



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0107030
(43) 공개일자 2024년07월08일

- (51) 국제특허분류(Int. Cl.)
G06T 7/11 (2017.01) A61B 6/00 (2024.01)
A61B 6/51 (2024.01) G06T 11/00 (2006.01)
G06T 5/70 (2024.01) G06T 7/136 (2017.01)
G16H 30/40 (2018.01)
- (52) CPC특허분류
G06T 7/11 (2017.01)
A61B 6/51 (2024.01)
- (21) 출원번호 10-2023-0194765
(22) 출원일자 2023년12월28일
심사청구일자 2023년12월28일
- (30) 우선권주장
1020220187246 2022년12월28일 대한민국(KR)

- (71) 출원인
조선대학교산학협력단
광주광역시 동구 조선대길 126(서석동, 산학협력단)
주식회사 에이치티코어
광주광역시 동구 조선대길 7, 2215호(서석동, 조선대학교 치과대학)
- (72) 발명자
이범식
광주광역시 남구 제석로80번길 5
가꾸어 무함마드 아프
광주광역시 동구 지산로52번길 11-3
(뒷면에 계속)
- (74) 대리인
특허법인아이엠

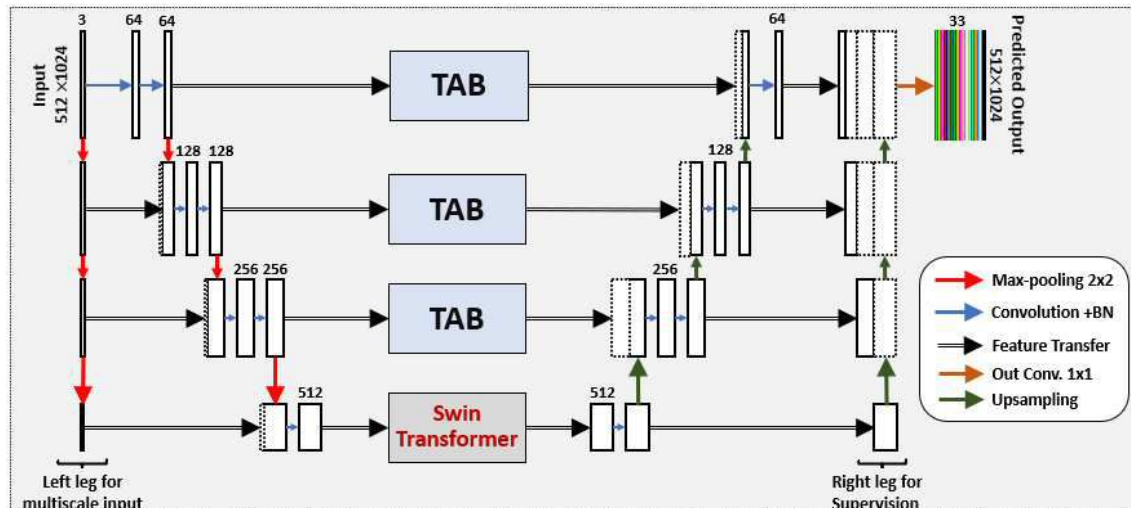
전체 청구항 수 : 총 24 항

(54) 발명의 명칭 치과 X-ray 이미지 다중클래스 분할방법 및 장치

(57) 요약

본 발명은 의료영상에 대한 영상분할기법 관련기술로서, 보다 구체적으로는 기존에 존재하는 U-Net의 기본 구조에 인코더 왼쪽에 형성된 왼쪽 다리 및 디코더 오른쪽에 형성된 오른쪽 다리를 포함하는 M-Net 유사 모델에서, 상기 인코더와 디코더 사이의 스킵 연결에 트랜스포머 및 TAB(Tooth Attention Block)를 통합하여 의료영상 특히 치과 X-ray 이미지를 정확하게 분할 할 수 있는 치과 X-ray 이미지 다중클래스 분할방법 및 장치에 관한 것이다.

대표도 - 도4



(52) CPC특허분류

A61B 6/5217 (2020.08)
G06T 11/003 (2013.01)
G06T 5/70 (2024.01)
G06T 7/136 (2017.01)
G16H 30/40 (2018.01)
G06T 2207/10116 (2013.01)
G06T 2207/20084 (2013.01)
G06T 2207/30036 (2013.01)
G06T 2210/12 (2013.01)

전시윤

광주광역시 동구 필문대로253번길 10-1 조대명품타운 101동 305호

(72) 발명자

이수지

광주광역시 광산구 송정로 80 온세계아이조움아파트 101동 803호

김규빈

광주광역시 북구 첨단연신로133번길 8 106동 303호

이 발명을 지원한 국가연구개발사업

과제고유번호	1345354483
과제번호	2022R1I1A3065473
부처명	교육부
과제관리(전문)기관명	한국연구재단
연구사업명	이공학학술연구기반구축
연구과제명	과거 영상 복원을 위한 영상 및 비디오 디올디파잉 연구
기 여 율	1/2
과제수행기관명	조선대학교
연구기간	2022.06.01 ~ 2023.02.28

이 발명을 지원한 국가연구개발사업

과제고유번호	2023002400
과제번호	ITAS0318220110020001000100100
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신산업진흥원
연구사업명	인공지능중신산업융합 집적단지 조성사업 AI융합대학 지원사업
연구과제명	AI융합대학지원사업
기 여 율	1/2
과제수행기관명	조선대학교 산학협력단
연구기간	2023.01.01 ~ 2023.12.31

명세서

청구범위

청구항 1

치과 X-ray 이미지를 입력받는 단계; 상기 입력된 치과 X-ray 이미지를 기 학습된 학습 모델에 적용하여 영상 분할을 수행하는 단계; 및 상기 영상 분할이 수행된 치과 X-ray 이미지를 제공하는 단계;를 포함하며,

상기 학습 모델은,

U-Net 모델에서 인코더 왼쪽에 형성된 왼쪽 다리 및 디코더 오른쪽에 형성된 오른쪽 다리를 포함하는 M-Net 유사 모델로서, 상기 인코더와 디코더 사이의 스킵 연결에 통합되어 형성된 TAB(Teeth Attention Block) 및 스윈 트랜스포머(Swin Transformer)를 포함하는, 치과 X-ray 이미지 다중클래스 분할방법.

청구항 2

제 1 항에 있어서,

상기 학습 모델에서 상기 TAB는 경계 인식 또는 경계 개선 필터 역할을 수행하여 복잡한 객체 경계 묘사 정확도가 향상되도록 어텐션 기능을 미세 조정하고,

상기 스윈트랜스포머는 입력 특징 맵을 겹치지 않는 로컬 창으로 나누어 효율적으로 처리하고 MSA(mult-head self-attention) 메커니즘을 활용하여 여러 관계를 동시에 학습하여 치과 X-ray 이미지의 장거리 종속성을 캡처하는 것을 특징으로 하는 치과 X-ray 이미지 다중클래스 분할방법.

청구항 3

제 2 항에 있어서,

상기 스윈트랜스포머는 상기 인코더와 디코더 사이의 스킵 연결 중 최하단레이어에 통합되어 병목현상을 방지하고, 상기 TAB는 최하단레이어를 제외한 나머지 레이어에 통합되어 상기 인코더의 각 레이어에서 출력되는 특징 맵에서 노이즈와 의미 없는 정보의 영향을 감소시키는 것을 특징으로 하는 치과 X-ray 이미지 다중클래스 분할방법.

청구항 4

제 1 항 내지 제 3 항 중 어느 한 항에 있어서,

상기 영상 분할을 수행하는 단계는 상기 치과 X-ray 이미지를 인코더로 입력하는 제1단계; 인코더에서 출력된 특징 맵을 TAB 또는 스윈트랜스포머로 입력하는 제2단계; 상기 TAB 또는 스윈트랜스포머에서 출력된 특징 맵을 상기 디코더로 입력하는 제3단계; 및 상기 디코더에서 출력된 결과값을 분류하는 제4단계;를 포함하는 것을 특징으로 하는 치과 X-ray 이미지 다중클래스 분할방법.

청구항 5

제 4 항에 있어서, 상기 제1단계는

상기 왼쪽다리의 제1레이어에서 X-ray 치아이미지에 대해 상기 인코더의 제1 레이어용 패치 이미지를 생성하고, 상기 제1레이어용 패치이미지에 대해 맥스풀링을 수행하여 그 결과값을 제2레이어의 입력으로 제공하는 단계;를 포함하고,

상기 인코더의 제1레이어에서 수행되는 상기 제1 레이어에 대한 패치 이미지에 대해 컨볼루션(convolution)을 수행하는 단계; 상기 컨볼루션의 결과값에 대해 맥스 풀링(max pooling)이 수행되는 단계; 및 상기 맥스풀링 결과값을 상기 인코더의 제2 레이어의 입력으로 제공하는 단계;를 포함하는 것을 특징으로 하는 치과 X-ray 이미지 다중클래스 분할방법.

청구항 6

제 5 항에 있어서, 상기 제1단계는

상기 왼쪽다리의 제2레이어에서 수행되는 입력된 사이즈 그대로 인코더의 제2 레이어용 패치이미지를 생성하고, 상기 제2레이어용 패치이미지에 대해 맥스풀링을 수행하여 그 결과값을 제3레이어의 입력으로 제공하는 단계;를 포함하며,

상기 인코더의 제2레이어에서 상기 입력된 제1 레이어에 대한 맥스풀링의 결과값 및 상기 제2레이어용 패치 이미지를 연결(concatenate)시켜 컨볼루션(convolution)을 수행하는 단계; 상기 컨볼루션의 결과값에 대해 맥스풀링(max pooling)이 수행되는 단계; 및 상기 맥스풀링 결과값을 상기 인코더의 제3 레이어의 입력으로 제공하는 단계;를 포함하는 것을 특징으로 하는 치과 X-ray 이미지 다중클래스 분할방법.

청구항 7

제 6 항에 있어서, 상기 제1단계는

상기 왼쪽다리의 제3레이어에서 수행되는 입력된 사이즈 그대로 인코더의 제3레이어용 패치이미지를 생성하고, 상기 제3레이어용 패치이미지에 대해 맥스풀링을 수행하여 그 결과값을 제4레이어의 입력으로 제공하는 단계;를 포함하며,

상기 인코더의 제3레이어에서 상기 입력된 제2 레이어에 대한 맥스풀링의 결과값 및 상기 제3레이어용 패치 이미지를 연결(concatenate)시켜 컨볼루션(convolution)을 수행하는 단계; 상기 컨볼루션의 결과값에 대해 맥스풀링(max pooling)이 수행되는 단계; 및 상기 맥스풀링 결과값을 상기 인코더의 제4 레이어의 입력으로 제공하는 단계;를 포함하는 것을 특징으로 하는 치과 X-ray 이미지 다중클래스 분할방법.

청구항 8

제 7 항에 있어서, 상기 제1단계는

상기 왼쪽다리의 제4레이어에서 수행되는 입력된 사이즈 그대로 인코더의 제4 레이어용 패치이미지를 생성하는 단계;를 포함하고,

상기 인코더의 제4레이어에서 상기 입력된 제3 레이어에 대한 맥스풀링의 결과값 및 상기 제4레이어용 패치 이미지를 연결(concatenate)시켜 컨볼루션(convolution)을 수행하는 단계;를 포함하는 것을 특징으로 하는 치과 X-ray 이미지 다중클래스 분할방법.

청구항 9

제 8 항에 있어서, 상기 제2단계는

상기 인코더의 각 레이어에서 출력되는 특징 맵이 해당 레이어에 위치한 TAB의 입력으로 제공되는 단계; 및

상기 인코더의 최하단 레이어에서 출력되는 특징 맵이 스윙트랜스포머의 입력으로 제공되는 단계;를 포함하는 것을 특징으로 하는 치과 X-ray 이미지 다중클래스 분할방법.

청구항 10

제 9 항에 있어서,

상기 TAB의 입력으로 제공되는 단계는 RPN(Region proposal network)에 의해 치아 구조를 포함하는 잠재적 영역을 나타내는 경계 상자(boundary boxes)가 감지되고, 상기 감지된 경계 상자는 CAB(Channel Attention Block)를 통과하면서 특징 맵의 어텐션이 미세 조정된 후 상기 인코더의 스킵 연결과 곱연산되어 수행되는 것을 특징으로 하는 치과 X-ray 이미지 다중클래스 분할방법.

청구항 11

제 10 항에 있어서, 상기 제3단계는

상기 스윈트랜스포머에서 출력되는 특징 맵을 상기 디코더의 최하단레이어의 입력으로 제공하고, 상기 디코더의 차하위 레이어에서 출력되는 특징 맵과 상기 TAB의 해당레이어에서 출력되는 특징 맵을 연결시켜 상기 디코더의 해당레이어의 입력으로 제공하는 단계;

상기 디코더의 해당 레이어에 대한 컨볼루션(convolution)을 수행하는 단계; 및

상기 디코더의 해당 레이어에서 출력되는 특징 맵을 오른쪽 다리의 해당 레이어의 입력으로 제공하면서, 업샘플링하여 상기 디코더의 차상위 레이어에 입력으로 제공하는 단계;를 포함하는 것을 특징으로 하는 치과 X-ray 이미지 다중클래스 분할방법.

청구항 12

제 11 항에 있어서, 상기 제3단계는

상기 디코더의 최하단레이어에서 출력되는 특징 맵을 입력받아 상기 오른쪽다리의 최하단 레이어에 대한 컨볼루션(convolution)을 수행하고, 상기 디코더의 최하단레이어를 제외한 각 레이어에서 출력되는 특징 맵과 차하위 레이어에서 얻어진 특징 맵을 연결시켜 상기 오른쪽다리의 해당 레이어에 대한 컨볼루션(convolution)을 수행하는 단계;

상기 오른쪽 다리의 해당 레이어에 대한 컨볼루션 결과값에 대해 업샘플링이 수행된 결과값이 차상위 레이어에 입력으로 제공되는 단계; 및

상기 오른쪽 다리의 2번째 레이어에 대한 컨볼루션 결과값에 대해 업샘플링이 수행된 결과값과 상기 디코더의 최상단 레이어에서 컨볼루션을 마친 특징 맵을 연결시켜 컨볼루션하는 단계; 를 포함하는 것을 특징으로 하는 치과 X-ray 이미지 분할 방법.

청구항 13

치과 X-ray 이미지를 입력받는 입력부;

상기 입력된 치과 X-ray 이미지를 기학습된 학습 모델에 적용하여 시맨틱 분할을 수행하는 영상 분할부; 및

상기 시맨틱 분할이 수행된 치과 X-ray 이미지를 제공하는 영상 제공부;를 포함하며,

상기 학습 모델은,

U-Net 모델에서 인코더 왼쪽에 형성된 왼쪽 다리 및 디코더 오른쪽에 형성된 오른쪽 다리를 포함하는 M-Net 유사 모델로서, 상기 인코더와 디코더 사이의 스킵 연결에 통합되어 형성된 TAB(Tooth Attention Block) 및 스윈트랜스포머(Swin Transformer)를 포함하는, 치과 X-ray 이미지 다중클래스 분할장치.

청구항 14

제 13 항에 있어서,

상기 학습 모델에서 상기 TAB는 경계 인식 또는 경계 개선 필터 역할을 수행하여 복잡한 객체 경계 묘사 정확도가 향상되도록 어텐션 기능을 미세 조정하고,

상기 스윈트랜스포머는 입력 특징 맵을 겹치지 않는 로컬 창으로 나누어 효율적으로 처리하고 MSA(multi-head self-attention) 메커니즘을 활용하여 여러 관계를 동시에 학습하여 치과 X-ray 이미지의 장거리 종속성을 캡처하는 것을 특징으로 하는 치과 X-ray 이미지 다중클래스 분할장치.

청구항 15

제 14 항에 있어서,

상기 스윈트랜스포머는 상기 인코더와 디코더 사이의 스킵 연결 중 최하단레이어에 통합되어 병목현상을 방지하고, 상기 TAB는 최하단레이어를 제외한 나머지 레이어에 통합되어 상기 인코더의 각 레이어에서 출력되는 특징 맵에서 노이즈와 의미 없는 정보의 영향을 감소시키는 것을 특징으로 하는 치과 X-ray 이미지 다중클래스 분할장치.

청구항 16

제 13 항 내지 제15항 중 어느 한 항에 있어서,

상기 영상분할부는 상기 치과 X-ray 이미지를 상기 인코더로 입력하고, 상기 인코더에서 출력된 특징 맵을 TAB 또는 스윈트랜스포머로 입력하며, 상기 TAB 또는 스윈트랜스포머에서 출력된 특징 맵을 상기 디코더로 입력하고, 상기 디코더에서 출력된 결과값을 분류하는 것을 특징으로 하는 치과 X-ray 이미지 다중클래스 분할장치.

청구항 17

제 16 항에 있어서,

상기 영상분할부는 상기 왼쪽다리의 제1레이어에서 X-ray 치아이미지에 대해 상기 인코더의 제1 레이어용 패치 이미지를 생성하고, 상기 제1레이어용 패치이미지에 대해 맥스풀링을 수행하여 그 결과값을 제2레이어의 입력으로 제공하는 것을 포함하고,

상기 인코더의 제1레이어에서 수행되는 상기 제1 레이어에 대한 패치 이미지에 대해 컨볼루션(convolution)을 수행하며, 상기 컨볼루션의 결과값에 대해 맥스 풀링(max pooling)을 수행한 후 상기 맥스풀링 결과값을 상기 인코더의 제2 레이어의 입력으로 제공하는 것을 특징으로 하는 치과 X-ray 이미지 다중클래스 분할장치.

청구항 18

제 17 항에 있어서,

상기 영상분할부는 상기 왼쪽다리의 제2레이어에서 수행되는 입력된 사이즈 그대로 인코더의 제2 레이어용 패치 이미지를 생성하고, 상기 제2레이어용 패치이미지에 대해 맥스풀링을 수행하여 그 결과값을 제3레이어의 입력으로 제공하는 것을 포함하며,

상기 인코더의 제2레이어에서 상기 입력된 제1 레이어에 대한 맥스풀링의 결과값 및 상기 제2레이어용 패치 이미지를 연결(concatenate)시켜 컨볼루션(convolution)을 수행하고 상기 컨볼루션의 결과값에 대해 맥스 풀링(max pooling)을 수행한 후 상기 맥스풀링 결과값을 상기 인코더의 제3 레이어의 입력으로 제공하는 것을 특징으로 하는 치과 X-ray 이미지 다중클래스 분할장치.

청구항 19

제 18 항에 있어서,

상기 영상분할부는 상기 왼쪽다리의 제3레이어에서 수행되는 입력된 사이즈 그대로 인코더의 제3레이어용 패치 이미지를 생성하고, 상기 제3레이어용 패치이미지에 대해 맥스풀링을 수행하여 그 결과값을 제4레이어의 입력으로 제공하는 것 포함하며,

상기 인코더의 제3레이어에서 상기 입력된 제2 레이어에 대한 맥스풀링의 결과값 및 상기 제3레이어용 패치 이미지를 연결(concatenate)시켜 컨볼루션(convolution)을 수행하고, 상기 컨볼루션의 결과값에 대해 맥스 풀링(max pooling)을 수행한 후 상기 맥스풀링 결과값을 상기 인코더의 제4 레이어의 입력으로 제공하는 것을 특징으로 하는 치과 X-ray 이미지 다중클래스 분할장치.

청구항 20

제 16 항에 있어서,

상기 영상분할부는 상기 왼쪽다리의 제4레이어에서 수행되는 입력된 사이즈 그대로 인코더의 제4 레이어용 패치 이미지를 생성하는 것을 포함하고,

상기 인코더의 제4레이어에서 상기 입력된 제3 레이어에 대한 맥스풀링의 결과값 및 상기 제4레이어용 패치 이미지를 연결(concatenate)시켜 컨볼루션(convolution)을 수행하는 것을 특징으로 하는 치과 X-ray 이미지 다중클래스 분할장치.

청구항 21

제 16 항에 있어서,

상기 영상분할부는 상기 인코더의 각 레이어에서 출력되는 특징 맵을 해당 레이어에 위치한 TAB의 입력으로 제공하고, 상기 인코더의 최하단 레이어에서 출력되는 특징 맵을 스윈트랜스포머의 입력으로 제공하는 것을 특징으로 하는 치과 X-ray 이미지 다중클래스 분할장치.

청구항 22

제 21 항에 있어서,

상기 영상분할부는 상기 TAB의 입력을 RPN(Region proposal network)에 의해 치아 구조를 포함하는 잠재적 영역을 나타내는 경계 상자(boundary boxes)를 감지하고, 상기 감지된 경계 상자가 CAB(Channel Attention Block)를 통과하면서 특징 맵의 어텐션을 미세 조정한 후 상기 인코더의 스킵 연결과 곱연산시켜 수행하는 것을 특징으로 하는 치과 X-ray 이미지 다중클래스 분할장치.

청구항 23

제 22 항에 있어서,

상기 영상분할부는 상기 스윈트랜스포머에서 출력되는 특징 맵을 상기 디코더의 최하단레이어의 입력으로 제공하고, 상기 디코더의 차하위 레이어에서 출력되는 특징 맵과 상기 TAB의 해당레이어에서 출력되는 특징 맵을 연결시켜 상기 디코더의 해당레이어의 입력으로 제공하며, 상기 디코더의 해당 레이어에 대한 컨볼루션(convolution)을 수행하고, 상기 디코더의 해당 레이어에서 출력되는 특징 맵을 오른쪽 다리의 해당 레이어의 입력으로 제공하면서, 업샘플링하여 상기 디코더의 차상위 레이어에 입력으로 제공하는 것을 특징으로 하는 치과 X-ray 이미지 다중클래스 분할장치.

청구항 24

제 23 항에 있어서,

상기 영상분할부는 상기 디코더의 최하단레이어에서 출력되는 특징 맵을 입력받아 상기 오른쪽다리의 최하단 레이어에 대한 컨볼루션(convolution)을 수행하고, 상기 디코더의 최하단레이어를 제외한 각 레이어에서 출력되는 특징 맵과 차하위 레이어에서 얻어진 특징 맵을 연결시켜 상기 오른쪽다리의 해당 레이어에 대한 컨볼루션(convolution)을 수행하며, 상기 오른쪽 다리의 해당 레이어에 대한 컨볼루션 결과값에 대해 업샘플링이 수행된 결과값을 차상위 레이어에 입력으로 제공하며, 상기 오른쪽 다리의 2번째 레이어에 대한 컨볼루션 결과값에 대해 업샘플링이 수행된 결과값과 상기 디코더의 최상단 레이어에서 컨볼루션을 마친 특징 맵을 연결시켜 컨볼루션하는 것을 특징으로 하는 치과 X-ray 이미지 다중클래스 분할장치.

발명의 설명

기술 분야

[0001] 본 발명은 의료영상에 대한 영상분할기법 관련기술로서, 보다 구체적으로는 기존에 존재하는 U-Net의 기본 구조에 인코더 왼쪽에 형성된 왼쪽 다리 및 디코더 오른쪽에 형성된 오른쪽 다리를 포함하는 M-Net 유사 모델에서, 상기 인코더와 디코더 사이의 스킵 연결에 트랜스포머 및 TAB(Tooth Attention Block)를 통합하여 의료영상 특히 치과 X-ray 이미지를 정확하게 분할 할 수 있는 치과 X-ray 이미지 다중클래스 분할방법 및 장치에 관한 것이다.

배경 기술

[0002] 복잡한 절차를 단순화하고 예측가능한 결과를 제공하는 동시에 의사가 환자 안전을 높이는 데 도움이 되면서 의료영상 분야의 인공 지능 애플리케이션이 성장하고 있다. 의료영상 분석은 이상 탐지, 영상 분할, 분류 등 여러 가지 이점을 제공하는 딥러닝 기술을 사용한다. 분할은 의료 영상 처리에서 광범위하게 다루어지는 문제이다. 좋은 분할 결과를 얻는 것은 어렵고 기존 방법은 정확한 이미지 경계를 찾는 데 심각한 한계가 있다. 인간의 치아와 같은 의료영상의 다중 클래스 라벨링은 덜 탐구된 문제이며 다중 라벨 분할을 수행하는 정확한 방법이 없다.

[0003] 연구자들은 의료영상 분할을 향상시키기 위해 여러 가지 방법을 사용했다. Yamanaknavar와 Lee는 자동화된 뇌 MRI 분할을 위해 글로벌 주목 CNN 모델을 사용한 새로운 M-SegNet 아키텍처를 제시했다. 기본 아키텍처인 M-Net은 뇌 이미지 분할에도 사용된다. Badrinarayanan et al. 는 심층 컨볼루션 인코더-디코더 아키텍처인 SegNet을 개발하여 의료 영상 애플리케이션에서 큰 성공을 거두었다. Gu et al. 은 2차원 의료영상 분할을 위한 상황 인코딩 네트워크인 CE-Net을 제안했다. Lin et al.은 의미론적 분할에서 심층 구조 모델에 대한 효율적인 조각별 학습 접근 방식을 제시했다. 태아 MRI를 사용하여 태반에 대한 최소 대화형 분할 접근 방식인 Slic-Seg는 Wang 등이 제안했다. Lee et al.는 자동화된 뇌 MRI 분할을 위한 패치별 U-Net 구조를 제안했다. 딥러닝 기반 기술은 다양한 의료 영상 관련 과제에서 탁월한 성능을 입증하며 그 잠재력을 강조하고 있다. 따라서 AI 시스템은 잠재적으로 건강 데이터 결과를 개선하고 의료 비용을 낮추며 의학 연구를 발전시킬 수 있다. 예를 들어, 치과 영상 분석에서 딥러닝 기술은 치아와 치아 구조를 분할하고 분류하는 데 있어 유망한 결과를 보여주었으며, 그 결과 다양한 치아 상태에 대한 진단 및 치료 계획이 향상되었다.

[0004] 치아 이미지 분할은 컴퓨터를 이용한 치과 진단에서 중요한 과정이며 치아 상태를 분석하는 초기 단계 역할을 한다. 치과의사는 파노라마 방사선 사진을 사용하여 상실 치아, 치아 발달, 매복 치아 및 인접 관계를 포함한 다양한 치아 상태를 평가할 수 있다. 이는 이미지 분할을 통해 달성된다. 현재 기술은 파노라마 X선 이미지에 대한 실측 식별 메커니즘을 사용하고 분할 아키텍처를 활용하여 잠재적으로 임상 진단을 용이하게 할 수 있는 정확한 분할 결과를 생성한다. 파노라마 치과 X-ray 스캔과 치아 이미지 분할 기술은 정확한 측정을 가능하게 하고 턱과 치아에 대한 포괄적인 보기를 제공하기 때문에 컴퓨터 보조 치과 진단의 필수 구성 요소이다. 예를 들어, 잠재적인 치아 상실로 이어지는 심각한 잇몸 감염인 치주염과 같은 심각한 치과 질환을 진단하려면 파노라마 이미지에서 정확한 치아 분할이 필수적이다. 상세한 세분화를 통해 치과의사는 치주 인대 공간 확대 또는 뼈 손실과 같은 치아 및 주변 구조의 이상을 식별할 수 있다. 그러나 초기 평가에는 세분화가 중요하지만 철저한 진단 및 치료 계획을 위해서는 교익 X-레이 또는 CBCT와 같은 보다 구체적인 영상 기술이 필요한 경우가 많다. 결과적으로 치아 분할은 치과 진단 도구의 초석으로 남아 있다.

[0005] 치아를 뚜렷한 그룹으로 정확하게 분류하는 것은 교정 치료 계획, 치과 임플란트 수술, 법의학 치아학과 같은 다양한 응용 분야에서 매우 중요함에도 불구하고 치과 이미지 분석에서 중요한 과제를 제기한다. 치과 이미지에

서 치아를 분할하기 위해 수동, 반자동 및 자동 접근 방식이 고안되었다. 이 분야의 발전에도 불구하고 치아 복원, 부정교합 및 병리학적 상태는 분할 성능에 영향을 미칠 수 있다. 이전에는 대부분 작업이 이진 분할을 기반으로 하며 치아 유형에 따라 치아의 다중 클래스 라벨링을 위한 감독 모델을 개발하기 위한 실측 데이터가 없다.

[0006] 또한, 다양한 의료 영상 분할 문제는 이기종 픽셀 강도, 노이즈/잘못 정의된 경계, 가변성이 높은 불규칙한 모양을 포함하여 중요한 기술적 문제를 제시한다. 의료 영상 분할에서 주요 문제는 데이터 세트의 가용성인데, 예를 들어 다른 클래스에 대한 치아 분할 데이터 세트를 사용할 수 없고, 의료영상은 인공물이 많아 일반 영상과 다르기 때문이다.

[0007] 인공물을 사용한 분할은 분할을 더 어렵고 어렵게 만든다. 도 1에 도시된 바와 같은 치아 영상 분할시에는 사람의 입에 다른 많은 조직과 뼈가 있어 분할하기가 더 어렵다.

[0008] CNN(Convolutional Neural Network)은 로컬 공간 데이터를 수집하고 특징 표현을 학습하는 기능으로 인해 치과 영상에서 이미지 분할을 위한 주요 기술로 등장했다. 그러나 Transformer와 같은 딥 러닝 아키텍처의 최근 발전은 장거리 종속성과 글로벌 컨텍스트를 효과적으로 모델링함으로써 여러 컴퓨터 비전 작업에서 CNN을 능가할 수 있는 잠재력을 보여주었다. 배경 소음으로부터 치아 및 기타 치아 구조를 정확하게 분할하는 CNN 기반 분할 모델의 발전으로 치아 상태에 대한 정밀한 분석 및 진단이 가능해졌다. 대규모 치과 이미지 데이터 세트의 가용성이 증가하고 딥러닝 기술이 지속적으로 발전함에 따라 CNN 기반 모델은 빠르고 정확하며 자동화된 이미지 분석을 촉진함으로써 치과 이미징에서 훨씬 더 중요한 역할을 할 것으로 예상된다.

[0009] 영역 제안 방법이라고도 알려진 경계 상자 필터는 특정 관심 영역에 초점을 맞춰 딥러닝 모델의 성능을 향상시키기 위해 의료 영상 분할 작업에 광범위하게 사용되었다. 폐 결절을 정확하게 분할할 때 경계 상자 필터를 사용하여 CT 스캔에서 결절을 식별했다. 유사하게, 궤장 분할 정확도를 높이기 위해 경계 상자 필터를 사용하거나 경계 상자 필터를 사용하여 대장내시경 이미지에서 병변을 성공적으로 분할했으며, 의료 이미지의 특정 관심 영역에 초점을 맞춰서 경계 상자 기술이 분할 정밀도를 향상시킬 수 있음을 입증했다. 치과 영상에서는 파노라마 방사선 사진과 치과용 콘빔 컴퓨터 단층촬영(CBCT)에서 관심 영역을 식별하여 치아와 기타 치아 구조를 분할하는 데 경계 상자 기술이 사용되었다. 이러한 연구는 경계 상자 필터가 치과 및 의료 이미지 분할의 정확성과 효율성을 크게 향상시켜 다양한 조건에 대한 진단 및 치료 계획을 향상시킬 수 있는 잠재력을 가지고 있음을 시사한다.

[0010] Transformer는 최근 이미지 분할을 포함한 다양한 컴퓨터 비전 애플리케이션에서 탁월한 결과를 제공하는 견고한 딥 러닝 아키텍처로 등장했다. Transformers는 self-attention 프로세스를 사용하여 장기적인 종속성과 글로벌 환경을 성공적으로 표현함으로써 지리적 거리에 관계없이 입력 이미지의 픽셀 또는 특징 간의 연결을 기록할 수 있다. 이러한 특성으로 인해 트랜스포머는 치아와 주변 구조 사이의 상황과 관계를 정확하게 캡처하는 것이 중요한 치과 이미지 분할 작업에 잠재적으로 적합한 선택이 된다.

선행기술문헌

특허문헌

[0011] (특허문헌 0001) 국내특허공개번호 제10-2022-0129406호

발명의 내용

해결하려는 과제

[0012] 따라서, 본 발명의 목적은 관심 영역에 초점을 맞추는 모델의 기능이 촉진되어 로컬 및 글로벌 컨텍스트를 모두 효과적으로 캡처할 수 있어 정확한 치아 분할을 제공하기 때문에 치아 모양의 변화나 병변의 모양을 식별하여 치주질환이나 치아 우식의 발견을 용이하여 다양한 치과 질환의 조기 발견 및 진단이 가능하고, 상세한 치료 계획을 세우는 데 큰 도움이 될 수 있는 치과 X-ray 이미지 다중클래스 분할 방법 및 장치를 제공하는 것이다.

[0013] 본 발명의 목적들은 이상에서 언급한 목적들로 제한되지 않으며, 언급되지 않은 또 다른 목적들은 아래의 기재로부터 당업자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

- [0014] 수술된 본 발명의 목적을 달성하기 위해, 본 발명은 치과 X-ray 이미지를 입력받는 단계; 상기 입력된 치과 X-ray 이미지를 기 학습된 학습 모델에 적용하여 영상 분할을 수행하는 단계; 및 상기 영상 분할이 수행된 치과 X-ray 이미지를 제공하는 단계;를 포함하며, 상기 학습 모델은, U-Net 모델에서 인코더 왼쪽에 형성된 왼쪽 다리 및 디코더 오른쪽에 형성된 오른쪽 다리를 포함하는 M-Net 유사 모델로서, 상기 인코더와 디코더 사이의 스킵 연결에 통합되어 형성된 TAB(Teeth Attention Block) 및 스윈트랜스포머(Swin Transformer)를 포함하는, 치과 X-ray 이미지 다중클래스 분할방법을 제공한다.
- [0015] 바람직한 실시예에 있어서, 상기 학습 모델에서 상기 TAB는 경계 인식 또는 경계 개선 필터 역할을 수행하여 복잡한 객체 경계 묘사 정확도가 향상되도록 어텐션 기능을 미세 조정하고, 상기 스윈트랜스포머는 입력 특징 맵을 겹치지 않는 로컬 창으로 나누어 효율적으로 처리하고 MSA(mult-head self-attention) 메커니즘을 활용하여 여러 관계를 동시에 학습하여 치과 X-ray 이미지의 장거리 종속성을 캡처한다.
- [0016] 바람직한 실시예에 있어서, 상기 스윈트랜스포머는 상기 인코더와 디코더 사이의 스킵 연결 중 최하단레이어에 통합되어 병목현상을 방지하고, 상기 TAB는 최하단레이어를 제외한 나머지 레이어에 통합되어 상기 인코더의 각 레이어에서 출력되는 특징 맵에서 노이즈와 의미 없는 정보의 영향을 감소시킨다.
- [0017] 바람직한 실시예에 있어서, 상기 영상 분할을 수행하는 단계는 상기 치과 X-ray 이미지를 인코더로 입력하는 제1단계; 인코더에서 출력된 특징 맵을 TAB 또는 스윈트랜스포머로 입력하는 제2단계; 상기 TAB 또는 스윈트랜스포머에서 출력된 특징 맵을 상기 디코더로 입력하는 제3단계; 및 상기 디코더에서 출력된 결과값을 분류하는 제4단계;를 포함한다.
- [0018] 바람직한 실시예에 있어서, 상기 제1단계는 상기 왼쪽다리의 제1레이어에서 X-ray 치아이미지에 대해 상기 인코더의 제1 레이어용 패치 이미지를 생성하고, 상기 제1레이어용 패치이미지에 대해 맥스폴딩을 수행하여 그 결과값을 제2레이어의 입력으로 제공하는 단계;를 포함하고, 상기 인코더의 제1레이어에서 수행되는 상기 제1 레이어에 대한 패치 이미지에 대해 컨볼루션(convolution)을 수행하는 단계; 상기 컨볼루션의 결과값에 대해 맥스폴딩(max pooling)이 수행되는 단계; 및 상기 맥스폴딩 결과값을 상기 인코더의 제2 레이어의 입력으로 제공하는 단계;를 포함한다.
- [0019] 바람직한 실시예에 있어서, 상기 제1단계는 상기 왼쪽다리의 제2레이어에서 수행되는 입력된 사이즈 그대로 인코더의 제2 레이어용 패치이미지를 생성하고, 상기 제2레이어용 패치이미지에 대해 맥스폴딩을 수행하여 그 결과값을 제3레이어의 입력으로 제공하는 단계;를 포함하며, 상기 인코더의 제2레이어에서 상기 입력된 제1 레이어에 대한 맥스폴딩의 결과값 및 상기 제2레이어용 패치 이미지를 연결(concatenate)시켜 컨볼루션(convolution)을 수행하는 단계; 상기 컨볼루션의 결과값에 대해 맥스 폴딩(max pooling)이 수행되는 단계; 및 상기 맥스폴딩 결과값을 상기 인코더의 제3 레이어의 입력으로 제공하는 단계;를 포함한다.
- [0020] 바람직한 실시예에 있어서, 상기 제1단계는 상기 왼쪽다리의 제3레이어에서 수행되는 입력된 사이즈 그대로 인코더의 제3레이어용 패치이미지를 생성하고, 상기 제3레이어용 패치이미지에 대해 맥스폴딩을 수행하여 그 결과값을 제4레이어의 입력으로 제공하는 단계;를 포함하며, 상기 인코더의 제3레이어에서 상기 입력된 제2 레이어에 대한 맥스폴딩의 결과값 및 상기 제3레이어용 패치 이미지를 연결(concatenate)시켜 컨볼루션(convolution)을 수행하는 단계; 상기 컨볼루션의 결과값에 대해 맥스 폴딩(max pooling)이 수행되는 단계; 및 상기 맥스폴딩 결과값을 상기 인코더의 제4 레이어의 입력으로 제공하는 단계;를 포함한다.
- [0021] 바람직한 실시예에 있어서, 상기 제1단계는 상기 왼쪽다리의 제4레이어에서 수행되는 입력된 사이즈 그대로 인코더의 제4 레이어용 패치이미지를 생성하는 단계;를 포함하고, 상기 인코더의 제4레이어에서 상기 입력된 제3 레이어에 대한 맥스폴딩의 결과값 및 상기 제4레이어용 패치 이미지를 연결(concatenate)시켜 컨볼루션(convolution)을 수행하는 단계;를 포함한다.
- [0022] 바람직한 실시예에 있어서, 상기 제2단계는 상기 인코더의 각 레이어에서 출력되는 특징 맵이 해당 레이어에 위치한 TAB의 입력으로 제공되는 단계; 및 상기 인코더의 최하단 레이어에서 출력되는 특징 맵이 스윈트랜스포머의 입력으로 제공되는 단계;를 포함한다.
- [0023] 바람직한 실시예에 있어서, 상기 TAB의 입력으로 제공되는 단계는 RPN(Region proposal network)에 의해 치아 구조를 포함하는 잠재적 영역을 나타내는 경계 상자(boundary boxes)가 감지되고, 상기 감지된 경계 상자는 CAB(Channel Attention Block)를 통과하면서 특징 맵의 어텐션이 미세 조정된 후 상기 인코더의 스킵 연결과 곱

연산되어 수행된다.

- [0024] 바람직한 실시예에 있어서, 상기 제3단계는 상기 스윈트랜스포머에서 출력되는 특징 맵을 상기 디코더의 최하단 레이어의 입력으로 제공하고, 상기 디코더의 차하위 레이어에서 출력되는 특징 맵과 상기 TAB의 해당레이어에서 출력되는 특징 맵을 연결시켜 상기 디코더의 해당레이어의 입력으로 제공하는 단계; 상기 디코더의 해당 레이어에 대한 컨볼루션(convolution)을 수행하는 단계; 및 상기 디코더의 해당 레이어에서 출력되는 특징 맵을 오른쪽 다리의 해당 레이어의 입력으로 제공하면서, 업샘플링하여 상기 디코더의 차상위 레이어에 입력으로 제공하는 단계;를 포함한다.
- [0025] 바람직한 실시예에 있어서, 상기 제3단계는 상기 디코더의 최하단레이어에서 출력되는 특징 맵을 입력받아 상기 오른쪽다리의 최하단 레이어에 대한 컨볼루션(convolution)을 수행하고, 상기 디코더의 최하단레이어를 제외한 각 레이어에서 출력되는 특징 맵과 차하위 레이어에서 얻어진 특징 맵을 연결시켜 상기 오른쪽다리의 해당 레이어에 대한 컨볼루션(convolution)을 수행하는 단계; 상기 오른쪽 다리의 해당 레이어에 대한 컨볼루션 결과값에 대해 업샘플링이 수행된 결과값이 차상위 레이어에 입력으로 제공되는 단계; 및 상기 오른쪽 다리의 2번째 레이어에 대한 컨볼루션 결과값에 대해 업샘플링이 수행된 결과값과 상기 디코더의 최상단 레이어에서 컨볼루션을 마친 특징 맵을 연결시켜 컨볼루션하는 단계;를 포함한다.
- [0026] 또한, 본 발명은 치과 X-ray 이미지를 입력받는 입력부; 상기 입력된 치과 X-ray 이미지를 기학습된 학습 모델에 적용하여 시멘틱 분할을 수행하는 영상 분할부; 및 상기 시멘틱 분할이 수행된 치과 X-ray 이미지를 제공하는 영상 제공부;를 포함하며, 상기 학습 모델은, U-Net 모델에서 인코더 왼쪽에 형성된 왼쪽 다리 및 디코더 오른쪽에 형성된 오른쪽 다리를 포함하는 M-Net 유사 모델로서, 상기 인코더와 디코더 사이의 스킵 연결에 통합되어 형성된 TAB(Tooth Attention Block) 및 스윈트랜스포머(Swin Transformer)를 포함하는, 치과 X-ray 이미지 다중클래스 분할장치를 제공한다.
- [0027] 바람직한 실시예에 있어서, 상기 학습 모델에서 상기 TAB는 경계 인식 또는 경계 개선 필터 역할을 수행하여 복잡한 객체 경계 묘사 정확도가 향상되도록 어텐션 기능을 미세 조정하고, 상기 스윈트랜스포머는 입력 특징 맵을 겹치지 않는 로컬 창으로 나누어 효율적으로 처리하고 MSA(multi-head self-attention) 메커니즘을 활용하여 여러 관계를 동시에 학습하여 치과 X-ray 이미지의 장거리 종속성을 캡처한다.
- [0028] 바람직한 실시예에 있어서, 상기 스윈트랜스포머는 상기 인코더와 디코더 사이의 스킵 연결 중 최하단레이어에 통합되어 병목현상을 방지하고, 상기 TAB는 최하단레이어를 제외한 나머지 레이어에 통합되어 상기 인코더의 각 레이어에서 출력되는 특징 맵에서 노이즈와 의미 없는 정보의 영향을 감소시킨다.
- [0029] 바람직한 실시예에 있어서, 상기 영상분할부는 상기 치과 X-ray 이미지를 상기 인코더로 입력하고, 상기 인코더에서 출력된 특징 맵을 TAB 또는 스윈트랜스포머로 입력하며, 상기 TAB 또는 스윈트랜스포머에서 출력된 특징 맵을 상기 디코더로 입력하고, 상기 디코더에서 출력된 결과값을 분류한다.
- [0030] 바람직한 실시예에 있어서, 상기 영상분할부는 상기 왼쪽다리의 제1레이어에서 X-ray 치아이미지에 대해 상기 인코더의 제1 레이어용 패치 이미지를 생성하고, 상기 제1레이어용 패치이미지에 대해 맥스풀링을 수행하여 그 결과값을 제2레이어의 입력으로 제공하는 것을 포함하고, 상기 인코더의 제1레이어에서 수행되는 상기 제1 레이어에 대한 패치 이미지에 대해 컨볼루션(convolution)을 수행하며, 상기 컨볼루션의 결과값에 대해 맥스 풀링(max pooling)을 수행한 후 상기 맥스풀링 결과값을 상기 인코더의 제2 레이어의 입력으로 제공한다.
- [0031] 바람직한 실시예에 있어서, 상기 영상분할부는 상기 왼쪽다리의 제2레이어에서 수행되는 입력된 사이즈 그대로 인코더의 제2 레이어용 패치이미지를 생성하고, 상기 제2레이어용 패치이미지에 대해 맥스풀링을 수행하여 그 결과값을 제3레이어의 입력으로 제공하는 것을 포함하며, 상기 인코더의 제2레이어에서 상기 입력된 제1 레이어에 대한 맥스풀링의 결과값 및 상기 제2레이어용 패치 이미지를 연결(concatenate)시켜 컨볼루션(convolution)을 수행하고 상기 컨볼루션의 결과값에 대해 맥스 풀링(max pooling)을 수행한 후 상기 맥스풀링 결과값을 상기 인코더의 제3 레이어의 입력으로 제공한다.
- [0032] 바람직한 실시예에 있어서, 상기 영상분할부는 상기 왼쪽다리의 제3레이어에서 수행되는 입력된 사이즈 그대로 인코더의 제3레이어용 패치이미지를 생성하고, 상기 제3레이어용 패치이미지에 대해 맥스풀링을 수행하여 그 결과값을 제4레이어의 입력으로 제공하는 것 포함하며, 상기 인코더의 제3레이어에서 상기 입력된 제2 레이어에 대한 맥스풀링의 결과값 및 상기 제3레이어용 패치 이미지를 연결(concatenate)시켜 컨볼루션(convolution)을 수행하고, 상기 컨볼루션의 결과값에 대해 맥스 풀링(max pooling)을 수행한 후 상기 맥스풀링 결과값을 상기 인코더의 제4 레이어의 입력으로 제공한다.

- [0033] 바람직한 실시예에 있어서, 상기 영상분할부는 상기 왼쪽다리의 제4레이어에서 수행되는 입력된 사이즈 그대로 인코더의 제4 레이어용 패치이미지를 생성하는 것을 포함하고, 상기 인코더의 제4레이어에서 상기 입력된 제3 레이어에 대한 맥스풀링의 결과값 및 상기 제4레이어용 패치 이미지를 연결(concatenate)시켜 컨볼루션(convolution)을 수행한다.
- [0034] 바람직한 실시예에 있어서, 상기 영상분할부는 상기 인코더의 각 레이어에서 출력되는 특징 맵을 해당 레이어에 위치한 TAB의 입력으로 제공하고, 상기 인코더의 최하단 레이어에서 출력되는 특징 맵을 스윈트랜스포머의 입력으로 제공한다.
- [0035] 바람직한 실시예에 있어서, 상기 영상분할부는 상기 TAB의 입력을 RPN(Region proposal network)에 의해 치아 구조를 포함하는 잠재적 영역을 나타내는 경계 상자(boundary boxes)를 감지하고, 상기 감지된 경계 상자가 CAB(Channel Attention Block)를 통과하면서 특징 맵의 어텐션을 미세 조정된 후 상기 인코더의 스킵 연결과 곱연산시켜 수행한다.
- [0036] 바람직한 실시예에 있어서, 상기 영상분할부는 상기 스윈트랜스포머에서 출력되는 특징 맵을 상기 디코더의 최하단레이어의 입력으로 제공하고, 상기 디코더의 차하위 레이어에서 출력되는 특징 맵과 상기 TAB의 해당레이어에서 출력되는 특징 맵을 연결시켜 상기 디코더의 해당레이어의 입력으로 제공하며, 상기 디코더의 해당 레이어에 대한 컨볼루션(convolution)을 수행하고, 상기 디코더의 해당 레이어에서 출력되는 특징 맵을 오른쪽 다리의 해당 레이어의 입력으로 제공하면서, 업샘플링하여 상기 디코더의 차상위 레이어에 입력으로 제공한다.
- [0037] 바람직한 실시예에 있어서, 상기 영상분할부는 상기 디코더의 최하단레이어에서 출력되는 특징 맵을 입력받아 상기 오른쪽다리의 최하단 레이어에 대한 컨볼루션(convolution)을 수행하고, 상기 디코더의 최하단레이어를 제외한 각 레이어에서 출력되는 특징 맵과 차하위 레이어에서 얻어진 특징 맵을 연결시켜 상기 오른쪽다리의 해당 레이어에 대한 컨볼루션(convolution)을 수행하며, 상기 오른쪽 다리의 해당 레이어에 대한 컨볼루션 결과값에 대해 업샘플링이 수행된 결과값을 차상위 레이어에 입력으로 제공하며, 상기 오른쪽 다리의 2번째 레이어에 대한 컨볼루션 결과값에 대해 업샘플링이 수행된 결과값과 상기 디코더의 최상단 레이어에서 컨볼루션을 마친 특징 맵을 연결시켜 컨볼루션한다.

발명의 효과

- [0038] 상술된 본 발명의 의료 영상 분할 방법 및 장치에 의하면, 의료영상 특히 치과 X-ray 이미지인 경우 관심 영역에 초점을 맞추는 모델의 기능이 촉진되어 로컬 및 글로벌 컨텍스트를 모두 효과적으로 캡처할 수 있어 정확한 치아 분할을 제공하기 때문에 치아 모양의 변화나 병변의 모양을 식별하여 치주질환이나 치아 우식의 발견을 용이하여 다양한 치과 질환의 조기 발견 및 진단이 가능하고, 상세한 치료 계획을 세우는 데 큰 도움이 될 수 있다.
- [0039] 본 발명의 효과는 이상에서 언급한 효과로 제한되지 않으며, 언급되지 않은 또 다른 효과들은 아래의 기재로부터 당업자에게 명확하게 이해될 수 있을 것이다.

도면의 간단한 설명

- [0040] 도 1은 종래 알려진 U-Net 아키텍처를 도시한 것이다.
- 도 2a는 파노라마 X-ray 치아 이미지를 도시한 것이다.
- 도 3a는 본 발명의 실시예에 따른 치과 X-ray 이미지 다중클래스 분할 장치의 구성도이고, 도 3b는 본 발명의 실시예에 따른 치과 X-ray 이미지 다중클래스 분할장치를 이용한 치과 X-ray 이미지 다중클래스 분할방법의 순서도이며, 도 3c는 전처리를 위한 블록다이어그램이다.
- 도 4는 본 발명의 실시예에 따른 치과 X-ray 이미지 다중클래스 분할방법에 사용되는 학습모델인 분할 네트워크 아키텍처의 일 구현예를 도시한 것이다.
- 도 5a는 도 4의 아키텍처에 도시된 TAB(Tooth Attention Block)의 간략 구조도이고, 도 5b는 도 5a의 각 블록에서 이루어지는 연산 과정을 도시한 것이다.
- 도 6은 도 4의 아키텍처에 도시된 스윈트랜스포머의 일 구현예를 도시한 것이다.
- 도 7a 및 도 7b는 본 발명의 실험예1에서 공지된 학습모델을 수행하여 얻어진 치과 X-ray 이미지 결과물과 본

발명의 제안된 학습모델을 수행하여 얻어진 치과 X-ray 이미지 결과물의 사진이다.

도 8a 및 도 8b는 본 발명의 실험예2에서 공지된 학습모델을 수행하여 얻어진 치과 X-ray 이미지 결과물과 본 발명의 제안된 학습모델을 수행하여 얻어진 치과 X-ray 이미지 결과물의 사진이다.

발명을 실시하기 위한 구체적인 내용

- [0041] 본 발명에서 사용되는 용어는 본 발명에서의 기능을 고려하면서 가능한 현재 널리 사용되는 일반적인 용어들을 선택하였으나, 이는 당 분야에 종사하는 기술자의 의도 또는 관례, 새로운 기술의 출현 등에 따라 달라질 수 있다. 또한, 특정한 경우는 출원인이 임의로 선정한 용어도 있으며, 이 경우 해당되는 발명의 설명 부분에서 상세히 그 의미를 기재할 것이다.
- [0042] 본 발명에서 언급한 '포함한다', '갖는다', '이루어진다' 등이 사용되는 경우 '~만'이 사용되지 않는 이상 다른 부분이 추가될 수 있다. 구성 요소를 단수로 표현한 경우에 특별히 명시적인 기재 사항이 없는 한 복수를 포함하는 경우를 포함한다.
- [0043] 구성 요소를 해석함에 있어서, 별도의 명시적 기재가 없더라도 오차 범위를 포함하는 것으로 해석한다.
- [0044] 제1 또는 제2 등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 이런 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 이해되어야 한다. 예를 들어 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소는 제1 구성요소로도 명명될 수 있다.
- [0045] 어떤 구성요소가 다른 구성요소에 "연결되어" 있다거나 "접속되어" 있다고 언급된 때에는, 그 다른 구성요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성요소가 존재할 수도 있다고 이해되어야 할 것이다. 반면에, 어떤 구성요소가 다른 구성요소에 "직접 연결되어" 있다거나 "직접 접속되어" 있다고 언급된 때에는, 중간에 다른 구성요소가 존재하지 않는 것으로 이해되어야 할 것이다. 구성요소들 간의 관계를 설명하는 표현들, 예를 들어 "~간의에"와 "바로~간의에" 또는 "~에 이웃하는"과 "~에 직접 이웃하는" 등도 마찬가지로 해석되어야 한다.
- [0046] 본 발명의 여러 구현예들 각각의 특징적인 부분들은 부분적으로 또는 전체적으로 서로 결합 또는 조합가능하고, 기술적으로 다양한 연동 및 구동이 가능하며, 각 구현예들은 서로에 대하여 독립적으로 실시 가능할 수도 있고 연관 관계로 함께 실시할 수도 있다.
- [0047] 이하, 첨부한 도면 및 바람직한 실시예들을 참조하여 본 발명의 기술적 구성을 상세하게 설명한다.
- [0048] 그러나, 본 발명은 여기서 설명되는 실시예에 한정되지 않고 다른 형태로 구체화 될 수도 있다. 명세서 전체에 걸쳐 본 발명을 설명하기 위해 사용되는 동일한 참조번호는 동일한 구성요소를 나타낸다.
- [0049] 본 발명은 치과 X-ray 이미지 다중클래스 분할방법 및 장치에 관한 것으로, 특히 치아 이미지를 정확하게 분할할 수 있는 새로운 학습모델 즉 치아 분할 아키텍처를 개발한 것에 그 기술적 특징이 있는데, 본 발명에서 개발된 학습모델은 U-Net 구조에서 인코더 왼쪽에 형성된 왼쪽 다리 및 디코더 오른쪽에 형성된 오른쪽 다리를 포함하여 M자형 구조가 형성된 M-Net 유사 모델의, 인코더와 디코더 사이의 스킵 연결에 TAB(Teeth Attention Block) 및 스윈트랜스포머(Swin Transformer)가 통합된 구조로서, 본 발명의 치아 분할 아키텍처에 의하면 로컬 및 글로벌 상황을 모두 캡처할 수 있고, 정보를 효과적으로 전달하여 치아와 주변 구조를 정확하게 묘사할 수 있다.
- [0050] 종래 알려진 U-Net 아키텍처가 도시된 도 1을 참조하면, 기존의 U-Net은 풀링 작업이 업샘플링 연산자로 대체되는 연속 계층으로 일반적인 계약 네트워크를 보완하고, 업샘플링 레이어는 출력의 해상도를 높이며, 연속적인 컨볼루션 계층은 이 정보를 기반으로 정확한 출력을 조립하는 방법을 학습할 수 있으며, 네트워크는 스킵 연결을 통해 컨텍스트 정보를 고해상도 계층으로 전파할 수 있고, 누락된 정보를 스킵 연결 보상할 수 있지만, 컨볼루션 기반 네트워크는 일반적으로 큰 이미지에 대한 장거리 종속성을 캡처하지 못하여 치아와 주변 구조를 정확하게 묘사할 수 없는 문제가 있었기 때문이다.
- [0051] 도 3a는 본 발명의 실시예에 따른 치과 X-ray 이미지 다중클래스 분할 장치의 구성도이고, 도 3b는 본 발명의 실시예에 따른 치과 X-ray 이미지 다중클래스 분할장치를 이용한 치과 X-ray 이미지 다중클래스 분할방법의 순서도이며, 도 3c는 전처리를 위한 블록다이어그램이다. 도 4는 본 발명의 실시예에 따른 치과 X-ray 이미지 다중클래스 분할방법에 사용되는 학습모델인 분할 네트워크 아키텍처의 일 구현예를 도시한 것이다. 도 5a는 도 4의 아키텍처에 도시된 TAB(Teeth Attention Block)의 간략 구조도이고, 도 5b는 도 5a의 각 블록에서 이루어지

는 연산 과정을 도시한 것이며, 도 6은 도 4의 아키텍처에 도시된 스윈트랜스포머의 일 구현예를 도시한 것이다.

- [0052] 도 3a 내지 도 6을 참조하여 본 발명의 실시예에 따른 치과 X-ray 이미지 다중클래스 분할방법 및 장치의 구성을 보다 구체적으로 살펴본다.
- [0053] 도 3a를 참조하면, 본 발명의 실시예에 따른 치과 X-ray 이미지 다중클래스 분할장치(100)는 영상 입력부(110), 영상 분할부(120) 및 영상 제공부(130)를 포함한다. 영상 입력부(110)는 도 3b에 도시된 바와 같이 의료영상을 입력받는다(S210). 여기서, 치과 X-ray 이미지는 도 2에 도시된 바와 같은 치과 파노라마 이미지일 수 있는데, 이 이미지는 모든 치아와 주변 구조를 포함하여 입 전체를 포착하는 포괄적인 X-ray 이미지로서, 파노라마 이미지는 구강 건강에 대한 폭넓은 시각을 위해 치과에서 일반적으로 사용된다. 다음으로, 영상 분할부(120)는 입력받은 치과 X-ray 이미지를 기학습된 학습 모델에 적용하여 시맨틱 분할(Semantic segmentation)을 수행한다(S220).
- [0054] 본 발명의 일 실시예에 따른 치과 X-ray 이미지 다중클래스 분할방법에서 사용되는 학습 모델 즉 딥러닝 네트워크는 도 4에 도시된 바와 같이 다운샘플링을 통해 특징 표현을 추출하는 인코더와 업샘플링을 통해 분할 마스크를 재구성하는 디코더로 구성되는 U-Net과 유사한 구조와, 공간 정보를 보존하기 위한 인코더와 디코더 레이어 사이의 스킵 연결과, 정확한 특징 표현 학습을 촉진하고 분할 성능을 향상시키기 위해 인코더의 왼쪽과 디코더의 오른쪽에 각각 사용된 왼쪽다리 및 오른쪽 다리와, 무엇보다도 스킵 연결에 통합되어 형성된 TAB(Tooth Attention Block) 및 스윈트랜스포머(Swin Transformer block)를 포함하는 구조이다.
- [0055] 이 때, TAB는 경계 인식 또는 경계 개선 필터 역할을 수행하여 복잡한 객체 경계 묘사 정확도가 향상되도록 어텐션 기능을 미세 조정하고, 스윈트랜스포머는 입력 특징 맵을 겹치지 않는 로컬 창으로 나누어 효율적으로 처리하고 MSA(multi-head self-attention) 메커니즘을 활용하여 여러 관계를 동시에 학습하여 치과 X-ray 이미지의 장거리 종속성을 효과적으로 캡처할 수 있다.
- [0056] 따라서, 스윈트랜스포머는 인코더와 디코더 사이의 스킵 연결 중 최하단레이어에 통합되어 병목현상을 방지하고, TAB는 최하단레이어를 제외한 나머지 레이어에 통합되어 인코더의 각 레이어에서 출력되는 특징 맵에서 노이즈와 의미 없는 정보의 영향을 감소시킬 수 있다.
- [0057] 본 발명의 학습모델에서 치과 X-ray 이미지 다중클래스 분할 과정은 인코더-디코더 기반 구조로 되어 있으며, 왼쪽다리(left leg) 및 오른쪽 다리(right leg)와 같은 두 가지 사이드 경로를 포함하고, 인코더와 디코더의 스킵 연결 사이에 TAB 및 스윈트랜스포머가 통합되어 있으므로, 치과 X-ray 이미지 다중클래스 분할 과정(S220)은 치과 X-ray 이미지를 인코더로 입력하는 제1단계; 인코더에서 출력된 특징 맵을 TAB 또는 스윈트랜스포머로 입력하는 제2단계; 상기 TAB 또는 스윈트랜스포머에서 출력된 특징 맵을 상기 디코더로 입력하는 제3단계; 및 디코더에서 출력된 결과값을 분류하는 제4단계;를 포함할 수 있다. 이하에서 사용되는 인코더, TAB, 스윈트랜스포머, 디코더 및 오른쪽 다리 등 각 구성요소들에 포함된 각 레이어에서 출력되는 결과값 또는 얻어진 결과값은 각 레이어에서 출력되는 특징 맵을 의미하는 것으로 해석될 수 있다.
- [0058] 본 발명의 왼쪽다리 및 인코더의 동작을 포함하는 다운샘플링 과정 즉 제1단계를 살펴본다.
- [0059] 본 발명에 따르면, 초기 입력 이미지는 일반적으로 높이 x 너비 x 채널(H x W x C)의 형태로 되어 있는데 도시된 바와 같이, 512x1024x3 크기의 아무 분할이 되지 않은 치과 X-ray 이미지가 입력으로 주어진다.
- [0060] 도 4에 도시된 치아 분할 아키텍처에서 왼쪽 다리(left leg)는 M-Net 유사 구조의 중요한 구성요소로서, 주요 역할은 입력 치과 파노라마 이미지가 다양한 규모로 처리되는 다운샘플링 프로세스를 통해 모델을 안내하는 것인데, 즉 모델의 가이드 역할을 하며 입력 이미지를 다양한 크기나 세부 수준으로 분류하는 데 도움이 되고, 이미지를 다양한 방식으로 보는 이러한 과정을 통해 모델은 더 미세한 특징과 전체 구조를 모두 이해할 수 있으며 최종 출력에서 보다 효과적인 치아 분할이 가능하다.
- [0061] 보다 구체적으로 왼쪽 다리의 역할을 살펴보면, 왼쪽 다리는 다중클래스 치아 분할 작업 중에 감독을 위해 지정되는데, 이는 치아를 정확하게 분할하기 위해 모델의 학습 과정을 안내하는 데 중추적인 역할을 한다는 것을 의미한다. 또한, 다운샘플링을 수행하는데 이 작업 중에 입력된 치과 이미지는 다양한 크기로 처리되며, 이는 다양한 세부 수준으로 이미지를 보는 것과 유사하고, 다운샘플링을 수행하는 동안 왼쪽 다리는 치과 이미지에서 계층적 특징을 추출하는데 이러한 기능은 기본 패턴부터 보다 복잡한 패턴까지 다양하며 이미지의 다양한 추상화 수준을 나타낼 수 있다. 또한, 왼쪽 다리는 멀티스케일 감독에 기여하여 모델이 여러 스케일에서 치과 이미

지를 이해할 수 있도록 한다. 이는 미세한 세부 사항과 더 광범위한 상황 정보를 모두 캡처하는 데 중요하다.

- [0062] 따라서, 왼쪽다리는 kernel 크기가 2x2 인 max-pooling 작업으로 구성되고, 각각의 최대 풀 출력(max pool output)은 해당 인코딩 레이어에 입력으로 출력한다. 즉 max pooling은 해당 인코딩 레이어에 입력으로 공급된다. 왼쪽 다리는 이와 같이 input이미지를 maxpooling 하는 것인데 MaxPooling은 특정 영역에서 최대값을 선택하여 해상도를 줄이는 연산이다.
- [0063] 왼쪽다리의 제1레이어에서는 의료영상이미지에 대해 입력된 사이즈 그대로 인코더의 제1 레이어용 패치이미지를 생성하고, 생성된 제1 레이어용 패치이미지를 인코더의 제1레이어에 입력으로 제공한다. 또한, 생성된 제1 레이어용 패치이미지를 size = 2, stride = 2 의 Maxpooling을 통해 크기를 줄여 왼쪽다리의 제2 레이어의 입력으로 제공한다.
- [0064] 동일하게, 왼쪽다리의 제2레이어는 입력된 사이즈 그대로 제2레이어용 패치이미지를 생성하고, 인코더의 제2레이어에 입력으로 제공된다. 또한, 생성된 제2레이어용 패치이미지를 size = 2, stride = 2 의 Maxpooling을 통해 크기를 줄여 왼쪽다리의 제3 레이어의 입력으로 제공한다.
- [0065] 동일한 과정을 수행하여 왼쪽다리의 최하단 레이어(도시된 구조에는 제4레이어)에서 최종레이어용 패치이미지를 생성하여 인코더의 최하단 레이어에 입력으로 제공할 수 있다.
- [0066] 다음으로, 인코더에 대해 살펴보면, 인코더는 치과 X-ray 이미지에서 계층적 특징 맵을 추출하는 역할을 하며, 다운샘플링은 이 프로세스의 핵심 측면으로 인코더에 의해 수행되는 다운샘플링 작업은 치아 분할 아키텍처에서 중요한 단계인데, 인코더가 수행하는 다운샘플링 작업에는 계층적 특징 추출, 배치 정규화, ReLU 활성화 및 맥스 풀링이 포함된 컨볼루션 레이어 사용이 포함된다. 이 프로세스는 계층적 특징 맵을 추출하여 모델이 치과 파노라마 이미지의 다양한 세부 수준을 이해할 수 있도록 한다. 얻은 특징 맵은 구조의 다양한 단계에서 정확한 치아 분할을 달성하는 데 근본적인 역할을 한다.
- [0067] 인코더는 계층적 특징 추출을 위해 각 레이어에서 커널 크기가 3x3인 여러 컨볼루션 레이어를 활용한다. 이러한 컨볼루션 레이어는 입력 이미지를 스캔하여 특징을 추출하는 필터 역할을 한다. 여러 레이어를 사용하면 치과 이미지 내에서 기본적인 것부터 복잡한 것까지 다양한 수준의 특징을 학습하는 데 도움이 된다. 모델의 안정성을 향상하고 학습 속도를 높이기 위해 각 컨볼루션 레이어 후에 배치 정규화가 수행된다. 모델의 매개변수가 일관된 범위 내에 유지되도록 보장하여 보다 효과적인 훈련을 촉진한다. ReLU(Rectified Linear Unit) 활성화 함수가 각 컨볼루션 후에 적용된다. ReLU는 모델에 비선형성을 도입하여 치과 이미지의 복잡한 패턴과 특징을 학습하는 능력을 향상시킨다. 특징 맵의 공간 차원 축소를 위한 맥스 풀링(Max-pooling) 레이어가 인코더에 포함된다. 맥스 풀링은 특징 맵의 공간 차원을 줄여 수용 필드를 효과적으로 늘리는 기술로서 입력 이미지에서 더 큰 패턴과 상황별 정보를 캡처하는 데 도움이 된다.
- [0068] 도 4에 도시된 치아 분할 아키텍처에서 인코더의 동작을 살펴보면, 인코딩 경로에서 인코더의 각 레이어는 3x3 컨볼루션, BatchNormalization, ReLU(rectified linear unit)가 순차적으로 일어나는 2개의 컨볼루션 블록이 연속되어 이루어진 구성으로, 컨볼루션 결과값에 대해 맥스 풀링(2x2)으로 영상의 사이즈를 줄이는 과정이 수행된다.
- [0069] 보다 구체적으로 살펴보면, 인코더의 제1레이어는 입력된 제1레이어용 패치이미지에 대해 연속된 2개의 컨볼루션블록에서 각각 3x3 컨볼루션, BatchNormalization 및 ReLU(rectified linear unit)가 순차적으로 적용되어, 입력된 이미지를 컨볼루션(3x3) 시켜 특징을 추출하고, 배치 정규화를 통해 변형된 분포가 나오지 않도록 정규화 하며, ReLU 함수를 이용해 이진화시켜 컨볼루션 결과값이 얻어진다. 인코더의 제1레이어에서 얻어진 컨볼루션 결과값은 스트라이드(stride)가 2인 2x2 맥스 풀링 연산을 사용하여 다운 샘플링된다. 풀링은 입력 데이터의 공간 차원을 줄이는 연산으로, 2x2 Max 풀링은 각 2x2 영역에서 최대값을 추출하는 방식으로 동작할 수 있다.
- [0070] 인코더의 제2레이어는 입력된 제1 레이어에 대한 맥스풀링의 결과값 및 제2레이어용 패치이미지를 연결시켜 상술된 바와 동일한 과정을 통해 컨볼루션 결과값을 얻은 후 맥스풀링 연산을 사용하여 다운 샘플링시킨다.
- [0071] 상술된 인코더의 제2레이어에서 수행되는 연산을 맥스풀링 결과 값이 최하단 레이어의 입력으로 제공될 때까지 동일하게 반복수행한다.
- [0072] 인코더의 최하단레이어는 입력된 제3 레이어에 대한 맥스풀링의 결과값 및 상기 제4레이어용 패치 이미지를 연결(concatenate)시켜 상술된 바와 동일한 과정을 통해 컨볼루션(convolution)을 수행하여 얻어진 컨볼루션 결과값을 스윙트랜스포머에 입력으로 제공한다. 도시된 바와 같이 입력 이미지는 인코더 경로에서 512개의 특징 맵

으로 다운 샘플링된다.

- [0073] 다음으로, 본 발명의 치아 분할 아키텍처에서 TAB(Tooth Attention Block) 및 스윈트랜스포머를 통해 일어나는 제2단계에 대해 살펴보면, 도 4에 도시된 바와 같이 인코더의 각 레이어에서 출력되는 특징 맵이 해당 레이어에 위치한 TAB의 입력으로 제공되는 단계 및 인코더의 최하단 레이어에서 출력되는 특징 맵이 스윈트랜스포머의 입력으로 제공되는 단계가 수행될 수 있다.
- [0074] 먼저, TAB(Tooth Attention Block)는 도시된 바와 같이 인코더와 디코더 레이어 사이의 치아 분할 아키텍처의 스킵 연결에 통합된 구성요소로서, 경계 인식 또는 경계 개선 필터 역할을 하며 이미지의 서로 다른 객체 또는 구조 간의 경계 인식을 향상시키기 위한 분할 작업에서 중요한 역할을 한다. TAB는 각 픽셀 주위의 국소 수용 필드를 관찰하여 치아 이미지에서 치아 분할 모델의 초점을 향상시킨다. 이 수용 필드에서 필터링 작업을 구현하면 개체의 경계, 특히 단일 치아의 가장자리와 인접 구조가 강조된다.
- [0075] TAB에서 사용하는 어텐션 메커니즘은 분할 마스크의 정확성과 정밀도를 향상시키는데 중요한 역할을 한다. 구체적으로 이는 개인 치아와 주변 치아의 분화를 효과적으로 개선함으로써 달성된다. 간단히 말해서, TAB은 치아 분할 작업에 도움이 되지 않는 관련 없는 특징이나 노이즈를 최소화하면서 치아의 특징과 경계를 선택적으로 최적화하여 치과 이미지에서 노이즈와 의미 없는 정보의 영향을 줄인다. 이러한 향상된 어텐션은 보다 정확하고 일관된 분할 결과를 달성하는 데 도움이 되며, 치과 파노라마 이미지에서 개별 치아와 그 경계를 보다 명확하고 명확하게 보여준다.
- [0076] self-attention 메커니즘을 사용하는 제안된 TAB를 사용하면 모델은 획득한 상황별 지식을 활용하여 이미지 내 다양한 치아 부분의 중요성을 평가할 수 있다. 훈련 단계에서 TAB는 개별 치아 주변의 경계 픽셀에 증가된 어텐션 점수를 할당한다. 이는 특징 맵을 생성하는 동안 경계 영역에 더 많은 어텐션을 기울임으로써 개별 치아를 구별하는 모델의 기능이 향상됨을 나타낸다. 전체 분할 성능의 향상은 이러한 경계 영역에 대한 TAB의 주의 기능에 의해 촉진된다. 구체적으로, 분할 마스크 내부의 치아 경계를 강화하고 인접 치아 간의 겹침 발생을 최소화하며 분할 프로세스의 정밀도와 정확성을 향상시키는 데 도움이 된다. TAB 방법을 우리 모델에 통합하면 특히 조밀하게 배열되거나 약간 겹치는 치아를 특징으로 하는 복잡한 치과 이미지에서 주목 할만한 이점을 얻을 수 있으므로 기존 분할 기술의 한계를 완화할 수 있다.
- [0077] TAB는 도 5a에서 볼 수 있듯이, 인코더와 디코더 레이어 사이의 치아 분할 아키텍처의 스킵 연결에 통합되었다. 보다 구체적으로 도 5b를 참조하여 TAB의 각 블록에서 일어나는 연산을 살펴보면, RPN(Region proposal network)에 의해 치아 구조를 포함하는 잠재적 영역을 나타내는 경계 상자(boundary boxes)가 감지되고, 상기 감지된 경계 상자는 CAB(Channel Attention Block)를 통과하면서 특징 맵의 어텐션이 미세 조정된 후 상기 인코더의 스킵 연결과 곱연산이 수행될 수 있다.
- [0078] 그 결과 TAB에서 특정 필터를 활용하면 복잡한 객체 경계, 특히 개별 치아 및 주변 구조와 관련된 객체 경계를 정확하게 묘사할 때 모델의 분할 정확도가 크게 향상된다. 업그레이드된 기능 맵은 이후 스킵 연결을 통해 인코더의 기능과 곱해진다. 결합 후, 병합된 엔터티는 디코더에 의해 처리되어 최종 분할을 달성한다. TAB는 각 픽셀 주위의 로컬 인접 수용 필드를 고려하고 치아의 경계를 강조하기 위한 필터링 작업을 구현하여 작동한다. TAB의 필터링 작업은 컨볼루션 레이어 및 어텐션 메커니즘과 같은 기술을 활용하여 객체 경계를 학습하고 타겟팅할 수 있다. 본 발명의 아키텍처와 같이 스킵 연결 내에 TAB를 통합하면 다음과 같이 치아 분할 아키텍처에 여러 가지 이점이 제공된다.
- [0079] 1. 향상된 분할 정확도: 모델은 개체 경계에 집중하여 각 치아와 주변 구조를 구별할 수 있으므로 보다 정확한 분할 마스크를 생성할 수 있다.
- [0080] 2. 더 부드럽고 선명한 객체 경계: TAB를 활용하면 분할 마스크 내에서 비정상적이거나 고르지 않은 가장자리의 존재를 줄여 더 세련되고 뚜렷한 객체 경계를 얻을 수 있다.
- [0081] 3. 겹치거나 인접한 개체의 처리 개선: 치아 이미지에서 치아가 근접하거나 겹치는 경우가 많다. TAB를 구현하면 인접하거나 겹치는 치아를 구별하는 모델의 능력이 향상되어 향상된 성능을 얻을 수 있다.
- [0082] 또한, 본 발명의 치아 분할 아키텍처에서 스윈트랜스포머에 대해 살펴보면, 스윈트랜스포머는 self-attention 메커니즘을 사용하여 로컬 및 글로벌 상황 정보를 효과적으로 캡처하기 위해 사용된 것으로, 본 발명의 치아 분할 아키텍처는 병목 지점에 전략적으로 배치된 스윈트랜스포머 블록을 통합하며, 이러한 블록은 장거리 종속성을 캡처하는 데 중요한 역할을 한다.

- [0083] 스윈트랜스포머의 MSA(multi-head self-attention) 레이어 뒤에는 위치별 피드포워드 레이어와 레이어 정규화가 이어지며, 이를 통해 Swin Transformer는 멀티클래스 치아 분할 작업을 성공적으로 관리할 수 있다. 특히, MSA 메커니즘은 치과 이미지의 여러 부분에 걸쳐 복잡한 공간 관계를 캡처하는 데 도움이 되는 반면, 위치별 피드포워드 레이어는 입력 특징 맵 내에서 겹치지 않는 로컬 창을 사용하여 로컬 표현을 향상시킨다. 도 6을 참조하면 본 발명의 아키텍처에 사용된 일 구현예에 따른 스윈트랜스포머 레이어를 알 수 있는데, 패치 임베딩은 이미지의 패치를 만드는 데 사용되며, 패치 추출에 사용되는 위치 인코딩(PE)과 결합된다. 도 6에 도시된 두 개의 연속 레이어가 있다. 첫 번째 계층(I_0)에는 계층 정규화, 윈도우 기반 다중 헤드 어텐션[W-MSA(multi-head self-attention)], MLP(multi-layer perceptron)가 포함된다. 두 번째 레이어(I_1)는 첫 번째 레이어와 동일한 기능을 수행하지만 W-MSA는 the shifting window based multi-self head attention(SW-MSA)으로 대체된다. 상기 SW-MSA는 attention을 덜 복잡하게 계산하는 데 사용되어 왔다. 재구성 프로세스에 유용한 필수 정보를 캡처하기 위해 스윈 트랜스포머 블록이 수행하는 세부 프로세스는 다음과 같다.
- [0084] 1. Self-Attention 메커니즘: 스윈 트랜스포머 블록은 Self-Attention 메커니즘을 사용한다. 이 메커니즘을 통해 모델은 중요도에 따라 치과 이미지의 다양한 부분에 초점을 맞출 수 있다. 이를 통해 모델은 먼 지역 간의 관계를 고려하여 장거리 종속성을 효과적으로 포착할 수 있다.
- [0085] 2. 비중첩 로컬 창: 입력 기능 맵은 비중첩 로컬 창으로 나뉜다. 이러한 분할을 통해 더 작은 영역 내에서 정보를 효율적으로 처리할 수 있어 치과 이미지에 대한 보다 자세한 분석이 가능해진다.
- [0086] 3. 창 병합 및 이동: 각 self-attention 레이어 후에 스윈 트랜스포머는 창 병합 및 이동을 차단한다. 이러한 동적 조정은 모델이 치과 이미지의 여러 부분 사이의 복잡한 연결을 이해할 수 있도록 보장하여 장거리 종속성을 캡처하는 데 중요하다.
- [0087] 4. 위치별 피드포워드 레이어: 스윈 트랜스포머 블록에는 위치별 피드포워드 레이어가 포함되어 있다. 이러한 레이어는 치아 이미지 내의 복잡한 비선형 관계를 학습하는 모델의 능력을 향상시킨다. 복잡한 패턴과 구조에 대한 전반적인 이해에 도움이 된다.
- [0088] 5. 적응성 및 비지역 정보 관리: 스윈 트랜스포머는 아키텍처의 병목 지점에 전략적으로 배치된다. 이를 통해 치아의 다양한 형태와 미묘한 패턴 및 먼 관계를 인식해야 하는 필요성을 고려하여 장거리 의존성을 효과적으로 포착할 수 있다. 즉 치아는 다양한 형태를 나타내기 때문에 치아 이미지에서 미묘한 패턴과 먼 관계를 인식하는 것이 필수적이므로 전체적으로 이러한 방법은 치과 이미지 분할 문제를 해결하는 데 있어 스윈트랜스포머의 효율성을 향상시킨다.
- [0089] 다음으로, 본 발명의 치아 분할 아키텍처의 디코더에서 일어나는 제3단계에 대해 살펴보면, 스윈트랜스포머에서 출력되는 특징 맵을 디코더의 최하단레이어의 입력으로 제공하고, 디코더의 차하위 레이어에서 출력되는 특징 맵과 TAB의 해당레이어에서 출력되는 특징 맵을 연결시켜 디코더의 해당레이어의 입력으로 제공하며, 디코더의 해당 레이어에 대한 컨볼루션(convolution)을 수행하고, 디코더의 해당 레이어에서 출력되는 특징 맵을 오른쪽 다리의 해당 레이어의 입력으로 제공하면서, 업샘플링하여 디코더의 차상위 레이어에 입력으로 제공하는 것을 포함할 수 있다.
- [0090] 이와 같이, 디코더는 특징 맵의 공간 해상도를 복구하고 분할된 치아를 재구성하는 프로세스인 업샘플링을 담당한다. 즉 디코더의 각 레이어가 ReLU 활성화 기능이 있는 3 x 3 크기의 컨볼루션 필터로 구성되어 입력 크기를 두 배로 하는 업 샘플링 연산이 디코더 출력에서 원래 입력 이미지를 재구성하는 데 사용될 수 있는데, 디코더에서 수행되는 업샘플링 동작을 구체적으로 설명하면 다음과 같다.
- [0091] 1. 업샘플링을 위한 전치된 컨볼루션: 업샘플링은 전치된 컨볼루션(부분 스트라이드 컨볼루션 또는 디컨볼루션이라고도 함)을 사용하여 달성된다. 이러한 작업은 깊이를 유지하면서 기능 맵의 높이와 너비를 늘린다. 일 구현예로서 디코더에서 일어나는 컨볼루션은 디코더의 차하위 레이어에서 출력되는 특징 맵과 TAB의 해당레이어에서 출력되는 특징 맵을 연결시켜 Conv2DTranspose (2x2)를 통해 영상 사이즈를 늘리고, 인코더에서 연결한 값을 받아 컨볼루션 (3x3) 시켜주며, 배치정규화를 통해 변형된 분포가 나오지 않도록 정규화하여 수행될 수 있다.
- [0092] 2. 분할 마스크 재구성: 업샘플링과 동시에 디코더는 업샘플링된 특징 맵에서 분할 마스크를 재구성한다. 이러한 분할 마스크는 입력 치과 이미지에 대한 픽셀별 클래스 예측을 나타낸다.
- [0093] 3. 공간 정보를 위한 스킵 연결: 스킵 연결은 인코더와 디코더 계층 사이의 간격을 연결하여 디코더에서 중추적인 역할을 한다. 이러한 연결은 중간 계층을 통해 인코더에서 해당 디코더 계층으로 기능 맵을 보낸다. 이를 통

해 인코더의 이전 레이어의 고해상도 세부 정보와 더 깊은 레이어의 추상적인 저해상도 기능을 통합할 수 있다.

[0094] 4. 로컬 세부 정보와 글로벌 컨텍스트의 융합: 스킵 연결을 통한 기능의 융합은 분할 마스크를 보다 정확하게 재구성하는 데 도움이 된다. 이 프로세스는 국소적인 세부 사항과 전체적인 맥락을 모두 포착하여 치아 분할의 정확성을 향상시킨다. 이는 모델이 치과 이미지의 미세한 특징과 더 넓은 패턴을 모두 고려하도록 보장한다.

[0095] 이와 같이, 디코더의 업샘플링 작업에는 공간 해상도를 높이기 위해 전치된 컨볼루션을 사용하는 작업이 포함된다. 동시에 분할 마스크가 재구성되어 픽셀별 클래스 예측을 나타낸다. 스킵 연결은 인코더의 고해상도 세부 정보와 디코더의 추상 기능을 융합하는 데 활용되어 치과 이미지에 대한 포괄적인 이해를 보장하고 정확한 치아 분할을 촉진한다.

[0096] 도시된 바와 같이 디코더의 최상단레이어의 최종 출력 단계에서 기능 맵이 512에서 64로 줄어든다.

[0097] 또한, 본 발명의 치아 분할 아키텍처의 오른쪽다리에서 일어나는 제3단계에 대해 살펴보면, 디코더의 최하단레이어에서 출력되는 특징 맵을 입력받아 상기 오른쪽다리의 최하단 레이어에 대한 컨볼루션(convolution)을 수행하고, 디코더의 최하단레이어를 제외한 각 레이어에서 출력되는 특징 맵과 차 상위 레이어에서 얻어진 특징 맵을 연결시켜 오른쪽다리의 해당 레이어에 대한 컨볼루션(convolution)을 수행하는 단계; 오른쪽 다리의 해당 레이어에 대한 컨볼루션 결과값에 대해 업샘플링이 수행된 결과값이 차 상위 레이어에 입력으로 제공되는 단계; 및 오른쪽 다리의 2번째 레이어에 대한 컨볼루션 결과값에 대해 업샘플링이 수행된 결과값과 상기 디코더의 최상단레이어에서 컨볼루션을 마친 특징 맵을 연결시켜 컨볼루션하는 단계;를 포함할 수 있다. 일 구현예로서, 오른쪽 다리는 도 4에 도시된 바와 같이 2번째 레이어에 대한 컨볼루션 결과값에 대해 업샘플링이 수행된 결과값과 디코더의 최상단 레이어에서 컨볼루션을 마친 특징 맵을 연결한 다음 마지막 컨볼루션(1×1)을 수행하여 입력된 치과 X-ray 이미지로부터 시맨틱 분할 이미지를 흑백으로 출력한다. 1x1 컨볼루션은 사람 치아에 따라 이미지를 32개 채널에 매핑하기 위해 사용된다.

[0098] 이와 같이 오른쪽 다리는 치아 분할 과정의 후반 단계에서 특징 맵을 디코딩하고 캡처하는 데 기여하며 전체 이미지 처리 파이프라인에서 모델이 효과적으로 감독되도록 보장하는 중요한 역할을 한다. 즉 오른쪽 다리는 왼쪽 다리와 유사하게 특히 U-Net 아키텍처의 업샘플링(디코딩) 구성요소 중에 감독 프로세스에 관여하는데, 오른쪽 다리에 수행되는 동작은 다음과 같다.

[0099] 1. 감독의 역할: 오른쪽 다리는 다중 클래스 치아 분할 작업 중에 감독을 위해 지정된다. 그 역할은 이미지 처리의 업샘플링 또는 디코딩 단계에서 특히 중요하다.

[0100] 2. 업샘플링(디코딩): 오른쪽 다리는 디코딩이라고도 알려진 업샘플링 작업에 적극적으로 참여한다. 이 프로세스에는 특징 맵의 공간 해상도를 재구성하고 치아의 최종 분할에 기여하는 작업이 포함된다.

[0101] 3. 특징 표현: 오른쪽 다리는 왼쪽 다리와 유사하게 특징 표현에 기여한다. 이는 이미지 처리의 후반 단계에서 특징을 캡처하여 인코더와 왼쪽 다리에서 얻은 계층적 특징을 보완한다.

[0102] 4. 멀티스케일 감독: U-Net 구조의 일부인 오른쪽 다리는 멀티스케일 감독에 기여한다. 이를 통해 다운샘플링(인코딩) 및 업샘플링(디코딩) 단계에서 모델이 여러 규모로 감독되어 객체를 정확하게 캡처하고 인식하는 능력이 향상된다.

[0103] 상술된 바와 같이, 본 발명의 치아 분할 아키텍처 구조 중 왼쪽다리 및 오른쪽 다리는 다중 클래스 치아 분할 작업 중 감독을 위해 사용되었다. 아키텍처의 다운샘플링(인코딩) 및 업샘플링(디코딩) 구성요소 모두에 다중 스케일 감독을 통합하여 모델은 여러 규모에서 감독된다. 이러한 기술은 물체를 잘 포착하고 인식하는 데 도움이 되며 치과 이미지에서 발견된 다양한 종류의 치아를 정확하게 분할하는 데 도움이 된다. 이 접근 방식은 다양한 규모에 걸쳐 특징 표현을 용이하게 할 뿐만 아니라 모델의 차별화 능력을 향상시켜 뚜렷한 치아 범주를 구별한다.

[0104] 마지막으로, 영상 제공부(130)는 시맨틱 분할이 수행된 의료영상 이미지를 제공한다(S230).

[0105] 구체적으로, 영상 제공부(130)는 유선 또는 무선 통신 방식으로 통신하여 사용자에게 시맨틱 분할이 수행된 의료영상 이미지를 제공할 수 있다. 이때, 무선 통신 방식은 비콘(Beacon), 와이파이(Wi-Fi), 블루투스(bluetooth), 지그비(zigbee), 적외선 통신(IrDA, Infrared Data Association), UWB(Ultra Wideband) 또는 RFID(Radio Frequency Identification) 등과 같은 근거리 통신 방식 또는 퍼블릭 블록체인 네트워크, 프라이빗 블록체인 네트워크 및 하이브리드 블록체인 네트워크 등과 같은 블록체인 네트워크를 포함할 수 있다.

[0106] 본 발명에서는 맞춤형 정사각형 주사위 손실 함수를 사용하여 모델을 훈련하였다. 의료 이미지 분할 작업에 일반적으로 사용되는 다이스 손실(Dice loss)은 예측 이미지와 실제 이미지 간의 중첩을 계산하므로 이러한 작업에서 자주 관찰되는 클래스 불균형을 해결하는 데 이상적이다.

[0107] 기존의 주사위 손실 함수에서 점수는 예측된 분할 맵과 실제 분할 맵 사이의 교차 영역을 두 맵의 총 픽셀 수로 나눈 값의 두 배로 계산된다. Dice 손실은 1에서 Dice 점수를 뺀 값으로 계산되어 예측된 분할과 실제 분할 간의 최상의 중첩을 얻는다.

[0108] 손실을 공식화하기 전에 픽셀 값을 제공하여 주사위 손실 함수를 변경했다. 정사각형 주사위 손실(Square Dice loss)로 설명되는 다음 수정 사항은 모든 픽셀에 더 큰 강조점을 부여하여 모델이 분할 경계 변경에 더 민감하게 만든다. 사소한 편차라도 결과 출력의 품질에 큰 영향을 미칠 수 있기 때문에 치과 영상에서 정확한 분할 경계의 중요성은 아무리 강조해도 지나치지 않다.

[0109] 정사각형 주사위 손실 함수에는 0으로 나누는 것을 방지하는 정규화 인자가 포함되어 있다. 계산에는 두 맵의 제공된 결합된 합을 고려하면서 실제 분할 맵과 예측 분할 맵 사이의 교차점을 결정하는 작업이 포함된다. 이로 인해 신경망이 각 픽셀에 대해 정확한 예측을 우선시하여 정밀도와 재현율이 향상된다.

$$L = 1 - \left(2 \sum (y_t \cdot y_p)^2 + \epsilon \right) / \left(2 \sum (y_t^2 + y_p^2) + \epsilon \right) \quad (1)$$

[0110] 여기서 L은 손실 함수를 나타낸다. y_t 와 y_p 는 각각 Ground Truth와 예측된 분할 맵을 나타내며, ϵ 은 0으로 나누기를 방지하기 위한 평활화 요소이다. 효율성으로 인해 제공 주사위 손실이 우리 방법에 선택되었으며 이는 우수한 분할 결과로 이어진다. 의료영상 분할 작업에 널리 사용되는 Soft Dice loss, Tversky loss 및 Log-Cosh Dice loss 함수와 같은 여러 손실 함수를 시도한 결과, 제공 주사위 손실 함수는 DSC 및 JI 등에서 훨씬 더 나은 분할 성능을 달성한 것으로 관찰되었다.

[0112] 실험예 1

[0113] 다음과 같이 세팅하여 본 발명의 아키텍처(Ours)를 훈련시켰다.

[0114] 1. 데이터세트

[0115] 본 발명에서 데이터 세트는 치과 대학과 학생들 간의 공동 노력인 치과 파노라마 이미지로 구성된다. 이러한 파노라마 이미지는 감독 플랫폼을 사용하여 꼼꼼하게 주석을 달았으며 그 결과 여러 등급에 걸쳐 별도의 치아를 상세하게 분류했다. 전체적으로 컬렉션은 540개의 주석이 달린 이미지로서, 계산 효율성을 보장하고 메모리 요구량을 줄이기 위해 중요한 해부학적 랜드마크를 보존하면서 이러한 이미지의 크기를 도 3c에 도시된 블록다이 어그램에 따라 전처리 단계를 하기와 같이 수행하여 1024×512 픽셀 크기로 조정했다.

[0116] 학습모델을 훈련하고 테스트하기 전에 파노라마 이미지의 전반적인 품질을 향상시키기 위해 사전 처리 단계를 도 3c에 도시된 것처럼 이미지 크기 조정, 픽셀 강도를 [0, 255] 범위로 정규화 및 다중 스케일 형태를 순서대로 수행한다. 크기 조정 단계는 이미지 크기를 1024×512 픽셀 크기로 조정하는 것으로 모델 훈련 및 테스트 중 계산 효율성에 기여한다. 이 특정 크기는 이미지 품질과 메모리 수요의 균형을 맞추고 계산 리소스 사용을 최적화하기 위해 전략적으로 선택될 수 있다. 정규화 단계는 이미지 내의 픽셀 강도는 특정 범위(일반적으로 [0, 255])로 정규화하는 것으로, 이 단계에서는 데이터가 일관되고 표준화된 규모 내에서 작동하도록 보장하여 모델이 패턴을 더 쉽게 학습할 수 있도록 한다. 다중 규모 형태학은 노이즈를 줄이고 대비를 개선하며 치과 이미지의 두드러진 특징을 강조하기 위해 일련의 형태학적 작업이 다양한 이미지 규모에 적용된다. 여기에는 이미지 내 개체의 구조와 모양을 조작하는 작업이 포함된다. 이러한 전처리는 모델의 입력 데이터를 정제하여 훈련 및 테스트 단계에서 향상된 성능을 보장하는 데 필수적으로 필요할 수 있다.

[0117] 2. 실험적 설정

[0118] 본 발명에 따른 치아 분할 아키텍처를 훈련하고 평가하기 위해 균일한 설정을 활용하여 기존 방법론과 공정하고 일관된 비교를 보장했다. Keras 프레임워크와 NVIDIA GeForce RTX 3090 그래픽 처리 장치(GPU)가 모델 훈련 및 평가에 활용되었다. 데이터세트는 훈련(70%), 검증(15%), 테스트(15%) 데이터세트로 나누어졌다. 50개 에포크에 걸쳐 훈련된 2×2 영역이 있는 Swin Transformer 모델을 적용했다. 초기 학습률은 10^{-4} 로 설정되었으며 이후 잠재적인 과적합 문제를 해결하기 위해 적응적으로 10^{-7} 로 감소되었다. 훈련 과정 전반에 걸쳐 배치 크기를 2개로 유지했다. 과적합을 극복하기 위해 인코더 컨볼루션 레이어 뒤에 드롭아웃 레이어가 추가된다. 또한 검증 손실

이 5번 연속으로 감소하지 않으면 학습률을 10% 감소시키는 전략을 사용했다.

[0119] 본 발명에 따른 치아 분할 아키텍처의 유효성은 Accuracy(ACC), Jaccard Index(JI), Precision, Recall 및 Specity 총 5가지 표준 평가 지표를 사용하여 정량적으로 평가되었다.]. 이러한 측정항목은 다음과 같이 정의된다.

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \quad (2)$$

$$\text{JI} = \frac{|P \cap G|}{|P \cup G|} \quad (3)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (4)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (5)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (6)$$

[0120]

[0121] 여기서 TP(참 긍정)와 TN(참 부정)은 각각 치아와 치아가 아닌 것으로 정확하게 분류된 픽셀 수를 나타낸다. 반대로 FP(False Positive)와 False Negative(FN)는 각각 치아와 비치아로 잘못 분류된 픽셀 수를 나타낸다. Jaccard 지수의 맥락에서 "P"는 일반적으로 모델이 예측한 요소 집합을 나타내고 "G"는 실제 값 집합을 나타낸다.

[0122] 이러한 지표는 0에서 1까지의 비교 가능한 척도를 제공하며, 1은 예측 값과 실제 값이 정확히 일치함을 나타낸다. 이러한 매개변수에서 점수가 높을수록 분할 성능이 더 우수함을 의미하며, 이는 모델이 치아를 정확하게 분할하고 치아 이미지의 다른 구조와 치아를 구별하는 데 효율적이라는 것을 나타낸다. 이러한 평가 지표를 적용함으로써 제안된 치아 분할 모델에 대한 포괄적인 성능 평가를 제공하여 현장에서 확립된 다른 방법과의 비교 분석을 용이하게 한다.

[0123] 3. 결과

[0124] 본 발명에 따른 치아 분할 아키텍처 즉 학습모델을 전통적인 U-Net[20], Attention U-Net[12], ResNet-50 Attention U-Net[39], Swin U-Net[40], BB-Unet과 동일한 Modified U-Net[10]을 포함하여 잘 확립된 여러 분할 모델과 비교했다. 본 발명에 따른 치아 분할 아키텍처 즉 제안된 모델의 성능은 ACC, DSC(Dicesimilarity Coefficient), JI, 정밀도, 재현율, 그리고 특이성을 포함하여 다중 중요 지표를 사용하여 평가되었다. 세계치과의사연맹(FDI) 표기법을 기반으로 X-ray 이미지의 각 치아를 32개 범주로 분할하는 것을 목표로 하며, 각 치아는 #11~#18, #21~#28, #31~#38 및 #41~ #48까지로 분류된다. 본 발명에 따른 치아 분할 아키텍처는 멀티클래스 32개 분류 픽셀을 사용하여 각 픽셀을 특정 숫자로 분류하므로 참양성(TP), 참음성(TN), 거짓양성(FP), 거짓음성(FN)도 측정하여 객관적인 평가 지표를 얻었다. 세분화를 위해, TP, TN, FP 및 FN은 치아 수를 하나의 클래스로 간주하고 다른 모든 치아를 다른 클래스로 고려하여 각 치아 유형에 대해 계산된다. 이 과정은 각 클래스의 각 치아 유형에 대해 반복되었다.

[0125] 또한, 계산 비용 분석을 수행하여 동일한 실험 조건에서 각 방법의 훈련 시간을 측정했다. 본 발명에 따른 치아 분할 아키텍처(Proposed)는 표 1(본 발명의 방법과 기존방법의 분할성능 비교)에 나타난 분할 성능[DSC 값: 0.9102]을 달성하기 위해 훈련하는데 약 60분이 필요했다.

표 1

MODELS	ACC	DSC	JI	PRECISION	RECALL	SPECIFICITY
U-Net [20]	0.9720	0.7602	0.6871	0.7458	0.8366	0.9725
ATTENTION U-Net [12]	0.9720	0.7846	0.7132	0.7557	0.8391	0.9725
RESNET-50 ATTENTION U-Net [39]	0.9721	0.7875	0.7172	0.7487	0.8544	0.9726
SWIN U-Net [40]	0.9712	0.6348	0.5296	0.6107	0.7192	0.9721
MODIFIED-U-Net [10]	0.9726	0.9004	0.8489	0.7898	0.9366	0.9728
PROPOSED	0.9726	0.9102	0.8501	0.8046	0.9389	0.9730

[0126]

[0127] 다른 방법 즉 U-Net[20], Attention U-Net[12], ResNet-50 Attention U-Net[39], Swin U-Net[40], Modified U-Net[10]은 각각 0.7602, 0.7846, 0.7875, 0.6348 및 0.9004의 DSC 값을 얻기 위해 45, 48, 48, 68 및 64분의

학습 시간을 필요로 했다. 여기서, U-Net[20]은 Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. arXiv preprint arXiv:1505.04597의 방법이고, Attention U-Net[12]는 O. Oktay et al., 'Attention U-Net: Learning Where to Look for the Pancreas', 04 2018의 방법이며, ResNet-50 Attention U-Net[39]는 R. Ji, L. Zhang, Y. Wang, C. Shen, and X. Wang, 'Attention U-Net: Learning Where to Look for Biomedical Image Segmentation', IEEE Transactions on Medical Imaging, vol. 38, no. 11, pp. 2993-3005, 2019.의 방법이고, Swin U-Net[40]은 Cao, H., Wang, Y., Chen, J., Jiang, D., Zhang, X., Tian, Q., & Wang, M. (2021). Swin-Unet: Unet-like pure transformer for medical image segmentation. arXiv preprint arXiv:2105.05537의 방법이며, Modified U-Net[10]은 Nader, R., Smorodin, A., La Fourniere, N. D., Amouriq, Y., & Autrusseau, F. (2022, May 18). Automatic teeth segmentation on panoramic X-rays using deep neural networks. Retrieved from 방법이다. 이 값은 위에서 언급한 실험 설정에서 얻은 것이다. 학습 중 중지 기준과 학습 속도는 각 방법마다 다양하므로 복잡성-성능 균형의 우월성을 판단하는 것은 어렵지 않다. 각 방법 별 분할 성능 변화를 조사하기 위해 중지 기준과 학습률 값을 조정하여 유사한 실행 시간을 설정했다. 표 2는 거의 동일한 실행 시간에서 DSC 및 JI와 같은 분할 성능을 비교한 것이다.

표 2

MODELS	RUN-TIMES (IN MINUTES)	DSC	JI
U-NET [20]	61	0.7556	0.6815
ATTENTION U-NET [12]	59	0.7796	0.7071
RESNET-50 ATTENTION U-NET [39]	61	0.7804	0.7096
SWIN U-NET [40]	60	0.5432	0.4302
MODIFIED-U-NET [10]	61	0.9003	0.8490
PROPOSED	60	0.9102	0.8501

[0128]

[0129]

표 2에서 볼 수 있듯이 본 발명의 학습모델을 사용한 다중클래스 분할 방법은 유사한 복잡성 하에서 훨씬 더 높은 분할 성능을 달성할 수 있다. 이는 또한 본 발명의 방법이 동일한 분할 성능을 달성하기 위해 훨씬 더 낮은 런타임이 필요함을 나타낸다.

[0130]

본 발명의 학습모델에 따른 정확도 점수(0.9726)는 다른 모델의 정확도 점수와 비슷하다. 그러나 Dice Coefficient(0.9102)와 JI(0.8501)를 기준으로 볼 때 본 발명에 따른 학습모델이 다른 모델보다 훨씬 뛰어난 성능을 보였다. 이러한 점수는 본 발명의 학습 모델이 거짓 긍정과 거짓 부정을 최소화하면서 참 긍정을 구별한다는 것을 보여준다. 또한 각각 0.8046과 0.9389의 정밀도와 재현율 점수가 추가 증거를 제공했다. 표 2는 본 발명의 학습 모델이 다른 모델에 비해 우수한 성능을 보여준다.

[0131]

각 모델의 성능을 비판적으로 분석했다. 전통적인 U-Net 모델은 ACC, Dice 점수 및 기타 지표 측면에서 칭찬할 만한 성능을 보였지만, 본 발명에서 제안된 학습모델의 성능을 능가하지는 못했다. 마찬가지로 Attention U-Net과 ResNet-50 Attention U-Net은 기존 U-Net 모델에 비해 향상된 성능을 보여주었지만 여전히 본 발명의 학습모델의 성능보다 떨어졌고, Swin U-Net 모델의 성능은 U-Net 모델보다 떨어졌으며, 반면 Modified U-Net은 본 발명의 학습모델 보다 다소 떨어지지만 유사한 성능을 보여주었다.

[0132]

본 발명의 학습 모델의 뛰어난 성능은 주로 Swin Transformer와 경계 상자의 통합과 제곱 주사위 손실을 활용하는 수정된 손실 함수의 적용에 기인할 수 있다. 이러한 설계 결정을 통해 본 발명의 학습 모델은 입력 이미지에서 치아 구조를 학습하고 분할할 수 있었으며, 이를 통해 모든 평가 지표 전반에 걸쳐 성능이 향상되었다.

[0133]

정량적 비교 외에도 도 7a 및 7b와 도 8a 및 8b에 도시된 바와 같이, 본 발명의 학습모델과 상술된 모든 모델의 분할 결과를 시각적으로 비교했다. 이 시각적 비교는 본 발명의 학습모델이 치아 구조를 정확하게 분할한다는 증거를 제공한다. 즉, 본 발명의 학습 모델은 스윙트랜스포머, 경계 상자 및 수정된 손실 함수의 이점을 통해 보다 정확하고 일관된 분할 결과를 일관되게 생성했으며, 다양한 치과 이미지에 사용되는 접근 방식의 향상된 일반화 가능성은 주로 TAB 및 제곱 주사위 손실 함수의 추가에 기인할 수 있다. 치과 영상에 TAB를 사용하면 경계 묘사의 정확성이 향상되고 개별 치아 분석의 세부 수준이 높아진다. 이 혁신적인 접근 방식은 치과 이미지에서 흔히 발견되는 노이즈와 아티팩트의 영향을 줄이는 효과를 입증했다.

[0134]

도 5b의 제1열에 대한 자세한 분석은 본 발명에 따른 학습모델의 효율성을 보여준다. 예를 들어 치아 #48(FDI 표기법)의 경우 대부분의 기존 모델은 경계를 정확하게 정의하는 데 어려움을 겪었다. 그러나 본 발명의 학습

모델은 탁월한 정밀도를 보여 지상 진실과 거의 일치하는 분할 결과를 제공한다. 이는 주로 TAB가 각 픽셀을 둘러싼 국소 수용 필드에 초점을 맞춰 개체의 경계를 강조하고 개별 치아와 인접 치아 간의 구별을 크게 향상시키는 기능 때문이다.

[0135] 본 발명의 학습모델에 따른 이 분석은 특히 복잡한 치과 구조를 다루고 정확하고 균일한 분할 결과를 얻을 때 다른 접근 방식에 비해 제안된 접근 방식의 잠재적 이점을 강조한다. 본 발명과 같이 학습모델에 TAB를 통합하면 전반적인 성능이 크게 향상되어 치과 영상 분석 및 진단 진행 시 자산으로서의 잠재력을 나타낸다.

[0136] 표 1 및 표 2 와 도 7a 내지 도 8b에 제시된 결과는 본 발명의 학습모델 즉 맞춤형 치과 세분화 모델이 평가 지표 및 시각적 비교 측면에서 여러 최첨단 모델보다 성능이 우수하다는 것을 보여준다. 이러한 성능 향상은 주로 고유한 TAB 덕분일 수 있는데, 경계 개선 필터로 TAB를 사용하면 각 치아와 인접 구조의 식별이 크게 향상된다. 스윈트랜스포머와 특정 손실 함수가 중요한 역할을 하지만 TAB을 사용하면 치과 이미지 분석의 정확성과 효율성이 더욱 향상된다. 따라서, 본 발명은 치과 이미지 분석의 미래 발전을 위한 강력한 플랫폼을 제공하고 진단, 치료 계획 및 환자 모니터링을 포함한 치과 절차의 잠재적 영향을 향상시킬 것으로 예측된다.

[0137] 실험예 2

[0138] 본 발명의 학습모델 즉 치아 분할 아키텍처(네트워크)에 대한 각 구성 요소의 기여도를 평가하기 위해 다음과 같이 절제 실험이 수행되었다. 유사한 훈련 조건에서 네트워크의 각 변형에는 기본 U-Net 모델의 필수 구성 요소가 연속적으로 포함되었다. 본 발명에서 조사한 구성요소에는 Deep Supervision, Swin Transformers 및 TAB가 포함된다. ACC, DSC, JI, 정밀도, 재현율, 특이성과 같은 몇 가지 중요한 지표를 사용하여 변형 A-D라는 각 모델과 본 발명의 학습 모델인 전체 네트워크(Proposed)의 효율성을 조사했다. 절제 연구의 변형 구성 요소를 표 3에 나타내고, 각 변수에 대한 성능 결과를 표 4에 나타내었다.

표 3

VARIATIONS	U-NET	DEEP SUPERVISION	SWIN TRANSFORMER	TAB
VARIATION A	✓	✗	✗	✗
VARIATION B	✓	✓	✗	✗
VARIATION C	✓	✗	✓	✗
VARIATION D	✓	✗	✗	✓
PROPOSED	✓	✓	✓	✓

[0139]

표 4

VARIATIONS	ACCURACY	DSC	JI	PRECISION	RECALL	SPECIFICITY
VARIATION A	0.9720	0.7602	0.6871	0.7458	0.8366	0.9725
VARIATION B	0.9722	0.7846	0.7132	0.7557	0.8391	0.9725
VARIATION C	0.9721	0.7644	0.6905	0.7569	0.8477	0.9726
VARIATION D	0.9725	0.9001	0.8476	0.7908	0.9312	0.9728
PROPOSED	0.9726	0.9102	0.8501	0.8046	0.9389	0.9730

[0140]

[0141] 표 3은 각 변형의 구성 요소에 대한 세부 정보를 제공한다. 변형 A는 기본 U-Net 구조이며 각 변형에는 Deep Supervision(변형 B), Swin Transformers(변형 C) 및 TAB(변형 D)라는 추가 구성 요소가 포함되어 있다. 표 4는 절제 연구에 대한 분할 성능을 보여준다. 표 4에서 볼 수 있듯이 분할 결과의 정확도는 변형 개수가 적을수록 높아진다. Deep Supervision이 사용되는 변형 B의 DSC 및 JI에서는 약간의 이득이 관찰되었다. 이러한 개선은 네트워크 전체에 걸쳐 더 나은 기능 전파로 인해 발생하며 치아 클래스 간의 모델 구별이 향상된다. 스윈트랜스포머(Swin Transformers)는 Variance C의 결과에서 볼 수 있듯이 DSC에서 약간의 향상을 가져온다. 이는 Self-Attention 메커니즘을 갖춘 Swin Transformers가 로컬 및 전역 상황 정보를 캡처할 수 있다는 사실에 기인한다. 각 치아가 인접 치아의 맥락에 영향을 미칠 수 있는 치과 이미지의 복잡한 구조를 분할한다. 제안된 TAB가 단독으로 수행되는 변형 D에서 성능 향상이 눈에 띄게 관찰되었다. TAB는 치아 경계에 선택적으로 초점을 맞춰 모델 성능을 향상시키고 분할 마스크의 정확성과 정밀도를 향상시킨다. TAB은 경계 픽셀에 더 높은 주의 점

수를 할당하여 치아와 주변 구조 간의 차별화를 개선하여 개별 치아의 가장자리를 더욱 뚜렷하게 만든다. 이는 모델의 전반적인 성능을 향상시켜 성능 지표에서 관찰할 수 있는 보다 정확하고 상세한 치아 분할 결과를 제공한다.

[0142] 절제 연구의 각 변형과 관련된 분할 성능을 보여주는 표 4로부터, 추가 구성 요소가 포함된 각 연속 변형이 성능 지표를 점진적으로 증가시키는 것으로 나타났다.

[0143] 상술된 바와 같이, 기본 U-Net 구조를 기반으로 설정했을 때, Deep Supervision 및 Swin Transformers와 같은 기능을 추가함에 따라 모델 성능이 향상된 결과를 보여준다. 추가된 기능 도구 중에서 본 발명에서 최초로 제안된 TAB에 대한 성능이 가장 크게 향상되었다. TAB는 복잡한 치아 해부학을 강조하고 치아 경계를 미세하게 묘사하는 데 탁월한 능력을 보여주는데, TAB에 내장된 새로운 어텐션 메커니즘은 복잡한 치아 구조를 정확하게 강조하여 매우 정확한 분할 결과를 제공하는 것으로 보인다.

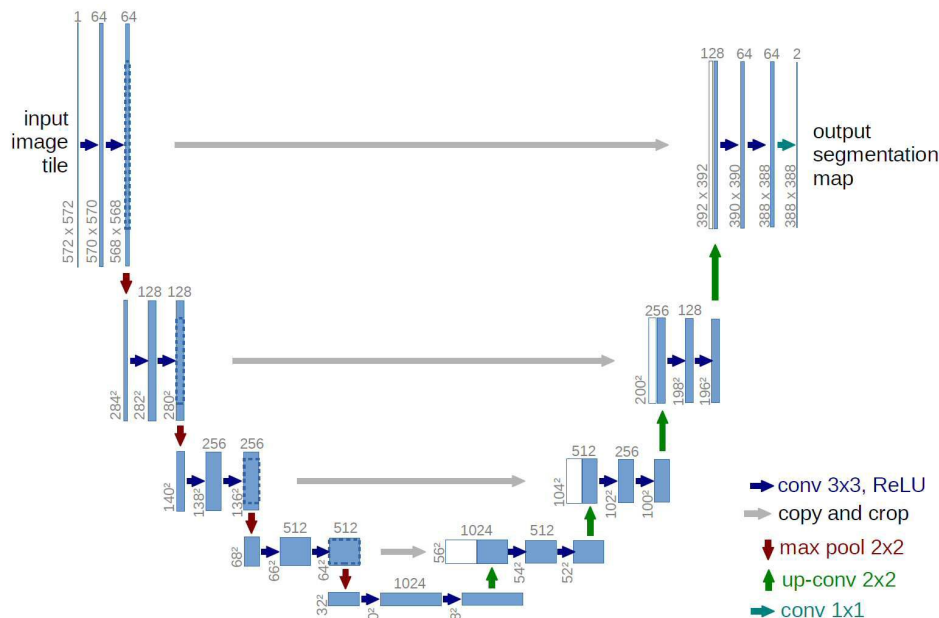
[0144] 이러한 결과로부터 본 발명의 학습모델이 적용된 치과 X-ray 이미지 다중클래스 분할방법 및 장치는 치과 X-ray 이미지에 대한 정확한 분할 성능으로 인해 치과 이미지 분석을 크게 향상시키고 치과 응용 분야의 발전에 기여할 수 있는 잠재력을 가지고 있다. 예를 들어 치아 모양의 변화나 병변의 모양 등을 파악해 치주질환이나 치아 우식증 등을 조기에 발견하는 등 자동화된 질병 진단 및 치료 계획의 타당성을 보여줄 수 있다.

[0145] 본 발명은 이상에서 살펴본 바와 같이 바람직한 실시 예를 들어 도시하고 설명하였으나, 상기한 실시 예에 한정되지 아니하며 본 발명의 정신을 벗어나지 않는 범위 내에서 당해 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 다양한 변경과 수정이 가능할 것이다.

[0146] 즉, 본 명세서에 개시된 다양한 실시예들은 순서에 관계없이 수행될 수 있으며, 동시에 또는 별도로 수행될 수 있다. 일 실시예에서, 본 명세서에서 설명되는 각 도면에서 적어도 하나의 단계가 생략되거나 추가될 수 있고, 역순으로 수행될 수도 있으며, 동시에 수행될 수도 있다. 본 명세서에 개시된 실시예들은 본 발명의 기술적 사상을 한정하기 위한 것이 아니라, 설명하기 위한 것이고, 이러한 실시예들에 의하여 본 발명의 범위가 한정되는 것은 아니다.

도면

도면1

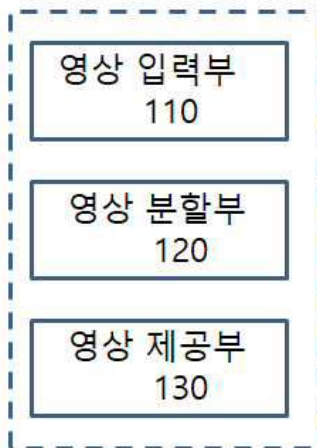


도면2

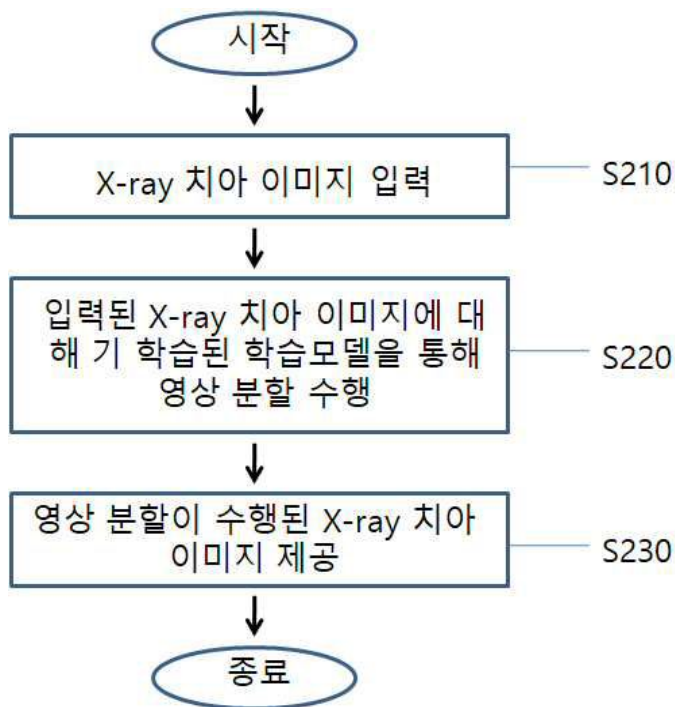


도면3a

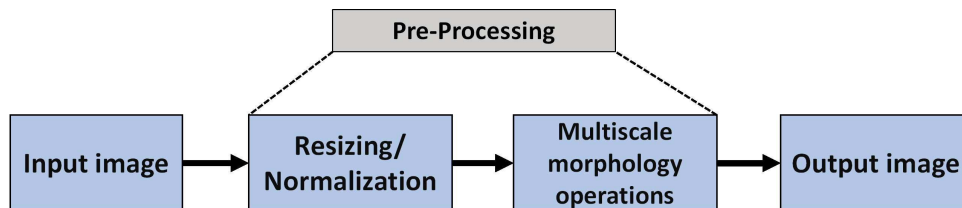
100



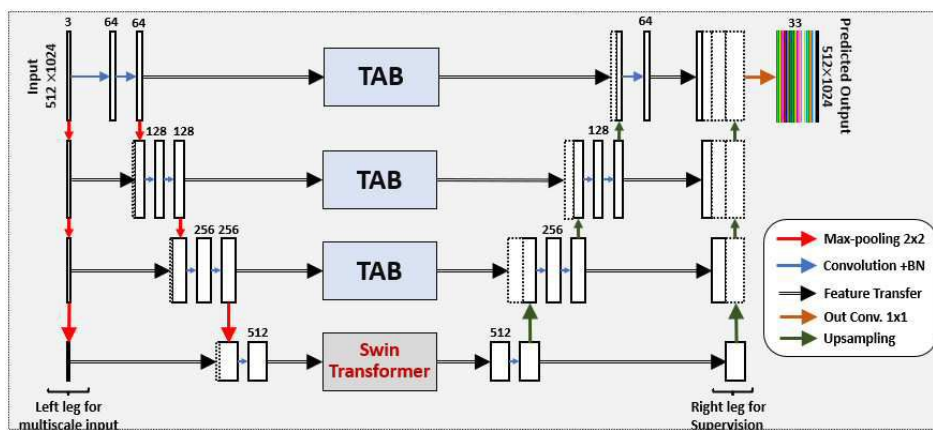
도면3b



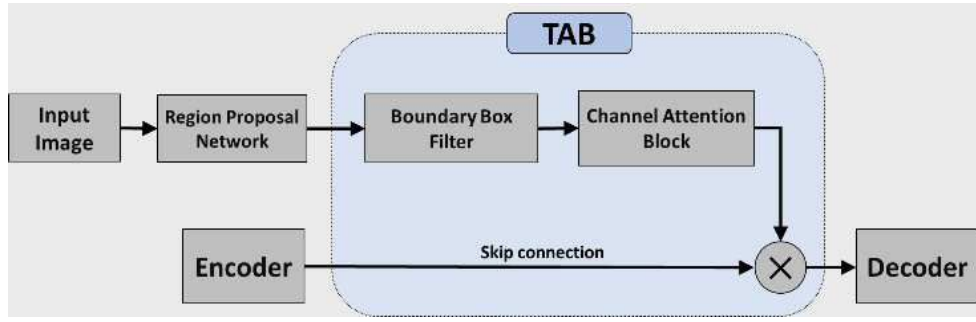
도면3c



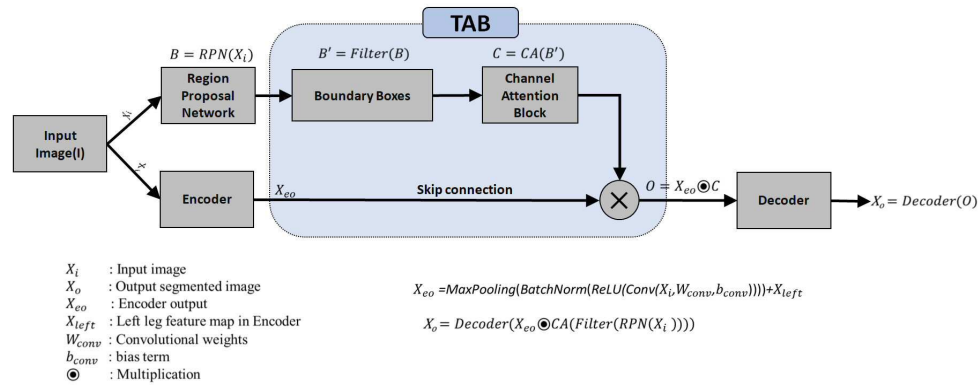
도면4



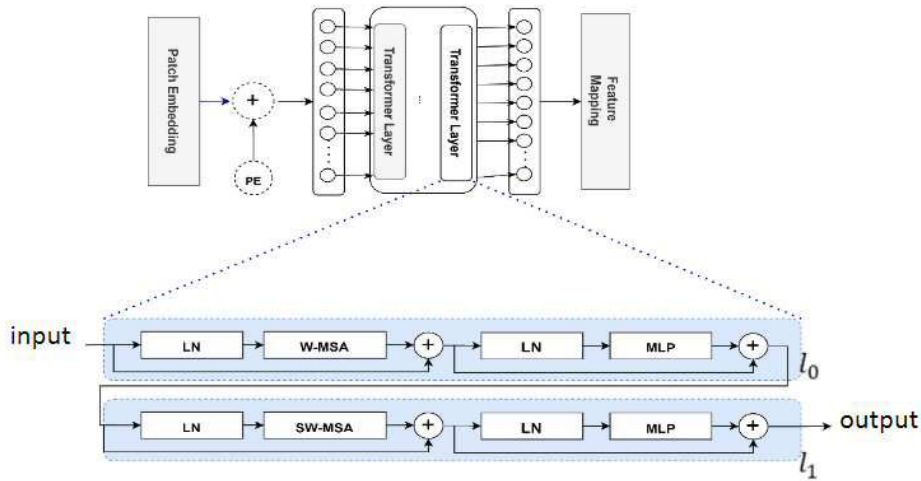
도면5a



도면5b

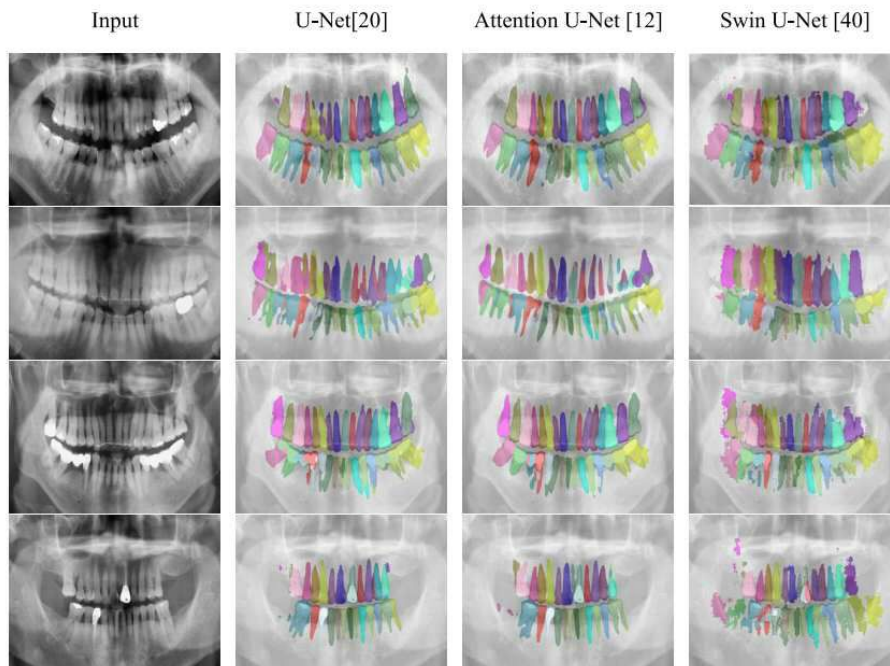


도면6

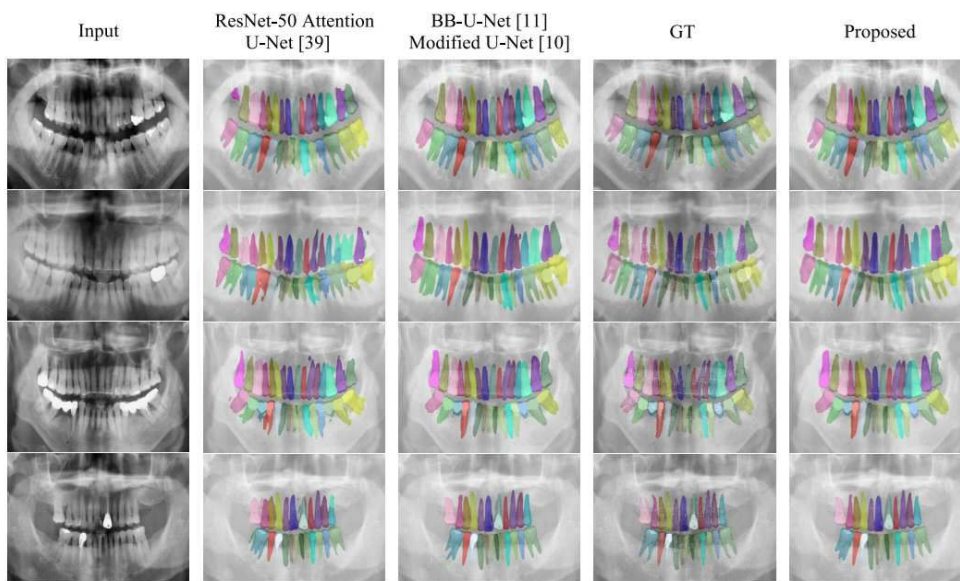


LN: Layer Normalization
MLP: Multi-Layer Perceptron
W-MSA: Window based multi-self head attention
SW-MSA: Shifting Window based multi-self head attention

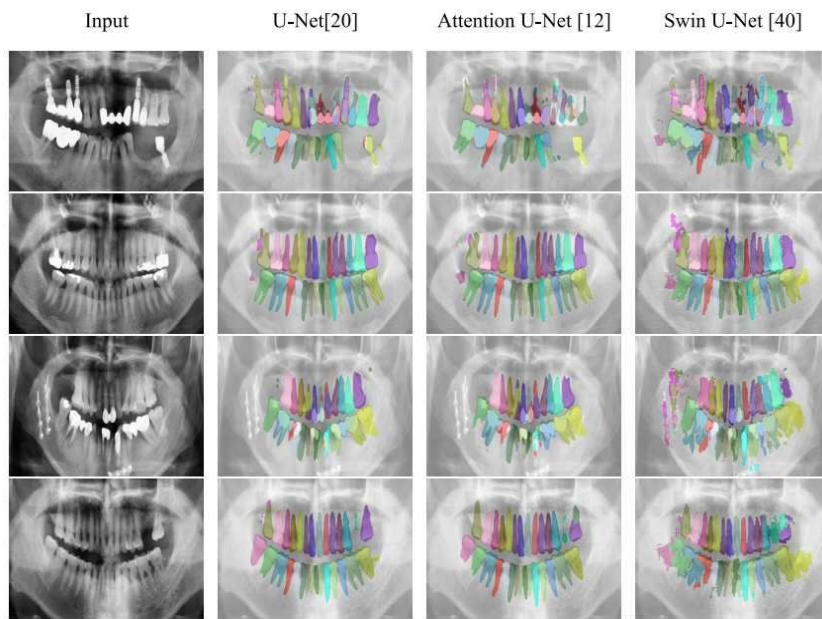
도면7a



도면7b



도면8a



도면8b

