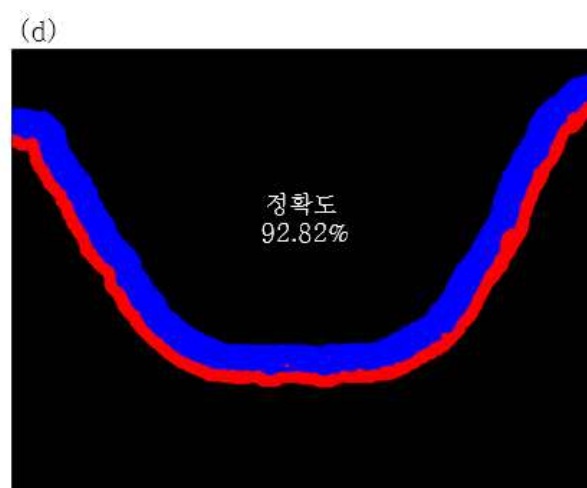
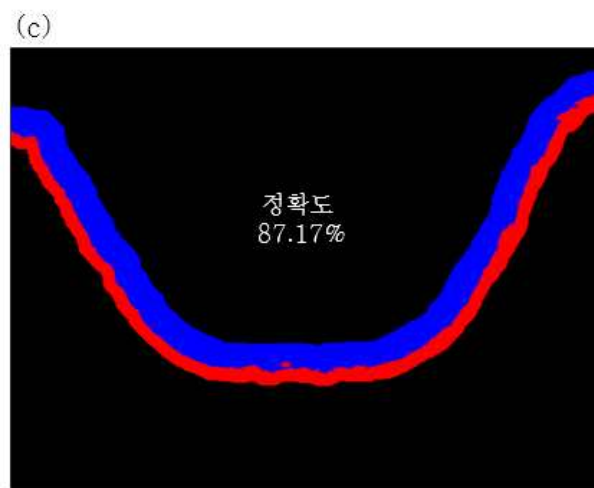
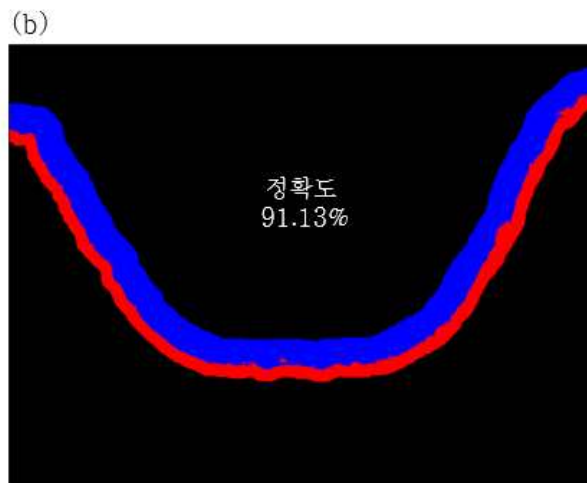
DSC	Test No.				평균
	1	2	3	4	
kNN	94.54	95.36	79.39	94.83	91.03
Augmentation 하지 않은 모델	93.83	93.15	95.48	94.65	94.17
Augmentation 진행한 모델	95.91	96.27	96.75	97.39	96.58

그림 15. 방법론 간 결과이미지 비교

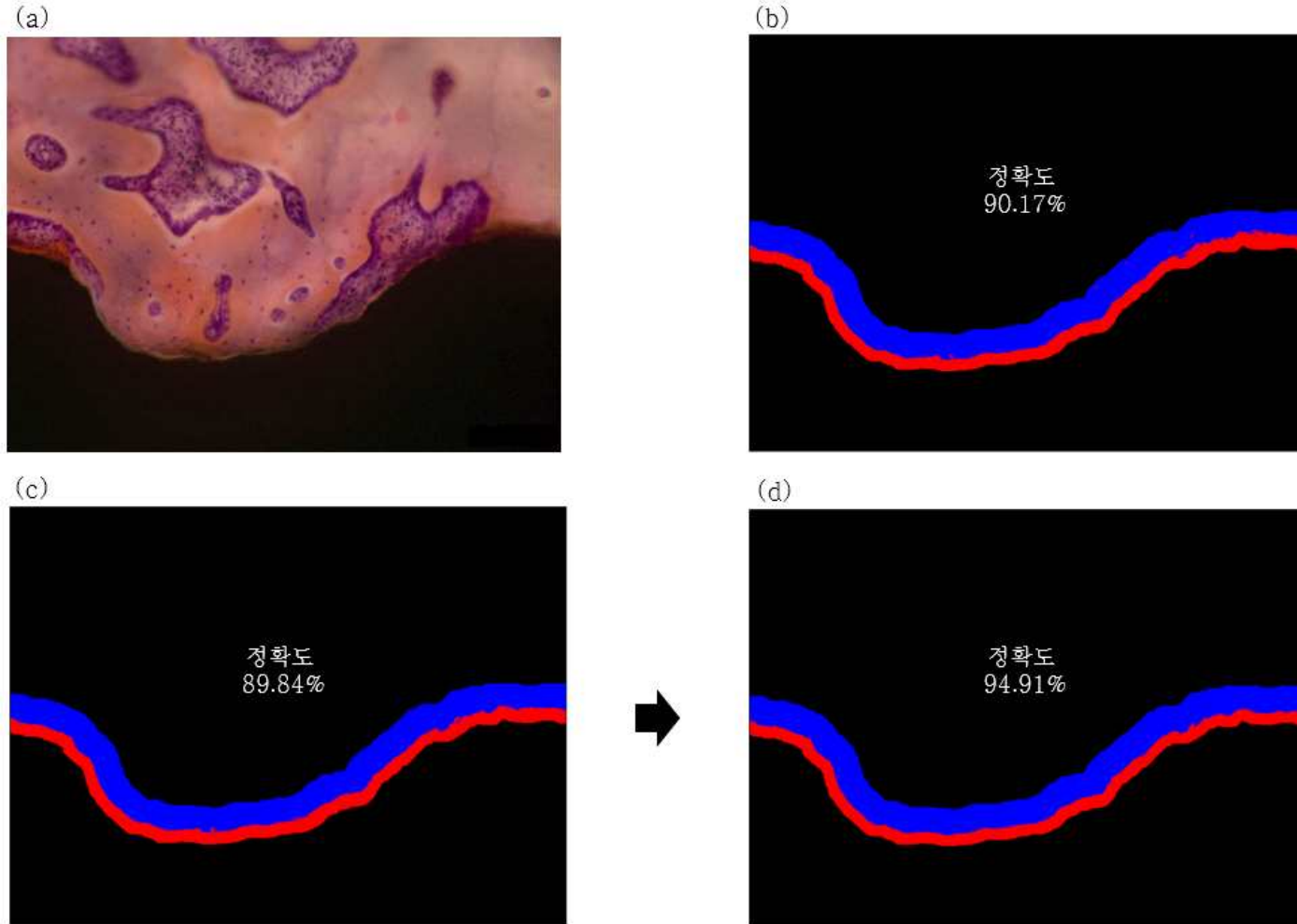
(A) 테스트 이미지 01 (a) 원본 (b) kNN (c) 딥러닝 data augmentation 전 (d) 딥러닝 data augmentation 후



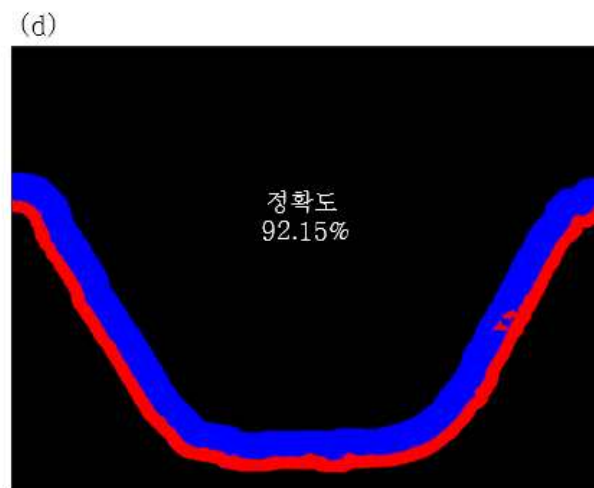
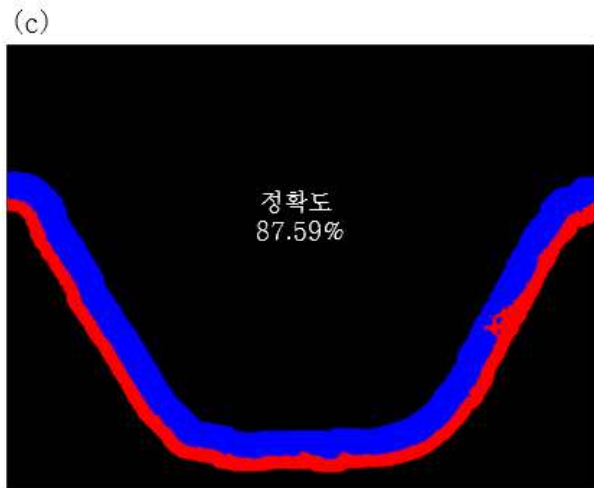
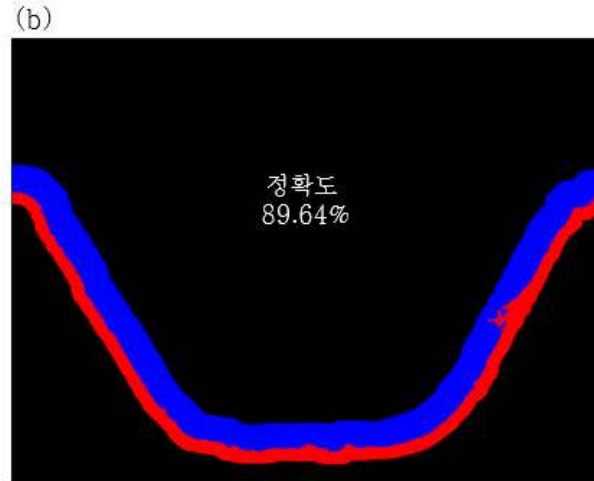
(B) 테스트 이미지 02 (a) 원본 (b) kNN (c) 딥러닝 data augmentation 전 (d) 딥러닝 data augmentation 후



(C) 테스트 이미지 03 (a) 원본 (b) kNN (c) data augmentation 전 (d) 딥러닝 data augmentation 후

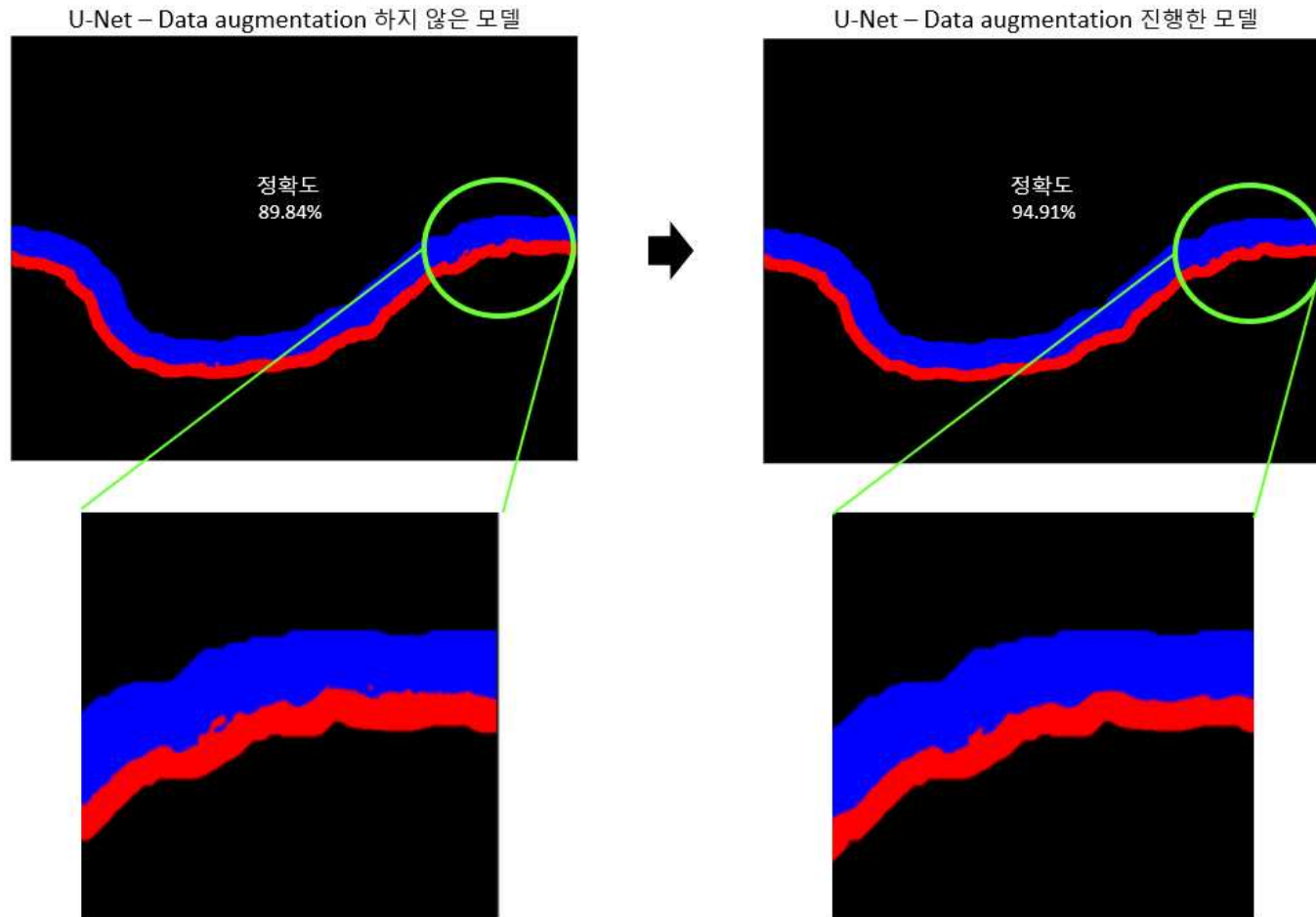


(D) 테스트 이미지 04 (a) 원본 (b) kNN (c) data augmentation 전 (d) 딥러닝 data augmentation 후



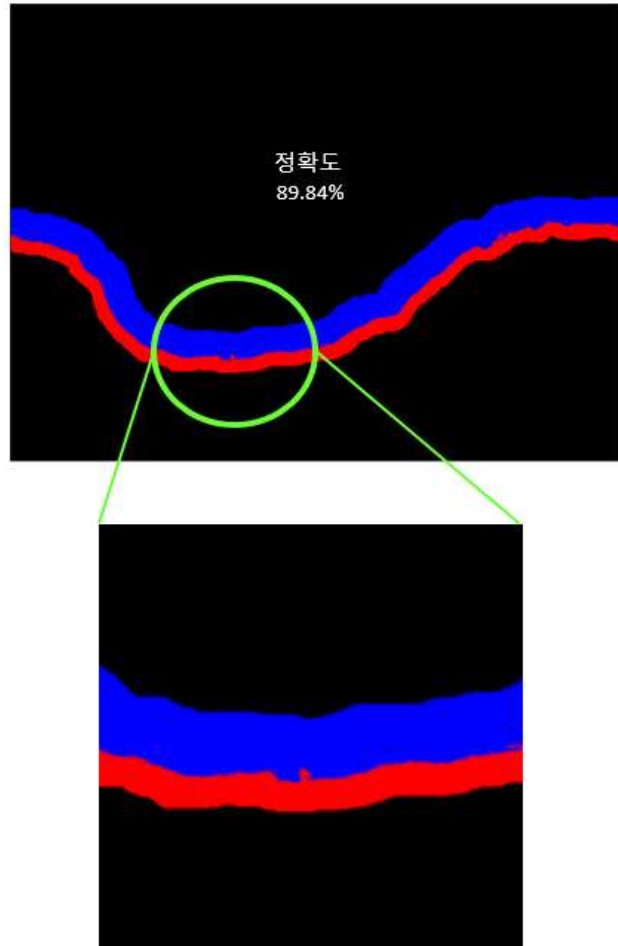
딥러닝 U-Net 모델에서 데이터 증강 기법을 이용하였을 때, 수치적인 정확도 향상과 함께 분할 이미지에서도 세부적으로 서로의 영역으로 침범이 줄어들고 경계가 명확해지는 것을 확인하였고, 추가적인 데이터 확보와 증강기법이 동반된다면 더욱 양호한 결과를 추출할 수 있을것이라 기대할 수 있었다[그림 16].

그림 16. 딥러닝 Data augmentation 전/후 결과에서 정성적으로 향상된 케이스 제시
A. 케이스 1 (a) Data augmentation 전 (b) Data augmentation 후

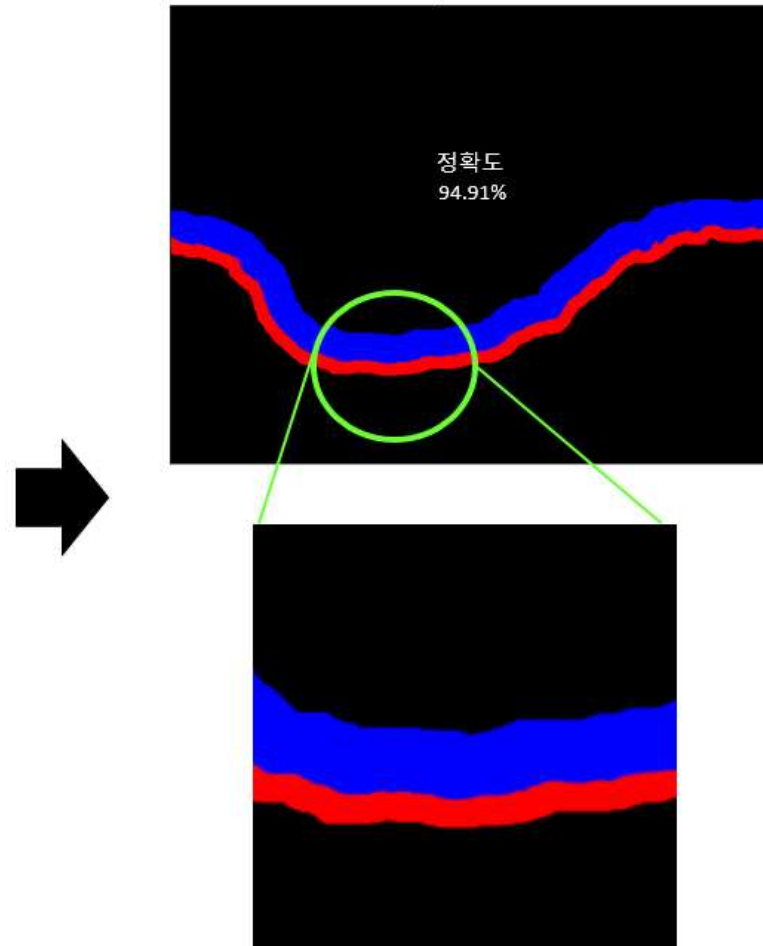


B. 케이스 2 (a) Data augmentation 전 (b) Data augmentation 후

U-Net – Data augmentation 하지 않은 모델



U-Net – Data augmentation 진행한 모델



제 4 장 고찰

치과용 임플란트의 세계시장 규모는 2021년 기준 2028년까지 연평균 성장률 11%로 성장할 것으로 예상되며[23], 지속적인 연구개발이 활발히 진행 중이다. 연구개발의 실험실 단계에서 가장 신뢰도가 높은 평가 방법은 조직학적 분석으로, 병리학자는 조직 슬라이드 이미지에서 중요 지표인 Bone to implant contact(BIC)와 Bone area(BA)를 측정하기 위해 임플란트와 뼈의 영역을 명확히 구분하여 수동으로 표시하여야 한다. 이때 병리학자 간 시각적 인식에 따라 나타날 수 있는 판단의 차이는 같은 표본에서 상이한 평가 결과로서 나타난다[24]. 그리고 극히 제한된 수의 전문가에 의해 수행된다는 점은 분석의 효율성과 일관성을 향상시키기 위해 반드시 극복해야 하는 한계점이다. 이를 보완할 수 있는 도구 개발을 위해 컴퓨터 기술의 도입이 지속적으로 이루어져 왔으나[25], 아직까지 한계점을 극복할 수 있을 정도의 역할을 하지 못하였다.

의료 영상 분석에서 인공지능 기반 컴퓨터 보조 진단 기술에 대한 연구가 활발히 이루어진 것은 컨볼루션 신경망(Convolutional Neural Network, CNN) 등장 이후로, 이 신경망의 컨볼루션 계층(convolutional layer)은 의료 이미지에서 중요한 특징들을 추출할 수 있음이 보고되었다[26]. 2017년 Granter, Scott [27]에 의해 디지털 병리 분야에서 인공지능 기술의 CNN 활용 가능성이 보고되었으며, 해당연도 기준 병리학계는 CNN 알고리즘을 이용한 의료계의 연구 수에서 선두를 차지하였다[14].

최초로 전체 슬라이드 영상(Whole Slide Imaging, WSI)를 이용한 연구[28]에서는 여성 유방암 환자의 헤마톡실린-에오신 염색 감시림프절 생검(Sentinel lymph node biopsy) 슬라이드 이미지를 사용하였다. 전이의 존재 유무를 분류하는 업무에 있어 시간 제한을 준 병리의사 11명, 시간

제한을 주지 않은 병리의사와 CNN 모델의 정확도를 비교하였을 때, 가장 우수한 5개 모델의 정확도 평균은 96%로 시간 제한을 주지 않은 병리의사의 정확도인 96.6%와 유사한 수치를 보였다. 이때 시간 제한을 준 병리의사 11명의 정확도 평균은 81%였다. 소수의 전문가에 집약된 업무와 시간 제약을 고려하였을 때 인공지능 딥러닝 기법은 이를 극복할 수 있는 가능성이 확인되었다. 하지만 아직까지 뼈 조직형태계측에 있어 CNN의 도입을 시도한 연구는 진행된 바 없다.

컴퓨터 보조 진단 기술을 이용한 뼈 조직분석에서 선행 연구개발된 방법들을 살펴보면, 모두 이미지에서 검색된 색상만을 기반으로 특정 영역을 탐지한다. 반자동 혹은 자동 방식을 비교한 연구[11]에서 보고된 소프트웨어는 뼈와 임플란트의 골융합을 평가하기 위해 관심영역을 수동으로 설정해 주면, 자동으로 임플란트 영역을 탐지하고 임플란트의 둘레와 특정 면적을 측정할 수 있다. 임플란트가 아닌 조직 영역에서 뼈에 대한 구분은 특정 검색으로 나타나는 색상의 임계값(threshold)에 의해 이루어지는 방식으로, 수동으로 매 사진마다 역치를 조정해야 하는 번거로움이 동반되었고 유사한 색상에서는 구분 능력이 떨어지는 것이 확인되었다. 더불어 품질이 좋지 않은 표본의 경우 더욱 수동 편집 요구사항이 증가하였기에 연구자들에게 채택되지 않았다.

임플란트가 식립된 뼈조직 슬라이드는 표본의 제작마다 나타날 수 있는 검색의 정도 차이, 임플란트와 조직의 경계가 험거워지며 검색이 질게 되는 경우, 카메라 촬영에서 나타나는 음영과 같이 변수가 많이 발생한다. 따라서 색상을 표현하는 이미지 픽셀값(pixel intensity)에만 의존하여 분할이 이루어지는 방식을 사용할 경우 또다른 한계점이 나타난다.

치과용 임플란트의 뼈 조직형태계측에서 컴퓨터 보조 진단 기술 개발의 가장 필수적인 단계는 임플란트와 조직으로의 명확한 분할이 이루어지는 것이다. 하지만 현재까지 제안된 색상 인식 기반의 방법들은 이 기본적인 단계에서부터 번거로움과 한계점이 드러났다. 이 모든 것을 극복하기 위해 본 연구에서는 인공지능 딥러닝 CNN 모델의 적용을 시도하였고, 결과적으로 소량의 데이터에서도 색상을 기반으로 분할이 이루어지는 인공지능 방법론 ‘kNN’과 비교하여 우수한 정확도와 함께 정성적으로도 우월한 성능을 나타냈다. 더불어 분석하는 이미지마다 대표하는 색상값, 혹은 임계값을 지정해줘야 하는 기존의 번거로움 없이 사용이 가능하다.

본 연구에서는 적은 데이터 수를 보완하기 위하여 데이터 증강 기법을 이용하였고, 이를 통해 더 향상된 연구 결과를 도출하였다. 높은 성능을 보이는 딥러닝 훈련을 위해서는 대규모의 학습 데이터를 사용해야 하는 것이 일반적이다. 그리고 학습된 모델은 학습된 데이터와 유사한 양상의 이미지에서 높은 성능을 나타내지만, 새로 입력 받은 데이터의 속성이 학습된 데이터와 차이가 있는 경우 일반적으로 낮은 성능을 보이기도 한다. 따라서 학습용 데이터는 다양한 특징들을 포함하고 있어야 한다. 하지만 의료 영상은 학습 데이터의 주석(annotation)작업에 해당 분야 전문가가 참여해야 하는 특성으로 다른 분야보다 많은 노력과 시간이 필요하며 다량의 데이터를 수집하기 어렵다. 이럴 때 데이터의 다양성을 확보하고 딥러닝 성능을 높이고자 사용하는 방법이 ‘데이터 증강’이다 [29].

데이터 증강에는 전통적 영상처리 기반의 증강 기법과 적대적생성네트워크(Generative Adversial Network, GAN) 기반의 증강기법, 그리고

최근에는 학습 영상 일부를 잘라내거나 영상들을 섞는 방식으로 간단하면서도 학습 성능을 크게 끌어올리는 기법이 소개되었다. 그중 전통적 영상처리 기반 증강 기법에는 무작위로 뒤집기, 회전하기, 자르기 등 영상의 물리적 형태를 변화시키는 기법부터 영상의 색상이나 밝기를 변화하는 기법까지 다양하다[30]. 본 연구에서는 이미지 분석에서 흔히 사용되는 전통적 영상처리 기반 데이터 증강 기법인 영상의 회전, 이동, 반전을 이용하여 성능을 향상시켰다.

이처럼 소규모의 데이터를 이용하면서 나타날 수 있는 문제를 극복하고 높은 성능을 도출하였음에도, 본 연구에는 몇가지 제한점이 존재한다. 첫 번째로 연구에 사용된 데이터는 한 중동물 실험을 통해 같은 조건 하 표본이 제작되고 이미지 촬영이 진행된 것으로, 표본의 종류, 염색과 촬영 과정에서 나타날 수 있는 조직 슬라이드 이미지의 다양한 경우의 수를 반영하지 못하였다. 두 번째로 임플란트의 뼈 조직학적 분석에서 중요한 평가지표인 Bone to implant contact(BIC)나 Bone area(BA)를 추출해 내기 위해서는 ①임플란트, ②뼈, ③뼈가 아닌 영역의 세가지 클래스에 대한 명확한 분할까지 달성해야 한다. 이를 위해서는 정확한 주석(annotation)이 포함된 데이터를 생성해야 하기에 전문가의 적극적인 협조가 동반되어야 한다. 따라서 추후 양적, 질적으로 보완된 데이터를 확보하고 검증하는 후속 연구가 필요하다.

기술적인 측면에서는 연구에 사용한 데이터 증강 방법 외에도 뼈 조직 슬라이드 이미지의 특성에 맞추어 최적의 방법을 탐색해 볼 필요가 있다[31]. 의료 영상에서 역시 앞서 언급한 GAN이나 그 외 최신 기법을 적용하는 연구가 이루어지고 있다. 임플란트 조직 이미지에서 전통적 영상처리 외에도 다양한 증강 기법을 적용해 보아야 한다. 더불어 U-Net

모델을 기본 골격으로 한 다양한 모델이 선행 연구로서 제안되었다. 다양한 모델에 적용하여 결과를 비교해볼 수 있다. 본 연구결과를 토대로 지속적인 연구를 통해 인공지능 딥러닝 기법은 뛰어난 성능과 함께 뼈조각형태계측 분야에서 수동 개입 의존도를 낮춘 효율적이고 일관적인 보조 도구로서 활용이 가능할 것으로 기대한다.

참 고 문 헌

- [1] C. E. Misch, "The importance of dental implants.," *Gen Dent*, vol. 49, no. 1, pp. 38 - 45, 2001.
- [2] "Medical device production and import/export statistics in 2020," Ministry of Food and Drug Safety, 2021.
- [3] H. Sarve, "Evaluation of osseointegration using image analysis and visualization of 2D and 3D image data," 2011.
- [4] 구강소화기과식품의약품안전처 의료기기심사부, "치과용임플란트의 조직반응 평가방법에 관한 가이드라인(민원인 안내서).pdf." 식품의약품안전처 의료기기심사부 구강소화기과, 2015.
- [5] C. Manresa, M. Bosch, M. C. Manzanares, P. Carvalho, and J. J. Echeverría, "A new standardized automatic method for bone to implant contact histomorphometric analysis based on backscattered scanning electron microscopy images," *Clin Oral Implan Res*, vol. 25, no. 6, pp. 702 - 706, 2014, doi: 10.1111/clr.12129.
- [6] D. Malhan et al., "An Optimized Approach to Perform Bone Histomorphometry," *Front Endocrinol*, vol. 9, p. 666, 2018, doi: 10.3389/fendo.2018.00666.
- [7] O. Kujan, A. Khattab, R. J. Oliver, S. A. Roberts, N. Thakker, and P. Sloan, "Why oral histopathology suffers inter-observer variability on grading oral epithelial dysplasia: An attempt to understand the sources of variation," *Oral Oncol*, vol. 43, no. 3, pp. 224 - 231, 2007, doi: 10.1016/j.oraloncology.2006.03.009.
- [8] A. C. Chi, N. Katabi, H.-S. Chen, and Y.-S. L. Cheng, "Interobserver Variation Among Pathologists in Evaluating Perineural Invasion for Oral Squamous Cell Carcinoma," *Head Neck Pathology*, vol. 10, no. 4, pp. 451 - 464, 2016, doi: 10.1007/s12105-016-0722-9.
- [9] S.-H. Hong, X. Jiang, L. Chen, P. Josh, D.-G. Shin, and D. Rowe, "Computer-Automated Static, Dynamic and Cellular Bone

- Histomorphometry,” *J Tissue Sci Eng*, vol. Suppl 1, no. 01, p. 004, 2012, doi: 10.4172/2157-7552.s1-004.
- [10] I. Arganda-Carreras et al., “Trainable Weka Segmentation: a machine learning tool for microscopy pixel classification,” *Bioinformatics*, vol. 33, no. 15, pp. 2424 - 2426, 2017, doi: 10.1093/bioinformatics/btx180.
- [11] R. J. van 't Hof, L. Rose, E. Bassonga, and A. Daroszewska, “Open source software for semi-automated histomorphometry of bone resorption and formation parameters,” *Bone*, vol. 99, 2017, doi: 10.1016/j.bone.2017.03.051.
- [12] K. Bera, K. A. Schalper, D. L. Rimm, V. Velcheti, and A. Madabhushi, “Artificial intelligence in digital pathology – new tools for diagnosis and precision oncology,” *Nat Rev Clin Oncol*, vol. 16, no. 11, pp. 703 - 715, 2019, doi: 10.1038/s41571-019-0252-y.
- [13] 대한병리학회의료정보연구회 et al., “디지털병리 가이드라인 권고안.” 대한병리학회의료정보연구회, 2020.
- [14] G. Litjens et al., “A survey on deep learning in medical image analysis,” *Med Image Anal*, vol. 42, pp. 60 - 88, 2017, doi: 10.1016/j.media.2017.07.005.
- [15] G. Litjens et al., “Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis,” *Sci Rep-uk*, vol. 6, no. 1, p. 26286, 2016, doi: 10.1038/srep26286.
- [16] S. Wang, D. M. Yang, R. Rong, X. Zhan, and G. Xiao, “Pathology Image Analysis Using Segmentation Deep Learning Algorithms,” *Am J Pathology*, vol. 189, no. 9, pp. 1686 - 1698, 2019, doi: 10.1016/j.ajpath.2019.05.007.
- [17] S. Christian et al., “Going Deeper with Convolutions,” 2015 Ieee Conf Comput Vis Pattern Recognit Cvpr, pp. 1 - 9, 2015, doi: 10.1109/cvpr.2015.7298594.
- [18] A. Rehman, N. Abbas, T. Saba, S. I. ur Rahman, Z. Mehmood,

- and H. Kolivand, "Classification of acute lymphoblastic leukemia using deep learning," *Microsc Res Techniq*, vol. 81, no. 11, pp. 1310 - 1317, 2018, doi: 10.1002/jemt.23139.
- [19] W.-H. Kim et al., "Dental implants with electrochemical nanopattern formation to increase osseointegration," *Journal of Industrial and Engineering Chemistry*.(submitted paper)
- [20] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *Lect Notes Comput Sc*, pp. 234 - 241, 2015, doi: 10.1007/978-3-319-24574-4_28.
- [21] J. Bertels et al., "Optimizing the Dice Score and Jaccard Index for Medical Image Segmentation: Theory & Practice," *Arxiv*, pp. 92 - 100, 2019, doi: 10.1007/978-3-030-32245-8_11.
- [22] L. R. Dice, "Measures of the Amount of Ecologic Association Between Species," *Ecology*, vol. 26, no. 3, pp. 297 - 302, 1945, doi: 10.2307/1932409.
- [23] G. V. Research, "Dental Implant Market Size, Share & Trends Analysis Report By Implants Type (Titanium, Zirconium), By Region (North America, Europe, Asia Pacific, Latin America, MEA), And Segment Forecasts, 2021 - 2028," 2021.
- [24] J. G. Elmore et al., "Diagnostic Concordance Among Pathologists Interpreting Breast Biopsy Specimens," *Jama*, vol. 313, no. 11, pp. 1122 - 1132, 2015, doi: 10.1001/jama.2015.1405.
- [25] J. Wang, H. Zhu, S.-H. Wang, and Y.-D. Zhang, "A Review of Deep Learning on Medical Image Analysis," *Mob Networks Appl*, vol. 26, no. 1, pp. 351 - 380, 2021, doi: 10.1007/s11036-020-01672-7.
- [26] H. Chen, X. Qi, L. Yu, Q. Dou, J. Qin, and P.-A. Heng, "DCAN: Deep contour-aware networks for object instance segmentation from histology images," *Med Image Anal*, vol. 36, pp. 135 - 146, 2017, doi: 10.1016/j.media.2016.11.004.

- [27] S. R. Granter, A. H. Beck, and D. J. P. Jr, "AlphaGo, Deep Learning, and the Future of the Human Microscopist," *Arch Pathol Lab Med*, vol. 141, no. 5, pp. 619 - 621, 2017, doi: 10.5858/arpa.2016-0471-ed.
- [28] B. E. Bejnordi et al., "Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer," *Jama*, vol. 318, no. 22, pp. 2199 - 2210, 2017, doi: 10.1001/jama.2017.14585.
- [29] H. R. Roth et al., "Anatomy-Specific Classification of Medical Images Using Deep Convolutional Nets," 2015 *Ieee 12th Int Symposium Biomed Imaging Isbi*, pp. 101 - 104, 2015, doi: 10.1109/isbi.2015.7163826.
- [30] A. Mikolajczyk and M. Grochowski, "Data augmentation for improving deep learning in image classification problem," 2018 *Int Interdiscip Phd Work Iiphdw*, pp. 117 - 122, 2018, doi: 10.1109/iiphdw.2018.8388338.
- [31] D. Tellez et al., "Quantifying the effects of data augmentation and stain color normalization in convolutional neural networks for computational pathology," *Med Image Anal*, vol. 58, p. 101544, 2019, doi: 10.1016/j.media.2019.101544.

ABSTRACT

A Study on the Applicability of Deep Learning Technique as an Assistant for Histological Analysis of Dental Implants

: A Consideration on the Separation
of the Interface with Surrounding Tissues

Min-Sun Lee

Department of Oral Pathology

The Graduate School

Seoul National University

Objectives

The purpose of this study is to examine whether deep learning technique can be used as an efficient and consistent aids in the bone histomorphometry of dental implant tissue images. To this end, the image segmentation accuracy of the implant and the surrounding tissue was evaluated compared to the previously developed staining color-based method.

Methods

In this study, images taken at 200x magnification from 16 dental implant

specimens placed in the tibiae of rabbits were used. Through image annotation processing for use as ground truth, an entire image was divided into two classes: an implant area and a tissue area. The region of interest(ROI) was set around the boundary line of the two classes. The artificial intelligence methodologies used in this study are kNN and deep learning U-Net models, where kNN is a method of segmenting color based on the representative pixel value(instance) of each class. The U-Net is a model actively used in AI-based medical image segmentation research. This study used the Vanilla U-Net, which is the most basic structure. To compare the results between the two methodologies, test images not used for deep learning training were used. The ground truth image and the predicted image were compared on a pixel-by-pixel basis, and the accuracy, sensitivity, and the area under the ROC curve (AUROC) were evaluated. Also, the segmentation accuracy was compared using two evaluation indicators: IOU (Intersection Over Union) and DSC (Dice Similarity Coefficient).

Results

An instance was designated by randomly selecting a single image from kNN, and the accuracy with ground truth was compared as an IOU index by extracting predicted results from all image data. The accuracy of ROI was 90.39% on average in the range of 65-96%. Of these, the lowest accuracy was 65.83%, which rose to 91.39% when the instances were reselected according to the image, and in the predicted images, it was observed that the boundaries of each class became clearer. Through this, it was confirmed that there is a limitation of requiring the observer's manual intervention to adjust the reference value each time the tissue image to be evaluated is changed, which is the same limitation shown in the previously developed staining color-based method. In the ROI, based on the implant, the average sensitivity was 90.68% for kNN and 98.32% for U-Net with

data augmentation. The accuracy was 94.89% in kNN and 97.86% in U-Net, and AUROC was 93.69% in kNN and 99.79% in U-Net. IOU was 84.19% in kNN and 93.4% in U-Net, and DSC was 91.03% in kNN and 96.58% in U-Net model. When the deep learning method was used, superior performance was confirmed compared to the existing color-based segmentation method. In addition, in the segmented image, it was shown that the segmentation into two classes was more clear than kNN without encroaching on each other.

Conclusions

Deep learning technique can compensate for differences in results between observers that may appear when measuring bone to implant contact (BIC) and bone area (BA), which are important indicators in histological analysis of dental implants. The potential of deep learning technique was confirmed for development as an auxiliary tool to improve the consistency and efficiency of analysis. If tissue images including various aspects are obtained and training based on expert annotations is accompanied, it is determined that it could be used as an assisting means of bone-tissue analysis in pathology laboratories by possibly achieving segmentation accuracy from tissue area to bone and non-bone area.

Key words: Bone Histomorphometry, Dental Implant Segmentation, Deep Learning, Medical Image Analysis

Student Number: 2020-24988