

건설 현장 CCTV 영상에서 딥러닝을 이용한 사물 인식 기초 연구

A Basic Study on the Instance Segmentation with Surveillance Cameras at Construction Sites using Deep Learning based Computer Vision

강 경 수*

조 영 운**

류 한 국***

Kang, Kyung-Su

Cho, Young-Woon

Ryu, Han-Guk

Abstract

The construction industry has the highest occupational fatality and injury rates related to accidents of any industry. Accordingly, safety managers closely monitor to prevent accidents in real-time by installing surveillance cameras at construction sites. However, due to human cognitive ability limitations, it is impossible to monitor many videos simultaneously, and the fatigue of the person monitoring surveillance cameras is also very high. Thus, to help safety managers monitor work and reduce the occupational accident rate, a study on object recognition in construction sites was conducted through surveillance cameras. In this study, we applied to the instance segmentation to identify the classification and location of objects and extract the size and shape of objects in construction sites. This research considers ways in which deep learning-based computer vision technology can be applied to safety management on a construction site.

키 워 드 : 건설안전, 딥러닝, 컴퓨터 비전, 인스턴스 분할

Keywords : construction safety, deep learning, computer vision, instance segmentation

1. 서 론

건설 현장은 작업자, 중장비, 가설시설물, 자재 등이 혼재하고 날씨와 같은 외부 환경에 노출되어 있어 안전 재해의 발생 위험이 매우 크다. 이에 따라 대규모 건설 현장의 경우, CCTV를 설치하여 안전 관리자가 여러 대의 카메라를 통해 작업자들의 상태를 관제하고 있다. 하지만 사람의 눈으로 여러 대의 CCTV 화면을 동시에 관제하는 것은 불가능하며 관리자의 업무 피로도 또한 매우 높다. 따라서 본 연구는 CCTV 화면에서 사물들의 정보를 실시간으로 제공하고자 딥러닝 기반 컴퓨터 비전 기술을 적용하여 성능을 분석하였다.

2. 본 론

2012년 AlexNet¹⁾을 시작으로 딥러닝이 컴퓨터 비전에 적용되면서 상당한 성능 향상을 이루었고 다양한 산업 분야에서 딥러닝을 적용하기 위해 노력하고 있다. 건설업은 안전성, 생산성, 검사 분야 등에 적용 연구가 활발하게 이루어지고 있다. 그러나 건설업에 적용된 기존 연구를 살펴보면 컴퓨터 비전 기술 발전 대비 오래된 모형을 사용하거나 단순히 객체를 식별하고 위치를 추출한 연구가 대부분이다. 건설 현장 CCTV에 적용하기 위해서는 식별한 객체의 위치 정보 뿐만 아니라 크기와 모양을 추출하고 이를 활용하는 것이 매우 중요하다. 이러한 특징들을 연고자 인스턴스 세그멘테이션(instance segmentation) 모형인 YOLACT²⁾를 적용하였다.

본 연구는 다수의 현장에서 영상, 사진 데이터를 수집하고 6개의 클래스를 폴리곤 어노테이션(polygon annotation)으로 레이블링(labeling) 작업을 수행하였다. 총 982개의 데이터를 8:2로 나누어 학습 및 검증용으로 사용하였다. 또한, CCTV 시점인 탑뷰(top view)뿐만 아니라 다양한 시점에서 활용 가능한 데이터셋을 생성하고자 하였다. 표 1은 MS COCO 데이터셋 표준 평가 기법을 요약한 정량 평가 결과이며 그림 1은 건설 데이터셋을 학습한 모형의 정성 평가 결과이다.

* 삼육대학교 건설기술및관리연구소, 경영학 박사

** 삼육대학교 건설기술및관리연구소, 데이터 사이언스 석사

*** 삼육대학교 건축학과 부교수, 교신저자(ryuhanguk333@gmail.com)

표 1. 건설 데이터셋을 적용한 레이블별 인스턴스 세그멘테이션 폴리곤 마스크 성능 결과

분류	클래스	AP	AP ₅₀	AP ₇₅	AP _S	AP _M	AP _L
작업자 관련	작업자(person)	0.662	0.883	0.750	0.279	0.584	0.734
	안전모(hardhat)	0.634	0.751	0.740	0.644	0.648	0.559
	안전조끼(safety vest)	0.437	0.560	0.479	0.000	0.523	0.391
중장비	굴착기(excavator)	0.740	0.908	0.834	0.303	0.534	0.780
	덤프 트럭(dump truck)	0.783	0.855	0.824	0.221	0.532	0.838
	콘크리트 믹서 트럭 (concrete mixer truck)	0.858	0.900	0.879	0.090	0.534	0.894

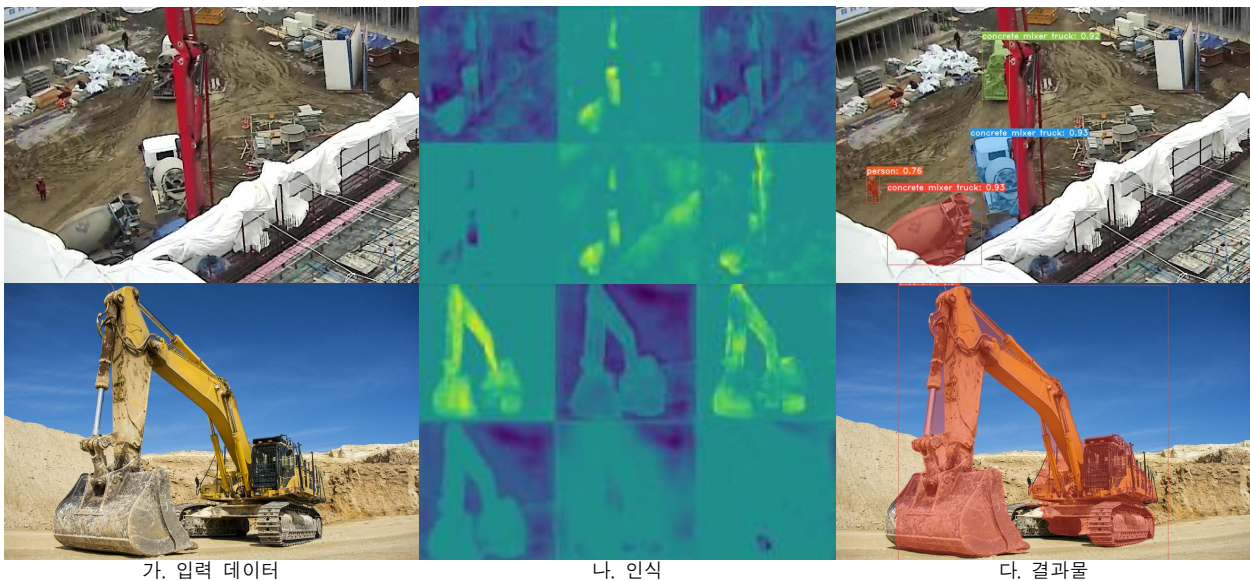


그림 1. YOLACT 모형 입출력 시각화: Protonet 인스턴스 분할 과정(나)과 바운딩 박스 및 폴리곤 마스크 결과물(다)

3. 결론

본 연구는 CCTV 뿐만 아니라 다양한 카메라 시점에서 활용 가능한 건설 데이터셋을 구축하고, 단순한 객체 식별과 위치 정보를 추출할 뿐만 아니라 크기와 형태까지 실시간으로 파악할 수 있는 인스턴스 세그멘테이션을 적용하여 정량 및 정성 평가를 수행하였다. 작은 객체(32²px 이하)의 성능이 안전모를 제외한 나머지 클래스에서 낮은 인식 성능을 보여주지만, 전체적인 모형의 성능이 적은 데이터셋임에도 불구하고 현장에 적용 가능한 수준인 것으로 나타났다. 향후 추가적인 데이터 수집 및 학습한 모형에서 추출된 정보를 활용할 수 있는 다양한 연구를 진행할 예정이다. 해당 연구가 건설의 다양한 분야에 적용될 표본을 제시하는 발판이 될 것으로 기대한다.

Acknowledgement

본 논문은 2020년 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임을 밝히며 이에 감사를 드립니다. (No. 2020R1A2B5B01001609)

참고 문헌

1. Bolya, D., Zhou, C., Xiao, F., & Lee, Y. J., Yolact: Real-time instance segmentation. In Proceedings of the IEEE international conference on computer vision, pp.9157~9166, 2019.
2. Krizhevsky, A., Sutskever, I., & Hinton, G. E., Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems, pp.1097~1105, 2012.