

DeepLabV3+를 이용한 고해상도 위성영상에서의 도시 변화탐지

송창우*, Wiratama Wahyu**

*주식회사 컨텍 주임연구원

**주식회사 컨텍 선임연구원

{changwoo.song, wiratama.wahyu}@contec.kr

Urban Change Detection for High-resolution Satellite Images using DeepLabV3+

Chang-Woo Song*, Wiratama Wahyu**

* Assistant Research Engineer, CONTEC Co., Ltd

** Senior Research Engineer, CONTEC Co., Ltd

요약

본 논문에서는 고해상도의 시계열 위성영상을 딥러닝 알고리즘으로 학습하여 도시 변화탐지를 수행한다. 고해상도 위성영상을 활용한 서비스는 4 차 산업혁명 융합 신사업 중 하나인 스마트시티에 적용하여 도시 노후화, 교통 혼잡, 범죄 등 다양한 도시 문제 해결 및 효율적인 도시를 구축하는데 활용이 가능하다. 이에 본 연구에서는 도시 변화탐지를 위한 딥러닝 알고리즘으로 DeepLabV3+를 사용한다. 이는 인코더-디코더 구조로, 공간 정보를 점진적으로 회복함으로써 더욱 정확한 물체의 경계면을 찾을 수 있다. 제안하는 방법은 DeepLabV3+의 레이어와 loss function 을 수정하여 기준보다 좋은 결과를 얻었다. 객관적인 성능평가를 위해, 공개된 데이터셋 LEVIR-CD 으로 학습한 결과로 평균 IoU 는 0.87, 평균 Dice 는 0.93 을 얻었다.

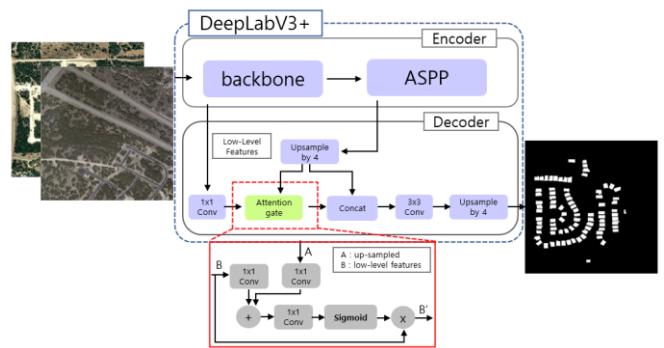
1. 서론

우주 관련 사업이 정부 주도에서 민간 주도의 사업으로 바뀌고 있다. 최근 민간 주도의 우주 산업 특징은 빅데이터와 인공지능(AI)를 활용하여 도심지, 대형 마트 등의 위성영상을 분석하는 서비스가 각광을 받고 있다. 특히 1m 이하의 고해상도 위성영상의 공급이 증대됨에 따라 이를 활용한 도시 변화 분석 기술의 연구가 활발히 이루어지고 있다.[1] 이는 4 차 산업 혁명 융합 신사업 중 하나인 스마트시티에 적용하여 도시 노후화, 교통 혼잡, 자연재해 등 다양한 도시 문제 해결 및 도시 계획 등에 활용이 가능하다. 위성 영상은 접근하기 어려운 지역에서 일어나는 변화정보를 얻을 수 있기에 원격 탐사 분야에서 각광을 받는 주제이다. 이는 도시 개발이 근접한 미개발지역으로 확산되는 현상이나 지진에 대한 모니터링 등 크고 복잡한 도시의 문제점들을 해결해 나가는데 도움을 주고 있다.[2]

2. DeepLabV3+

DeepLabV3+는 그림 1 과 같이 인코더-디코더 구조

이다. 인코더 과정에서는 입력 영상을 다운-샘플링 하여 특징을 추출하며, 디코더 과정에서는 업-샘플링을 통해 labeled map 을 만든다.



(그림 1) 제안하는 A-DeepLabV3+ 전체 과정.

DeepLabV3+는 기존 convolution 과 달리 커널 사이에 간격을 두는 atrous convolution 을 사용한다. atrous convolution 은 다운 샘플링 과정에서 일어날 수 있는 상세한 정보의 유실을 줄여 모델의 성능을 높일 수 있다. 또한, ASPP(Atrous Spatial Pyramid Pooling)를 적용

하여 다양한 크기의 정보를 담을 수 있게 구성하였다.[3] 마지막으로 디코더 과정을 통해 원본 이미지와 동일한 해상도로 복원되며, 최종 결과를 예측한다.

3. 제안하는 방법

제안하는 방법인 A-DeepLabV3+은 기존의 디코더 과정에 그림 1과 같이 Attention gate[4]을 추가한다. Attention gate은 입력 영상에 두드러진 특징을 강조하면서 관련 없는 영역을 억제하는 방법이다.[4] 그림 1에서 표현된 Attention gate는 인코더 과정에서 추출된 low-level features을 정제하는 과정을 나타낸다.

BCE (Binary Cross Entropy)는 binary segmentation에서 자주 쓰이는 loss function이다. 또한, 일반적으로 segmentation에서 쓰이는 지표는 Dice 와 IoU (Intersection of Union)이다. 더욱 정확한 map을 획득하기 위해 식 1과 같이 다양한 지표를 통한 loss function을 사용한다.

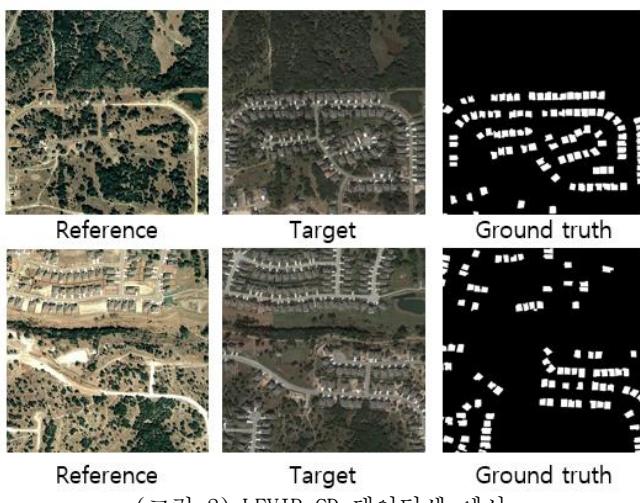
$$\mathcal{L} = \mathcal{L}_{cls} + (1 - Dice) + (1 - IoU)$$

<식 1> 제안하는 Loss function

식 1에서 \mathcal{L}_{cls} 는 BCE를 의미한다.

4. 실험 결과 및 결론

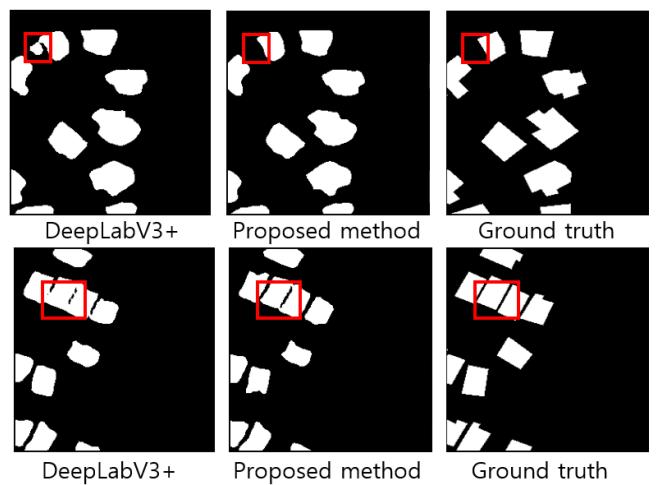
객관적인 성능평가를 위해, 본 연구에서는 LEVIR-CD[5] 데이터셋을 사용한다. 그림 2는 LEVIR-CD 데이터셋의 예시를 나타낸다. 데이터셋은 동일 영역의 시계열 영상들과 정답 영상들로, 1024 x 1024의 해상도를 가진다. 본 연구에서는 메모리의 효율적인 사용을 위해 256 x 256으로 나누어진 데이터셋을 사용한다.



(그림 2) LEVIR-CD 데이터셋 예시.

주관적인 성능 평가를 나타낸 그림 3의 빨간색 영역은 제안하는 방법이 기존보다 향상되었음을 나타낸다. 또한, 객관적인 성능평가를 위해 표 1은 평균 Dice, IoU를 비교한 것으로 제안하는 방법은 기존의

방법보다 향상됨을 보였다.



(그림 3) 실험 결과.

<표 1> mean Dice, mean IoU 비교

Method	DeepLabV3+	Proposed
mean Dice	0.8739	0.8774
mean IoU	0.9287	0.9310

참고문헌

- [1] 송창우, Wiratama Wahyu, 정지훈, 홍성재, 김대희, 강주형, “SPADEF 기반 U-Net을 이용한 고해상도 위성영상에서의 도시 변화탐지”, 대한원격탐사학회지, 36, 6-2, 1579-1590, 2020.
- [2] 송창우, 정지훈, 김대희, 홍성재, 강주형, “고해상도 시계열 위성영상을 이용한 객체 영역 기반 도시 변화탐지”, 대한원격탐사학회 추계학술대회, 2020.
- [3] 이성혁, 이명진, “위성영상을 활용한 토지피복 분류 항목별 딥러닝 최적화 연구”, 대한원격탐사학회지, 36, 6-2, 1591-1604, 2020.
- [4] Ozan Oktay et al, “Attention U-Net: Learning Where to Look for the Pancreas”, arXiv preprint arXiv:1804.03999, 2018.
- [5] Chen, Hao, and Zhenwei Shi. "A spatial-temporal attention-based method and a new dataset for remote sensing image change detection." Remote Sensing, 12, 10, 2020.