creative commons

COMMONS DEED

저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

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

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

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

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

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

공학석사 학위논문

삼차원 의료영상의 혈관 영역추출을 위한 순환신경망이 활용된 심층학습 알고리즘 개발 및 검증

Development and validation of a deep learning algorithm using recurrent neural network for blood vessel segmentation of 3D medical image

2020년 12월

서울대학교 대학원

협동과정 바이오엔지니어링전공

이 겨 레

삼차원 의료영상의 혈관 영역추출을 위한 순환신경망이 활용된 심층학습 알고리즘 개발 및 검증

지도교수 Sungwan Kim

이 논문을 공학석사 학위논문으로 제출함

2020년 12월

서울대학교 대학원

협동과정 바이오엔지니어링전공

이 겨 레

이겨레의 공학석사 학위논문을 인준함

2020년 12월

위 원 장 박 지 훈 (인)

부 위 원 장 Sungwan Kim (인)

부 위 원 장 이 경 호 (인)

위 원 이 경 준 (인)

국문초록

삼차원 의료영상의 혈관 영역추출을 위한 순환신경망이 활용된 심층학습 알고리즘 개발 및 검증

이겨레

서울대학교 대학원

협동과정 바이오엔지니어링 전공

전산화단층촬영, 자기공명영상 등의 삼차원 의료영상에서 혈관 영역추출은 다양한 심혈관계 질환의 진단 및 수술계획 수립에 요구되는 선행 절차이다. 특히 혈관은 타 장기에 비해 좁고 복잡하다는 특징이 있어 영역추출의 난도가 높고 소모되는 인력 및 시간 또한 상당하다. 그렇기 때문에 혈관 영역추출을 위한 자동화 기술 연구들이 활발히 이루어지고 있다.

최근에는 그래픽 처리 장치와 소프트웨어 기술이 발전함에 따라 인공지능으로 의료영상을 분석해 판독 및 진단을 돕는 연구분야가 크게 활성화 되고 있다. 그 중, 합성곱신경망을 사용한 심층학습 접근법은 영상의 공간적 특징을 자동 추출함으로써 뛰어난 성능을 기록했고, 도약연결부를 지닌 특수한 형태의 합성곱신경망인 U-Net은 인공지능 영역추출 모델의 표본으로

인식되고 있다.

삼차원 혈관 영역추출을 위한 심층학습 모델 또한 U-Net 기반으로 구성되는 경우가 일반적이다. 특히 타 장기 영역추출에도 범용적으로 사용되는 이차원 U-Net이나 삼차원 U-Net을 채택하는 사례가 많다. 하지만 혈관은 일반적인 장기와 다르게 신체 내에서 매우 얇고 복잡하게 얽혀 있다는 구조적 특징이 있어 범용 모델을 기반으로 접근할 경우 타 장기 영역추출에 비해 상대적인 어려움이 발생한다.

본 연구에서는 삼차원 혈관 영역추출의 성능을 향상시키기 위해 U-Net 기반의 새로운 맞춤형 모델인 '거미형 U-Net'을 제안한다. 이 모델의 핵심 알고리즘은 기존 이차원 U-Net 접근법에 순환신경망을 삽입하는 것이다. 순환신경망은 본래 음성, 자연어 등 순서가 있는 자료의 문맥적 연결성을 학습하는 일에 특화된 심층신경망의 특수한 형태이다. 따라서 삼차원 의료영상을 이미지 시퀀스 (image sequence)로 간주한다면 순환신경망을 적용해 인접한 이미지와의 문맥적 연결성을 모델에 효과적으로 반영시킬 수 있다. 특히 혈관은 기본적으로 서로 연결되어 있으므로 이미지 간 뚜렷한 문맥이 존재하기 때문에 순환신경망을 사용할 경우 학습의 효과가 좋아질 수 있다.

본 연구는 제시한 접근법의 효과 검증을 위해 뇌혈관, 간혈관 및 좌심실 영역추출을 위한 세 종류의 삼차원 혈관 데이터베이스를 확보해 학습 및 평가를 진행한다. 또한 성능의 비교를 위해 기존의 대표 모델인 이차원 U-Net과 삼차원 U-Net을 비롯하여 순환신경망을 다른 방식으로 활용한 FCN-RNN에 대해 같은 조건 하에서 추가 학습시켜 비교군으로 설정한다.

마지막으로, 영역추출의 성능 평가에 일반적으로 사용되는 다이스 계수 (DSC) 및 자카드 계수 (IoU)의 평균치를 보고한다. 대표적으로 뇌혈관 데이터베이스에 대해 이차원 U-Net은 0.745, 삼차원 U-Net은 0.716, FCN-RNN은 0.752, 그리고 거미형 U-Net은 0.807의 평균 DSC를 기록하며 기존 모든 모델의 성능을 능가함을 확인할 수 있다.

주요어 : 딥러닝, 삼차원 혈관 영역추출, 순환신경망, 합성곱신경망, 인접영상 간 문맥학습

학 번 : 2019-22708

목차

국문초록 i

목차 iv

그림 목차 vi

표 목차 ix

1. 서론 10

1.1. 배경 10

1.2. 연구의 목적 13

2. 본론 15

2.1. 알고리즘 15

2.1.1. U-Net 기저모델 16

2.1.2. 날실부 19

2.1.3. 씨실부 20

2.2. 데이터베이스 22

2.2.1. 뇌 자기공명혈관조영영상 22

2.2.2. 복부 전산화단층촬영영상 26

2.2.3. 심장 자기공명영상 27

2.3. 비교 모델 28

2.4. 구현 및 실험 29

3. 결과 및 분석 30

3.1. 성능평가지표 30

3.2. 정량적 성능평가 30

3.3. 정성적 성능평가 31
3.3.1. 뇌 자기공명혈관조영영상 32
3.3.2. 복부 전산화단층촬영영상 34
3.3.3. 심장 자기공명영상 35
4. 고찰 37
4.1. 실험 결과 고찰 37
4.2. 알고리즘 고찰 38
5. 결론 39
참고 문헌 40
Abstract 42

그림 목차

그림 1. TIME-OF-FLIGHT 기법으로 촬영한 뇌혈관 자기공명혈관조영영상 (MRA) 예시. MRA의 특성 상 혈관은 높은 강도로 표현되며(붉은 화살표) 한 장 당 혈관 영역이 차지하는 비율이 1% 미만으로 매우 작기 때문에 심층신경망 학습 시 비혈관 영역에의 압도 현상이 일어난다. 12
그림 2. 거미형 U-NET의 전체 구조. 삼차원 영상을 수평면 영상에 대한 이미지 시퀀스로 변환해 처리한다. 해당 모델은 마치 직물을 방적기로 짜듯이 날실부(검정색 실선)와 그 가운데에 위치한 씨실부(검정색 막대)로 이루어져 있다. 먼저 변환된 수평면 영상 시퀀스를 날실부의 부호기에 넣어 공통된 공간 특징(회색 육면체)들의 추출을 마무리 짓는다. 다음으로 순환신경망으로 이루어진 씨실부로 문맥을 아는 공간 특징으로 변화시킨다. 마지막으로, 날실부의 복호기를 통해 문맥이 반영된 혈관 영역 추출을 진행한다.....15
그림 3. 본 연구에서 사용한 U-NET의 얼개. 전방의 하행부는 부호기로 영상의 공간적 특징을 추출하며, 후방의 상행부는 복호기로 부호기와 도약 연결부로부터 공간적 특징을 받아 분할 영역을 예측한다.....16
그림 4. 날실부의 얼개. 각각의 BN에 해당하는 굵은 화살표는 $\mathrm{U-NET}$을 의미하며, T개의 $\mathrm{U-NET}$이 병렬로 나열되어 있다. 이로부터 여러 장의 연속된 영상을 동시에 받아 처리할 수 있게 된다. 그림 상으로 알 수 있듯이, 각각의 날실들은 아직 서로 어떠한 연결도 없기 때문에 인접 문맥이 반영되지 못한다.....19
그림 5. 날실부에 씨실부를 적용한 거미형 U-NET의 완성된 얼개. 씨실부에서 존재했던 T개의 연결되지 않은 공간적 특징(SPATIAL FEATURE, 흰 막대기)들이 씨실부의 합성곱 장단기메모리를 통해 연결되어 문맥을 아는 공간적 특징(CONTEXT-AWARE SPATIAL FEATURE, 붉은 막대기)들로 변화된다. 이는 복호기의 예측 단계에서 앞뒤로 인접한 영상의 특징을 참고하게 만들어 보다 정확한 결과를 얻을 수 있게 한다.....20
그림 6. TIME-OF-FLIGHT 자기공명혈관조영술로 촬영한 뇌의 수평면 영상(좌) 및 지도학습을 위해 전문의가 혈관 영역을 표기한 레이블(우).....22
그림 7. 문턱값 방식을 적용해 추출한 혈관 영역(좌) 및 영상의학과 전문의에 의해 정돈된 최종 레이블(우). 문턱값 방식만을 적용할 경우 위양성 영역(붉은 화살표)이 표현되는 경우가 잦으며, 관심영역이 아닌 혈관(검은 화살표)을 걸러낼 수 없다.....23

그림 8. 레이블 작성을 위해 활용한 MEDICAL IMAGING INTERACTION TOOLKIT (MITK) 프로그램의 인터페이스(상) 및 레이블링에 활용된 태블릿의 모습(하). MITK에서는 총 세 가지 뷰(AXAL, SAGITTAL, CORONAL)로 영상을 확인하며 혈관 영역을 그릴 수 있고, 이를 삼차원으로 재구성해 편리하게 결과 검토가 가능하다. 24
그림 9. 레이블링 대상의 대표 혈관을 정의하기 위한 TOF 데이터의 MAXIMUM INTENSITY PROJECTION (MIP) 예시 영상. 중앙의 대뇌동맥륜 (붉은 원)을 중심으로 퍼져나는 형태로 주요 혈관을 그려나갔고 붉은 선을 초과하는 영역인 A2, M2 그리고 P2 (각각 A, B, C)는 레이블로 취급하지 않았다. 25
그림 10. 복부 전산화단층촬영 영상의 수평면 이미지 예시(좌) 및 간 내부 혈관을 표현한 레이블(우). 좌측 그림의 붉은 영역은 간의 대략적인 범위를 의미한다. 26
그림 11. 심장 자기공명영상의 수평면 이미지 예시(좌) 및 좌심실 벽 영역을 표현한 레이블(우) 27
그림 12. 실험에 사용한 모든 U-NET 기반 모델의 간소화된 구조. 각각 이차원 U-NET (좌상), 삼차원 U-NET (좌하), FCN-RNN (우상) 그리고 거미형 U-NET (우하)을 의미한다. 푸른 화살표는 RELU 함수로 활성화 된 합성곱 계층에 배치 정규화가 이루어진 연산이며, 붉은 화살표는 최대값 풀링, 노란색 화살표는 상행 합성곱(UPCONVOLUTION) 계층, 녹색 화살표는 소프트맥스 함수를 의미한다. LSTM은 붉은 상자로 표현되었다. 28
그림 13. 주요 뇌동맥에 대한 세 종류의 비교모델 및 거미형 U-NET의 삼차원 예측 결과 예시. 거미형 U-NET은 가장 높은 성능을 보였고 주요 혈관을 타 모델에 비해 정밀하게 예측하였다. 32
그림 14. 전대뇌동맥의 영역 추출 결과를 확대하여 나타낸 그림. 좌측의 흑백 그림은 원본 영상을 MAXIMUM INTENSITY PROJECTION을 통해 표현한 이차원 그림이며, 우측 그림은 위에서부터 순서대로 이차원 U-NET, 삼차원 U-NET, FCN-RNN 그리고 거미형 U-NET으로 예측한 전대뇌동맥 영역을 나타낸다. 33
그림 15. 간 혈관에 대한 세 종류의 비교모델 및 거미형 U-NET의 삼차원 예측 결과 예시. 거미형 U-NET은 타 모델들에 비해 크게 높은 성능을 기록하였으며, 세부적인 혈관을 가장 잘 예측하였다. 34
그림 16. 네 종류의 비교모델 및 거미형 U-NET의 삼차원 모델링 결과 예

시. 노란색은 정답과 일치하는 영역, 붉은색은 위양성 영역, 그리고 파란색은 위음성 영역이다. 각 모델명 하단의 숫자는 왼쪽부터 각각 평균 다이스 계수 및 평균 자카드 계수를 의미한다.....35

표 목차

표 1. 학습 및 평가에 사용한 세 종류의 데이터베이스. 모델의 일반화 능력을 확인하기 위해 모달리티, 대상 장기 및 혈관이 모두 다르게 구성되었다. 28
표 2. 거미형 U-NET 및 모든 비교모델에 대한 다이스 계수(좌) 및 자카드 계수(우) 성능. 모든 데이터베이스에 대해 기존의 모든 비교모델들의 성능을 능가한다. 30

1. 서론

1.1. 배경

의료영상에 대한 혈관 영역추출은 여러 심혈관계 질환에 대해 진단, 수술 계획수립을 비롯한 다양한 임상 절차를 수행하기 위해 요구되는 선행 단계이다. 예를 들어, 좌심실의 영역추출은 심혈관 관련 건강상태를 분석하기 위한 지표인 심장의 수축기 및 확장기 용적 분석에 활용되며 [1], 뇌혈관의 영역추출은 혈관의 협착 및 폐쇄 검출을 도와 뇌경색 진단을 수월하게 한다 [2].

혈관 영역추출은 분야의 특수성 때문에 전문인력이 소모되며, 또한 혈관의 복잡한 특성 상 많은 시간과 노력을 필요로 한다. 따라서 혈관 영역추출의 자동화 기술 개발은 중요한 연구 주제이다. 기존 혈관 영역추출 자동화를 위한 전통적인 방식은 혈관 강조 기법 [3, 4], 경계선 기반 기법 [5, 6], 그리고 구역 기반 기법 [7] 등이 존재한다. 먼저 혈관 강조 기법은 사전 정의된 영상 필터를 통해 혈관 영역과 비혈관 영역의 대조도를 높인 뒤, 문턱값을 적용하여 영역을 추출하는 방식이다. 경계선 기반 기법은 곡선에 대한 수학적 특징 요소들을 고려해 혈관 경계선을 검출함으로써 영역추출을 이루어 낸다. 마지막으로, 구역 기반 기법은 곡선이 아닌 특정 관심영역에 대해 사전 정의된 특징 요소들을 적용해 혈관 영역을 추출한다. 나열한 전통적 방식들은 대상 영상의 종류에 따라 최적의 매개변수 값이 달라지며 예외에 민감하게 반응하기 때문에 안정적인 성능 확보가 어렵다는 한계가 존재한다.

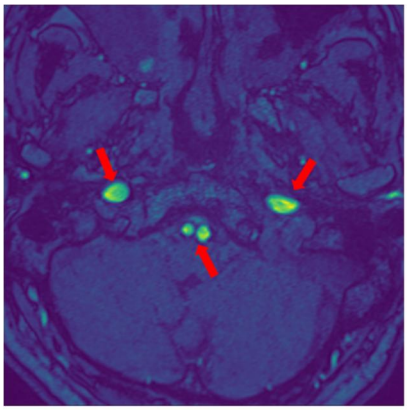
최근, 상기 방식들의 한계를 뛰어넘기 위해 합성곱 연산

기반의 심층신경망으로 의미론적 분할 (semantic segmentation) 문제를 풀어냄으로써 혈관 영역추출을 해결하는 인공지능 연구가 활발히 진행되고 있다 [8, 9]. 그 중 부호기-복호기 (encoder-decoder) 구조의 합성곱신경망은 부호기를 통해 영상의 특징을 자동으로 추출하고, 복호기를 통해 원본 영상과 같은 크기의 분할 마스크 (segmentation mask)를 예측함으로써 혈관을 비롯한 모든 의료영상 분할 문제에 적용되고 있다. 특히 부호기-복호기 구조의 특수한 형태 중 하나인 U-Net은 부호기와 복호기의 같은 층위를 직접적으로 연결하는 도약 연결부 (skip connection)의 도입을 통해 높은 성능 향상을 이루었다. 본래 부호기-복호기 구조에서 복호기는 낮은 해상도의 특징들만 가지고 원본 해상도의 마스크를 복구해야 했기에 정밀한 예측이 어렵다는 문제가 있었지만, 도약 연결부는 같은 층위의 특징을 직접적으로 복호기에게 참조시킴으로써 이 문제를 완화시켰고, 의료영상 분할을 위한 대표적인 모델로 인정받고 있다.

삼차원 혈관 영역추출 분야 또한 위와 같은 이유로 U-Net 기반의 모델을 설계하는 경우가 일반적이며, 사용하는 합성곱 연산의 차원에 따라 이차원 U-Net [10] 혹은 삼차원 U-Net [11]을 활용한다. 예를 들어, 삼차원 자기공명혈관조영영상에서 두개내혈관을 추출한 연구가 존재하는데 [12], 이 연구에서는 이차원 U-Net을 사용하였으며 고전적인 방식인 Graph Cut과 비교했을 때 더 높은 성능을 달성하였다. 하지만 삼차원 영상에서 이차원 접근법을 사용할 경우, 각각의 평면에 대한 특징이 추출될 뿐, 부피에 대한 삼차원적 특징은 추출되지 않기 때문에 최고의 성능이 아닐 수 있다는 가능성이 존재한다. 삼차원 U-Net을

활용해 실험용 쥐의 뇌혈관을 추출함으로써 알츠하이머 질병분석에 활용한 연구 또한 존재한다 [13]. 이 연구에서는 모델의 구조적 변경을 통해 원래의 삼차원 U-Net보다 높은 DSC를 달성하였다. 하지만 이러한 삼차원 접근법은 대단히 많은 양의 메모리를 요구하기 때문에 영상 학습에 최적화 된 모델을 자유롭게 설계하기 힘들다는 어려움이 있다.

또한 혈관은 타 장기에 비해 굉장히 복잡하고 얇다는 구조적 특징을 가지고 있기 때문에 이러한 범용 U-Net들을 똑같이 적용하더라도 성능 확보에 어려움이 발생한다. 먼저 그림 1에서 확인할 수 있듯이, 전체 영상 대비 혈관이 차지하는 영역이 매우 적으므로 모델의 학습이 비혈관 영역 위주로 진행되는 압도(domination) 현상이 발생한다. 이를 극복하기 위해 혈관 영역 위주로 영상을 쪼개 학습시키는 방식을 도입하곤 하는데, 이 역시 모델의 시야를 좁히기 때문에 의료영상 분할 시 중요한 단서로 작용할 수 있는 전체적인 해부학적 구조를 잃게 된다.



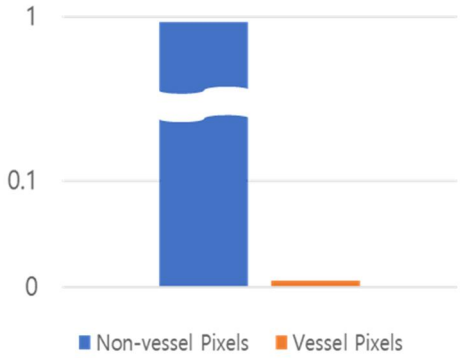


그림 1. Time-of-flight 기법으로 촬영한 뇌혈관 자기공명혈관조영영상(MRA) 예시. MRA의 특성 상 혈관은 높은 강도로 표현되며(붉은 화살표) 한 장 당 혈관 영역이 차지하는 비율이 1% 미만으로 매우 작기 때문에 심층신경망 학습 시 비혈관 영역에의 압도 현상이 일어난다.

이를 극복하기 위해 혈관의 또 다른 뚜렷한 특징인 연결성을 활용하는 것이 주요할 수 있다. 혈관은 기능 상 혈액을 운반하는 것이 주 역할이기 때문에 기본적으로 서로 연결되어 있다. 즉, 삼차원 영상을 이미지 시퀀스로 생각하였을 때, 인접한 프레임들에 표현된 혈관 영역들은 서로 참조할 만한 유사성을 지닌 문맥을 형성하고 있다. 이를 시계열 데이터의 문맥 학습에 특화된 심층신경망의 일종인 순환신경망을 활용해 강조한다면 부족한 혈관 영역에 대한 정보를 보강함으로써 영역추출의 정확도를 높일 수 있을 것이다.

이와 유사한 철학을 공유하는 접근법이 등장하기도 했다. FCN-RNN [14] 은 삼차원 의료영상에서 병변 분할을 위해 도입된 접근법으로 임의의 부호기-복호기 뒤에 임의의 순환신경망을 추가 배치할 경우 여러 질병의 병변 분할에 대한 성능이 향상됨을 실험적으로 보여주었다. 본 연구는 이러한 근거를 바탕으로 순환신경망이 추가된 U-Net 구조를 제시함으로써 삼차원 혈관영역 추출을 위한 범용 모델을 구현하고 그 효과를 실험적으로 검증하고자 한다.

1.2. 연구의 목적

본 연구는 삼차원으로 표현된 혈관의 내재적 특징인 연결성을 강조해 삼차원 혈관 영역추출의 성능을 높이기 위한 새로운 심층신경망 구조를 제안한다. 이를 위해 대표 모델로 널리 사용되는 이차원 U-Net을 기반으로 날실부와 씨실부라는 고유

요소를 구축해 마치 방적기로 직물을 짜듯이 혈관 영역을 학습한다. 모델의 전체적인 얼개와 작동 방식이 거미의 외관과 행동양식을 닮았기에 이 모델을 '거미형 U-Net'으로 명명하였다.

날실부는 U-Net이 여러 겹으로 병렬 연결 되어 있으며 여러 장의 영상으로부터 각각의 공간 특징 (spatial feature)들을 추출한다. 다음으로, 씨실부는 연속적인 데이터의 문맥 학습에 활용되는 순환신경망 [15]으로 구현되어 날실부의 부호기와 복호기 사이에 수직으로 배치되어 있으며, 앞서 추출된 공간 특징들에게 문맥적 연결성을 추가적으로 부여함으로써 문맥을 아는 공간 특징 (context-aware spatial feature)으로 변화시킨다. 이러한 구조는 부족한 혈관 관련 정보를 인접한 이미지로부터 보강해 성능을 높이는 동시에, 이차원 U-Net의 한계점인 삼차원 공간 특징의 부재를 완화하고, 삼차원 U-Net의 한계점인 메모리의 과도한 요구량을 줄일 수 있다는 장점이 존재한다.

본 연구는 소개된 거미형 U-Net 구조의 효과를 확인하기 위한 실험을 진행했다. 먼저 여러 혈관과 모달리티 (modality)에 대한 적용가능성을 보이기 위해 세 종류의 서로 다른 혈관영상 데이터베이스에 대해 학습 및 평가를 수행했고, 성능을 비교하기 위해 기존에 사용되던 세 가지 대표 모델인 이차원 U-Net, 삼차원 U-Net, 그리고 FCN-RNN을 공평한 환경에서 학습시켜 그 결과를 평균 다이스 계수 (DSC)로 제시했다.

거미형 U-Net을 통해 심근경색, 뇌경색 등 치명적인 심혈관계 질환의 정확한 진단을 도움으로써 국민건강증진 및 환자들의 삶의 질 향상에 일조하는 것이 본 연구의 궁극적인 목표이다.

2. 본론

2.1. 알고리즘

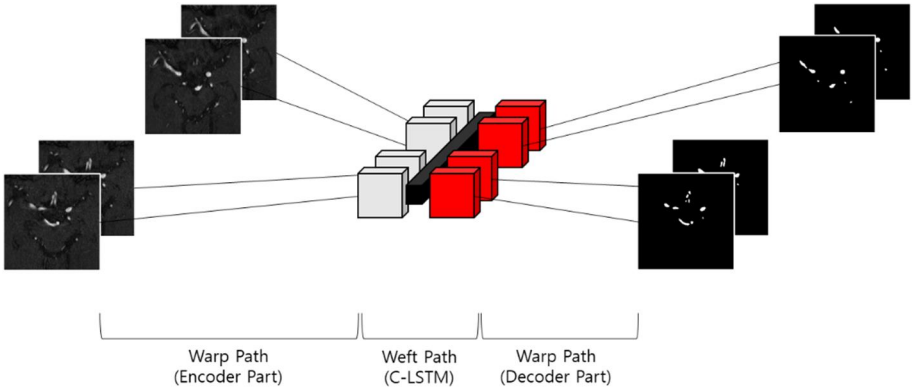


그림 2. 거미형 U-Net의 전체 구조. 삼차원 영상을 수평면 영상에 대한 이미지 시퀀스로 변환해 처리한다. 해당 모델은 마치 직물을 방적기로 짜듯이 날실부(검정색 실선)와 그 가운데에 위치한 씨실부(검정색 막대)로 이루어져 있다. 먼저 변환된 수평면 영상 시퀀스를 날실부의 부호기에 넣어 공통된 공간 특징(회색 육면체)들의 추출을 마무리 짓는다. 다음으로 순환 신경망으로 이루어진 씨실부로 문맥을 아는 공간 특징으로 변화시킨다. 마지막으로, 날실부의 복호기를 통해 문맥이 반영된 혈관 영역 추출을 진행한다.

본 연구에서 제안하는 거미형 U-Net은 고유 요소인 날실부 (warp path)와 씨실부 (weft path)를 도입해 각각 영상의 공간적 특징과 문맥적 연결성을 학습한다. 날실부는 부호기와 복호기로 구성되어 있으며, 씨실부는 두 요소 사이에 위치한다.

이 알고리즘은 이차원 접근법의 대표 모델인 U-Net [10]을 기반으로 구축되었다. U-Net, 날실부, 그리고 씨실부에 대해서는 다음 하위 목차를 통해 세부적으로 설명한다.

2.1.1. U-Net 기저모델

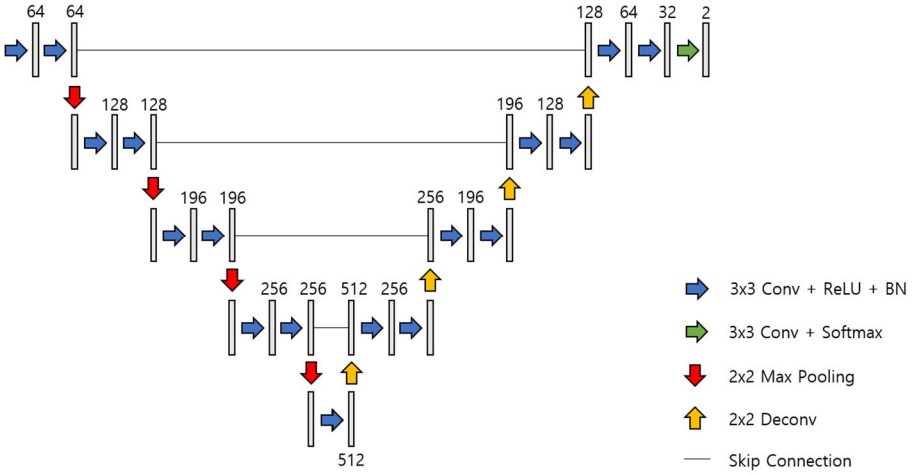


그림 3. 본 연구에서 사용한 U-Net의 얼개. 전방의 하행부는 부호기로 영상의 공간적 특징을 추출하며, 후방의 상행부는 복호기로 부호기와 도약 연결부로부터 공간적 특징을 받아 분할 영역을 예측한다.

U-Net은 영상분할을 위해 주로 사용되는 심층신경망 기반 부호기-복호기 구조의 특수한 형태이다. 전방의 부호기는 영상으로부터 공간 특징을 추출하며, 후방의 복호기는 추출된 특징들로부터 분할 결과를 예측한다.

부호기 (encoder)는 합성곱신경망으로 구성되어 있으며, 한 장의 영상에 대해 층위를 낮추어 가며 공간 특징을 추출한다. 구체적으로는 층위가 낮아져 감에 따라 부호기의 두 가지 요소가 영상에 대해 작동하게 된다. 첫째로, 주어진 영상을 2x2 픽셀로 균일하게 조각낸 뒤, 각 조각들로부터 가장 큰 값을 취하는 2x2 최대값 풀링 (max pooling)을 통해 영상의 가로 및 세로 길이를 반감시킨다. 이는 영상의 크기를 줄일 때, 가장 뚜렷한 값을 유지하여 강력한 특징을 유지하게 하며, 동시에 고정된 크기의

합성곱 필터가 영상의 더 넓은 부분을 잡아낼 수 있게 하기 때문에 층위가 낮아짐에 따라 다양한 특징 추출이 가능해질 수 있게 돕는 역할을 수행한다. 둘째로, 적용되는 합성곱 필터의 갯수가 점점 늘어난다. 상위 층위에서 추출된 특징은 상대적으로 단순한 규칙을 지니고 있는데, 이 때 합성곱 필터의 갯수를 늘리게 되면 모델은 해당 필터를 통해 단순한 규칙으로부터 더 복잡한 특징을 재조합하며 학습하게 되며, 이는 모델의 전체적인 수용력을 높임으로써 영상에 대한 넓은 이해도를 갖출 수 있게 한다.

복호기 (decoder)는 단어의 본래 의미 그대로 단서로부터 원하는 결과를 복구하는 역할을 한다. 특히 의미론적 분할 문제에서는 원본 영상 크기와 같은 분할 마스크 (segmentation mask)를 복구하는 것이 그 목적이다. 그렇기 때문에 복호기는 부호기와는 완전히 반대로 작동하도록 설계되어 있고, 따라서 진행될수록 합성곱 필터의 갯수는 줄어들며 영상의 해상도는 늘어난다. 해상도를 늘릴 때, 단순히 선형 보간법을 통해 비어있는 픽셀을 메꾸는 것이 아니라 특수한 합성곱 연산을 사용해 영상의 특징을 바탕으로 픽셀을 복구하게 된다. 그렇다고 하더라도 단순한 부호기-복호기 구조만을 사용할 경우, 이미 부호기를 통해 해상도가 대단히 낮아진 공간 특징으로부터 원본 해상도의 분할 마스크를 정확히 예측하는 것은 어렵다. 그러나 U-Net의 복호기는 다음 문단에서 설명할 도약 연결부를 통해 충분한 정보를 제공받아 정밀한 복구능력을 갖추게 되며 결과적으로 높은 성능향상을 이룰 수 있게 되었다.

도약 연결부 (skip connection)는 U-Net에서 가장 특이할

부분이며 도식 상 같은 층위의 부호기와 복호기를 연결하는 여러 개의 실선으로 표현된다. 이 선의 실제 작동 방식은 같은 층위에 해당하는 부호기의 공간 특징을 그대로 복사하여 복호기에 전달하는 것이다. 복호기는 도약 연결부로부터 직접적으로 참조된 공간 특징과 한 단계 아래 층위에서 합성곱 연산을 통해 복구된 공간 특징을 종합해 복구를 수행하게 되고, 이는 풍부하고 유의미한 정보량을 바탕으로 예측 정확도를 높이는 데에 주요한 역할을 한다.

본 연구에서는 그림 2와 같이 사전 정의된 U-Net을 구현하여, 공정성을 부여하기 위해 실험에 사용된 모든 모델의 공통 기본구조로 활용했다. 층위는 총 5단계이며, 단계가 깊어짐에 따라 추출하는 공간 특징의 해상도는 반절되지만, 합성곱 필터 수는 각각 64, 128, 196, 256, 그리고 512개로 늘어난다. 모든 합성곱 연산은 정류된 선형 함수 (rectified linear unit, ReLU)에 의해 활성화되어 모델에 비선형성을 부여하며, 그 뒤에 배치 정규화(batch normalization) 계층을 추가해 학습의 안정성을 제공한다. 마지막 층은 소프트맥스 (softmax) 함수를 통해 각 픽셀마다 이진 클래스(혈관 및 비혈관)에 대한 확률을 나타내게 된다. 이후 각 픽셀 별로 더 높은 확률을 가진 클래스를 선택하여 최종 분할 마스크를 완성시킨다.

2.1.2. 날실부

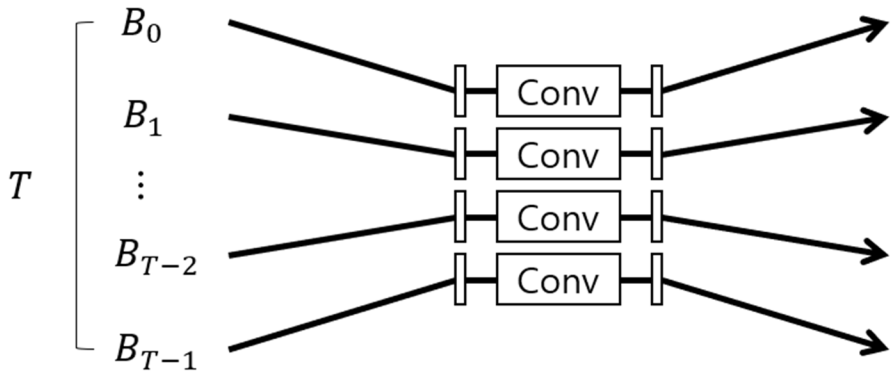


그림 4. 날실부의 얼개. 각각의 B_N에 해당하는 굵은 화살표는 U-Net을 의미하며, T개의 U-Net이 병렬로 나열되어 있다. 이로부터 여러 장의 연속된 영상을 동시에 받아 처리할 수 있게 된다. 그림 상으로 알 수 있듯이, 각각의 날실들은 아직 서로 어떠한 연결도 없기 때문에 인접 문맥이 반영되지 못한다.

날실부 (그림3)는 병렬로 쌓은 T개의 U-Net 기저모델 ($B_N, N = \{0, 1, \ldots, T-1\}$)로 구성된다. 이 때, 각 U-Net의 모든 학습 가능한 매개변수는 공유되도록 설계하여, 모든 U-Net이 일관성을 갖춘 채로 공간 특징을 추출하게 만든다. 이 구조는 한 장의 영상만을 처리하는 U-Net과는 다르게 T장의 연속된 영상에 대한 동시 처리를 가능하게 한다. 특히 날실부 내부에서는 데이터 증강을 위해 일반적으로 사용되는 무작위 섞기 (random shuffle) 기법을 사용하지 않는다. 이는 영상의 순서가 섞일 경우, 혈관의 연결성이라는 문맥이 파괴되므로 이를 학습하기 위해 도입된 씨실부가 정상적으로 학습되지 못하기 때문이다. 씨실부에 대해서는 2.1.3.에서 구체적으로 설명한다.

날실부의 부호기는 씨실부를 위해 연속된 영상의 공간적

특징들을 추출해 일렬로 나열하며, 복호기는 씨실부로부터 문맥을 아는 공간적 특징 (context-aware spatial feature)을 제공받아 보다 풍부한 정보량을 바탕으로 혈관 영역을 예측한다.

T는 설계자에 의해 자유롭게 설정될 수 있다. 하지만 T가 너무 크게 설정될 경우 메모리 요구량이 늘어나며, 또한 일반적으로 두 이미지의 거리가 멀어질수록 문맥적 연결성은 약화되기 때문에 비효율적일 수 있다. 선행실험을 통해 5를 초과하는 T값은 성능 향상을 이루지 못함을 확인했고, 이 결과를 바탕으로 모든 실험에서 T를 5로 설정했다.

2.1.3. 씨실부

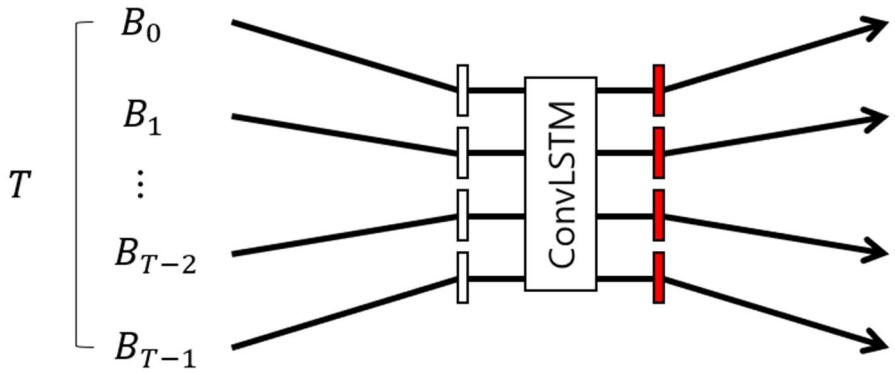


그림 5. 날실부에 씨실부를 적용한 거미형 U-Net의 완성된 얼개. 씨실부에서 존재했던 T개의 연결되지 않은 공간적 특징(spatial feature, 흰 막대기)들이 씨실부의 합성곱 장단기메모리를 통해 연결되어 문맥을 아는 공간적 특징(context-aware spatial feature, 붉은 막대기)들로 변화된다. 이는 복호기의 예측 단계에서 앞뒤로 인접한 영상의 특징을 참고하게 만들어 보다 정확한 결과를 얻을 수 있게 한다.

씨실부는 기 설명된 날실부의 T 개의 최하위 계층을

순환신경망 (recurrent neural network)으로 대체함으로써 구성(그림4)된다. 순환신경망은 본래 연결성이 존재하는 자연어, 음성 데이터 등에 대한 문맥 추출에 능한 신경망이다. 혈관 또한 혈액을 나르는 기능적 특성이 있기 때문에 기본적으로 서로 연결되어 있다는 연결성이 존재한다. 따라서 순환신경망을 수평면에 따라 적용하면 혈관의 연결성이라는 문맥 정보를 추가할 수 있다. 결과적으로 임의의 프레임의 혈관 영역을 예측할 때, 앞뒤로 인접한 영상의 정보까지 고려하게 되기 때문에 보다 부드럽고 정확한 추출이 가능해진다.

구체적으로는 순환신경망의 한 종류인 합성곱 기반의 장단기메모리 (LSTM, long short-term memory)를 활용했다 [16]. 먼저, 합성곱 기반의 순환신경망을 사용한 이유는, 해당 신경망은 합성곱 연산을 내부적으로 수행함으로써 관심영역 분할 시 필수적으로 요구되는 공간정보를 유지할 수 있도록 설계되었기 때문이다. 다음으로, 장단기메모리를 사용한 이유는, 해당 구조는 기억 소자 내 메모리의 안정적인 유지능력을 통해 순환신경망의 대표적인 모델군으로 자리잡았기 때문이다.

씨실부의 작동 과정은 다음과 같다. 먼저 날실부의 부호기로부터 연속된 T개 영상에 대한 공간적 특징들을 전달받는다. 이 공간적 특징들은 개별적으로 추출되었기 때문에 서로에 관한 문맥을 알지 못하는 상태이다. 씨실부는 이들을 장단기메모리에 통과시키고 문맥을 아는 공간 특징들로 변화시킨 뒤 복호기에 전달한다.

2.2. 데이터베이스

거미형 U-Net의 성능을 평가하기 위해 세 종류의 혈관 데이터베이스를 대상으로 실험을 진행했다. 특히 범용 혈관 영역추출 모델로의 가능성을 알아보기 위해 각각의 데이터베이스는 모달리티와 혈관 종류에 있어 다양성을 갖도록 구성했다. 모든 데이터베이스는 삼차원이며 수평면 (axial) 영상을 사용하였다.

2.2.1. 뇌 자기공명혈관조영영상

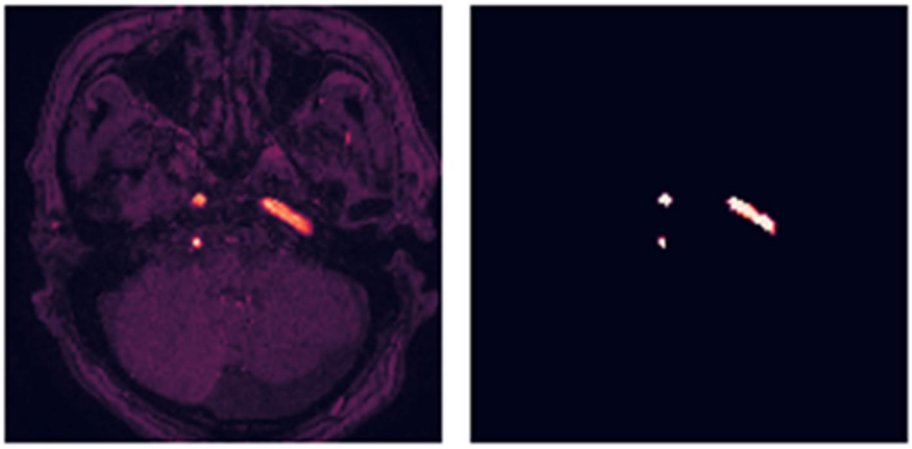


그림 6. Time-of-Flight 자기공명혈관조영술로 촬영한 뇌의 수평면 영상 (좌) 및 지도학습을 위해 전문의가 혈관 영역을 표기한 레이블(우).

두개 내 동맥의 협착 여부를 진단하기 위해 Time-of-Flight (TOF) 기법으로 촬영한 뇌 자기공명혈관조영 (MRA) 영상이다. 이 데이터베이스는 분당서울대학교병원 영상의학과에서 구축하였다. 먼저 2014년 7월부터 2019년 8월까지의 TOF MRA 촬영 명단을 수집하였다. 이후, 한 명의 영상의학과 전문의 및 한 명의 신경외과 전문의가 해당 촬영의 판독문을 재검토 하였고, 그 중 정상으로

구분된 경우만을 남겼다. 이후 무작위 추출을 통해 26건의 MRA 영상을 확보하였고 학습, 검증 및 평가를 위해 각각 18, 4, 4건으로 구분했다.

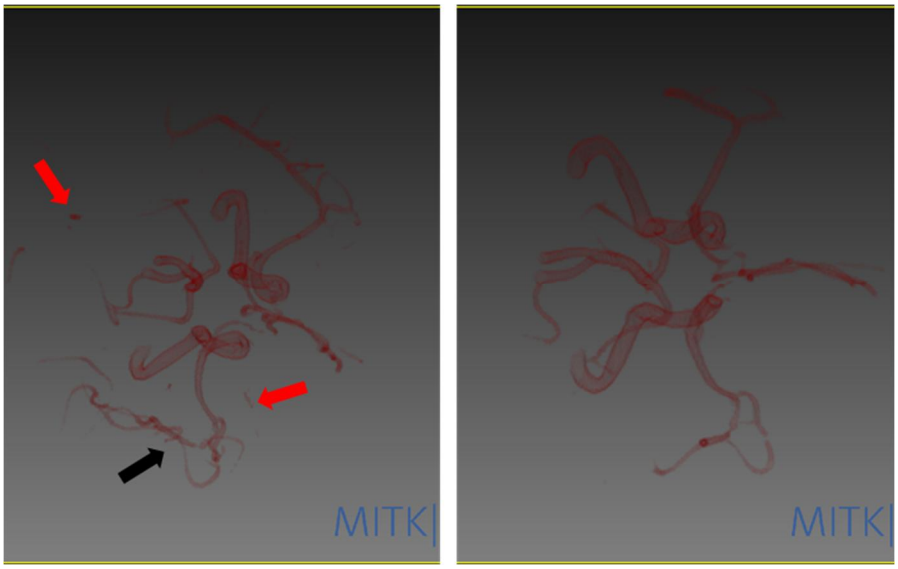


그림 7. 문턱값 방식을 적용해 추출한 혈관 영역(좌) 및 영상의학과 전문의에 의해 정돈된 최종 레이블(우). 문턱값 방식만을 적용할 경우 위양성 영역(붉은 화살표)이 표현되는 경우가 잦으며, 관심영역이 아닌 혈관(검은 화살표)을 걸러낼 수 없다.

제안할 모델의 학습 방법은 지도학습 (supervised learning)을 사용할 것이기 때문에 이에 해당하는 레이블을 마련하였다. 먼저 처음부터 혈관 영역을 그리는 수고를 덜기 위해 기존 혈관 영역추출에 사용되던 전통적 방법론 중 가장 간단한 문턱값 (threshold) 방식을 사용해 초벌 레이블링을 수행했다. 이는 특정 밝기 이상의 픽셀을 혈관으로, 그렇지 못한 픽셀을 비혈관으로 구분하는 단순한 알고리즘이므로 정확도가 높지 못하며 다음과 같은 문제점이 존재한다. 먼저 문맥과 전혀 관련없이 밝기가 높은 픽셀을 무조건적으로 혈관으로 인식하기 때문에 위양성 (false

positive) 픽셀들이 다수 발생한다. 다음으로 관심 영역이 아닌 미세한 혈관들도 일괄적으로 표현되기 때문에 후처리가 필요하다.

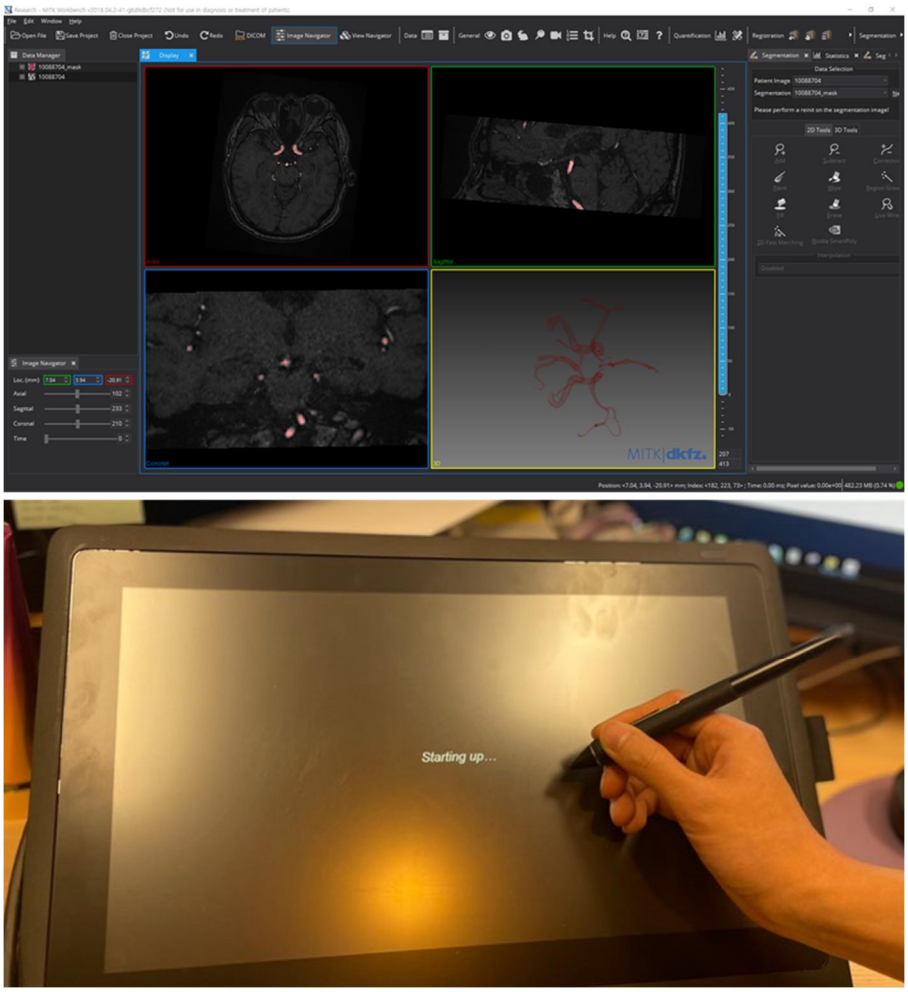


그림 8. 레이블 작성을 위해 활용한 Medical Imaging Interaction Toolkit (MITK) 프로그램의 인터페이스(상) 및 레이블링에 활용된 태블릿의 모습(하). MITK에서는 총 세 가지 뷰(axal, sagittal, coronal)로 영상을 확인하며 혈관 영역을 그릴 수 있고, 이를 삼차원으로 재구성해 편리하게 결과 검토가 가능하다.

위와 같이 초벌 레이블이 마련한 뒤, 영상의학과 전문의 두 명이 이를 바탕으로 관심영역에 해당하는 주요 혈관들을 그렸다. 이를 위해 Medical Imaging Interaction Toolkit (MITK) 이라는

의료영상 레이블링 목적의 공개 프로그램을 인터페이스로 채택해 사용하였고, 생산성 및 정확도를 높이기 위해 태블릿 두 대를 활용한 레이블링을 진행하였다.

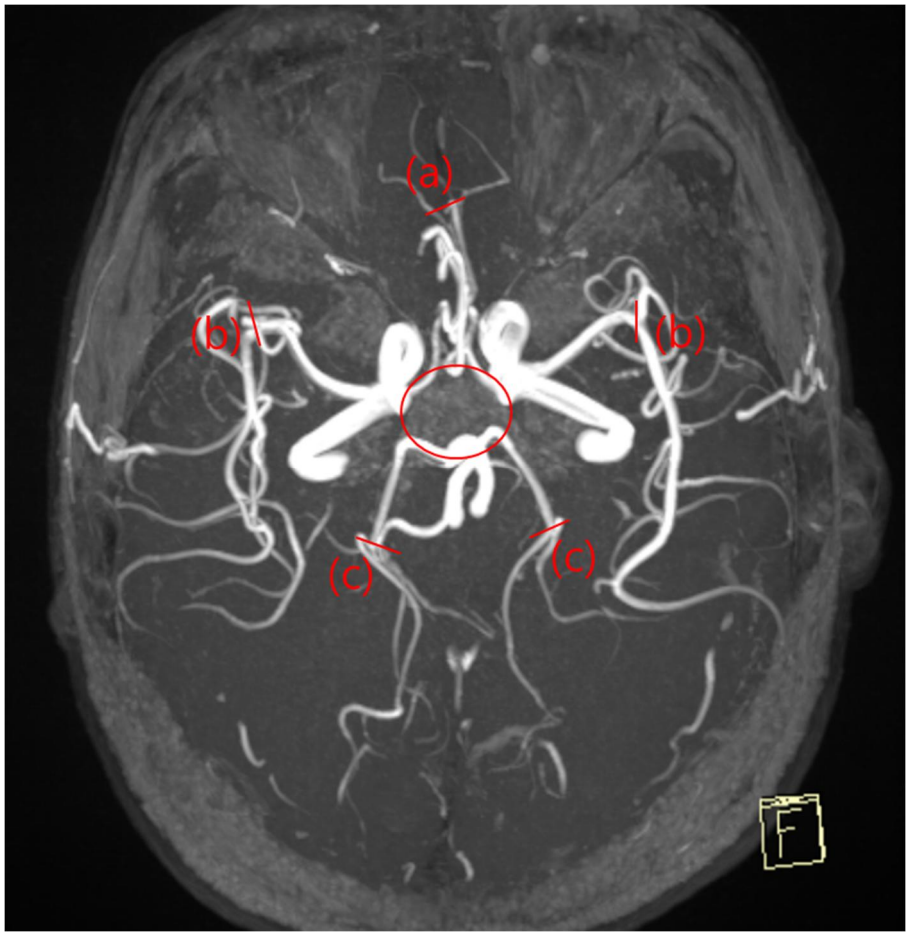


그림 9. 레이블링 대상의 대표 혈관을 정의하기 위한 TOF 데이터의 Maximum Intensity Projection (MIP) 예시 영상. 중앙의 대뇌동맥륜 (붉은 원)을 중심으로 퍼져나는 형태로 주요 혈관을 그려나갔고 붉은 선을 초과하는 영역인 A2, M2 그리고 P2 (각각 a, b, c)는 레이블로 취급하지 않았다.

모든 대표 혈관의 범위는 대뇌동맥륜을 중심으로 연결된 전대뇌동맥, 중대뇌동맥, 후대뇌동맥, 내경동맥 및 뇌기저동맥으로 한정지었다. 특히 전대뇌동맥, 중대뇌동맥 그리고 후대뇌동맥이

최초로 분기하는 지점인 A2, M2 그리고 P2 혈관 이후의 영역은 학습에 사용하지 않았다. 학습 시에는 혈관 영역을 중심으로 75%의 사각형 영역만을 잘라 사용했다.

2.2.2. 복부 전산화단층촬영영상

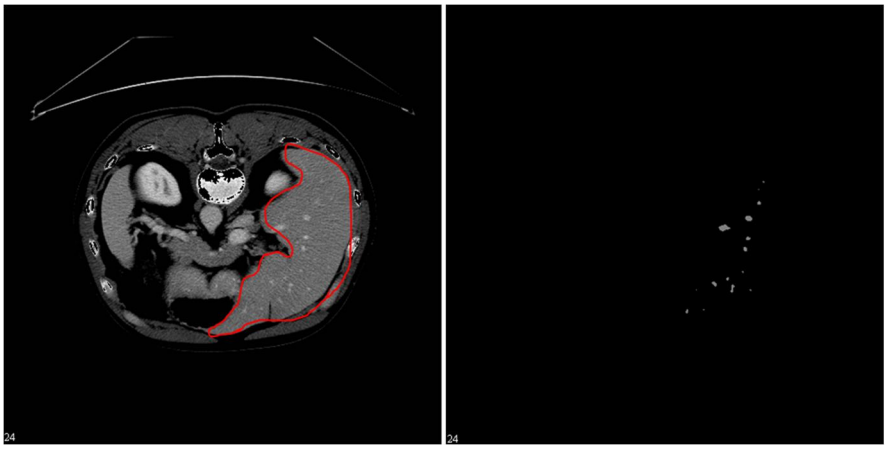


그림 10. 복부 전산화단층촬영 영상의 수평면 이미지 예시(좌) 및 간 내부 혈관을 표현한 레이블(우). 좌측 그림의 붉은 영역은 간의 대략적인 범위를 의미한다.

간 내부 종양 및 해당 병변 주위의 혈관을 진단하기 위해 시행된 복부 전산화단층촬영 (CT) 영상이다. 이 데이터베이스는 Decathlon Challenge 대회를 위해 제공된 다기관 공개 데이터베이스의 일부이다 [17]. 총 303건의 삼차원 CT 영상이 주어졌으며 이를 각각 학습, 검증 및 평가 용도로 사용하기 위해 253, 25, 25건으로 구분했다. 대회에서는 혈관을 비롯해 간 내부 종양 영역 또한 레이블로 제공되었지만 이 연구의 범위가 혈관 영역추출이기 때문에 종양에 해당하는 레이블은 제외하고 사용하였다. 또한, 혈관은 간 내부에 존재하기 때문에 불필요하게

학습에 사용되는 영역을 제거하기 위해 기존 512x512 크기의 영상을 간의 영역을 중심으로 잘라내었으며, 이를 256x256 크기로 일괄 변경하여 사용했다.

2.2.3. 심장 자기공명영상

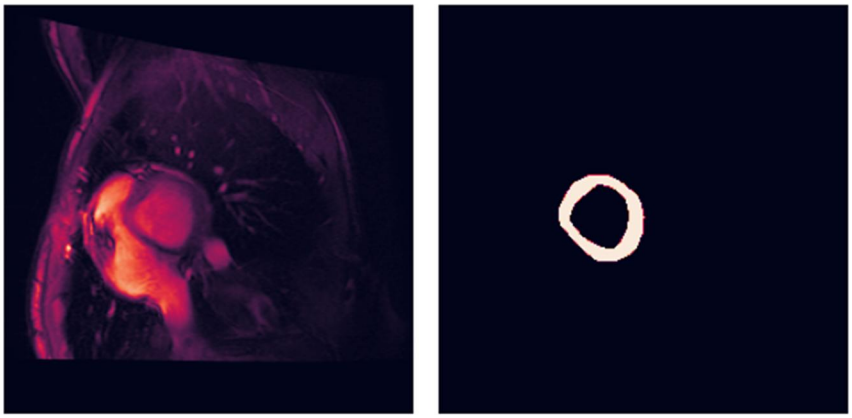


그림 11. 심장 자기공명영상의 수평면 이미지 예시(좌) 및 좌심실 벽 영역을 표현한 레이블(우).

심장의 건강상태 확인 및 비정상 심장박동의 관찰을 위해 시행된 자기공명영상 (MRI)이다. 이 데이터베이스는 York University에서 연구용으로 공개한 자료이다 [18]. 총 33건의 삼차원 MRI로 구성되어 있으며, 각 영상은 1회의 심장박동주기동안 촬영되었으며 20 프레임으로 구성되어 있다. 이를 27, 3, 3건으로 나누어 각각 학습, 검증 및 평가 용도로 사용하였다. 기 제공된 레이블은 좌심실의 외막과 내막의 경계를 나타내는 좌표였는데, 이로부터 외막과 내막 사이의 좌심실 벽 영역을 구해 최종 레이블로 사용하였다. 이후 불필요한 영역을 학습에서 제거하기 위해 심장을 중심으로 256x256 크기로 잘라 사용했다.

표 1. 학습 및 평가에 사용한 세 종류의 데이터베이스. 모델의 일반화 능력을 확인하기 위해 모달리티, 대상 장기 및 혈관이 모두 다르게 구성되었다.

Provider	분당서울대병원 영상의학과	King's College London 외 다수	York University
Modality	Brain MRA	Liver CT	Cardiac MRI
Target	두개 내 동맥	간혈관	좌심실벽
# of volumes	26	303	33

2.3. 비교 모델

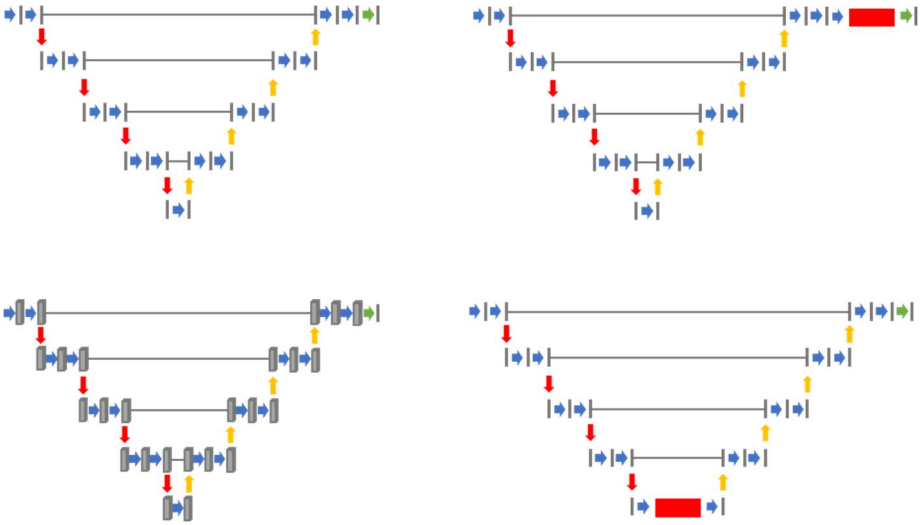


그림 12. 실험에 사용한 모든 U-Net 기반 모델의 간소화된 구조. 각각 이차원 U-Net (좌상), 삼차원 U-Net (좌하), FCN-RNN (우상) 그리고 거미형 U-Net (우하)을 의미한다. 푸른 화살표는 ReLU 함수로 활성화 된 합성곱 계층에 배치 정규화가 이루어진 연산이며, 붉은 화살표는 최대값 풀링, 노란색 화살표는 상행 합성곱(Upconvolution) 계층, 녹색 화살표는 소프트맥스 함수를 의미한다. LSTM은 붉은 상자로 표현되었다.

거미형 U-Net의 성능을 기존 모델과 비교하기 위한 세 가지의 비교 모델을 선정했다. 이차원 U-Net [10], 삼차원 U-Net [11], 그리고 FCN-RNN [14]이 그것이다. 먼저 앞의 두 모델은

영역분할을 위한 대표적인 이차원 및 삼차원 접근법으로 활용되고 있기에 선택하였다. FCN-RNN은 거미형 U-Net과 유사하게 의료영상에 대한 삼차원 영역분할을 위해 순환신경망을 다른 방식으로 도입한 기존의 접근법이다. 모든 모델은 2.1.1에서 정의한 U-Net 기저모델을 기반으로 구현되었다.

2.4. 구현 및 실험

모든 구현 및 실험은 Python 언어로 만들어진 심층신경망 라이브러리인 Keras와 Tensorflow를 활용해 진행되었다. 의료영상의 처리를 위해 오픈소스 패키지인 Pydicom과 Nibabel을 사용했다. 그래픽 처리 장치를 통한 학습 가속을 위해 Nvidia TITAN RTX 한 대를 사용했으며, 그래픽 처리 장치용 전문 소프트웨어인 CUDA 및 cuDNN를 활용했다.

손실 함수로는 가중치 조정을 통해 클래스 불균형이 있는 데이터베이스 학습에 효과적인 Focal Dice Loss [19]를 채택했고, 초기 학습률이 0.005인 Adam 최적화 방식을 적용했다. 타 데이터베이스로 초벌 학습을 시킨 뒤 본 학습을 진행하는 선행학습기법은 사용하지 않았으며, 10 에포크 (epoch) 이내에 검증 데이터베이스의 성능향상이 없을 경우 학습을 종료한 뒤, 가장 높은 성능을 기록한 에포크를 평가용 모델로 최종 선정하였다.

3. 결과 및 분석

3.1. 성능평가지표

정량적 성능평가를 위해 분할 모델의 평가에 일반적으로 사용되는 다이스 계수 (DSC) 및 자카드 계수 (IoU)를 각각 평가용 데이터베이스 전체에 대해 평균 내어 단일값으로 보고했다.

3.2. 정량적 성능평가

표 2. 거미형 U-Net 및 모든 비교모델에 대한 다이스 계수(좌) 및 자카드 계수(우) 성능. 모든 데이터베이스에 대해 기존의 모든 비교모델들의 성능을 능가한다.

	2D U-Net	3D U-Net	FCN-RNN	거미형 U-Net
뇌 MRA	0.745|0.673	0.716|0.617	0.752|**0.758**	**0.793**|0.743
복부 CT	0.327|0.277	0.145|0.152	0.351|0.271	**0.456**|**0.375**
심장 MRI	0.772|0.681	0.726|0.624	0.812|0.707	**0.833**|**0.728**

거미형 U-Net은 모든 데이터베이스에 대해 기존 비교모델들의 DSC를 능가했으며, 뇌 MRA의 IoU에 대해서만 FCN-RNN에 비해 약 1.5% 낮았다. 3D U-Net의 경우 모든 경우에서 가장 낮은 성능을 보였는데, 이는 수집한 의료영상 데이터베이스의 특성 상 영상마다 수평면 축에 대한 복셀 단위가 다르기 때문에 부피에 대한 이해에 혼동이 초래된 결과로 여겨진다 [20]. 특히 복셀 단위가 가장 다양하게 분포된 복부 CT에 대한 성능이 가장 낮다는 점을 보아도 이를 유추해볼 수 있다. 반대로 거미형 U-Net은 이에 관계 없이 높은 성능을 기록한 것으로 보아,

복셀 단위의 변화에 강인한 모습을 보이고 있다.

또한 순환신경망이 적용된 FCN-RNN의 성능이 거미형 U-Net의 성능에 버금가는 것으로 보았을 때, 삼차원 혈관 영역추출에 있어 순환신경망 활용이 효과적임을 미루어 짐작할 수 있다. 다양한 모델 조합으로부터 순환신경망과의 조합이 성능 상 가장 뛰어났다는 결과 또한 보고된 바 있다 [14].

3.3. 정성적 성능평가

정성적 성능평가를 위해 각 데이터베이스에 대한 평가용 영상 내 하나의 예시를 시각화하여 표현하였다. 모든 하위 항목의 그림에 대해 노란색은 영상의학과 전문의의 레이블과 일치하는 영역을, 붉은색은 위양성 영역을, 그리고 푸른색은 위음성 영역을 의미한다. 또한 그림에 삽입된 수치는 순서대로 해당하는 예시에 대한 평균 DSC 및 평균 IoU이다.

3.3.1. 뇌 자기공명혈관조영영상

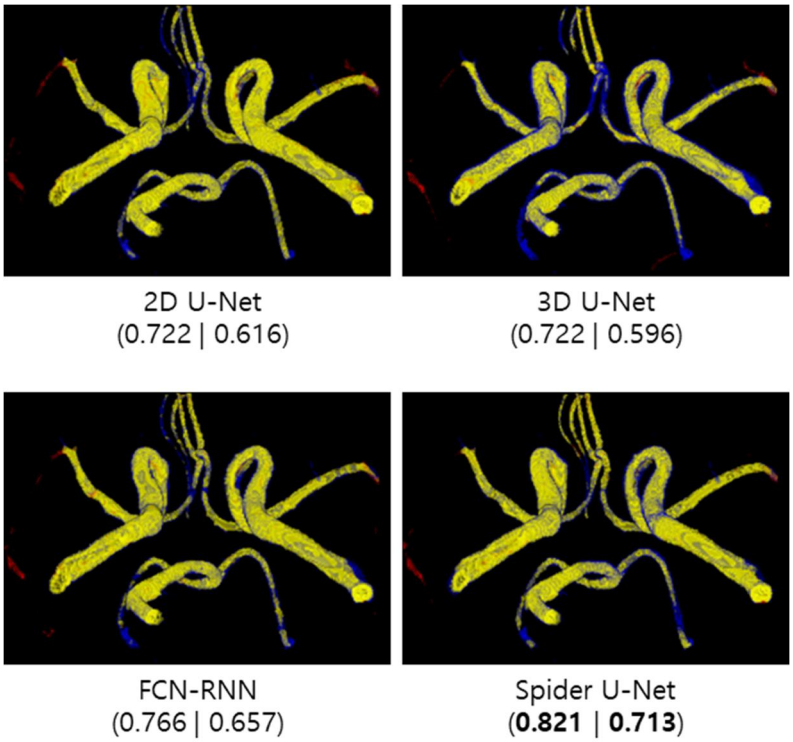


그림 13. 주요 뇌동맥에 대한 세 종류의 비교모델 및 거미형 U-Net의 삼차원 예측 결과 예시. 거미형 U-Net은 가장 높은 성능을 보였고 주요 혈관을 타 모델에 비해 정밀하게 예측하였다.

그림 13은 뇌 자기공명혈관조영영상으로부터 2.2.1.에서 정의한 주요 뇌동맥을 삼차원 영상으로 표현한 예시이다. 모든 모델은 가장 두꺼운 내경동맥 및 뇌기저동맥을 전체적으로 정확히 예측하였지만, 후대뇌동맥의 하행부에 대해서는 공통적으로 부정확한 모습을 보였다. 거미형 U-Net은 가장 높은 성능을 달성하였고 주요 혈관을 비교적 정밀하게 예측하였다.

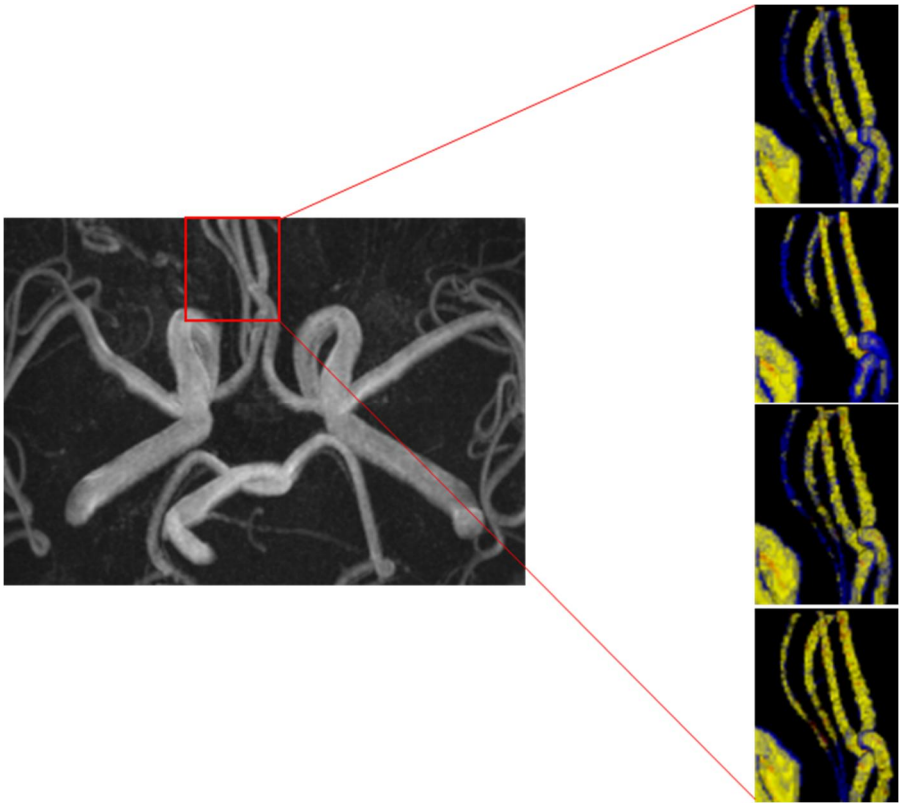


그림 14. 전대뇌동맥의 영역 추출 결과를 확대하여 나타낸 그림. 좌측의 흑백 그림은 원본 영상을 Maximum Intensity Projection을 통해 표현한 이차원 그림이며, 우측 그림은 위에서부터 순서대로 이차원 U-Net, 삼차원 U-Net, FCN-RNN 그리고 거미형 U-Net으로 예측한 전대뇌동맥 영역을 나타낸다.

특히, 주어진 혈관 중 전대뇌동맥은 가늘고 희미하게 표현되는 경향이 있어 특히 영역추출이 어려운 케이스이다. 그림 13을 보았을 때, 다른 모델들은 이 영역을 제대로 추출하지 못했으나, 거미형 U-Net은 유일하게 네 갈래의 전대뇌동맥을 모두 예측했을 뿐만 아니라 끊기지 않는 모습을 보여주며 해당 모델이 혈관의 연결성을 기반으로 인접 프레임 간 정보를 강화시키며 혈관영역을 추출하였다고 유추할 수 있는 근거가 된다.

3.3.2. 복부 전산화단층촬영영상

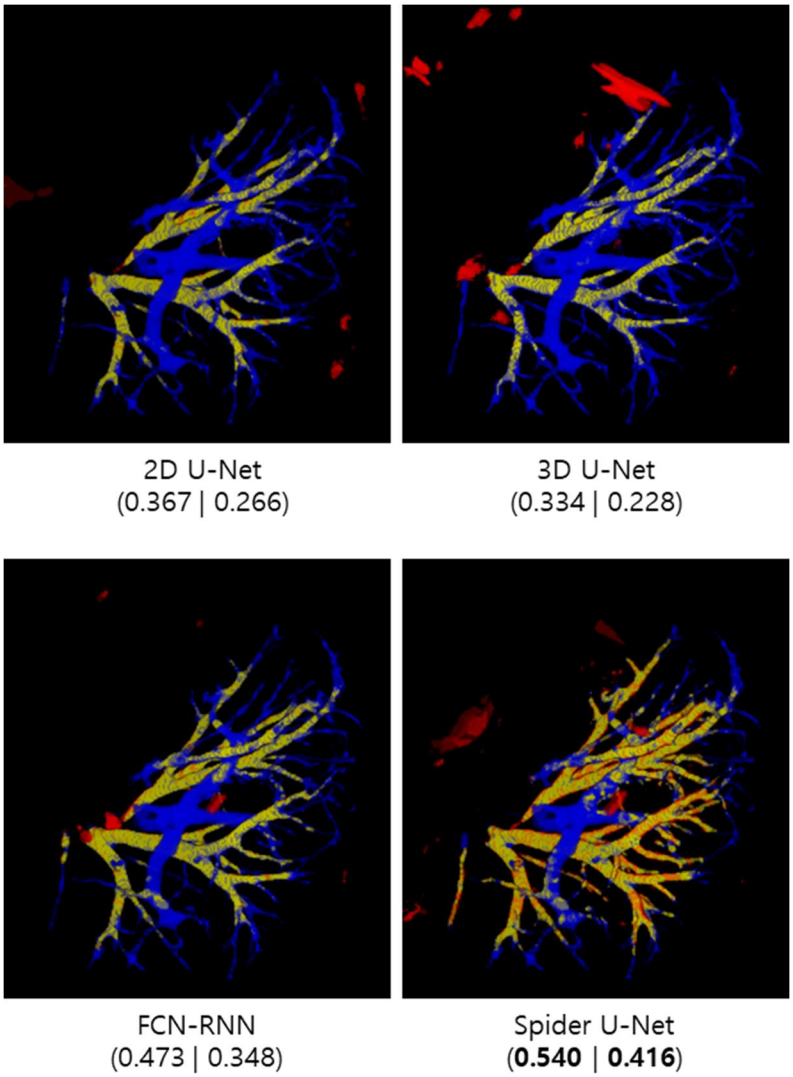


그림 15. 간 혈관에 대한 세 종류의 비교모델 및 거미형 U-Net의 삼차원 예측 결과 예시. 거미형 U-Net은 타 모델들에 비해 크게 높은 성능을 기록하였으며, 세부적인 혈관을 가장 잘 예측하였다.

그림 15는 복부 전산화단층촬영 영상에서 간 내부 혈관을 예측한 결과를 삼차원 영상으로 표현한 예시이다. 세 가지 비교 모델은 모두 두꺼운 혈관 위주로 예측이 진행되었으며, 미세한

혈관의 경우 배경으로 인식하며 수많은 위음성 영역을 남겼다. 또한 미세한 혈관을 옳게 예측한 경우에도 주위 혈관과의 연결이 끊어져 있는 모습이 다수 발견되어 영상의 문맥적 특징을 제대로 학습하지 못했다는 점을 드러낸다. 반면, 거미형 U-Net은 복잡한 간 혈관을 가장 세부적으로 예측하는 모습을 보이며 모든 모델에 대한 성능 상의 우위를 이루었다.

3.3.3. 심장 자기공명영상

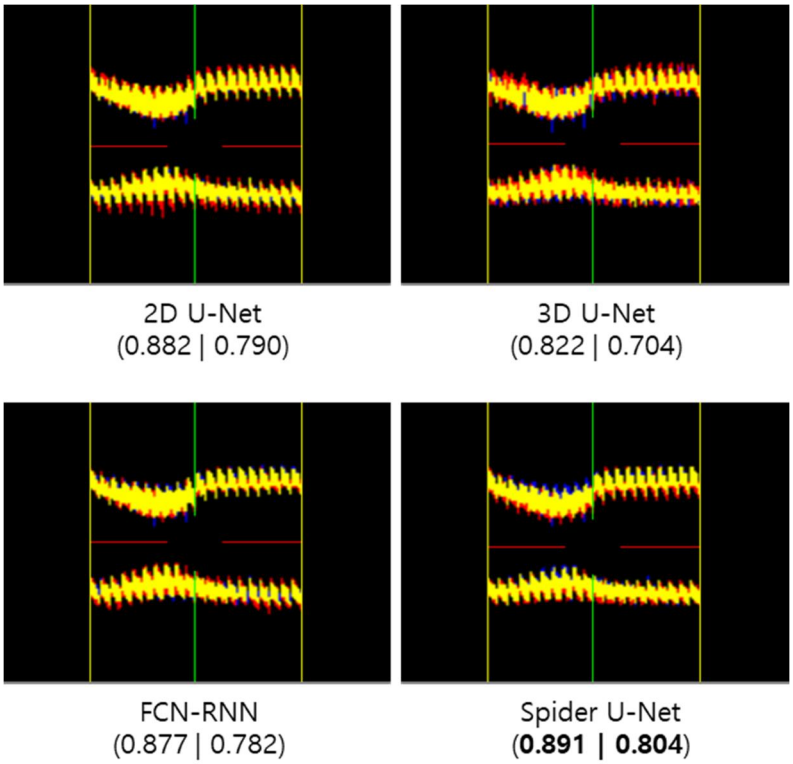


그림 16. 네 종류의 비교모델 및 거미형 U-Net의 삼차원 모델링 결과 예시. 노란색은 정답과 일치하는 영역, 붉은색은 위양성 영역, 그리고 파란색은 위음성 영역이다. 각 모델명 하단의 숫자는 왼쪽부터 각각 평균 다이스 계수 및 평균 자카드 계수를 의미한다.

그림 16은 심장 자기공명영상에서 단일 심장박동주기에

대한 좌심실 벽 영역의 변화를 수직면에 대한 단면으로 표현한 예시이다. 본 데이터베이스의 심장박동주기는 20장의 프레임으로 주어졌기 때문에 각 프레임에 대한 좌심실 벽 단면의 변화를 시간 순으로 연속되게 나타내었다. 삼차원 U-Net은 많은 위양성 영역을 예측하며 가장 불안정한 모습을 보였고, 나머지 모델들은 비교적 안정적인 모습을 보였다. 그 중에서도 거미형 U-Net은 가장 높은 성능을 기록했다.

4. 고찰

4.1. 실험 결과 고찰

삼차원 혈관영상추출을 심층신경망 기법으로 접근할 때에 LSTM을 활용할 경우 성능 향상을 노릴 수 있음을 알 수 있다. 표 2의 결과를 보았을 때 거미형 U-Net이 DSC 상 가장 높은 성능을 기록했고, 이와 유사하게 LSTM을 사용한 FCN-RNN이 그 뒤를 이으며 LSTM이 사용되지 않은 기존 모델들의 성능을 능가했다.

또한 LSTM을 사용할 경우 그것의 배치 또한 성능에 변화를 유발한다는 점을 알 수 있다. 본 연구에서 사용된 비교모델 중 하나인 FCN-RNN은 FCN 부분은 날실부로, RNN 부분을 씨실부로 구현하였다. 따라서 거미형 U-Net과의 유일한 차이점은 씨실부의 위치 뿐이다. 자세히 말하자면 FCN-RNN은 씨실부가 복호기의 후위에 배치되었으며 거미형 U-Net은 부호기와 복호기 사이에 배치되었다. 실험 결과로부터 거미형 U-Net의 성능이 전체적으로 높았다는 사실에서 미루어 볼 때, 씨실부에 해당하는 LSTM을 기존 모델의 후위에 배치하는 것 보다는 중심에 배치하는 것이 효과적임을 알 수 있다.

뇌 자기공명혈관조영영상 데이터베이스에 대한 거미형 U-Net의 정성적, 정량적 결과가 양호했기 때문에, 임상에의 적용 가능성을 추가로 확인해 볼 필요가 있다. 이를 위해 거미형 U-Net의 혈관영역추출 결과를 사용했을 때와 그렇지 않았을 때, 영상의학과 의료진들의 판독 정확도에 어떠한 영향을 미치는지를 밝혀내는 후속 연구를 진행할 예정이다.

복부 전산화단층활영 데이터베이스는 Z축에 대한 복셀

단위가 매우 다르다는 특징이 존재한다. 삼차원 U-Net의 성능이 굉장히 낮은 원인이 여기에 기인하는 것으로 보인다. 반면, 거미형 U-Net이 해당 데이터베이스에 대해서도 가장 높은 성능을 기록한 것으로 보았을 때, 이 모델은 복셀 단위에 대한 비등방성에 강인한 특징이 있음을 알 수 있다. 이는 촬영 장비나 환경에 따라 복셀 단위가 달라지는 경우가 매우 잦아지는 특징을 가진 삼차원 의료영상을 다룰 때 큰 장점으로 발휘될 수 있다.

4.2. 알고리즘 고찰

거미형 U-Net은 구조적으로 이차원 U-Net 및 삼차원 U-Net의 중간 단계에 위치한다고 말할 수 있다. 다시 말해, 거미형 U-Net은 기본적으로 평면 정보를 추출하되 부피 정보를 놓치지 않기 위해 LSTM을 Z축에 대해 연결해 성긴 (Sparse) 부피 정보를 획득할 수 있는 구조를 갖는다. 따라서 평면 정보만을 추출하는 이차원 U-Net보다 풍부한 특징을 추출할 수 있으며, 빽빽한 (Dense) 부피 정보를 추출하는 삼차원 U-Net보다 적은 메모리 사용량을 요구한다. 동시에 복셀 단위에 강인한 특징을 가지며, 실험적으로 가장 높은 성능 또한 기록했기에 삼차원 혈관영역 추출에 도입할 수 있는 효율적인 모델이 될 수 있을 것이다.

5. 결론

본 연구는 심층신경망을 활용한 삼차원 혈관 영역추출의 성능향상을 위해 새로운 모델을 제시했다. 제안된 거미형 U-Net은 영역이 좁고 복잡하지만 문맥적 연결성을 가졌다는 혈관의 내재적인 특징을 반영함으로써 혈관 영역추출을 효과적으로 수행할 수 있는 모델이다.

이를 구현하기 위한 핵심 요소는 합성곱 LSTM으로 구현된 씨실부이다. 씨실부는 날실부의 부호기와 복호기 사이에 삽입되어, 혈관 마스크의 예측을 시작할 때 문맥을 아는 공간적 특징을 복호기에 제공함으로써 보다 정확한 결과를 만들어내는 데에 일조한다.

거미형 U-Net은 씨실부를 통해 이차원 U-Net의 한계인 부피 정보의 부재를 완화했고, 삼차원 U-Net보다 적은 메모리를 요구하며, 유사한 접근법을 취한 FCN-RNN보다 전체적인 성능이 높은 모습을 보여주었다. 이러한 이유로 삼차원 혈관영역 추출에 있어 효과적인 모델임을 알 수 있다.

참고 문헌

[1] S. Dangi, Z. Yaniv, and C. A. Linte, "Left Ventricle Segmentation and Quantification from Cardiac Cine MR Images via Multi-task Learning," (in eng), *Stat Atlases Comput Models Heart,* vol. 11395, pp. 21-31, 2019.

[2] T. Matsushige *et al.*, "Ventricular Microaneurysms in Moyamoya Angiopathy Visualized with 7T MR Angiography," *American Journal of Neuroradiology,* vol. 37, no. 9, pp. 1669-1672, 2016, doi: 10.3174/ajnr.A4786.

[3] M. Javidi, H.-R. Pourreza, and A. Harati, "Vessel segmentation and microaneurysm detection using discriminative dictionary learning and sparse representation," *Computer methods and programs in biomedicine,* vol. 139, pp. 93-108, 2017.

[4] B. Zhang, F. Karray, Q. Li, and L. Zhang, "Sparse Representation Classifier for microaneurysm detection and retinal blood vessel extraction," *Information Sciences,* vol. 200, pp. 78-90, 2012/10/01/ 2012, doi: https://doi.org/10.1016/j.ins.2012.03.003.

[5] Y. Cheng, X. Hu, J. Wang, Y. Wang, and S. Tamura, "Accurate Vessel Segmentation With Constrained B-Snake," *IEEE Transactions on Image Processing,* vol. 24, no. 8, pp. 2440-2455, 2015, doi: 10.1109/TIP.2015.2417683.

[6] S.-H. Lee and S. Lee, "Adaptive Kalman snake for semi-autonomous 3D vessel tracking," *Computer Methods and Programs in Biomedicine,* vol. 122, no. 1, pp. 56-75, 2015/10/01/ 2015, doi: https://doi.org/10.1016/j.cmpb.2015.06.008.

[7] V. Caselles, F. Catté, T. Coll, and F. Dibos, "A geometric model for active contours in image processing," *Numerische Mathematik,* vol. 66, no. 1, pp. 1-31, 1993/12/01 1993, doi: 10.1007/BF01385685.

[8] R. Phellan, A. Peixinho, A. Falcão, and N. D. Forkert, *Vascular Segmentation in TOF MRA Images of the Brain Using a Deep Convolutional Neural Network*. 2017, pp. 39-46.

[9] B. Wang, S. Qiu, and H. He, "Dual Encoding U-Net for Retinal Vessel Segmentation," Cham, 2019: Springer International Publishing, in Medical Image Computing and Computer Assisted Intervention – MICCAI 2019, pp. 84-92.

[10] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," Cham, 2015: Springer International Publishing, in Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, pp. 234-241.

[11] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O.

Ronneberger, "3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation," Cham, 2016: Springer International Publishing, in Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016, pp. 424-432.

[12] M. Livne *et al.*, "A U-Net Deep Learning Framework for High Performance Vessel Segmentation in Patients With Cerebrovascular Disease," (in English), *Frontiers in Neuroscience,* Original Research vol. 13, no. 97, 2019-February-28 2019, doi: 10.3389/fnins.2019.00097.

[13] M. Haft-Javaherian, L. Fang, V. Muse, C. B. Schaffer, N. Nishimura, and M. R. Sabuncu, "Deep convolutional neural networks for segmenting 3D in vivo multiphoton images of vasculature in Alzheimer disease mouse models," *PLOS ONE,* vol. 14, no. 3, p. e0213539, 2019, doi: 10.1371/journal.pone.0213539.

[14] J. Chen, L. Yang, Y. Zhang, M. Alber, and D. Z. Chen, "Combining fully convolutional and recurrent neural networks for 3d biomedical image segmentation," in *Advances in neural information processing systems*, 2016, pp. 3036-3044.

[15] Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," *IEEE Transactions on Neural Networks,* vol. 5, no. 2, pp. 157-166, 1994, doi: 10.1109/72.279181.

[16] S. Xingjian, Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, and W.-c. Woo, "Convolutional LSTM network: A machine learning approach for precipitation nowcasting," in *Advances in neural information processing systems*, 2015, pp. 802-810.

[17] A. L. Simpson *et al.*, "A large annotated medical image dataset for the development and evaluation of segmentation algorithms," *arXiv preprint arXiv:1902.09063,* 2019.

[18] A. Andreopoulos and J. K. Tsotsos, "Efficient and generalizable statistical models of shape and appearance for analysis of cardiac MRI," *Medical Image Analysis,* vol. 12, no. 3, pp. 335-357, 2008/06/01/ 2008, doi: https://doi.org/10.1016/j.media.2007.12.003.

[19] P. Wang and A. C. S. Chung, "Focal Dice Loss and Image Dilation for Brain Tumor Segmentation," Cham, 2018: Springer International Publishing, in Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, pp. 119-127.

[20] F. Isensee *et al.*, "nnu-net: Self-adapting framework for u-net-based medical image segmentation," *arXiv preprint arXiv:1809.10486,* 2018.

Abstract

Development and validation of a deep learning algorithm using recurrent neural network for blood vessel segmentation of 3D medical image

Kyeo Rye Lee
Interdisciplinary Program in Bioengineering
The Graduate School
Seoul National University

3D blood vessel segmentation (3D BVS) is an important prerequisite for various cardiovascular disease diagnosis. 3D BVS is difficult because the blood vessel has complicated structure and is narrower than other organs. We propose a novel deep learning framework to improve performance by emphasizing inter-slice context that the blood vessels are connected through adjacent slices. We implemented a framework from U-Net with two structural modifications. First,

the baseline is duplicated several times in parallel to extract features of adjacent slices simultaneously. Second, to weave unconnected contexts between extracted features, a LSTM layer is incorporated between encoder and decoder. Consequently, the spatial information from x−y plain and the inter−slice context from z−axis make the segmentation masks smooth and accurate. Experiment with three 3D BVS datasets shows Spider U−Net outperforms whole representative models in average dice score.

Keywords: Deep learning, 3D blood vessel segmentation, convolutional neural network, recurrent neural network, inter−slice context learning

Student Number: 2019−22708