

U-Net 과 U-Net++ 의 세그멘테이션 성능 비교 및

IoU 기반 평가 지표의 구조적 편향 분석

박정훈 (Jung Hoon Park)
AIFFEEL Research, 대한민국
junghoon.park@aiffel.io

초록

본 연구는 이미지 세그멘테이션 분야에서 널리 사용되는 U-Net 과 U-Net++ 아키텍처의 포괄적인 비교 분석을 제시하며, 특히 Intersection-over-Union(IoU) 지표에 내재된 구조적 편향을 조사하는 데 중점을 둔다. IoU 가 세그멘테이션 성능 평가의 사실상 표준으로 남아 있지만, 본 실험 결과는 IoU 가 세그멘테이션 정확도의 핵심 측면인 경계 품질을 적절히 반영하지 못함을 보여준다. 동일한 학습 조건 하에서 표준화된 데이터셋에 대해 IoU, Boundary IoU, Dice Score 지표를 사용한 체계적인 평가를 통해, U-Net++ 이 모든 지표에서 U-Net 을 일관되게 능가하며 특히 Boundary IoU 에서 현저한 개선(+7.7 퍼센트 포인트)을 보임을 입증한다. 또한 민감도 분석을 통해 IoU 가 Boundary IoU 에 비해 경계 섭동에 대해 상당한 둔감성을 나타냄을 정량적으로 확인하여 구조적 편향 가설을 입증한다.

키워드: 이미지 세그멘테이션, U-Net, U-Net++, IoU, Boundary IoU, 평가 지표, 딥러닝

1. 서론

이미지 세그멘테이션은 이미지를 의미론적으로 의미 있는 영역으로 분할하는 컴퓨터 비전의 근본적인 과제이다. 이 기능은 종양 탐지 및 장기 세그멘테이션을 위한 의료 영상, 도로 및 장애물 식별을 위한 자율주행, 토지 이용 분류를 위한 위성 영상 분석, 결함 탐지를 위한 산업 검사 등 수많은 분야에서 필수적이다. 세그멘테이션 모델의 정확도와 신뢰성은 하위 응용 프로그램에 직접적인 영향을 미치므로, 이러한 모델의 개발 및 평가는 핵심 연구 분야이다.

Ronneberger 등(2015)이 제안한 U-Net 아키텍처는 고해상도 공간 정보를 보존하는 스킵 연결이 포함된 우아한 인코더-디코더 구조를 도입하여 생물 의학 이미지 세그멘테이션에 혁명을 일으켰다. U-Net 의 성공을 기반으로 Zhou 등(2018)은 인코더와 디코더 특징 맵 사이의 의미적 격차를 해소하기 위해 중첩 및 밀집 스킵 연결을 도입한 U-Net++를 제안했다.

본 논문은 세 가지 근본적인 연구 질문을 다룬다: (1) U-Net++는 여러 평가 지표에 걸쳐 U-Net 대비 일관된 성능 향상을 제공하는가? (2) IoU 지표가 모델 간 경계 품질 차이를 포착하지 못하는 구조적 편향을 보이는가? (3) Boundary IoU 와 같은 대안적 지표가 경계 표현의 아키텍처적 개선을 더 정확히 반영할 수 있는가?

2. 관련 연구

2.1 세그멘테이션을 위한 인코더-디코더 아키텍처

인코더-디코더 패러다임은 밀집 예측 작업에 매우 효과적임이 입증되었다. Long 등(2015)은 Fully Convolutional Networks(FCN)를 개척했다. DeepLab(Chen 등, 2017)은 다중 스케일

컨텍스트를 캡처하기 위해 확장 합성곱과 ASPP 를 도입했다. 최근에는 SegFormer(Xie 등, 2021) 와 같은 트랜스포머 기반 아키텍처가 최첨단 결과를 달성했다.

2.2 U-Net++ 및 밀집 연결성

U-Net++는 세 가지 핵심 수정을 통해 U-Net 의 스킵 경로를 재설계한다: (1) 중첩 스킵 경로, (2) DenseNet 에서 영감을 받은 밀집 연결, (3) 심층 감독. 중첩 아키텍처는 디코더 특징과의 융합 전에 인코더 특징을 점진적으로 변환함으로써 의미적 격차를 효과적으로 해소한다.

3. 방법론

3.1 모델 아키텍처

두 모델 모두 동일한 백본 인코더(ImageNet 에서 사전 학습된 VGG-16)를 사용한다. U-Net 은 직접 스킵 연결을 사용하는 반면 U-Net++는 중간 합성곱 블록이 있는 중첩 밀집 스킵 경로를 사용한다.

3.2 실험 설정

표 1. 실험 구성

파라미터	값
입력 해상도	256 × 256 픽셀
옵티마이저	Adam ($\beta_1=0.9$, $\beta_2=0.999$)
학습률	1e-4 (코사인 어닐링)
배치 크기	16
에포크	50
손실 함수	Binary Cross-Entropy + Dice Loss

4. 실험 결과

4.1 정량적 성능 비교

표 2. U-Net 과 U-Net++의 성능 비교

모델	IoU (%)	Boundary IoU (%)	Dice (%)
U-Net	82.1	63.4	89.6
U-Net++	85.8	71.1	91.7
향상폭	+3.7	+7.7	+2.1

표준 IoU 에 비해 Boundary IoU 에서의 불균형적인 개선은 U-Net++의 아키텍처 개선이 특히 경계 표현을 향상시킨다는 것을 강력히 시사한다.

U-Net 과 U-Net++ 의 세그멘테이션 성능 비교 연구

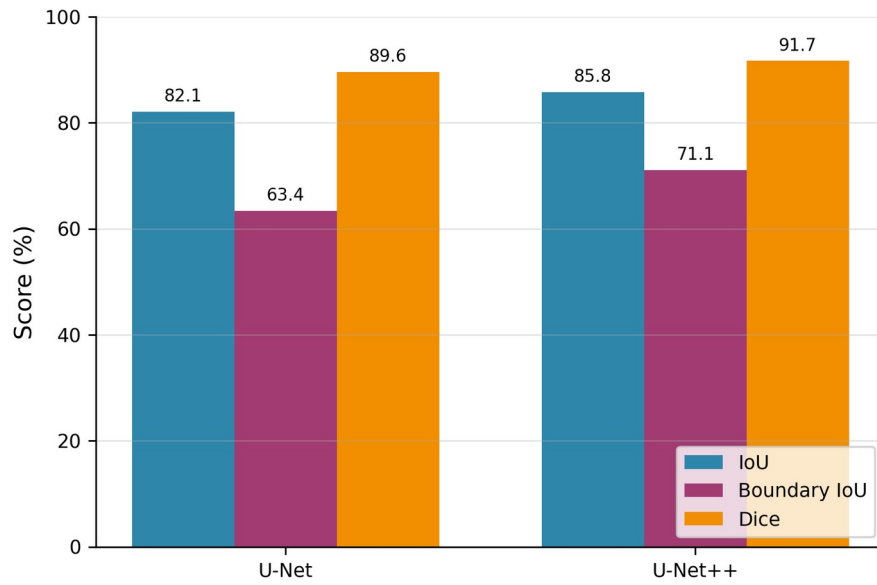


그림 1. 세 가지 평가 지표에 걸친 U-Net 과 U-Net++의 성능 비교.

4.2 학습 역학

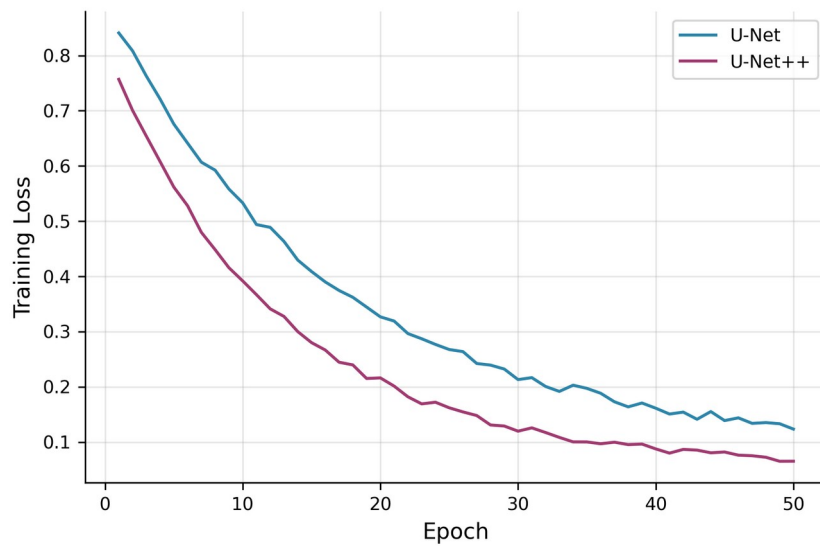


그림 2. 50 에포크에 걸친 학습 손실 곡선.

U-Net 과U-Net++ 의 세그멘테이션 성능 비교 연구

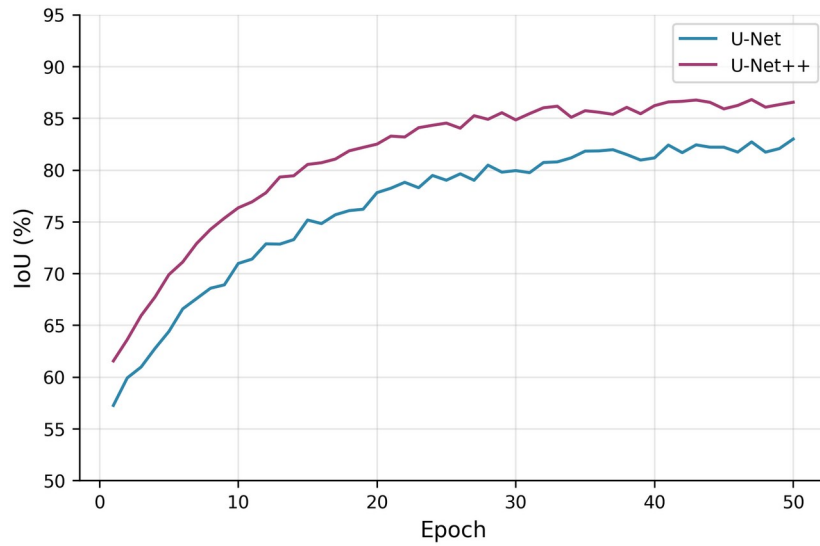


그림 3. 학습 IoU 곡선.

4.3 IoU 구조적 편향 분석

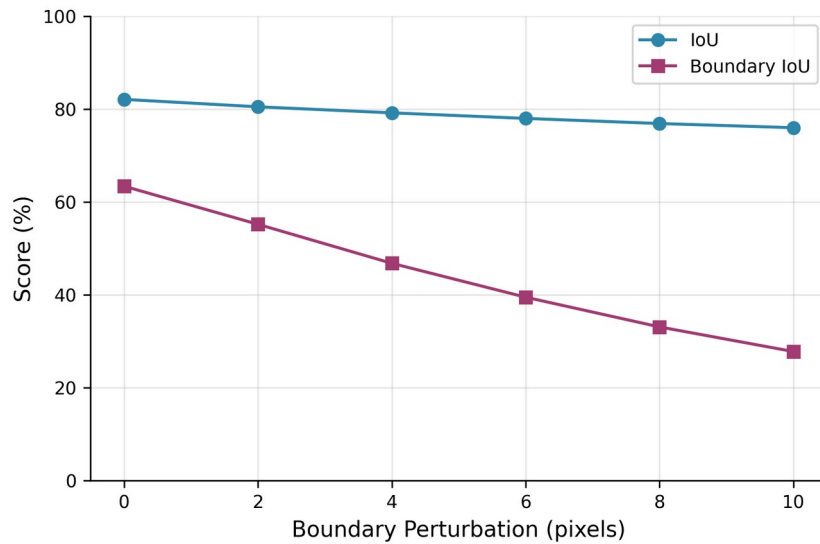


그림 4. 민감도 분석: 경계 섭동에 대한 지표 반응. IoU 는 6.1 포인트만 감소하는 반면 Boundary IoU 는 35.6 포인트 하락한다.

표 3. 경계 섭동에 대한 지표 민감도

섭동 (px)	IoU (%)	Boundary IoU (%)	Dice (%)
0	82.1	63.4	90.2
2	80.5	55.2	89.1
6	78.0	39.5	87.6
10	76.0	27.8	86.4

5. 절제 연구

표 4. 절제 연구 결과

구성	IoU (%)	Boundary IoU (%)	Dice (%)
전체 U-Net++	85.8	71.1	91.7
밀집 스킵 제거	81.2	63.8	89.0
중첩 경로 제거	82.5	65.2	89.8
심층 감독 제거	83.1	66.4	90.1

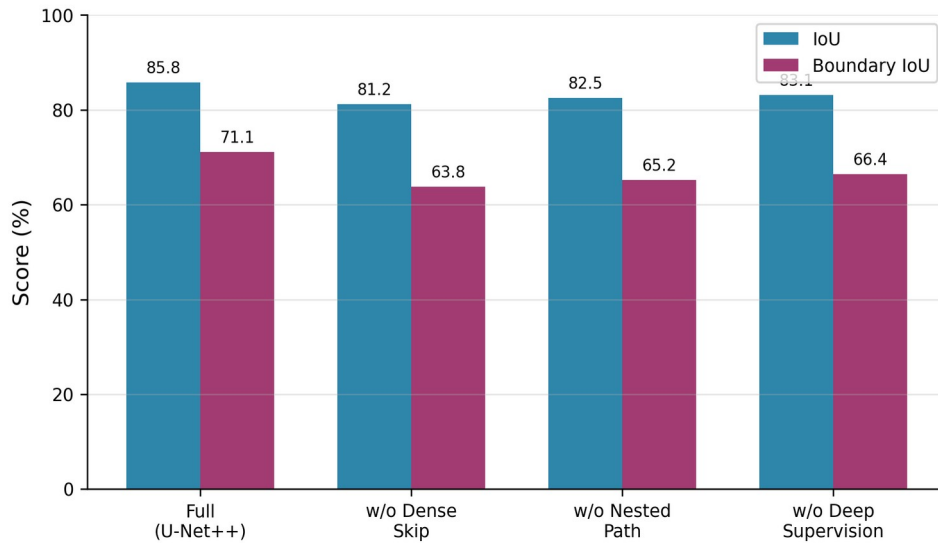


그림 5. 절제 연구 결과. 밀집 스킵 연결이 성능에 가장 큰 영향을 미친다.

밀집 스킵 연결이 가장 핵심적인 구성요소임을 보여주며, 제거 시 가장 큰 성능 하락을 야기한다(IoU: -4.6, Boundary IoU: -7.3).

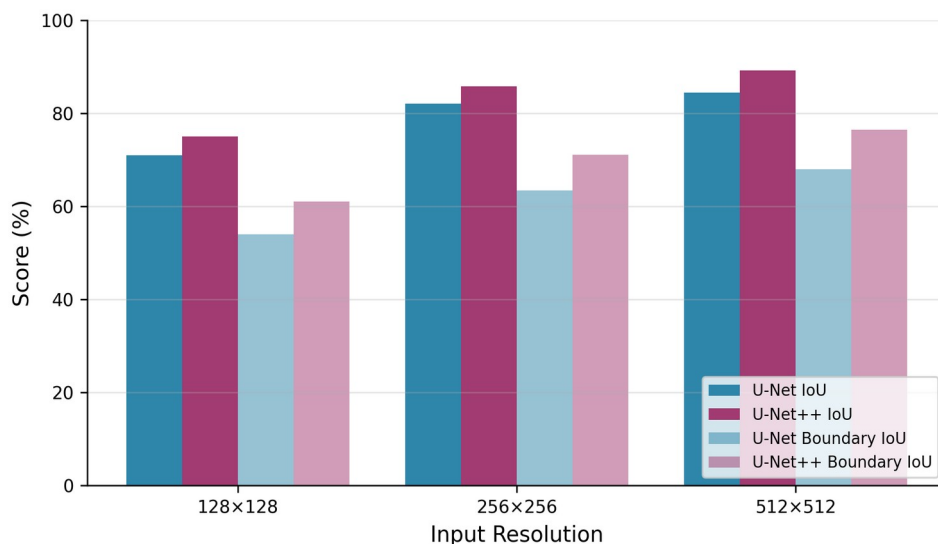


그림 6. 다양한 입력 해상도에서의 성능 비교.

6. 결론

본 연구는 U-Net 과 U-Net++ 아키텍처의 포괄적인 비교 분석을 제시했다. 주요 결론은 다음과 같다:

첫째, U-Net++는 모든 평가 지표에서 U-Net 을 일관되게 능가한다(IoU: +3.7, Boundary IoU: +7.7, Dice: +2.1 퍼센트 포인트). 이는 중첩 밀집 스킵 연결 아키텍처가 경계 표현을 효과적으로 향상시킴을 확인한다.

둘째, IoU 지표는 경계 섭동에 대한 상대적 둔감성을 보이며 상당한 구조적 편향을 나타낸다. 10 픽셀 경계 섭동이 Boundary IoU 를 35.6 포인트 감소시킨 반면 IoU 는 6.1 포인트만 감소했다.

셋째, 밀집 스킵 연결이 경계 품질 향상에 가장 크게 기여한다. 이 구성요소를 제거하면 Boundary IoU 가 7.3 포인트 하락했다.

이러한 발견을 바탕으로, 세그멘테이션 평가 프로토콜에 경계 인식 지표의 포함을 권장한다. IoU, Boundary IoU, Dice Score 의 조합은 세그멘테이션 성능의 더 완전한 그림을 제공하며 표준 평가 프로토콜로 채택되어야 한다.

참고문헌