

건설현장 내 주요객체 영역분할 성능 향상 기법에 관한 연구

나종호*, 추엔*, 윤일동**, 신휴성*

*한국건설기술연구원 미래스마트건설연구본부

**한국외국어대학교 컴퓨터공학과

e-mail:najongho@kict.re.kr, pnchuyen@kict.re.kr, yun@hufs.ac.kr,

hyushin@kict.re.kr

A study on the performance improvement technique for Instance segmentation of objects in the construction site

Jong-Ho Na*, Chuyen Pham*, Il-Dong Yun**, Hyu-Soung Shin*

*Dept. of Future and Smart Construction Research, Korea Institute of Civil Engineering and Building Technology

**Dept. of Computer Engineering, Hankuk University of Foreign Studies

요약

객체인식 및 영역분할 기법을 이용하여 건설현장의 주요한 객체를 높은 정확도로 검출하기 위해서는 검출대상 객체와 가능한 배경환경을 대표할 수 있는 영상 학습데이터가 필요하지만 이러한 학습 데이터를 구축하는 과정은 많은 인력과 시간이 소요되며, 특정 기상조건에 대한 대상객체 영상의 확보 가능 여부는 장담하기 어렵다. 따라서, 본 논문에서는 한정된 데이터를 활용하여 다양한 기상조건에 대해 영역분할 성능을 향상시키는 방법을 제안한다. 딥러닝 기반 스타일 전이 기법을 활용하여 일반 기상 조건의 객체 영상을 동일 객체이지만 맞이하기 어려운 열악한 기상 조건의 영상으로 변환하여 재생산하였다. 또한, 생산된 가상 영상은 실제 영상과 함께 학습시켰으며, 실제 열악기상 조건에 대한 추론 성능 향상 효과를 실험적으로 고찰하였다. 실제 열악기상 조건의 데이터로 평가 결과, 인공데이터의 학습 자료 반영만으로 모델 추론 성능 향상에 기여하는 결과를 확인할 수 있었으며, 다양한 조건을 추가로 반영할 시 더욱 범용성 높은 영역분할 모델이 생성 가능할 것으로 판단된다.

1. 서론

최근 정부는 산업안전보건법 및 중대재해기업처벌법 제정을 통해 산업현장의 사망 및 안전사고 대응을 지속적으로 강화하고 있다. 그럼에도, 2020년 사고사망자 현황에 의하면 건설 분야에서 458명의 사망자가 발생하여 전체에서 51.9%의 비율을 차지하며 사고율이 크게 줄지 않고 있다[1]. 이러한 이유로 사고 재해율을 감소시키기 위해 기술적 접근 방식으로 현장에 적용하려는 많은 연구가 수행되고 있다. 한국건설기술연구원에서는 택지개발 공사현장 확보를 통해 CCTV 원천 데이터를 취득하였으며, 건설 현장 내 취득가능한 주요 객체를 공사유형, 공종, 객체 종류에 따라 17개 객체를 분류하였다. 이를 기반으로 건설 현장 내 주요 객체의 영역 분할 인공지능 학습용 데이터를 생성 및 가공하고 Instance Segmentation의 대표적인 모델인 Mask R-CNN 모델[2]을 기반으로 추론 성능 평가 실험을 진행하였다. 건설 현장에서는 기상 조건에 따라 추론 성능의 차이가 매우 클 것으로 예상되며, 일관성 있는 추론 성능을 유지하기 위해

서는 다양한 열악한 기상 조건에서 취득한 데이터 확보가 필수적이다. 하지만 실제 열악한 기상 조건의 데이터를 취득하는데 시간적 그리고 비용적 어려움이 따른다. 따라서, 본 연구에서는 스타일 전이(style transfer) 기법을 통해 매우 사실적이며 열악한 기상 조건의 영상을 인위적으로 생성하여 학습시키고, 미확보 기상조건의 현장에 대한 추론 성능 향상에 미치는 정도를 실험적으로 고찰하였다.

2. 관련 이론

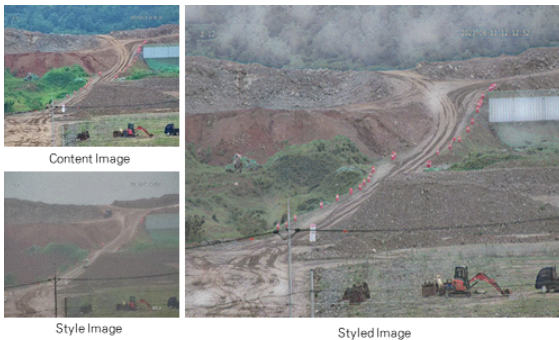
2.1 스타일 전이(Style Transfer) 기법

스타일 전이는 딥러닝 기반 이미지 합성 기술 중 하나로 타겟 이미지와 스타일 이미지가 주어져 있을 때, 형태는 타겟 이미지와 유사하게 유지하며 스타일만 원하는 스타일 이미지로 바꾸는 방법이다. 스타일 전이 기법은 “Artistic Style Transfer[3]”와 “Photorealistic Style Transfer[4]” 방법으로 구분되는데, 본 논문에서는 사실적인 데이터 생성이 목적이므로 “Photorealistic Style Transfer” 방법을 도입하였다. 웨

이블릿 기법을 활용하여 매우 사실적인 스타일 전이를 수행하는 WCT(Wavelet Corrected Transform) 모델[5]을 활용하였다.

2.2 스타일 전이 기반 인공 데이터 생성

스타일 전이 기법을 통해 건설 현장의 날씨 조건을 변환하는 작업을 진행하였다. 그림 1처럼 맑은 날씨의 건설현장 영상에 실제 비/안개 날씨의 스타일 영상을 적용하여 우측과 같이 매우 사실적인 인공 비/안개 날씨 조건 데이터셋을 인위적으로 생성하였다.



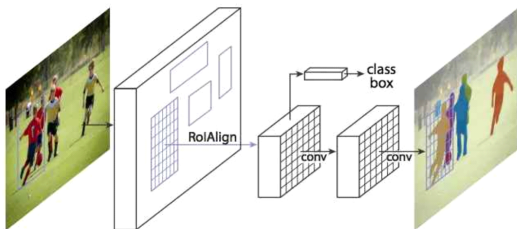
[그림 1] [좌상] 타겟영상, [좌하] 스타일영상, [우측] 인위적 비/안개 조건 영상 생성 결과

이렇게 스타일 전이 기법을 통해 생성된 영상은 기존 원본 데이터의 라벨링 정보의 재활용이 가능하기 때문에, 확보된 영상을 다양한 배경으로 변환하여 추가 레이블링 작업 없이 무한정 확장 시킬 수 있다.

2.3 영역 분할 기법(Instance Segmentation)

Mask R-CNN은 2017년 Facebook AI 팀이 개발한 영역 분할된 이미지를 마스킹하는 알고리즘으로 Faster R-CNN[6]에 각 픽셀이 오브젝트에 해당하는지 아닌지를 판단하는 Mask Branch를 추가한 알고리즘이다.

Classification, Localization, Segmentation이 동시에 가능한 장점을 가지고 있다.



[그림 2] Mask R-CNN 개념도

3. 실험

본 논문에서는 가용한 한정된 영상으로 보다 폭 넓은 환경 조건에서 영역 분할 추론 성능을 높이고자 하였다. 이에 대한 고찰을 위해 열악한 기상조건의 데이터를 생성하였으며, 생성된 인공 데이터를 활용해 실제 열악 기상 데이터에서 효과를 보이는지 검증 실험을 진행하였다.

3.1 실험 데이터 설명

실제 토공현장에 CCTV 설치를 통해 수개월간 데이터를 취득했으며, 일부 영상의 샘플링을 통해 데이터를 가공 완료하였다. 택지개발 공사의 주요 공종인 토공사, 지반개량 및 지반보강공사 등에서 자주 운용되는 건설 현장 내 주요한 대상 장비와 지형 객체를 선정하였으며, 선정된 객체 리스트는 다음 표 1와 같다.

[표 1] 건설 현장 내 주요객체 현황

객체 종류	주요 객체 리스트 현황		
정형 객체	굴삭기	덤프트럭	로더
	불도저	롤러	천공기
	항타기	크레인	자동차
비정형 객체	토사면	암사면	웅덩이
	토사더미	암더미	이동가능영역

정형 객체는 일반적으로 보편적인 형상을 갖고 있는 객체로 건설 현장에서 운용되고 있는 건설장비 및 일반 자동차들이 이에 해당한다. 그리고 비정형 객체는 특정 형상이 정해져 있지 않는 객체로 사면, 더미, 웅덩이 등이 이에 해당한다. 사면은 암사면, 토사면 그리고 더미는 암더미, 토사더미로 세분화하였다. 또한, 이동가능영역은 건설장비와 현장인력이 다닐 수 있는 상대적으로 안전한 구역으로 정의 된다.

데이터 가공 진행 시 비정형 객체의 라벨링 기준은 다음과 같이 정의된다.

1. 사면(slope)의 기준은 면 객체로 사면의 상단부와 하단부 사이에 평면이 존재하는 경우를 사면이라 정의한다.
2. 더미(mound)의 기준은 양각의 지형 객체이며, 사면과 다르게 양각의 상단부에 운용성(상단면이 없는 경우)이 없는 영역을 더미라 정의한다.
3. 웅덩이(fit)의 기준은 음각의 지형 객체로, 사면으로 둘러쌓여있고 지반 운용성이 있는 저면이 없는 객체로 정의한다.



[그림 3] 학습 데이터셋 예시



[그림 4] 평가 데이터셋 예시

실험에 사용된 데이터셋은 다음 표 2처럼 3종류로 구성하였다.

[표 2] 데이터셋 종류

Case	데이터셋 종류	데이터셋 설명	정지영상 개수
case 1	Train	맑은 날씨	618
		총계	618
case 2	Train	맑은 날씨	618
		인공 비/안개	309
		총계	927
case 3	Train	맑은 날씨	618
		실제 비/안개	180
		총계	798
	Test	실제 비/안개	300

Case 1은 맑은 날씨로만 구성된 데이터셋이며 Case 2는 Case 1에서 50%만 무작위로 추출하여 스타일 전이를 통해 인위적으로 생성한 가상 비/안개 데이터를 추가로 반영한 데이터셋이다. 마지막으로 Case 3은 실제 비/안개 데이터가 반영된 데이터셋이다. 학습에 사용된 데이터 예시는 그림3과 같다.

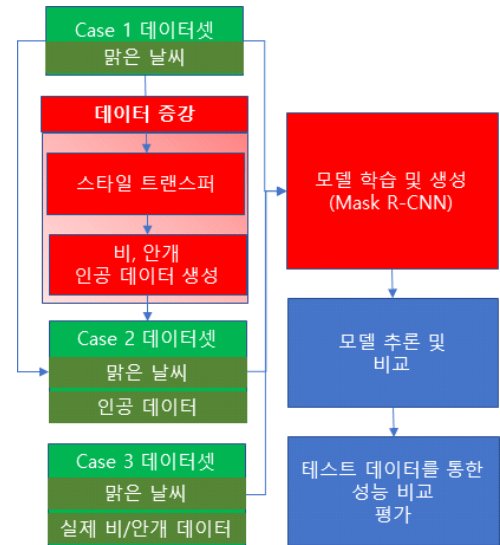
3.2 실험 방법

전체 실험 구성도는 그림 5와 같다. 앞서 설명한 3종류의 데이터셋 Case에 데이터를 Mask-RCNN 모델에서 학습을 진행하고 모델을 생성한다. 이후에 학습에 사용되지 않은 테스트 데이터셋으로 각각 생성된 모델 추론을 진행하고 결과를 분석하였다.

3.3 실험 결과 및 분석

영역 분할 모델의 성능 평가는 시각적인 추론 결과와 수치적 결과를 통해 비교분석을 진행하였다. 수치적 평가는 일반적으로 널리 사용되고 있는 Mean Average Precision(mAP)와 Dice score를 활용하였다. mAP는 객체를 얼마나 잘 맞췄

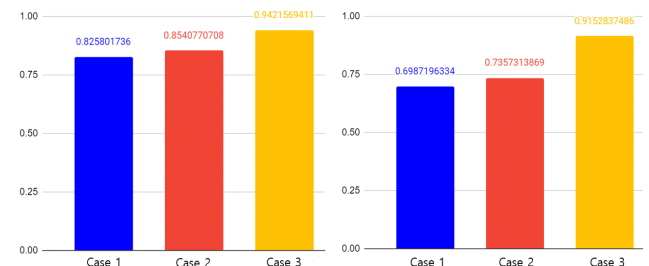
는지 평가하는 지표이며, Dice score는 추론된 영역 면적이 Ground Truth와 얼마나 일치하는지 수치화한 지표이다. IoU(Intersection over union) 값은 0.5로 설정하였다.



[그림 5] Mask R-CNN 개념도

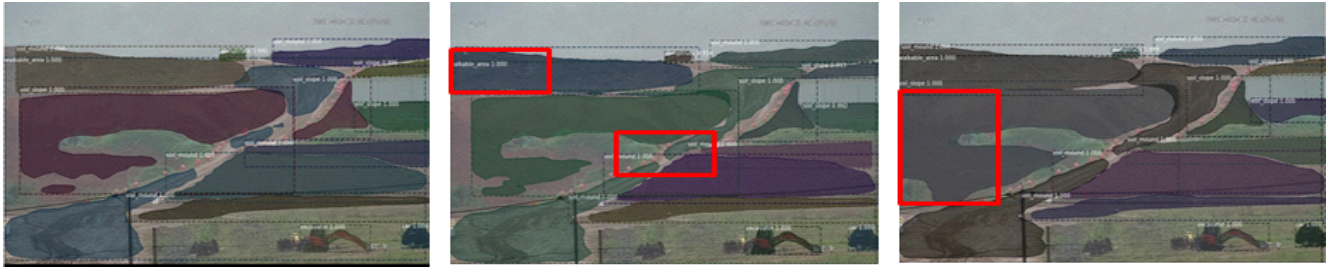
3.3.1 수치적 성능평가 결과

맑은 날씨로만 구성된 Case 1과 스타일 전이 기법을 활용하여 추가 인공데이터를 생성해 학습시킨 Case 2를 비교했을



[그림 6] [좌측] mAP 결과, [우측] Dice Score 결과

때, 실제 열악 기상 조건의 데이터로 구성된 Case 3보다는 낮은 결과를 보이지만 0.82에서 0.85로 소폭 향상된 mAP 검출 정확도를 보인다. Dice score 결과 또한 0.69에서 0.73으로 향상된 영역 분할 추론 결과를 나타낸다. 즉, 스타일 전이 기법



[그림 7] [좌측] Case 1 추론 결과, [중간] Case 2 샘플 결과, [우측] Case 3 샘플 결과

으로 열악한 기상 조건의 인공 데이터를 생산하여 학습 자료에 추가해 반영된다면 검출 정확도가 향상된다는 것이 수치적으로 확인하였다. 현재는 한 개의 기상환경 스타일만을 적용한 실험 결과이므로 더욱 다양한 조건의 스타일 이미지를 반영한다면 Case 1 데이터만으로 Case 3에 근접한 검출 정확도를 얻을 수 있을 것이라 예상된다.

3.3.2 시각 평가 결과

다음 그림 7은 동일한 샘플 이미지에서 설정된 3개 Case의 추론결과를 시각적으로 표현하였다. Case 1의 경우 맑은 날씨로만 학습을 진행하고 추론하였기 때문에 다른 Case에 비해 상대적으로 낮은 영역 분할 추론 성능을 보인다. 그에 반해 Case 1 데이터를 활용하여 인공 데이터를 생산시킨 Case 2는 Case 1에 비해 영역 분할 성능이 향상된 결과를 보였다. 하지만 실제 통합 데이터셋인 Case 3에 비해 결과는 상대적으로 낮게 나타났다.

시각적인 추론 결과와 수치적 결과를 통해 인공 데이터의 생산 및 학습 반영이 보다 포괄적인 환경에 대한 우수한 추론 성능을 보였다. 또한, 보다 다양한 기상 스타일 영상을 반영한다면 보다 환경 대응력 높은 강인한 검출 모델을 생성하리라 예상된다.

4. 결론

본 논문에서는 가용한 한정된 영상을 기반으로 보다 폭 넓은 환경조건에 대한 영역분할 모델의 추론 성능을 향상시키는 연구를 제안하였다. 실제 택지개발 공사 현장에서 CCTV 기반의 데이터를 수집하고 정형객체 및 비정형 객체를 선정하여 가공을 완료하였다. 더불어 딥러닝 기반의 스타일 전이 모델을 통해 매우 사실적인 인공 열악 환경 데이터를 생성하였으며, 학습 자료로 재 반영하였다. 이러한 학습 데이터를 3 종류의 Case로 구성하여 Mask R-CNN 모델을 통해 학습을 진행하고 평가를 완료하였다. 실험 결과, 실제 열악 환경의 테스트 데이터를 통해 각 모델의 평가지표를 비교분석을 통해 실제로 인공 데이터의 학습자료 재반영 만으로 모델 추론 성능 향상에 기여하는 결과를 확인할 수 있었으며, 다양한 조건

을 추가로 반영할 시 매우 강인한 영역 분할 모델의 구현 가능성을 확인하였다. 본 연구는 건설 현장 내 주요한 객체를 정의하고 영상 기반으로 객체 인식을 위한 핵심정보를 판단하는 요소 기술로 위험 영역의 구획화에 활용 가능하며, 지속적인 영상 빅데이터 생산을 통해 건설 현장 내 영상 기반 모니터링을 위한 주요대상물 공간정보 구축 시스템과의 연계에 기여하리라 전망한다.

감사의 글

본 연구는 과학기술정보통신부 한국건설기술연구원 연구운 영비지원(주요사업)사업으로 수행되었습니다 (과제번호 20220124-001, (22주요-대2-BIG)극한건설 환경 구현 인프라 및 TRL6 이상급 극한건설 핵심기술개발(7/9)).

참고 문헌

- [1] 고용노동부, “e-나라지표 산업재해 현황”
- [2] He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask r-cnn. In Proceedings of the IEEE international conference on computer vision (pp. 2961-2969).
- [3] Gatys, L. A., Ecker, A. S., & Bethge, M. (2016). Image style transfer using convolutional neural networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2414-2423).
- [4] Luan, F., Paris, S., Shechtman, E., & Bala, K. (2017). Deep photo style transfer. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4990-4998).
- [5] Yoo, J., Uh, Y., Chun, S., Kang, B., & Ha, J. W. (2019). Photorealistic style transfer via wavelet transforms. In Proceedings of the IEEE/CVF International Conference on Computer Vision (pp. 9036-9045).
- [6] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. Advances in neural information processing systems, 28, 91-99.