

인코더-디코더 구조의 CNN을 이용한 위성 영상에서의 변화탐지

박원희, 진동권, 김창수

고려대학교

whpark@mcl.korea.ac.kr, dongkwonjin@mcl.korea.ac.kr, changsukim@korea.ac.kr

Change Detection in Satellite Images Using Encoder-Decoder CNN

Won-Hui Park, Dongkwon Jin, Chang-Su Kim

Korea University

요 약

본 논문에서는 Convolutional Neural Network(CNN)를 이용한 위성 영상 변화탐지 알고리즘을 제안한다. 우선 EfficientNet 기반의 대칭 인코더-디코더 구조의 변화탐지 네트워크를 구성한다. 그리고 디코더 단계 ASPP 모듈을 추가하여 넓은 수용영역을 갖는 특징 정보를 통해 변화지도(change map)를 복원한다. 실험 결과를 통해 검출 성능 및 연산 효율성이 기존 기법보다 우수함을 보인다.

1. 서론

위성 영상의 변화탐지는 지표면에서 발생하는 변화를 원격으로 모니터링하기 위해 널리 활용되고 있다. 넓은 지역의 지형 변화탐지는 환경 모니터링, 재난 평가, 도시 발전 연구 등에 활용된다. 숲, 바다와 같은 자연 지형의 변화탐지가 주였던 과거에 비해 카메라 촬영 기술이 발달함에 따라 빌딩, 차량 등의 작은 객체가 섞인 도심 지역에서의 변화탐지가 중요해졌고, 더욱 정밀한 변화탐지 기술의 필요성이 증가되었다.

위성 영상 변화탐지는 그림 1과 같이 시간 차가 발생하는 위성 영상 쌍에 대해 의미적으로 변화한 영역을 탐지하는 기술이다. 이를 위해서는 위치적으로 동일한 대상에 대해 변화여부를 판단해야 하므로 위성 영상 쌍의 정합이 요구된다. 그러나 고층 건물 영역에서는 카메라의 촬영 각도에 따라 모습이 바뀌어 위치적 정합 오류가 크게 발생한다. 또한 카메라 센서 차이나 계절 변화에 따른 색상 왜곡이 발생한다. 이를 해결하기 위해 정사 보정,

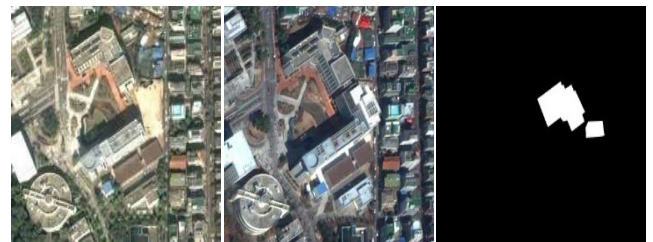


그림 1. 시계열 위성 영상 쌍 및 변화탐지

방사 보정과 같은 전처리 기법들이 제안되었다. 정사 보정은 시점에 따른 건물의 모습을 일정하게 보정하는 영상 정합 기법이고, 방사 보정은 영상의 색 왜곡을 보정하는 기법이다. 그러나 이러한 전처리 과정들은 복잡하고 계산적 비용이 크기 때문에 변화탐지에 적용하는데 한계가 있다.

최근 CNN의 도입으로, 객체 추적, 검출, 인식, 분할 등 컴퓨터 비전 분야에 상당한 발전을 가져왔다 [1, 2]. 변화탐지 분야에도 CNN을 기반으로 하는 연구가 활발히

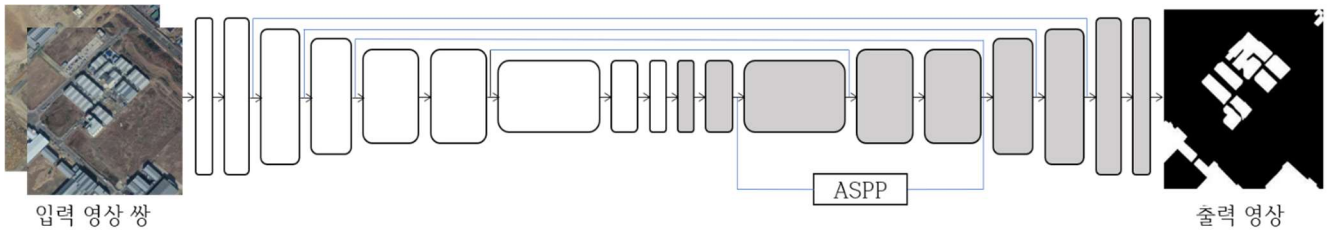


그림 2. 변화탐지 네트워크 구조

이루어지고 있다. 기존 연구[3]는 앞서 언급된 복잡한 전처리 과정들을 거치지 않고 CNN을 통해 변화탐지를 수행하는 기법을 제안하였다.

본 논문에서는 그림 1과 같은 시계열 위성 영상 쌍에 대해 CNN을 활용한 변화탐지 알고리즘을 제안한다. 우선 EfficientNet[2] 기반의 인코더-디코더 네트워크를 설계하여 기존 기법[3] 대비 성능 및 연산 효율성을 향상시킨다. 또한 의미적 분할(semantic segmentation) 분야에서 일반적으로 사용되는 ASPP[4] 모듈을 디코더 단에 추가하여 검출 정확도를 향상시킨다.

본 논문의 구성은 다음과 같다. 2장에서는 구성된 네트워크의 구조에 대해 설명한다. 3장에서는 변화탐지 데이터셋 및 네트워크 학습 과정에 대해 설명한다. 4장에서는 제안 기법에 대한 실험 결과를 분석한다. 5장에서는 본 논문의 결론을 맺는다.

2. 변화탐지를 위한 네트워크 구조

2.1 EfficientNet 기반의 인코더-디코더 네트워크

EfficientNet[2]은 네트워크의 깊이, 너비, 입력 해상도의 비율을 최적화해 성능과 연산 효율성을 모두 향상시키기 위해 제안된 네트워크이다. 본 논문에서는 그림 2와 같이 EfficientNet 기반의 인코더-디코더 구조의 변화탐지 네트워크를 구성한다. 네트워크는 시계열 위성 영상 쌍을 입력으로 받아 변화/비변화 이진 영상 형태의 변화지도를 출력한다. 이때 인코더는 영상을 입력으로 받아 고차원의 특징 벡터를 추출하고, 디코더는 인코더에서의 특징 벡터를 입력으로 받아 영상을 복원해 변화지도를 출력한다. 제안하는 네트워크의 구체적인 구조는 다음과 같다.

본 논문에서는 EfficientNetB4를 변형하여 변화탐지 네트워크를 설계한다. 완전연결계층을 제외한 EfficientNet B4를 변화탐지 네트워크의 인코더 단으로 사용하고, 대칭되는 구조로 디코더 단을 구성한다. 그리고 디코더 이전 단에서 출력된 정보와 대응하는 인코더 단 특징 정보를 결합하고, 이를 up-sampling 과정에 사용하여 영상 해상도를 복원한다. 디코더 마지막 단에 소프트맥스

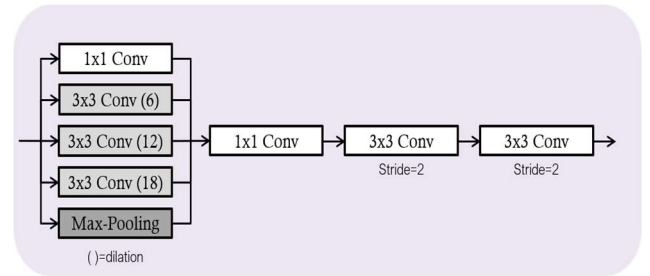


그림 3. ASPP 모듈 구조

연산을 적용하여 각 픽셀 값이 0과 1 사이의 확률 값을 가지는 변화지도를 출력한다.

2.2 ASPP(Atrous Spatial Pyramid Pooling) 모듈

본 논문에서는 그림 2와 같이 ASPP[4] 모듈을 변화탐지 네트워크 디코더에 적용한다. ASPP는 의미론적 분할 기법에서 사용되는 모듈로, 그림 3과 같이 병렬구조의 atrous 컨볼루션 연산 및 풀링 연산과 일련의 컨볼루션 연산으로 구성되어 있고, 동일 연산량 대비 넓은 수용 영역을 갖는다. ASPP 모듈에서 추출된 특징 정보와 이전 디코더 단에서 출력된 정보를 결합하고, 이를 up-sampling 과정에 사용함으로써 넓은 수용 영역을 갖는 특징 정보를 통해 영상을 복원한다.

3. 변화탐지 데이터셋 및 네트워크 학습

본 논문에서는 변화탐지 데이터셋[3]을 통해 네트워크를 학습 및 실험한다. 변화탐지 데이터셋은 Google Earth를 통해 취득한 965쌍의 600×600 고해상도 위성 영상 쌍 및 변화영역 라벨지도로 구성된다. 이 중 865쌍은 학습데이터, 100쌍을 실험데이터로 나뉘어져 있다. 라벨지도의 변화 영역은 1, 비변화 영역은 0으로 각 픽셀 값이 라벨링 되어 있다.

네트워크 학습을 위해 출력된 변화지도와 라벨지도간의 크로스-엔트로피를 계산하고 Adam optimizer[5]를 사용하여 파라미터를 최적화한다. 초기 학습율은 0.0001로 설정하고, 150번의 epoch마다 3/4배 감소시킨다. 영상 배치의 크기는 4로 설정한다.

4. 실험 결과

본 장에서는 변화탐지 실험 결과를 분석한다. 변화탐지 네트워크의 성능 측정을 위해, 네트워크로부터 예측된 변화지도의 각 픽셀을 라벨지도와 비교한다. True Positive (TP)와 True Negative(TN)는 각각 변화한 영역과 비변화한 영역을 맞게 예측한 픽셀의 개수를 의미한다. False Alarm(FA)은 비변화한 영역을 변화한 영역으로 잘못 탐지한 픽셀의 개수를 의미한다. Miss Alarm(MA)은 변화한 영역을 비변화한 영역으로 잘못 탐지한 픽셀의 개수를 의미한다. 정확도(Precision)와 재현율(Recall)은 다음 수식을 통해 정의된다.

$$\text{Precision} = \frac{TP}{TP + FA}, \quad \text{Recall} = \frac{TP}{TP + MA}$$

또한, 평가 지표로 사용되는 F1-score와 F2-score는 아래의 F-score 수식을 따른다.

$$F_{\beta} = (1 + \beta^2) \cdot \frac{\text{Precision} \cdot \text{Recall}}{(\beta^2 \cdot \text{Precision}) + \text{Recall}}$$

여기서 β 는 정확도와 재현율의 비율을 결정한다. $\beta=1$ 일 때 얻어지는 F1-score는 정확도와 재현율의 조화평균으로 정확도와 재현율을 같은 가중치로 평가한다. $\beta=2$ 일 때 얻어지는 F2-score는 정확도보다 재현율에 더 높은 가중치를 두고 평가한다.

표 1은 변화탐지 네트워크의 성능을 나타낸다. 제안 네트워크는 기존 기법[3]보다 F1-score, F2-score가 각각 3.81, 4.61 높은 결과를 보인다. 또한 깊이, 너비 그리고 입력 해상도의 비율이 최적화된 EfficientNet을 통해 네트워크의 파라미터 수를 줄여 연산 효율성이 증가하였다. 그림 4는 정성적 결과를 나타내며 기존 기법과 비교하여 검출 정확도가 향상된 것을 확인할 수 있다.

5. 결론

본 논문에서는 CNN을 활용한 위성 영상에서의 변화탐지 알고리즘을 제안하였다. EfficientNet 기반으로 대칭되는 인코더-디코더 구조의 변화탐지 네트워크를 구성하였다. ASPP 모듈을 디코더 단에 추가하여 넓은 수용영역을 갖는 특징 정보를 통해 변화지도를 복원하였다. 실험 결과, 기존 기법과 비교하여 성능이 개선되었으며, 네트워크의 파라미터 수를 줄여 연산 효율성을 증가시켰다.

	F1-score	F2-score	파라미터 수
기존 기법[3]	66.74	71.02	8.4M
제안 기법	70.55	75.63	3.3M

표 1. 변화탐지 성능 측정

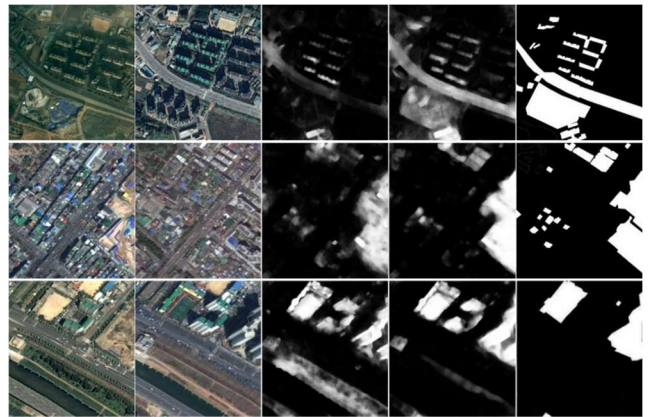


그림 4. 변화탐지 정성적 결과

감사의 글

이 논문은 2018년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. NRF-2018R1A2B3003896)

참고문헌

[1] K. Lim, W.-D. Jang, and C.-S. Kim, "Background Subtraction Using Encoder-Decoder Structured Convolutional Neural Network," In *AVSS*, 2017.

[2] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks.", In *ICML*, 2019.

[3] K. Lim, D. Jin, and C.-S. Kim, "Change Detection in High Resolution Satellite Images Using an Ensemble of Convolutional Neural Networks," In *APSIPA*, 2018.

[4] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs," In *TPAMI*, 2017.

[5] D. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," In *ICLR*, 2015.