



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0105313
(43) 공개일자 2024년07월05일

(51) 국제특허분류(Int. Cl.)
G06T 7/11 (2017.01) A61B 6/00 (2024.01)
A61B 6/51 (2024.01) G06T 3/40 (2024.01)
G06T 5/20 (2024.01) G16H 30/40 (2018.01)
(52) CPC특허분류
G06T 7/11 (2017.01)
A61B 6/51 (2024.01)
(21) 출원번호 10-2023-0194764
(22) 출원일자 2023년12월28일
심사청구일자 2023년12월28일
(30) 우선권주장
1020220187245 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
황인선
광주광역시 서구 화개1로78번길 8 501동 1002호
(74) 대리인
특허법인아이엠

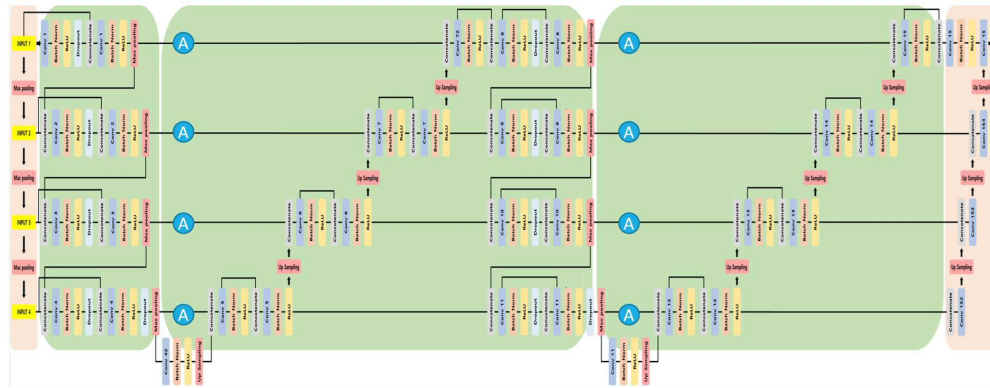
전체 청구항 수 : 총 22 항

(54) 발명의 명칭 MAM-net 구조 기반 의료영상 분할 방법 및 장치

(57) 요약

본 발명은 의료영상에 대한 영상분할기법 관련기술로서, 보다 구체적으로는 기존에 존재하는 M-Net의 기본 구조에 어텐션 모듈(Attention module)을 추가한 MAM-net구조를 통해 보다 효과적으로 영상 분할(image segmentation)을 할 수 있는 MAM-Net을 이용한 의료영상 분할 방법 및 장치에 관한 것이다.

대표도 - 도2b



(52) CPC특허분류

A61B 6/5217 (2020.08)

G06T 3/40 (2024.01)

G06T 5/20 (2024.01)

G16H 30/40 (2018.01)

G06T 2207/20081 (2013.01)

G06T 2207/20084 (2013.01)

G06T 2210/41 (2013.01)

(72) 발명자

김유진

광주광역시 광산구 수등로 76번길 74 중흥S클래스
1차 112동 1903호

박성연

광주광역시 남구 봉선로175번길 14 203동 204호

이 발명을 지원한 국가연구개발사업

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

명세서

청구범위

청구항 1

의료영상 이미지를 입력받는 단계; 상기 입력된 의료영상 이미지를 기학습된 학습 모델에 적용하여 영상 분할을 수행하는 단계; 및 상기 영상 분할이 수행된 의료영상 이미지를 제공하는 단계;를 포함하며,

상기 학습 모델이 컨볼루션(Convolution)과 맥스풀링(Maxpooling)을 수행하는 인코더; 및 업-샘플링(Up-sampling)과 컨볼루션(Convolution)을 수행하는 디코더를 포함하는 M-Net 모델로서, 상기 인코더와 상기 디코더를 연결하는 스킵연결 사이에 2개의 어텐션 모듈과 상기 2개의 어텐션 모듈 사이에 내부디코더 및 내부인코더가 형성된 IM(Inter Mediate)구조가 결합된 아키텍처를 갖는 의료영상 분할 방법.

청구항 2

제 1 항에 있어서,

상기 IM구조는 상기 인코더의 출력이 입력되는 제1어텐션모듈, 상기 제1어텐션모듈의 출력과 하위 레이어에서 얻어진 특징 맵을 연결(Concatenate)시켜 컨볼루션 및 업샘플링하는 내부디코더, 상기 내부디코더에서 출력된 특징 맵과 상위 레이어에서 얻어진 특징 맵을 연결시켜 컨볼루션 및 맥스풀링하는 내부인코더 및 상기 내부인코더의 출력이 입력되는 제2어텐션모듈이 상기 인코더 및 상기 디코더를 구성하는 각 레이어에 대응되게 형성되는 것을 특징으로 하는 의료영상 분할 방법.

청구항 3

제 2 항에 있어서,

상기 영상 분할을 수행하는 단계는 상기 의료영상 이미지를 상기 인코더로 입력하는 제1단계; 상기 인코더에서 출력된 특징 맵을 IM구조로 입력하는 제2단계; 및 상기 IM구조에서 출력된 특징 맵을 상기 디코더로 입력하는 제3단계; 및 상기 디코더에서 출력된 결과값을 분류하는 제4단계;를 포함하는 것을 특징으로 하는 의료영상 분할 방법.

청구항 4

제 3 항에 있어서, 상기 제1단계는

왼쪽다리의 제1레이어에서 수행되는 상기 의료영상이미지에 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 입력된 상기 제1 레이어용 패치 이미지에 대해 제1컨볼루션(convolution)을 수행하는 단계; 상기 제1컨볼루션결과값과 상기 제1레이어용 패치이미지를 연결(concatenate)시켜 제2컨볼루션을 수행하는 단계; 상기 제2컨볼루션 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 의료영상 분할 방법.

청구항 5

제 4 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 제1컨볼루션(convolution)을 수행하는 단계; 상기 제1컨볼루션결과값과 상기 제2레이어용 패치이미지를 연결(concatenate)시켜 제2컨볼루션을 수행하는 단계; 상기 제2컨볼루션의 결과값에 대해 맥스 풀링(max pooling)이 수행되는 단계; 및 상기 맥스풀링 결과값을 상기 인코더의 제3 레이어의 입력으로 제공하는 단계;를 포함하는 것을 특징으로 하는 의료영상 분할 방법.

청구항 6

제 5 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 제1컨볼루션(convolution)을 수행하는 단계; 상기 제1컨볼루션결과값과 상기 제3레이어용 패치이미지를 연결(concatenate)시켜 제2컨볼루션을 수행하는 단계; 상기 제2컨볼루션의 결과값에 대해 맥스 풀링(max pooling)이 수행되는 단계; 및 상기 맥스풀링 결과값을 상기 인코더의 제4 레이어의 입력으로 제공하는 단계;를 포함하는 것을 특징으로 하는 의료영상 분할 방법.

청구항 7

제 6 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 제1컨볼루션(convolution)을 수행하는 단계; 상기 제1컨볼루션결과값과 상기 제4레이어용 패치이미지를 연결(concatenate)시켜 제2컨볼루션을 수행하는 단계; 상기 제2컨볼루션의 결과값에 대해 맥스 풀링(max pooling)수행되는 단계; 및 상기 맥스 풀링 결과값에 기반한 특징 맵을 상기 IM구조의 입력으로 제공하는 단계;를 포함하는 것을 특징으로 하는 의료영상 분할 방법.

청구항 8

제 7 항에 있어서, 상기 제2단계는

상기 인코더의 각 레이어에서 출력되는 특징 맵이 해당 레이어에 위치한 제1어텐션모듈의 입력으로 제공되는 단계;

차하위 레이어에서 출력되는 특징 맵과 상기 제1어텐션모듈의 해당레이어에서 출력되는 특징 맵을 연결시켜 상기 내부디코더의 입력으로 제공하는 단계;

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

상기 내부디코더의 해당 레이어에서 출력되는 특징 맵을 상기 내부인코더의 해당 레이어의 입력으로 제공하면서, 업샘플링하여 상기 내부디코더의 차상위 레이어에 입력으로 제공하는 단계;를 포함하는 것을 특징으로 하는 의료영상 분할 방법.

청구항 9

제 8 항에 있어서, 상기 제2단계는

상기 내부디코더의 각 레이어에서 출력된 특징 맵이 차상위 레이어에서 얻어진 특징 맵과 연결되어 상기 내부인코더의 해당 레이어에 대한 입력으로 제공되는 단계;

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

상기 내부인코더의 해당 레이어에 대한 컨볼루션 결과값에 대해 맥스 풀링(max pooling)이 수행되는 단계; 및

상기 내부인코더의 해당 레이어에서 출력되는 특징 맵이 해당 레이어에 위치한 제2어텐션모듈의 입력으로 제공되는 단계;를 포함하는 것을 특징으로 하는 의료영상 분할 방법.

청구항 10

제 9 항에 있어서, 상기 제3단계는

차하위 레이어에서 출력되는 특징 맵과 상기 제2어텐션모듈의 해당레이어에서 출력되는 특징 맵을 연결시켜 상기 디코더의 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 의료영상 분할 방법.

청구항 11

제 10 항에 있어서, 상기 제3단계는

상기 디코더의 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 컨볼루션하는 단계; 를 포함하는 것을 특징으로 하는 의료영상 분할 방법.

청구항 12

의료영상 이미지를 입력받는 입력부;

상기 입력된 의료영상 이미지를 기학습된 학습 모델에 적용하여 시맨틱 분할을 수행하는 영상 분할부; 및

상기 시맨틱 분할이 수행된 의료영상 이미지를 제공하는 영상 제공부;를 포함하며,

상기 학습 모델은 컨볼루션(Convolution)과 맥스풀링(Maxpooling)을 수행하는 인코더; 및 업-샘플링(Up-sampling)과 컨볼루션(Convolution)을 수행하는 디코더를 포함하는 M-Net 모델로서, 상기 인코더와 상기 디코더를 연결하는 스킵연결 사이에 2개의 어텐션 모듈과 상기 2개의 어텐션 모듈 사이에 내부디코더 및 내부인코더가 형성된 IM(Inter Mediate)구조가 결합된 아키텍처를 갖는 의료영상 분할 장치.

청구항 13

제 12 항에 있어서,

상기 학습 모델에서 상기 IM구조는 상기 인코더의 출력이 입력되는 제1어텐션모듈, 상기 제1어텐션모듈의 출력과 하위 레이어에서 얻어진 특징 맵을 연결(Concatenate)시켜 컨볼루션 및 업샘플링하는 내부디코더, 상기 내부디코더에서 출력된 특징 맵과 상위 레이어에서 얻어진 특징 맵을 연결시켜 컨볼루션 및 맥스풀링하는 내부인코더 및 상기 내부인코더의 출력이 입력되는 제2어텐션모듈이 상기 인코더 및 상기 디코더를 구성하는 각 레이어에 대응되게 형성되는 것을 특징으로 하는 의료영상 분할 장치.

청구항 14

제 12 항 또는 제 13 항에 있어서,

상기 영상 분할부는 상기 의료영상 이미지를 상기 인코더로 입력하고, 상기 인코더에서 출력된 특징 맵을 IM구조로 입력하고, 상기 IM구조에서 출력된 특징 맵을 상기 디코더로 입력하며, 상기 디코더에서 출력된 결과값을 분류하는 것을 특징으로 하는 의료영상 분할 장치.

청구항 15

제 14 항에 있어서, 상기 영상 분할부는

왼쪽다리의 제1레이어에서 상기 의료영상이미지에 대해 상기 인코더의 제1 레이어용 패치 이미지를 생성하고, 상기 제1레이어용 패치이미지에 대해 맥스풀링을 수행하여 그 결과값을 제2레이어의 입력으로 제공하고,

상기 인코더의 제1레이어에서 입력된 상기 제1 레이어용 패치 이미지에 대해 제1컨볼루션(convolution)을 수행하고, 상기 제1컨볼루션결과값과 상기 제1레이어용 패치이미지를 연결(concatenate)시켜 제2컨볼루션을 수행하며, 상기 제2컨볼루션 결과값에 대해 맥스 풀링(max pooling)을 수행하고, 상기 맥스풀링 결과값을 상기 인코더의 제2 레이어의 입력으로 제공하는 것을 특징으로 하는 의료영상 영상 분할 장치.

청구항 16

제 15 항에 있어서, 상기 영상 분할부는

상기 왼쪽다리의 제2레이어에서 입력된 사이즈 그대로 인코더의 제2 레이어용 패치이미지를 생성하고, 상기 제2레이어용 패치이미지에 대해 맥스풀링을 수행하여 그 결과값을 제3레이어의 입력으로 제공하며,

상기 인코더의 제2레이어에서 상기 입력된 제1 레이어에 대한 맥스풀링의 결과값 및 상기 제2레이어용 패치 이미지를 연결(concatenate)시켜 제1컨볼루션(convolution)을 수행하며, 상기 제1컨볼루션결과값과 상기 제2레이어용 패치이미지를 연결(concatenate)시켜 제2컨볼루션을 수행하고, 상기 제2컨볼루션의 결과값에 대해 맥스 풀링(max pooling)을 수행하여 그 결과값을 상기 인코더의 제3 레이어의 입력으로 제공하는 것을 특징으로 하는 의료영상 분할 장치.

청구항 17

제 16 항에 있어서, 상기 영상분할부는

상기 왼쪽다리의 제3레이어에서 입력된 사이즈 그대로 인코더의 제3레이어용 패치이미지를 생성하고, 상기 제3레이어용 패치이미지에 대해 맥스풀링을 수행하여 그 결과값을 제4레이어의 입력으로 제공하며,

상기 인코더의 제3레이어에서 상기 입력된 제2 레이어에 대한 맥스풀링의 결과값 및 상기 제3레이어용 패치 이미지를 연결(concatenate)시켜 제1컨볼루션(convolution)을 수행하며, 상기 제1컨볼루션결과값과 상기 제3레이어용 패치이미지를 연결(concatenate)시켜 제2컨볼루션을 수행하고, 상기 제2컨볼루션의 결과값에 대해 맥스 풀링(max pooling)을 수행하여 그 결과값을 상기 인코더의 제4 레이어의 입력으로 제공하는 것을 특징으로 하는 의료영상 분할 장치.

청구항 18

제 17 항에 있어서, 상기 영상 분할부는

상기 왼쪽다리의 제4레이어에서 입력된 사이즈 그대로 인코더의 제4 레이어용 패치이미지를 생성하며,

상기 인코더의 제4레이어에서 상기 입력된 제3 레이어에 대한 맥스풀링의 결과값 및 상기 제4 레이어용 패치 이미지를 연결(concatenate)시켜 제1컨볼루션(convolution)을 수행하며, 상기 제1컨볼루션결과값과 상기 제4레이어용 패치이미지를 연결(concatenate)시켜 제2컨볼루션을 수행하고, 상기 제2컨볼루션의 결과값에 대해 맥스 풀링(max pooling)을 수행하여 상기 맥스 풀링 결과값에 기반한 특징 맵을 상기 IM구조의 입력으로 제공하는 것을 특징으로 하는 의료영상 분할 장치.

청구항 19

제 18 항에 있어서, 상기 영상 분할부는

상기 인코더의 각 레이어에서 출력되는 특징 맵을 해당 레이어에 위치한 제1어텐션모듈의 입력으로 제공하고, 차하위 레이어에서 출력되는 특징 맵과 상기 제1어텐션모듈의 해당레이어에서 출력되는 특징 맵을 연결시켜 상기 내부디코더의 입력으로 제공하며, 상기 내부디코더의 해당 레이어에 대한 컨볼루션(convolution)을 수행하고, 상기 내부디코더의 해당 레이어에서 출력되는 특징 맵을 상기 내부인코더의 해당 레이어의 입력으로 제공하면서, 업샘플링하여 상기 내부디코더의 차상위 레이어에 입력으로 제공하는 것을 특징으로 하는 의료영상 분할 장치.

청구항 20

제 19 항에 있어서, 상기 영상 분할부는

상기 내부디코더의 각 레이어에서 출력된 특징 맵을 차상위 레이어에서 얻어진 특징 맵과 연결시켜 상기 내부인코더의 해당 레이어에 대한 입력으로 제공하고, 상기 내부인코더의 해당 레이어에 대한 컨볼루션(convolution)을 수행하며, 상기 내부인코더의 해당 레이어에 대한 컨볼루션 결과값에 대해 맥스 풀링(max pooling)을 수행하고, 상기 내부인코더의 해당 레이어에서 출력되는 특징 맵을 해당 레이어에 위치한 제2어텐션모듈의 입력으로 제공하는 것을 특징으로 하는 의료영상 분할 장치.

청구항 21

제 20 항에 있어서, 상기 영상 분할부는

차하위 레이어에서 출력되는 특징 맵과 상기 제2어텐션모듈의 해당레이어에서 출력되는 특징 맵을 연결시켜 상기 디코더의 입력으로 제공하고, 상기 디코더의 해당 레이어에 대한 컨볼루션(convolution)을 수행하며, 상기 디코더의 해당 레이어에서 출력되는 특징 맵을 오른쪽 다리의 해당 레이어의 입력으로 제공하면서, 업샘플링하여 상기 디코더의 차상위 레이어에 입력으로 제공하는 것을 특징으로 하는 의료영상 분할 장치.

청구항 22

제 19 항에 있어서, 상기 영상 분할부는

상기 디코더의 각 레이어에서 출력되는 특징 맵과 차하위 레이어에서 얻어진 특징 맵을 연결시켜 상기 오른쪽다리의 해당 레이어에 대한 컨볼루션(convolution)을 수행하고, 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 컨볼루션하는 것을 특징으로 하는 의료영상 분할 장치.

발명의 설명

기술 분야

- [0001] 본 발명은 의료영상에 대한 영상분할기법 관련기술로서, 보다 구체적으로는 기존에 존재하는 M-Net의 기본 구조에 어텐션 모듈(Attention module)을 추가한 MAM-net구조를 통해 보다 효과적으로 영상 분할(image segmentation)을 할 수 있는 MAM-Net 구조 기반 의료영상 분할 방법 및 장치에 관한 것이다.

배경 기술

- [0002] 영상 분할(Image segmentation)은 디지털 영상을 복수의 픽셀 또는 복셀 집합으로 나누는 과정을 의미한다. 분할의 목적은 영상의 표현을 좀 더 의미있고 해석하기 쉬운 것으로 단순화하거나 변환하는 것이다. 영상 분할은 영상에서 물체와 경계를 찾는 데 사용된다. 분할의 결과는 전체 영상을 집합적으로 포함하는 영역의 집합이거나, 영상으로부터 추출된 윤곽의 집합이다.
- [0003] 인공지능이 급격히 발전함에 따라 여러 분야에서 딥러닝 기술이 적용되고 있다. 특히 의료분야에 있어 딥러닝 기술은 임상 및 진료에 중요한 부분으로 차지하고 있는데, 의료 영상 분할에도 인공지능 기술이 적용되고 있다.
- [0004] 의료영상 분할 분야에 사용되는 기존 아키텍처 중 M-Net 모델의 구조는 Left Leg, Right Leg, Encoding, Decoding으로 이루어져 있고 스킵 연결(skip connection)이 있는 계단식 컨볼루션 필터(cascade convolution filters)가 있으며 이는 네트워크가 더 유용한 기능에 집중할 수 있도록 도움을 준다.
- [0005] 여기서, Skip-Connection은 하나의 layer의 출력을 몇 개의 layer를 건너뛰고 다음 layer의 입력에 추가하는 방식으로서, 레이어가 깊어질수록 발생하는 기울기 소실 문제를 방지하는 효과와 미시적 정보를 보존해 더 좋은 특징 맵을 추출하는 효과가 있다. 그러나 M-Net의 스킵 연결만으로는 고성능의 정확도를 갖는 영상 분할 성능을 달성하는데 한계가 있으며, 복잡한 영역에서 분할 성능이 감소한다는 단점이 존재한다.
- [0006] 따라서, 딥러닝 네트워크 중 M-Net을 이용하여 의료 영상 분할 성능을 높이는 의료영상 분할 방법의 개발이 필요하다.

선행기술문헌

특허문헌

- [0007] (특허문헌 0001) 국내특허공개번호 제10-2022-0012991호

발명의 내용

해결하려는 과제

- [0008] 따라서, 본 발명의 목적은 Original M-Net의 복잡한 영역에서 분할 성능이 감소한다는 단점을 해결하기 위해, M-Net의 기본 구조에서 계층(layer)을 추가하고 인코딩 및 디코딩 레이어 사이의 Skip-Connection에 어텐션 모듈(Attention module)을 포함하는 IM구조를 사용하여 정확도가 크게 향상될 수 있는 MAM-Net(M and Attention M-Net) 구조 기반 의료영상 분할 방법 및 장치를 제공하는 것이다.
- [0009] 본 발명의 목적들은 이상에서 언급한 목적들로 제한되지 않으며, 언급되지 않은 또 다른 목적들은 아래의 기재로부터 당업자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

- [0010] 상술된 본 발명의 목적을 달성하기 위해, 본 발명은 의료영상 이미지를 입력받는 단계; 상기 입력된 의료영상 이미지를 기학습된 학습 모델에 적용하여 영상 분할을 수행하는 단계; 및 상기 영상 분할이 수행된 의료영상 이미지를 제공하는 단계;를 포함하며, 상기 학습 모델이 컨볼루션(Convolution)과 맥스풀링(Maxpooling)을 수행하는 인코더; 및 업-샘플링(Up-sampling)과 컨볼루션(Convolution)을 수행하는 디코더를 포함하는 M-Net 모델로서, 상기 인코더와 상기 디코더를 연결하는 스킵연결 사이에 2개의 어텐션 모듈과 상기 2개의 어텐션 모듈 사이에 내부디코더 및 내부인코더가 형성된 IM(Inter Mediate)구조가 결합된 아키텍처를 갖는 의료영상 영상

분할 방법을 제공한다.

- [0011] 바람직한 실시예에 있어서, 상기 IM구조는 상기 인코더의 출력이 입력되는 제1어텐션모듈, 상기 제1어텐션모듈의 출력과 하위 레이어에서 얻어진 특징 맵을 연결(Concatenate)시켜 컨볼루션 및 업샘플링하는 내부디코더, 상기 내부디코더에서 출력된 특징 맵과 상위 레이어에서 얻어진 특징 맵을 연결시켜 컨볼루션 및 맥스풀링하는 내부인코더 및 상기 내부인코더의 출력이 입력되는 제2어텐션모듈이 상기 인코더 및 상기 디코더를 구성하는 각 레이어에 대응되게 형성된다.
- [0012] 바람직한 실시예에 있어서, 상기 영상 분할을 수행하는 단계는 상기 의료영상 이미지를 상기 인코더로 입력하는 제1단계; 상기 인코더에서 출력된 특징 맵을 IM구조로 입력하는 제2단계; 및 상기 IM구조에서 출력된 특징 맵을 상기 디코더로 입력하는 제3단계; 및 상기 디코더에서 출력된 결과값을 분류하는 제4단계;를 포함한다.
- [0013] 바람직한 실시예에 있어서, 상기 제1단계는 왼쪽다리의 제1레이어에서 수행되는 상기 의료영상이미지에 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 입력된 상기 제1레이어용 패치 이미지에 대해 제1컨볼루션(convolution)을 수행하는 단계; 상기 제1컨볼루션결과값과 상기 제1레이어용 패치이미지를 연결(concatenate)시켜 제2컨볼루션을 수행하는 단계; 상기 제2컨볼루션 결과값에 대해 맥스 풀링(max pooling)이 수행되는 단계; 및 상기 맥스풀링 결과값을 상기 인코더의 제2 레이어의 입력으로 제공하는 단계;를 포함한다.
- [0014] 바람직한 실시예에 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 제1컨볼루션(convolution)을 수행하는 단계; 상기 제1컨볼루션결과값과 상기 제2레이어용 패치이미지를 연결(concatenate)시켜 제2컨볼루션을 수행하는 단계; 상기 제2컨볼루션의 결과값에 대해 맥스 풀링(max pooling)이 수행되는 단계; 및 상기 맥스풀링 결과값을 상기 인코더의 제3 레이어의 입력으로 제공하는 단계;를 포함한다.
- [0015] 바람직한 실시예에 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 제1컨볼루션(convolution)을 수행하는 단계; 상기 제1컨볼루션결과값과 상기 제3레이어용 패치이미지를 연결(concatenate)시켜 제2컨볼루션을 수행하는 단계; 상기 제2컨볼루션의 결과값에 대해 맥스 풀링(max pooling)이 수행되는 단계; 및 상기 맥스풀링 결과값을 상기 인코더의 제4 레이어의 입력으로 제공하는 단계;를 포함한다.
- [0016] 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 제3레이어에 대한 맥스풀링의 결과값 및 상기 제4레이어용 패치 이미지를 연결(concatenate)시켜 제1컨볼루션(convolution)을 수행하는 단계; 상기 제1컨볼루션결과값과 상기 제4레이어용 패치이미지를 연결(concatenate)시켜 제2컨볼루션을 수행하는 단계; 상기 제2컨볼루션의 결과값에 대해 맥스 풀링(max pooling)수행되는 단계; 및 상기 맥스 풀링 결과값에 기반한 특징 맵을 상기 IM구조의 입력으로 제공하는 단계;를 포함한다.
- [0017] 바람직한 실시예에 있어서, 상기 제2단계는 상기 인코더의 각 레이어에서 출력되는 특징 맵이 해당 레이어에 위치한 제1어텐션모듈의 입력으로 제공되는 단계; 차하위 레이어에서 출력되는 특징 맵과 상기 제1어텐션모듈의 해당레이어에서 출력되는 특징 맵을 연결시켜 상기 내부디코더의 입력으로 제공하는 단계; 상기 내부디코더의 해당 레이어에 대한 컨볼루션(convolution)을 수행하는 단계; 및 상기 내부디코더의 해당 레이어에서 출력되는 특징 맵을 상기 내부인코더의 해당 레이어의 입력으로 제공하면서, 업샘플링하여 상기 내부디코더의 차상위 레이어에 입력으로 제공하는 단계;를 포함한다.
- [0018] 바람직한 실시예에 있어서, 상기 제2단계는 상기 내부디코더의 각 레이어에서 출력된 특징 맵이 차상위 레이어에서 얻어진 특징 맵과 연결되어 상기 내부인코더의 해당 레이어에 대한 입력으로 제공되는 단계; 상기 내부인코더의 해당 레이어에 대한 컨볼루션(convolution)을 수행하는 단계; 상기 내부인코더의 해당 레이어에 대한 컨볼루션 결과값에 대해 맥스 풀링(max pooling)이 수행되는 단계; 및 상기 내부인코더의 해당 레이어에서 출력되는 특징 맵이 해당 레이어에 위치한 제2어텐션모듈의 입력으로 제공되는 단계;를 포함한다.

- [0019] 바람직한 실시예에 있어서, 상기 제3단계는 차하위 레이어에서 출력되는 특징 맵과 상기 제2어텐션모듈의 해당 레이어에서 출력되는 특징 맵을 연결시켜 상기 디코더의 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.
- [0020] 바람직한 실시예에 있어서, 상기 제3단계는 상기 디코더의 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 컨볼루션하는 단계;를 포함한다.
- [0021] 또한, 본 발명은 의료영상 이미지를 입력받는 입력부; 상기 입력된 의료영상 이미지를 학습된 학습 모델에 적용하여 시맨틱 분할을 수행하는 영상 분할부; 및 상기 시맨틱 분할이 수행된 의료영상 이미지를 제공하는 영상 제공부;를 포함하며, 상기 학습 모델은 컨볼루션(Convolution)과 맥스풀링(Maxpooling)을 수행하는 인코더; 및 업-샘플링(Up-sampling)과 컨볼루션(Convolution)을 수행하는 디코더를 포함하는 M-Net 모델로서, 상기 인코더와 상기 디코더를 연결하는 스킵연결 사이에 2개의 어텐션 모듈과 상기 2개의 어텐션 모듈 사이에 내부디코더 및 내부인코더가 형성된 IM(Inter Mediate)구조가 결합된 아키텍처를 갖는 의료영상 영상 분할 장치를 제공한다.
- [0022] 바람직한 실시예에 있어서, 상기 학습 모델에서 상기 IM구조는 상기 인코더의 출력이 입력되는 제1어텐션모듈, 상기 제1어텐션모듈의 출력과 하위 레이어에서 얻어진 특징 맵을 연결(Concatenate)시켜 컨볼루션 및 업샘플링하는 내부디코더, 상기 내부디코더에서 출력된 특징 맵과 상위 레이어에서 얻어진 특징 맵을 연결시켜 컨볼루션 및 맥스풀링하는 내부인코더 및 상기 내부인코더의 출력이 입력되는 제2어텐션모듈이 상기 인코더 및 상기 디코더를 구성하는 각 레이어에 대응되게 형성된다.
- [0023] 바람직한 실시예에 있어서, 상기 영상 분할부는 상기 의료영상 이미지를 상기 인코더로 입력하고, 상기 인코더에서 출력된 특징 맵을 IM구조로 입력하고, 상기 IM구조에서 출력된 특징 맵을 상기 디코더로 입력하며, 상기 디코더에서 출력된 결과값을 분류한다.
- [0024] 바람직한 실시예에 있어서, 상기 영상 분할부는 왼쪽다리의 제1레이어에서 상기 의료영상이미지에 대해 상기 인코더의 제1 레이어용 패치 이미지를 생성하고, 상기 제1레이어용 패치이미지에 대해 맥스풀링을 수행하여 그 결과값을 제2레이어의 입력으로 제공하고, 상기 인코더의 제1레이어에서 입력된 상기 제1 레이어용 패치 이미지에 대해 제1컨볼루션(convolution)을 수행하고, 상기 제1컨볼루션결과값과 상기 제1레이어용 패치이미지를 연결(concatenate)시켜 제2컨볼루션을 수행하며, 상기 제2컨볼루션 결과값에 대해 맥스 풀링(max pooling)을 수행하고, 상기 맥스풀링 결과값을 상기 인코더의 제2 레이어의 입력으로 제공한다.
- [0025] 바람직한 실시예에 있어서, 상기 영상 분할부는 상기 왼쪽다리의 제2레이어에서 입력된 사이즈 그대로 인코더의 제2 레이어용 패치이미지를 생성하고, 상기 제2레이어용 패치이미지에 대해 맥스풀링을 수행하여 그 결과값을 제3레이어의 입력으로 제공하며, 상기 인코더의 제2레이어에서 상기 입력된 제1 레이어에 대한 맥스풀링의 결과값 및 상기 제2레이어용 패치 이미지를 연결(concatenate)시켜 제1컨볼루션(convolution)을 수행하며, 상기 제1 컨볼루션결과값과 상기 제2레이어용 패치이미지를 연결(concatenate)시켜 제2컨볼루션을 수행하고, 상기 제2컨볼루션의 결과값에 대해 맥스 풀링(max pooling)을 수행하여 그 결과값을 상기 인코더의 제3 레이어의 입력으로 제공한다.
- [0026] 바람직한 실시예에 있어서, 상기 영상분할부는 상기 왼쪽다리의 제3레이어에서 입력된 사이즈 그대로 인코더의 제3레이어용 패치이미지를 생성하고, 상기 제3레이어용 패치이미지에 대해 맥스풀링을 수행하여 그 결과값을 제4레이어의 입력으로 제공하며, 상기 인코더의 제3레이어에서 상기 입력된 제2 레이어에 대한 맥스풀링의 결과값 및 상기 제3레이어용 패치 이미지를 연결(concatenate)시켜 제1컨볼루션(convolution)을 수행하며, 상기 제1컨볼루션결과값과 상기 제3레이어용 패치이미지를 연결(concatenate)시켜 제2컨볼루션을 수행하고, 상기 제2컨볼루션의 결과값에 대해 맥스 풀링(max pooling)을 수행하여 그 결과값을 상기 인코더의 제4 레이어의 입력으로 제공한다.
- [0027] 바람직한 실시예에 있어서, 상기 영상 분할부는 상기 왼쪽다리의 제4레이어에서 입력된 사이즈 그대로 인코더의 제4 레이어용 패치이미지를 생성하며, 상기 인코더의 제4레이어에서 상기 입력된 제3 레이어에 대한 맥스풀링의

결과값 및 상기 제4 레이어용 패치 이미지를 연결(concatenate)시켜 제1컨볼루션(convolution)을 수행하며, 상기 제1컨볼루션결과값과 상기 제4레이어용 패치이미지를 연결(concatenate)시켜 제2컨볼루션을 수행하고, 상기 제2컨볼루션의 결과값에 대해 맥스 풀링(max pooling)을 수행하여 상기 맥스 풀링 결과값에 기반한 특징 맵을 상기 IM구조의 입력으로 제공한다.

[0028] 바람직한 실시예에 있어서, 상기 영상 분할부는 상기 인코더의 각 레이어에서 출력되는 특징 맵을 해당 레이어에 위치한 제1어텐션모듈의 입력으로 제공하고, 차하위 레이어에서 출력되는 특징 맵과 상기 제1어텐션모듈의 해당레이어에서 출력되는 특징 맵을 연결시켜 상기 내부디코더의 입력으로 제공하며, 상기 내부디코더의 해당 레이어에 대한 컨볼루션(convolution)을 수행하고, 상기 내부디코더의 해당 레이어에서 출력되는 특징 맵을 상기 내부인코더의 해당 레이어의 입력으로 제공하면서, 업샘플링하여 상기 내부디코더의 차상위 레이어에 입력으로 제공한다.

[0029] 바람직한 실시예에 있어서, 상기 영상 분할부는 상기 내부디코더의 각 레이어에서 출력된 특징 맵을 차상위 레이어에서 얻어진 특징 맵과 연결시켜 상기 내부인코더의 해당 레이어에 대한 입력으로 제공하고, 상기 내부인코더의 해당 레이어에 대한 컨볼루션(convolution)을 수행하며, 상기 내부인코더의 해당 레이어에 대한 컨볼루션 결과값에 대해 맥스 풀링(max pooling)을 수행하고, 상기 내부인코더의 해당 레이어에서 출력되는 특징 맵을 해당 레이어에 위치한 제2어텐션모듈의 입력으로 제공한다.

[0030] 바람직한 실시예에 있어서, 상기 영상 분할부는 차하위 레이어에서 출력되는 특징 맵과 상기 제2어텐션모듈의 해당레이어에서 출력되는 특징 맵을 연결시켜 상기 디코더의 입력으로 제공하고, 상기 디코더의 해당 레이어에 대한 컨볼루션(convolution)을 수행하며, 상기 디코더의 해당 레이어에서 출력되는 특징 맵을 오른쪽 다리의 해당 레이어의 입력으로 제공하면서, 업샘플링하여 상기 디코더의 차상위 레이어에 입력으로 제공한다.

[0031] 바람직한 실시예에 있어서, 상기 영상 분할부는 상기 디코더의 각 레이어에서 출력되는 특징 맵과 차하위 레이어에서 얻어진 특징 맵을 연결시켜 상기 오른쪽다리의 해당 레이어에 대한 컨볼루션(convolution)을 수행하고, 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 컨볼루션하는 한다.

발명의 효과

[0032] 상술된 본 발명의 M-Net을 이용한 의료영상 분할 방법 및 장치에 의하면, Original M-Net의 기본 구조에서 계층(layer)을 추가하고 인코딩 및 디코딩 레이어 사이의 Skip-Connection에 어텐션 모듈(Attention module)을 포함하는 IM구조를 추가함으로써 Original M-Net의 복잡한 영역에서 분할 성능이 감소한다는 단점을 해결할 수 있다.

[0033] 또한, 본 발명에 의하면 달성해야 할 목적과의 연관성에 따라 추출한 특징에 가중치를 부여함으로써 네트워크 스스로 집중해야 할 영역을 선별하도록 하여 의료 영상을 분할하였고 제안하는 모델의 성능이 기존 Original M-Net에 비해 향상된 것을 확인할 수 있다.

[0034] 또한, 본 발명에 의하면 치아 영상 및 피부 영상 등의 의료영상을 입력으로 하여, Se block - Attention을 활용함으로써 건강 상태를 진단하고 그를 위한 환부 영역의 자동 분할(segmentation)을 효과적으로 할수있는 방법을 제시하며, Multiscale Attention을 활용하지 않았을 때보다 정확도(accuracy), 정밀도(precision) 등의 성능이 향상되는 효과가 있으며, 단순한 이진 분류가 아닌 다중 분류를 사용하여 입력 영상의 세밀한 분할(segmentation)이 가능하기 때문에 훨씬 정확하고 정밀한 진단이 가능하다.

[0035] 본 발명의 효과는 이상에서 언급한 효과로 제한되지 않으며, 언급되지 않은 또 다른 효과들은 아래의 기재로부터 당업자에게 명확하게 이해될 수 있을 것이다.

도면의 간단한 설명

[0036] 도 1a는 Original M-Net의 네트워크 구조도이고, 도 1b는 도 1a에 도시된 Original M-Net의 상세 구조도이며, 도 1c는 도 1b에 도시된 Left Leg의 구조도이고, 도 1d는 도 1b에 도시된 Encoding 구조도이며, 도 1e는 도 1b에 도시된 Decoding 구조도이며, 도 1f는 도 1b에 도시된 Right Leg 구조도이다.

도 2a는 Original M-Net에 추가된 IM구조 중 Skip-Connection Attention만 도시된 MAM-net(M and Attention M-Net) 네트워크를 도시한 것이고, 도 2b는 Original M-Net에 추가된 IM구조가 모두 도시된 MAM-net(M and

Attention M-Net) 네트워크에 대한 상세 구조도이며, 도 2c는 도 2b에 도시된 Left Leg + Encoding 구조를 도시한 것이며, 도 2d는 도 2b에 도시된 IM(Inter Mediate) 구조를 도시한 것이고, 도 2e는 도 2b에 도시된 Decoding + Right Leg 구조를 도시한 것이며, 도 2f는 도 2a 및 도 2b에 도시된 구조에서 사용되는 Attention 모듈의 일 예를 도시한 것이다.

도 3a는 Original M-Net을 이용한 치아 영상 분할 결과를 도시한 것이고, 도 3b는 본 발명의 MAM-net(M and Attention M-Net)을 이용한 치아 영상 분할 결과를 도시한 것이다.

도 4a는 Original M-Net을 이용한 스킨 영상 분할 결과를 도시한 것이고, 도 4b는 본 발명의 MAM-net(M and Attention M-Net)을 이용한 스킨 영상 분할 결과를 도시한 것이다.

도 5a는 본 발명의 실시예에 따른 의료영상 분할 장치의 구성도이고, 도 5b는 본 발명의 실시예에 따른 의료영상 분할 장치를 이용한 의료영상 분할 방법의 순서도이다.

발명을 실시하기 위한 구체적인 내용

- [0037] 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.
- [0038] 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.
- [0039] 구성 요소를 해석함에 있어서, 별도의 명시적 기재가 없더라도 오차 범위를 포함하는 것으로 해석한다.
- [0040] 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.
- [0041] 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 인접하는"과 "~에 직접 인접하는" 등도 마찬가지로 해석되어야 한다.
- [0042] 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.
- [0043] 이하, 첨부한 도면 및 바람직한 실시예들을 참조하여 본 발명의 기술적 구성을 상세하게 설명한다.
- [0044] 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.
- [0045] 본 발명은 M-Net의 기본 구조에서 계층을 추가하고 어텐션 모듈(Attention module)을 포함하는 IM구조를 사용함으로써 M-Net의 성능을 향상시킨 의료 영상 분할 기술에 관한 것이다.
- [0046] 따라서, 본 발명의 MAM-net(M and Attention M-Net) 네트워크 구조를 살펴보면, Original M-net의 인코딩 및 디코딩 부분과 attention M-Net을 조합하여 성능을 개선시키기 위해 만든 네트워크 구조로서, 기존 M-Net의 구조에서 영상에서 집중해야 하는 영역을 파악하는 기능을 하는 self attention block을 스킵 연결(skip connection)에 추가하여 레이어를 깊이 있게 만들어 성능을 높이하고자 하였다.
- [0047] 공지된 바와 같이, M-Net을 이용한 의료영상 분할 방법에서 사용되는 의료영상 분할을 위한 딥 콘볼루션 신경망은 하나 이상의 콘볼루션 계층(Convolution layer)을 사용할 수 있고, 영상을 분할 하기 위한 출력 정보를 단순화하여 특징 맵을 받아오기 위해 풀링층(pooling layer)를 사용할 수 있다. 일 구현예로서, 의료 영상 분할을 위해 M-Net구조를 기반으로 훈련하는 하나 이상의 콘볼루션 계층(Convolution layer)으로 구성된 딥 콘볼루션

신경망일 수 있다.

- [0048] 도 1a 내지 도 1f를 참조하여 Original M-Net에 대해 살펴보면, 도1a 및 도 1b에 도시된 바와 같이 M-Net 모델의 구조에는 skip connection이 있는 계단식 컨볼루션 필터(cascade convolution filters)가 있다. 각 구조영역의 일 구현예를 살펴보면 도 1c에 도시된 왼쪽다리(Left Leg)은 kernel 크기가 2x2인 max-pooling 작업으로 구성되고, 각각의 최대 풀 출력(max pool output)은 해당 인코딩 레이어에 입력으로 공급된다. 도 1d에 도시된 인코더(Encoder)에서 각각의 인코딩 경로는 ReLU 활성화 기능이 있는 컨볼루션 kernel 3 x 3 레이어로 구성되고, 필터 크기가 2 x 2 크기인 최대 풀링 작업(max-pooling)으로 구성될 수 있다. 도 1e에 도시된 디코더(Decoder)에서 디코딩 레이어는 ReLU 활성화 기능이 있는 3 x 3 크기의 컨볼루션 필터로 구성되며, 입력 크기를 두 배로 하는 업 샘플링 연산은 디코더 출력에서 원래 입력 이미지를 재구성하는 데 사용될 수 있다. 도 1f에 도시된 오른쪽 다리(Right Leg)에서 오른쪽 다리(right leg)의 출력과 디코딩 계층이 1x1 크기로 추가 처리되고, Up-sampling마다 채널의 수가 절반으로 줄어든다. 하지만 M-Net의 스킵 연결만으로는 고성능의 정확도를 갖는 영상 분할 성능을 달성하는 데 한계가 있으며, 복잡한 영역에서 분할 성능이 감소한다는 단점이 존재하기 때문에 본 발명은 M-Net의 기본구조에 계층을 추가하고 인코더와 디코더 사이를 연결하는 스킵연결 사이에 어텐션 모듈(Attention module)을 포함하는 IM구조를 추가하여 네트워크가 더 유용한 기능에 더욱더 집중할 수 있도록 변형된 M-Net 구조를 통해 기존 M-Net 보다 분할성능이 향상된 새로운 네트워크구조를 개발하였다.
- [0049] 도 2a는 Original M-Net에 추가된 IM구조 중 Skip-Connection Attention만 도시된 MAM-net(M and Attention M-Net) 네트워크를 도시한 것이고, 도 2b는 Original M-Net에 추가된 IM구조가 모두 도시된 MAM-net(M and Attention M-Net) 네트워크에 대한 상세 구조도이며, 도 2c는 도 2b에 도시된 Left Leg + Encoding 구조를 도시한 것이며, 도 2d는 도 2b에 도시된 IM(Inter Mediate) 구조를 도시한 것이고, 도 2e는 도 2b에 도시된 Decoding + Right Leg 구조를 도시한 것이며, 도 2f는 도 2a 및 도 2b에 도시된 구조에서 사용되는 Attention 모듈의 일 예를 도시한 것이다. 도 5a는 본 발명의 실시예에 따른 의료영상 분할 장치의 구성도이고, 도 5b는 본 발명의 실시예에 따른 의료영상 분할 장치를 이용한 의료영상 분할 방법의 순서도이다.
- [0050] 도 2a 내지 도 2f, 도 5a 및 도 5b를 참조하여 본 발명의 실시예에 따른 의료영상분할방법 및 의료영상분할장치의 구성을 보다 구체적으로 살펴본다.
- [0051] 도 5a는 본 발명의 실시예에 따른 의료영상분할 장치의 구성도이다. 도 5a를 참조하면, 본 발명의 실시예에 따른 의료영상분할 장치(100)는 영상 입력부(110), 영상 분할부(120) 및 영상 제공부(130)를 포함한다. 영상 입력부(110)는 도 5b에 도시된 바와 같이 의료영상을 입력받는다(S210). 여기서, 의료영상은 X-ray치아 이미지 또는 피부병변이미지를 포함하는 다양한 의료영상이미지일 수 있다. 다음으로, 영상 분할부(120)는 입력받은 의료영상이미지를 기학습된 학습 모델에 적용하여 시맨틱 분할(Semantic segmentation)을 수행한다. 본 발명의 일 실시예에 따른 의료영상 분할방법에서 사용되는 학습 모델 즉 딥러닝 네트워크는 도 2b에 도시된 바와 같이 도 1a에 도시된 기존 M-Net 모델에서 영상에서 집중해야 하는 영역을 파악하는 기능을 하는 self attention block을 스킵 연결(skip connection)에 추가하여 레이어를 깊이 있게 만들어 성능을 높이고자 인코더와 디코더를 연결하는 스킵연결 사이에 2개의 어텐션 모듈과 상기 2개의 어텐션 모듈 사이에 내부디코더 및 내부인코더가 순차적으로 형성된 IM(Inter Mediate)구조가 추가되어 구성된 MAM-Net(M and Attention M-Net) 모델이다.
- [0052] 도 2a 및 도 2b에 도시된 바와 같이 Original M-Net의 인코딩 및 디코딩 레이어 사이의 IM(Inter Mediate)구조가 추가된 스킵 연결(skip connection)은 네트워크가 후처리 작업 없이 이미지에 대해 정확한 라벨링(labeling)을 달성하기 위해 필요한 정보를 캡처하는 데 도움이 된다. 상술된 바와 같이 M-Net의 스킵 연결(skip connection)만으로는 고성능의 정확도를 갖는 영상 분할 성능을 달성하는 데 한계가 있어 복잡한 영역에서 분할 성능이 감소하므로, IM(Inter Mediate)구조 특히 어텐션 모듈을 인코딩 및 디코딩 레이어 사이의 스킵 연결에 추가하여 달성해야 할 목적과의 연관성에 따라 추출한 특징에 가중치를 부여함으로써 네트워크 스스로 집중해야 할 영역을 선별하도록 하여 의료 영상을 분할하였다.
- [0053] 본 발명에서 의료영상이미지 분할 과정은 인코더-디코더 기반 구조로 되어 있으며, I 자형 왼쪽다리(left leg)와 오른쪽 다리(right leg)와 같은 두 가지 사이드 경로를 포함하고, 인코더와 디코더의 스킵연결 사이에 IM구조가 포함되어 있으므로, 의료영상이미지 분할 과정(S220)은 의료영상 이미지를 인코더로 입력하는 제1단계; 인코더에서 출력된 특징 맵을 IM구조로 입력하는 제2단계; IM구조에서 출력된 특징 맵을 상기 디코더로 입력하는 제3단계; 및 디코더에서 출력된 결과값을 분류하는 제4단계;를 포함할 수 있다. 이하에서 사용되는 인코더, 제1어텐션모듈, 내부디코더, 내부인코더, 제2어텐션모듈, 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.

- [0054] 본 발명의 왼쪽다리 및 인코더의 동작을 포함하는 다운샘플링 과정 즉 제1단계를 도 2c를 참조하여 살펴본다.
- [0055] 본 발명에 따르면, 초기 입력 이미지는 일반적으로 높이 x 너비 x 채널($H \times W \times C$)의 형태로 되어 있는데 도시된 바와 같이, $256 \times 256 \times 3$ 크기의 아무 분할이 되지 않은 의료영상 이미지가 입력으로 주어진다. 여기서, 의료영상 이미지는 치아 x-ray 이미지 또는 피부병변이미지일 수 있는데, 치아 x-ray 이미지는 비공개적인 데이터셋이고, 피부 데이터셋은 공개적으로 사용할 수 있는 ISIC 2018 Challenge 데이터셋을 사용하였고 흑색종 피부를 분할하는 데 사용하였다.
- [0056] 먼저, 왼쪽다리(left leg)는 도 2c에 도시된 바와 같이 kernel 크기가 2×2 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 input 이미지를 maxpooling 하는 것인데 MaxPooling은 특정 영역에서 최대값을 선택하여 해상도를 줄이는 연산이다.
- [0057] 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.
- [0058] 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.
- [0059] 동일한 과정을 수행하여 왼쪽다리의 최하단 레이어(도시된 구조에는 제4레이어)에서 최종레이어용 패치이미지를 생성하여 인코더의 최하단 레이어에 입력으로 제공할 수 있다.
- [0060] 다음으로, 인코더의 동작을 살펴보면, 도 2c에 도시된 바와 같이, 인코딩 경로에서 인코더의 각 레이어는 3×3 컨볼루션, BatchNormalization, ReLU(rectified linear unit)가 순차적으로 일어나는 제1컨볼루션블록과 3×3 컨볼루션, BatchNormalization, ReLU(rectified linear unit)가 순차적으로 일어나는 제2컨볼루션블록이 연속되어 이루어진 2개의 컨볼루션 블록으로 구성되는데, 제1컨볼루션블록은 ReLU 후에 Dropout되어 정규화(regularization)되는 과정을 더 포함한다. 제2컨볼루션블록 이후에 맥스 풀링(2×2)으로 영상의 사이즈를 줄이는 과정이 수행된다.
- [0061] 보다 구체적으로 살펴보면, 인코더의 제1레이어에 입력된 제1레이어용 패치이미지에 대해 제1컨볼루션블록에서 3×3 컨볼루션, BatchNormalization 및 ReLU(rectified linear unit)가 순차적으로 적용되어, 입력된 이미지를 컨볼루션(3×3) 시켜 특징을 추출하고, 배치 정규화를 통해 변형된 분포가 나오지 않도록 정규화 하며, ReLU 함수를 이용해 이진화 해준다. 그 후 Dropout되어 제1컨볼루션의 결과값이 얻어진다. 제2컨볼루션블록에서 제1컨볼루션의 결과값과 왼쪽다리의 제1레이어용 패치이미지를 연결(concatenate)한 값에 대해 3×3 컨볼루션, BatchNormalization 및 ReLU(rectified linear unit)가 순차적으로 적용되어 제2컨볼루션 결과값이 얻어진다.
- [0062] 인코더의 제1레이어에서 얻어진 제2컨볼루션 블록의 출력 즉 제2컨볼루션 결과값은 스트라이드(stride)가 2인 2×2 맥스 풀링 연산을 사용하여 다운 샘플링된다. 풀링은 입력 데이터의 공간 차원을 줄이는 연산으로, 2×2 Max 풀링은 각 2×2 영역에서 최대값을 추출하는 방식으로 동작할 수 있다.
- [0063] 상술된 인코더의 제1레이어에서 수행되는 연산을 맥스풀링 결과 값이 IM구조의 입력으로 제공될 때까지 동일하게 반복수행한다.
- [0064] 다음으로, 본 발명의 MAM-net(M and Attention M-Net) 네트워크 중 IM구조에서 일어나는 과정 즉 제2단계를 도 2d를 참조하여 살펴본다.
- [0065] IM(Inter Mediate)구조는 인코더의 출력이 입력되는 제1어텐션모듈, 제1어텐션모듈의 출력과 하위 레이어에서 얻어진 특징 맵을 연결(Concatenate)시켜 컨볼루션 및 업샘플링하는 내부디코더, 내부디코더에서 출력된 특징 맵과 상위 레이어에서 얻어진 특징 맵을 연결시켜 컨볼루션 및 맥스풀링하는 내부인코더 및 내부인코더의 출력이 입력되는 제2어텐션모듈이 인코더와 디코더를 구성하는 각 레이어에 대응되게 형성된다.
- [0066] 도 2d에 도시된 바와 같이 IM(Inter Mediate)구조는 concatenation 연산으로 인코더의 중간 특징 정보를 디코더에서 활용하여 정보 손실을 최소화하는데, Up-sampling을 사용하여 불규칙적으로 샘플링 된 측정 데이터를 조밀

하게 나누어 규칙적인 형태의 데이터로 변환하였고 변환한 값을 Max Pooling된 값과 concatenate 하여 convolution 하도록 구성될 수 있다.

[0067] 보다 구체적으로 살펴보면, 도시된 바와 같이 인코더의 오른쪽에 형성된 제1어텐션모듈과 내부디코더 사이에서 다음과 같은 과정이 이루어진다. 즉 인코더의 각 레이어에서 출력되는 특징 맵이 해당 레이어에 위치한 제1어텐션모듈의 입력으로 제공되고, 차하위 레이어에서 출력되는 특징 맵과 상기 제1어텐션모듈의 해당레이어에서 출력되는 특징 맵을 연결시켜 상기 내부디코더의 입력으로 제공하며, 상기 내부디코더의 해당 레이어에 대한 컨볼루션(convolution)을 수행하고, 상기 내부디코더의 해당 레이어에서 출력되는 특징 맵을 상기 내부인코더의 해당 레이어의 입력으로 제공하면서, 업샘플링하여 상기 내부디코더의 차상위 레이어에 입력으로 제공하는데, 이 과정이 최상단 레이어까지 반복적으로 수행될 수 있다. 여기서, 내부디코더의 최하단 레이어에 입력으로 제공되는 차하위 레이어에서 출력되는 특징 맵은 인코더의 최하단 레이어에서 출력되는 맥스풀링 결과값을 컨볼루션하여 업샘플링한 것이다.

[0068] 내부디코더의 업 샘플링 과정에서는 convolution layer를 통해 입력 데이터를 확장하는 연산을 수행하며, 내부디코더에서는 이를 통해 고해상도의 특성 맵을 얻어내어 세분화된 정보를 복원한다. 따라서, 내부디코더에서 일어나는 컨볼루션은 입력 데이터를 확장하기 위해 일반적으로 3x3 크기의 커널을 사용하고, 스트라이드 2의 Convolution Transpose를 사용하여 공간적인 해상도를 두 배로 증가시키는 연산일 수 있다.

[0069] 또한, 순차적으로 형성된 내부디코더, 내부인코더 및 제2어텐션모듈사이에서는 다음과 같은 과정이 이루어질 수 있다.

[0070] 내부디코더의 각 레이어에서 출력된 특징 맵이 차상위 레이어에서 얻어진 특징 맵과 연결되어 상기 내부인코더의 해당 레이어에 대한 입력으로 제공되고, 상기 내부인코더의 해당 레이어에 대한 컨볼루션(convolution)을 수행하며, 상기 내부인코더의 해당 레이어에 대한 컨볼루션 결과값에 대해 맥스 풀링(max pooling)이 수행되고, 상기 내부인코더의 해당 레이어에서 출력되는 특징 맵이 해당 레이어에 위치한 제2어텐션모듈의 입력으로 제공될 수 있다. 이때, 내부인코더의 최상단 레이어에 제공되는 입력은 내부디코더의 최상단레이어에서 출력된 특징 맵일 수 있다.

[0071] 여기서, 내부인코더는 인코더와 동일하게 각 레이어가 3x3 컨볼루션, BatchNormalization, ReLU(rectified linear unit)가 순차적으로 일어나는 제1컨볼루션블록과 3x3 컨볼루션, BatchNormalization, ReLU(rectified linear unit)가 순차적으로 일어나는 제2컨볼루션블록이 연속되어 이루어진 2개의 컨볼루션 블록으로 구성되는데, 제1컨볼루션블록은 ReLU 후에 Dropout되어 정규화(regularization)되는 과정을 더 포함한다. 제2컨볼루션블록 이후에 맥스 풀링(2x2)으로 영상의 사이즈를 줄이는 과정이 수행된다. 따라서, 내부인코더의 해당 레이어에서 얻어진 컨볼루션 결과값은 제2컨볼루션 결과값을 의미한다. 따라서, 내부인코더의 해당 레이어에 대한 컨볼루션(convolution)은 상기 내부인코더의 해당 레이어에 대한 입력에 대해 제1컨볼루션을 수행한 결과값과 3x3 컨볼루션 결과값을 연결시켜 제2컨볼루션을 수행하는 것을 의미할 수 있다.

[0072] 또한, 제1어텐션 모듈 및 제2어텐션 모듈은 어텐션메카니즘이 일어나는 구성요소로서 어텐션 메카니즘은 '영상의 어느 영역에 집중하면 되는지 파악하는 능력'에 해당하는데, 구체적으로는 '달성해야 할 목적과의 연관성에 따라 추출한 특징에 가중치를 부여함으로써 네트워크 스스로 집중해야 할 영역을 선별하는 기술'이다. Squeeze-Excitation (SE) block은 CNN 기반 모델에 부착하여 사용하며 유연성을 지니고 있어 Residual 모델 또는 Inception 모델과 함께 사용할 수 있다. 또한 SE block은 Squeeze와 Excitation의 두 가지 과정으로 구성되어 있는데, Squeeze는 각 채널을 1차원으로 만드는 역할을, Excitation은 Squeeze에서 생성된 ($1 \times 1 \times C$) 벡터를 정규화하여 가중치를 부여하는 역할을 수행한다.

[0073] 도 2f에 도시된 Attention Module에서 SE-block은 Squeeze-and-Excitation(SE) block 채널 간의 상호의존을 명시적으로 모델링함으로써 channel-wise feature response를 적응하게 재보정해주는 단위로서, 네트워크에서 생성된 representation의 품질 향상을 목표로 하며 네트워크가 Feature recalibration을 수행할 수 있도록 해주는 메커니즘으로 유익한 feature를 선택적으로 강조하고, 덜 유용한 feature 억제한다.

[0074] 제1어텐션 모듈 및 제2어텐션 모듈에서 이루어지는 "Feature Recalibration"은 주로 입력 특징(feature)에 대한 가중치를 동적으로 조절하여 모델이 특정 위치나 채널에 더 집중하도록 하는 과정을 나타낸다. 이를 위해 일반적으로 사용되는 연산은 다음과 같다.

[0075] 1. Query, Key, Value 설정: 어텐션 메커니즘은 주로 Query(Q), Key(K), Value(V)의 세 가지 요소로 구성되고, 이는 주어진 입력에 대해 어떤 부분에 주목할지 결정하는 데 사용된다.

[0076] 2. 스케일링(Scaled) 닷 프로덕트 어텐션(Attention):

[0077] (1)Query와 Key의 유사도 계산:

[0078] 주로 닷 프로덕트를 사용하여 Query와 Key 간의 유사도를 계산한다.

[0079]
$$\text{Attention Score} = \frac{Q \cdot K}{\sqrt{d}}$$
 여기서 d는 Query 또는 Key의 차원이다.

[0080] (2) 소프트맥스(softmax) 함수 적용:

[0081] 계산된 유사도에 소프트맥스 함수를 적용하여 각 위치의 가중치를 얻는다. 이 가중치는 해당 위치에 대한 중요성을 나타낸다.

[0082]
$$\text{Attention Weight} = \text{softmax}(\text{Attention Score})$$

[0083] 3. Value와의 가중합(Weighted Sum): 얻어진 가중치를 사용하여 Value에 가중합을 수행한다. 이는 최종적으로 모델이 주어진 입력에서 중요한 부분에 집중하도록 한다.

[0084]
$$\text{Recalibrated Feature} = \text{Attention Weight} \cdot V$$

[0085] 이러한 어텐션 메커니즘을 사용하면 모델이 주어진 입력의 특정 부분에 더 집중하게 되어 성능이 향상될 수 있다.

[0086] 다음으로, 본 발명의 MAM-net(M and Attention M-Net) 네트워크 중 디코더 및 오른쪽다리에서 일어나는 과정 즉 IM구조에서 출력된 특징 맵을 디코더로 입력하는 제3단계를 도 2e를 참조하여 살펴본다.

[0087] 도 2e는 본 발명의 학습모델 MAM-net에 적용한 Decoder + Right Leg의 구조를 나타낸 것인데, 디코더 및 오른쪽 다리의 동작을 살펴보면, 상술된 내부디코더와 동일하게 convolution layer를 통해 입력 데이터를 확장하는 연산을 수행하는 업 샘플링이 이루어지므로, 디코더 및 오른쪽 다리에서는 이를 통해 고해상도의 특성 맵을 얻어 내어 세분화된 정보를 복원한다. 즉, 디코딩 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 있고, Right Leg의 경우 Right leg의 출력과 디코딩 계층이 1x1 크기로 추가 처리되고 Up-sampling마다 채널의 수가 절반으로 줄어든 수 있다.

[0088] 따라서, 차하위 레이어에서 출력되는 특징 맵과 상기 제2어텐션모델의 해당레이어에서 출력되는 특징 맵을 연결시켜 디코더의 입력으로 제공되면, 디코더에서는 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 과정을 수행하게 된다. 일 구현예로서 디코더에서 일어나는 컨볼루션은 하위 레이어에서 출력되는 특징 맵과 제2어텐션모델의 해당레이어에서 출력되는 특징 맵을 연결시켜Conv2DTranspose (2x2)를 통해 영상 사이즈를 늘리고, 인코더에서 연결한 값을 받아 컨볼루션 (3x3) 시켜주며, 배치정규화를 통해 변형된 분포가 나오지 않도록 정규화하여 수행될 수 있다.

[0089] 또한, 오른쪽 다리에서는 디코더의 각 레이어에서 출력되는 특징 맵과 차하위 레이어에서 얻어진 특징 맵을 연결시켜 상기 오른쪽다리의 해당 레이어에 대한 컨볼루션(convolution)을 수행하고, 오른쪽 다리의 해당 레이어에 대한 컨볼루션 결과값에 대해 업샘플링이 수행된 결과값이 차상위 레이어에 입력으로 제공되며, 오른쪽 다리의 2번째 레이어에 대한 컨볼루션 결과값에 대해 업샘플링이 수행된 결과값과 상기 디코더의 최상단 레이어에서 컨볼루션을 마친 특징 맵을 컨볼루션한 다음 마지막 컨볼루션을 수행하여 입력된 의료영상 이미지로부터 시맨틱 분할 이미지를 출력한다.

[0090] 디코더 및 오른쪽 다리에서 수행하는 컨볼루션은 입력 데이터를 확장하기 위해 일반적으로 3x3 크기의 커널을 사용하고, 스트라이드 2의 Convolution Transpose를 사용하여 공간적인 해상도를 두 배로 증가시키는데, 수식으로 표현하면 다음과 같다.

[0091]
$$Y = f(W^T * X + b)$$
 여기서, Y는 출력 특징 맵이고, X는 입력 특징 맵이며, W^T 는 Convolution Transpose 레이어의 커널(가중치)의 전치 행렬이고, b는 편향(bias)이며, f는 활성화 함수이다. 이렇게 convolution한 이미지를 다음과 같이 업샘플링의 연산과정을 거치게 된다.

[0092] 업샘플링은 일반적으로 Convolutional Transpose(또는 Fractionally Strided Convolution 또는 Deconvolution) 레이어 또는 특정 업샘플링 레이어를 사용한다.

[0093] Convolutional Transpose 레이어를 사용한 업샘플링 연산은 다음과 같이 정의된다.

[0094] $Y = f(X * W + b)$ 여기서, $X * W$ 는 입력 데이터에 대한 Convolutional Transpose 연산이고, $+$ 는 편향을 더하는 연산이며, f 는 활성화 함수이다.

[0095] Convolutional Transpose 연산은 입력 데이터를 확장하여 업샘플링된 결과를 얻는데, 고해상도 이미지를 생성하거나 세그멘테이션 맵을 확장하는데 활용된다.

[0096] 제4단계는 디코더에서 출력된 결과값을 분류하는 단계로서 분류 레이어에서 이루어질 수 있다. 분류레이어는 재구성된 이미지를 출력하는 시그모이드 함수가 있는 1x1 컨벌루션 레이어로 구성될 수 있으며, 보통 1x1 컨벌루션을 거친 후 시그모이드 함수를 적용하여 생성된다. 시그모이드 함수는 이진 분류 문제에 적합한 활성화 함수로, 출력을 0과 1 사이의 값으로 제한한다.

[0097] 영상 제공부(130)는 시맨틱 분할이 수행된 의료영상 이미지를 제공한다(S230).

[0098] 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.

[0099] 상술된 구성을 갖는 본 발명은 의료영상 분할을 위해 다양한 사전 훈련된 영상 분할(pre-trained segmentation) 네트워크를 이용하여 영상 분할(segmentation) 정보를 추출할 수 있고, 여러 형태의 집중 메커니즘(attention mechanism)을 구현하는 네트워크 형태를 통해 메인 네트워크가 그 정보를 활용할 수 있다.

[0100] 또한, 의료영상 분할을 위해 IM구조가 추가된 skip connection을 사용하여 네트워크의 단계를 거칠수록 사라지는 공간적인 정보들을 복구하는데 도움을 줄 수 있고, 학습 속도의 개선 및 layer가 이동될 때마다 이전 layer의 학습에 의해 가중치 값이 바뀌어 전달되는 입력 데이터의 분포와 현재 layer에서 학습했던 시점의 분포와 차이를 완화시키기 위해서 정규화(Batch Normalization)를 사용할 수 있다.

[0101] 실험예

[0102] 피부 데이터셋은 공개적으로 사용할 수 있는 ISIC 2018 Challenge 데이터셋 을 사용하였고 흑색종 피부를 분할하는데 사용하였다. 본 발명에서는 피부와 치아 데이터를 이용하여 M-Net의 복잡한 영역에서 분할 성능이 감소한다는 단점을 해결하기 위해, M-Net의 기본 구조의 Skip-Connection에 어텐션 모듈(Attention module)을 포함하는 IM구조를 추가하여 구성된 MAM-net의 성능이 향상됨을 확인하였다. 성능을 평가하기 위하여 분류성능평가 지표를 활용했다. 정확도(Accuracy)는 모델의 성능을 나타낼 수 있는 평가 지표이다. 정확도는 전체 예측 건에서 모델이 예측한 것이 정답인 경우의 비율로, 다음과 같은 수식으로 정의한다

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \quad (1)$$

[0103]

[0104] 치아영상에 대한 결과는 표 1, 도 3a 및 도 3b에, 피부영상에 대한 결과는 표 2, 도 4a 및 도 4b에 각각 나타내었다. 표 1 및 2에서 Original M-Net Skip-Connection Attention은 본 발명의 MAM-net 모델을 이용한 결과이다.

표 1

	Original M-Net	Original M-Net Skip-Connection Attention
Accuracy	0.93302	0.95486
F1	0.79584	0.87279
Jaccard	0.66725	0.77851
Recall	0.70838	0.85110
Precision	0.92838	0.90642

[0105]

표 2

	Original M-Net	Original M-Net Skip-Connection Attention
Accuracy	0.83918	0.86216
F1	0.56500	0.59400
Jaccard	0.46870	0.49738
Recall	0.61705	0.59872
Precision	0.71307	0.77494

[0106]

[0107] 표 1에서 보는 것처럼 정확도가 Original M-Net보다 Original M-Net Skip-Connection Attention 모델 즉 본 발명의 MAM-net 모델이 0.02184정도 성능이 더 높게 나온 것을 볼 수 있다. 표 2는 피부 영상을 Original M-Net과 본 발명의 MAM-net 모델을 이용한 결과이다. 표 2에서 보는 것처럼 정확도가 Original M-Net보다 본 발명의 MAM-net 모델이 0.02298정도 성능이 더 높게 나온 것을 볼 수 있다.

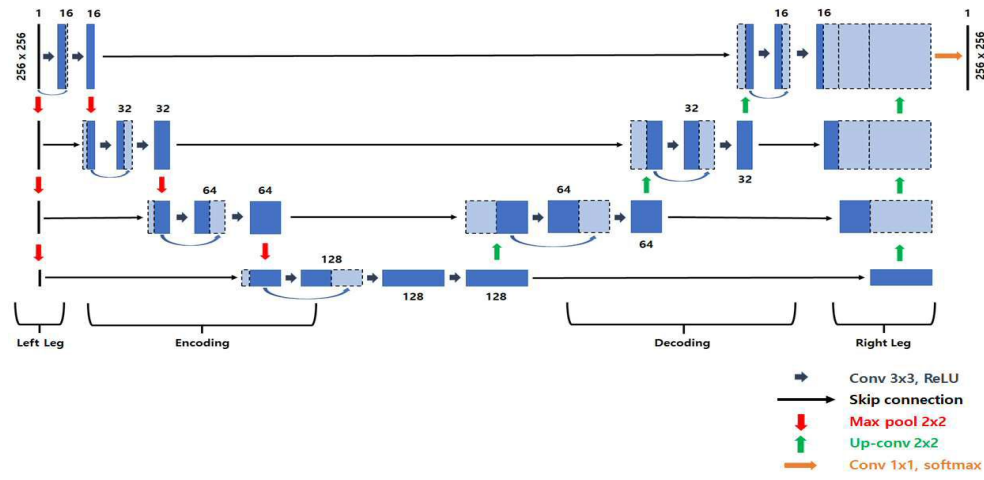
[0108] 도 3a는 Original M-Net을 이용한 치아 영상 결과이고 도 3b는 본 발명에 따라 MAM-net(Original M-Net Skip-Connection Attention)을 이용한 치아 영상 분할 결과이다. 도 4a는 Original M-Net을 이용한 피부 영상 분할 결과이고 도 4b는 MAM-net(Original M-Net Skip-Connection Attention)을 이용한 피부 영상 분할 결과이다.

[0109] 이러한 결과로부터 본 발명에 의하면 달성해야 할 목적과의 연관성에 따라 추출한 특징에 가중치를 부여함으로써 네트워크 스스로 집중해야 할 영역을 선별하도록 하여 의료 영상을 분할하였고 이러한 분할 결과로 얻어진 치아 영상 및 피부 영상이 기존 Original M-Net에 비해 향상된 것을 확인할 수 있으므로, 아키텍처를 효과적으로 활용하여 정확도가 향상된 의료영상 분할 방법을 제공할 수 있는 것을 알 수 있다.

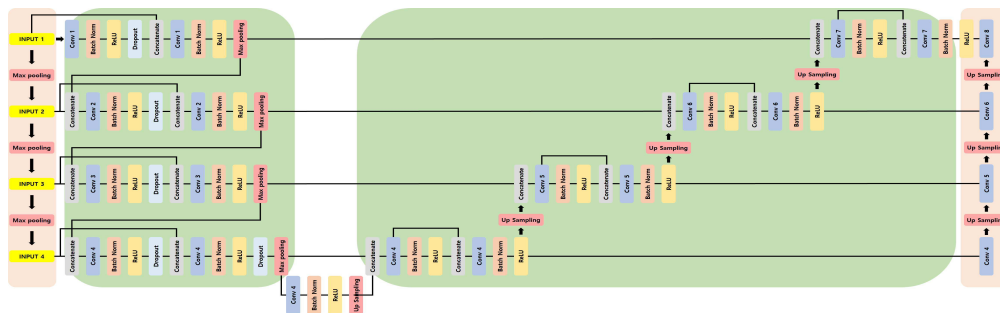
[0110] 본 발명은 이상에서 살펴본 바와 같이 바람직한 실시 예를 들어 도시하고 설명하였으나, 상기한 실시 예에 한정되지 아니하며 본 발명의 정신을 벗어나지 않는 범위 내에서 당해 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 다양한 변경과 수정이 가능할 것이다.

도면

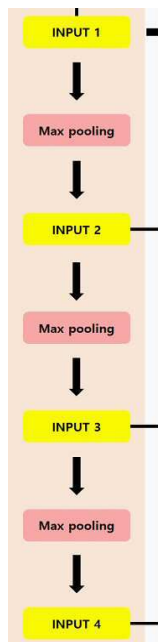
도면1a



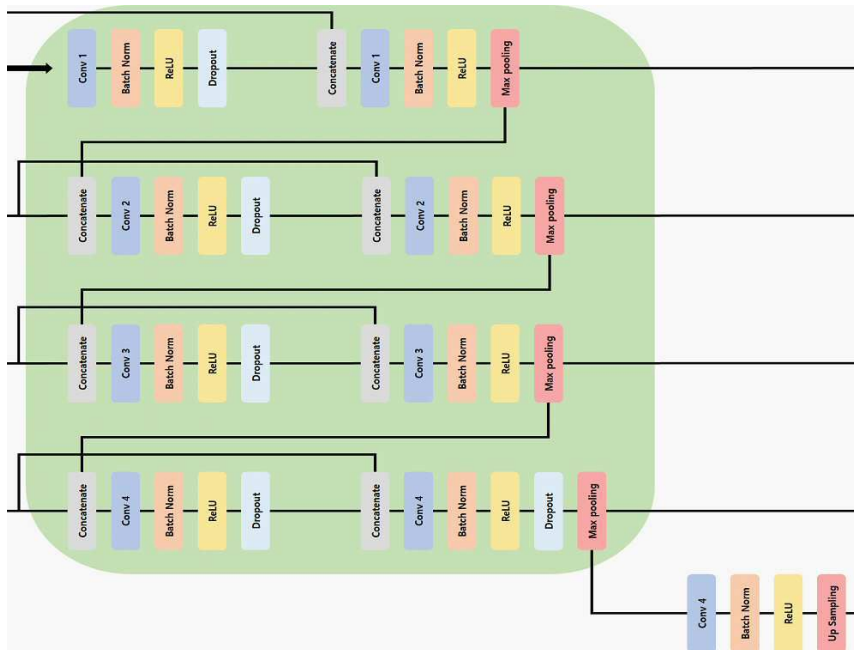
도면1b



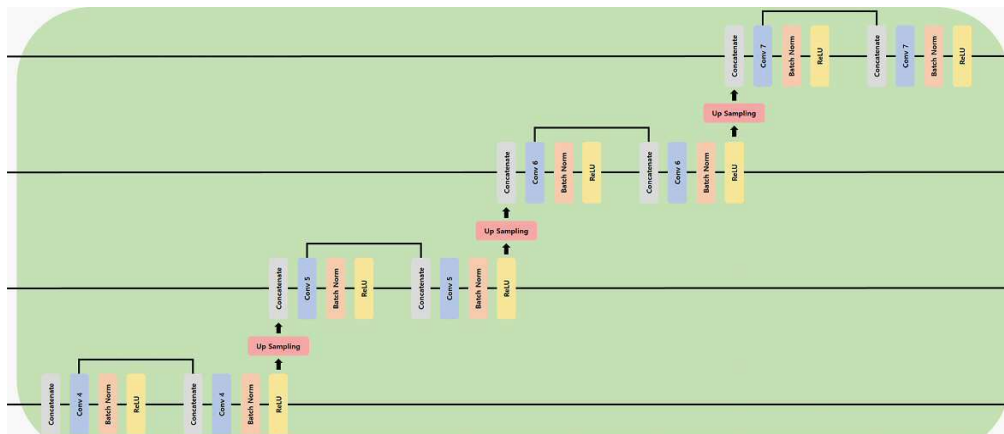
도면1c



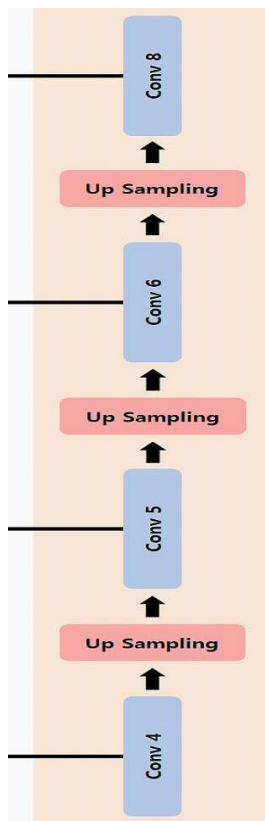
도면1d



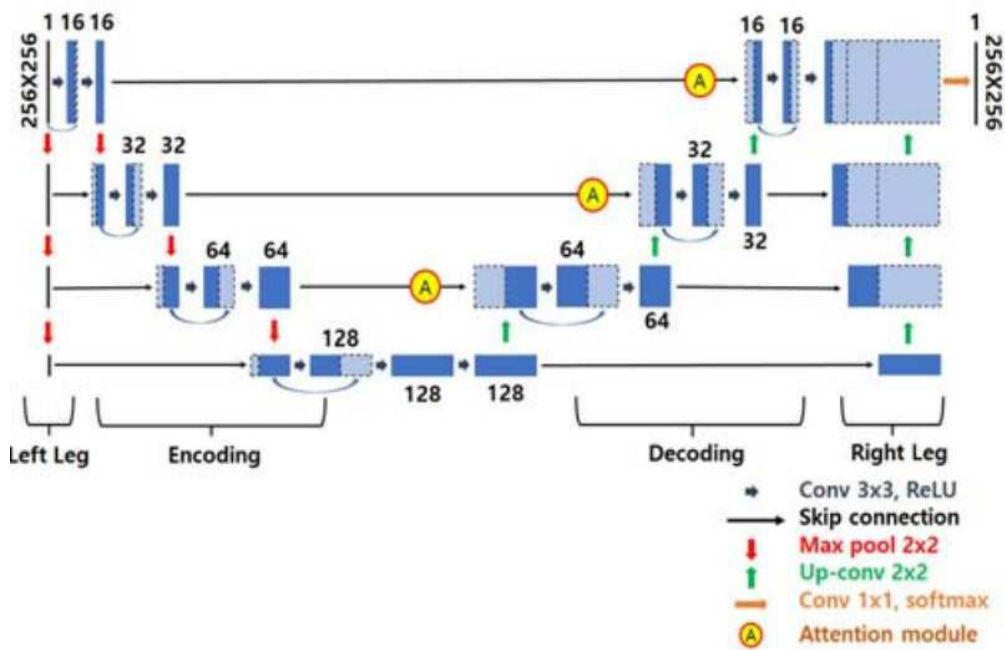
도면1e



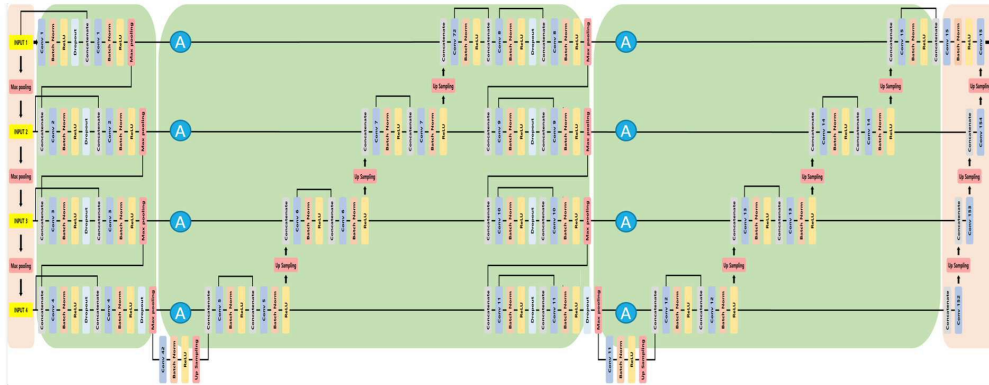
도면1f



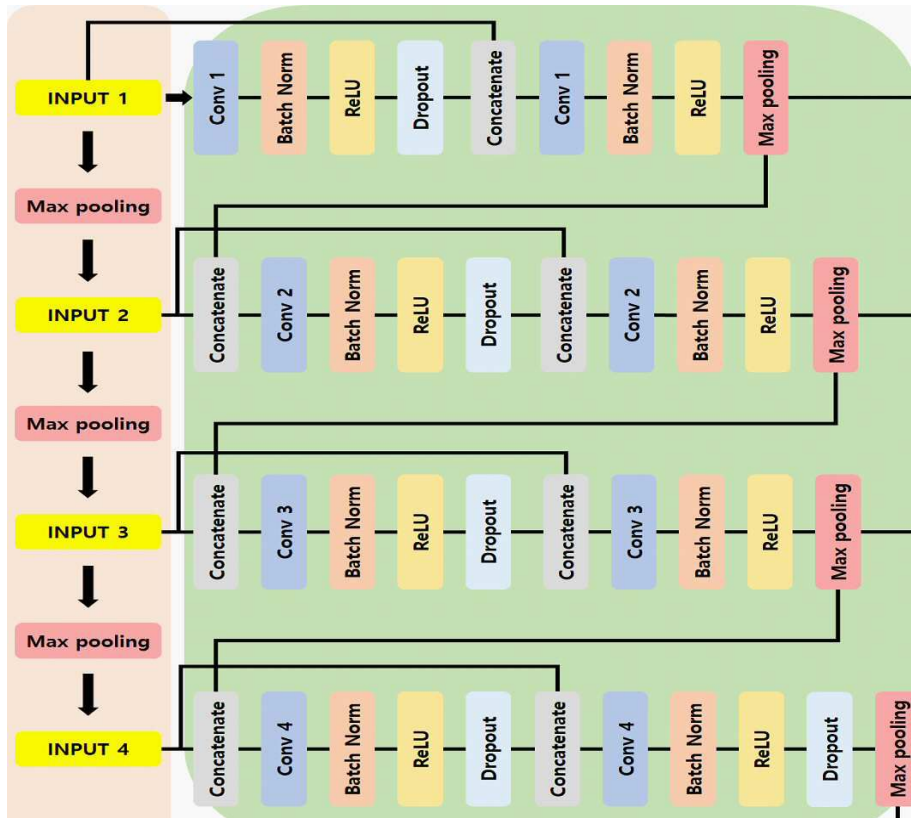
도면2a



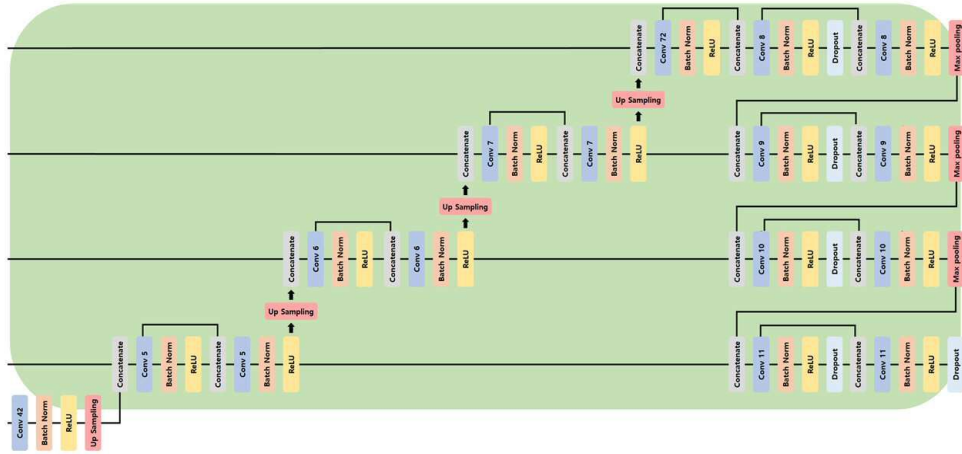
도면2b



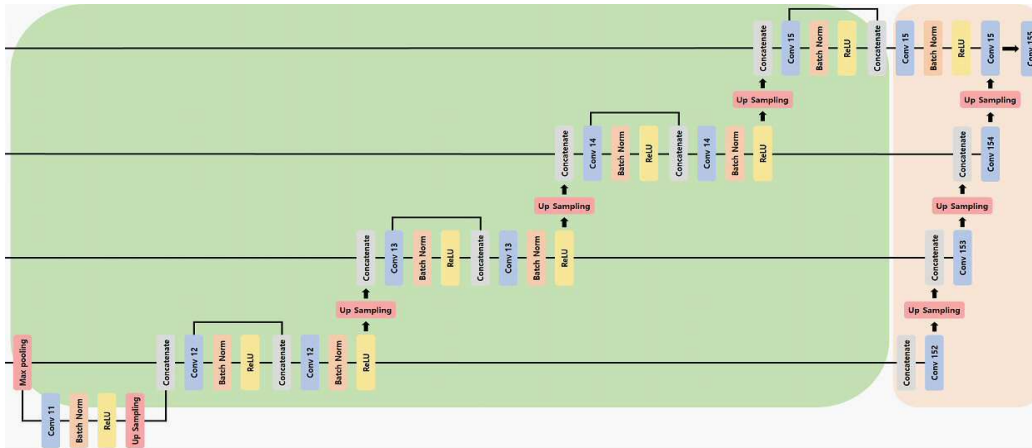
도면2c



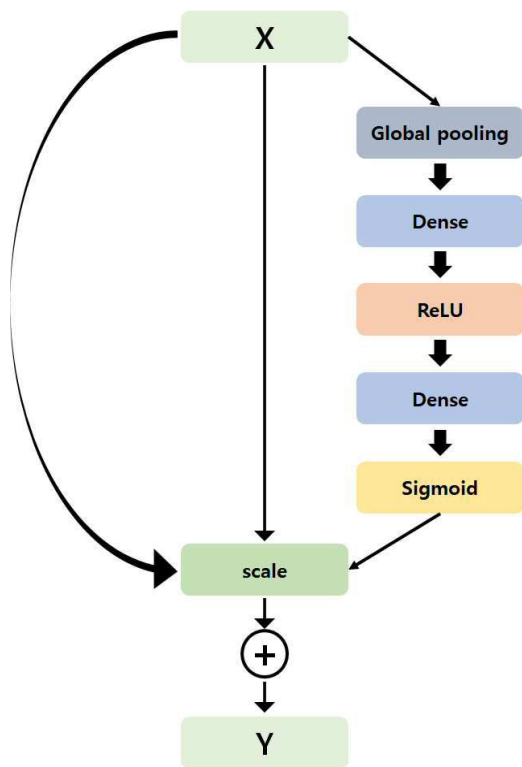
도면2d



도면2e



도면2f



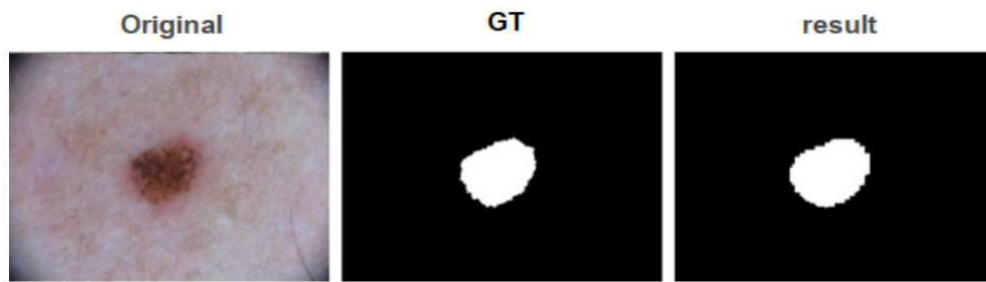
도면3a



도면3b



도면4a

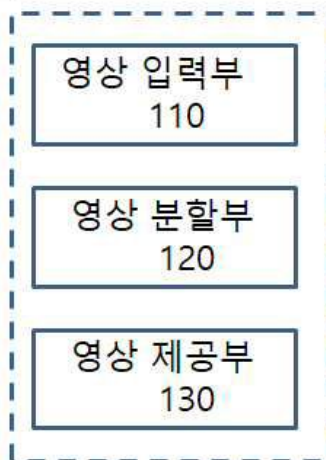


도면4b



도면5a

100



도면5b

