

YOLO-v3을 활용한 건설 장비 주변 위험 상황 인지 알고리즘 개발

심승보, 최상일*

한국건설기술연구원 차세대 인프라연구센터

Development on Identification Algorithm of Risk Situation around Construction Vehicle using YOLO-v3

Seungbo Shim, Sang-Il Choi*

Future Infrastructure Research Center, KICT

요약 최근 정부는 건설 산업의 재해율과 사고 사망률이 전체 산업 중 높은 비율을 차지한다는 점을 개선하기 위하여 새로운 대책을 강구하고 있다. 특히 4차 산업혁명의 시대적 흐름에 맞춰 ICT 기술과 융합된 건설 기술 개발에 집중적으로 투자하고 있다. 이런 상황에 대응하고자 본 논문에서는 건설 기계를 사용하는 작업에서 작업자의 안전성 향상을 위한 방법으로, 건설 기계 운전자와 주변 작업자 간의 작업 상황 정보를 공유하고 인지할 수 있는 개념을 제시하였다. 그리고 해당 개념의 일부를 실현하고자 카메라를 이용한 인공 지능 기반 영상처리 기술을 활용하여 토공 작업에 적용시켰다. 그 중에서도 다짐 장비를 이용한 실험을 통해 YOLO-v3 기반의 영상 처리 알고리즘으로 토공 작업 중에 주변 작업자 상황을 인지하고 위험 상황 여부를 판단할 수 있는 알고리즘을 구현하였다. 그 결과 본 알고리즘은 동영상에서 초당 15.06프레임을 처리하며 90.48%의 정확도로 건설 기계 주변 위험 상황을 인지할 수 있다. 향후 이 같은 기술을 활용하여 건설 현장의 안전사고 예방에 기여하고자 한다.

Abstract Recently, the government is taking new approaches to change the fact that the accident rate and accident death rate of the construction industry account for a high percentage of the whole industry. Especially, it is investing heavily in the development of construction technology that is fused with ICT technology in line with the current trend of the 4th Industrial Revolution. In order to cope with this situation, this paper proposed a concept to recognize and share the work situation information between the construction machine driver and the surrounding worker to enhance the safety in the place where construction machines are operated. In order to realize the part of the concept, we applied image processing technology using camera based on artificial intelligence to earth-moving work. Especially, we implemented an algorithm that can recognize the surrounding worker's circumstance and identify the risk situation through the experiment using the compaction equipment, and image processing algorithm based on YOLO-v3. This algorithm processes 15.06 frames per second in video and can recognize danger situation around construction machine with accuracy of 90.48%. We will contribute to the prevention of safety accidents at the construction site by utilizing this technology in the future.

Keywords : Deep Learning, Image Processing, Construction Safety, Sensor, Construction Equipment

본 연구는 한국건설기술연구원 주요사업 <블록체인 기반 도시 데이터 수집 및 분석 모델 개발>의 지원을 받아 수행되었음

*Corresponding Author : Sang-Il Choi(KICT)

email: sangilchoi@kict.re.kr

Received May 13, 2019

Accepted July 5, 2019

Revised June 7, 2019

Published July 31, 2019

1. 건설 산업 재해 현황과 기술 동향

1.1 건설 산업의 재해현황 분석

4차 산업혁명의 흐름과 함께 건설 현장에서 많은 인력과 시간이 투입되던 분야에 생산성이 획기적으로 향상되고 있는 추세이다. 이는 건설 안전 분야도 예외가 아닌데, 최근 드론과 무인화 장비의 비약적인 발전으로 인하여 다양한 기술 개발이 진행되고 있다. 이 같은 시대적 흐름에 맞춰 정부 차원에서도 건설 안전 분야에 새로운 기술을 도입하고자 노력하고 있다. 국토교통부에서 발표한 스마트건설기술 로드맵에 따르면 건설의 재해율은 타 산업과 비교하여 높은 수준이나, 그 동안 건설안전 강화를 위한 노력에도 불구하고 안전도가 정체되어 있다[1]. 이는 Fig. 1과 Fig. 2에서도 확인할 수 있다. 2013년부터 2017년까지의 통계 자료에 의하면 건설업의 재해율은 전체 산업에 비해 평균적으로 약 1.6배 정도 높은 것으로 나타났다. 또 건설업의 사고 사망률은 연평균 1.58명이지만, 전체 산업은 0.57명으로 건설업의 경우가 약 2.8배 정도 높은 것으로 드러났다.

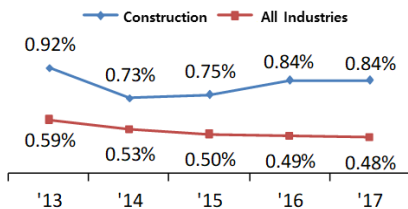


Fig. 1. Hazard rate (%) [1]

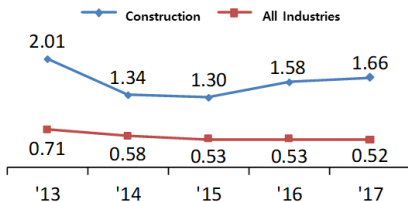


Fig. 2. Accident death rate per 10,000 persons [1]

이러한 실정 탓에 정부는 스마트 건설기술을 활용하여 2030년에는 건설자동화를 완성하겠다는 목표로 설계, 시공, 유지관리 분야 ICT 기술을 접목하여 기술 개발에 집중적으로 투자하려고 한다[1]. 그 가운데 ICT 기반 현장 안전사고 예방 기술을 핵심 기술로 선정하고 건설 현장의 장비, 인력을 통합 관제하여 안전사고를 예방하고자 노력하고 있다.

1.2 건설 기계 안전사고 예방 기술 동향

안전사고 예방 기술로서 건설 기계에 대한 안전 기술은 지속적으로 연구되고 있으며 그 중에서도 특히 센서 기술은 핵심 분야로 볼 수 있다. 센서는 용도 및 환경에 따라 다양하지만 이 중 작업자의 안전을 목적으로 사용하기 위하여 고려해야 할 사항들은 총 5가지로 Table 1과 같이 요약된다[2]. 이들 중 영상 센서를 사용했을 경우의 장점은 작업자 식별 기능을 제공한다는 점이며, 단점은 장애물과의 거리 측정이 어렵다는 점이다. 하지만 이 같은 단점은 스테레오 카메라를 사용하거나 깊이 카메라 (Depth Camera)를 사용하는 방법 등으로 보완이 가능하여 다양하게 연구가 진행되어 왔다.

실제로 건설 기계 분야에 적용된 기술을 살펴보면 우선 Soh et al.은 굴삭기를 이용한 작업 수행 중 안전사고를 예방하기 위하여 레이저 센서를 사용하였다[3]. 이를 통하여 굴삭기 주변에 장애물의 존재 유무를 판단하고 사고가 나지 않게 제어할 수 있도록 하였다.

Jo et al.은 Around View Monitor (AVM) 기술을 개발하여 카메라와 초음파 센서를 융합하여 영상 센서의 단점인 거리 측정 기능을 초음파로 보완하였다[2]. 총 4대의 AVM을 굴삭기에 장착하여 작업자의 근접 여부를 파악할 수 있도록 하였다. 하지만 영상을 통하여 주변에 작업자의 존재 여부를 판단하는 알고리즘은 제시되지 않았고, 디스플레이를 통해 굴삭기 운전자가 육안으로 판단할 수 있도록 시스템을 구축하였다.

Table 1. Analysis of the sensor's function for construction safety [2]

Considerations	Camera	Laser	UWB
Real-time Detection	○	○	○
Distance from Obstacle	X	○	○
Availability in Outdoor Condition	○	△	△
Economics & Convenience	△	○	○
Distinction between Worker & Obstacle	○	X	X

※ UWB: Ultra-Wideband

영상처리에 관한 연구도 활발하게 진행되었다. Seo et al.은 건설 현장에서 생길 수 있는 사고를 예방하기 위하여 안전 모니터링을 영상으로 해결하기 위한 연구를 소개하였다[4]. 그의 논문에서는 크게 객체 인식, 객체 추적, 그리고 상황 판단 단계로 구성되어 있다. 객체 인식과 추적 단계에서는 영상 분야에서 주로 사용되는 특징점

기반의 알고리즘을 사용하였고, 이를 바탕으로 필요한 대상 객체를 추적하였다. 그리고 이렇게 추적된 정보를 바탕으로 대상 객체의 행동 패턴을 분석하여 위험 여부를 판단하도록 하였다. Na et al.의 경우는 건설 장비에 부착된 카메라를 통하여 주변 환경 인식 기술을 제안하였다[5]. 작업자가 착용한 안전 조끼나 안전모 같은 세부적인 특징을 영상처리를 통하여 인식하고, 사전에 검교정된 정보를 바탕으로 거리까지 추정할 수 있는 방법을 제안하였다. Han et al. 또한, 이와 유사하게 비전을 활용한 건설 안전 관리 방법을 제안하였다[6]. 카메라를 통해 작업자를 인식하는 방법은 동일하지만, Body Skeleton Model까지 분석하는 알고리즘은 큰 차이점이다. 이 분석을 통하여 사다리 작업을 수행하는 작업자의 행동을 인식하여 위험 상황을 사전에 방지하는 기술을 제안하였다.

Kim et al.은 영상을 이용하여 수집한 정보를 바탕으로 인공지능을 활용한 건설 장비 인식과 분류에 대하여 연구하였다[7]. 적용된 방법은 Faster R-CNN과 유사한 것으로 완전 합성곱 신경망(Fully Convolutional Network)을 사용하였다[8]. ImageNet에서 대략 3,000여 장의 영상을 확보하여, 5가지로 분류할 수 있도록 학습용 데이터 세트를 구축하였다[9]. 알고리즘은 인식 성능 mAP 기준으로 96.33%를 보이고, 5가지 종류 모두 우수한 인식 성능을 보였다.

이 같은 연구 성과물은 건설 현장에서 발생할 수 있는 상황을 인지하여 중장비를 활용한 건설 작업의 효율성과 안전성을 향상시킬 수 있다. 더 나아가 최근 비약적으로 발전한 인공지능 기술을 접목할 경우 인식 성능 향상의 가능성이 높아질 것이다.

1.3 AI 기반 영상 객체 인식 기술 동향

영상에서 객체 인식을 하는 방법으로 최근에는 인공지능을 활용한 방식이 빈번하게 사용되고 있다. 그 중 가장 폭넓게 활용하는 방식은 R-CNN 계열의 알고리즘을 들 수 있다[10]. 영역 기반의 합성곱을 활용한 방식으로 여러 개의 신경망이 연속적으로 이어져 있고, 각 신경망에 해당하는 가중치 값들이 존재한다. 이 가중치 값은 학습에 의하여 결정되고, 이렇게 결정된 값들은 객체 인식에 활용한다. 가장 대표적인 방법으로 발표된 논문의 순서에 따라 Fast R-CNN, Faster R-CNN, Mask R-CNN이 있다[8,11,12]. 첫 번째 Fast R-CNN은 단일 단계 학습을 통해 Backbone Network를 확보할 수 있다[11]. 3가지 방법 중에 유일하게 기 학습된 가중치를 사용하지 않고, Xavier 초기값을 사용할 수 있다[13]. 단일 단계

학습을 할 수 있어 간단하게 인공지능 모델을 확보할 수 있는 장점이 있지만, 다수의 관심 영역 경계 상자 후보군을 생성하는데 시간이 소요된다는 단점이 있다. 이는 주로 긴 연산 시간을 갖는 Selective Search 알고리즘 수행에 기인한다[14]. 이 같은 이유로 실시간 객체 인식 기술로 사용하기에는 한계가 있다.

이 점을 개선하기 위하여 Faster R-CNN에서는 새로운 방법을 제안하였다[8]. 특히 관심 영역 경계 상자 후보군을 생성하는데, 기존의 알고리즘을 사용하는 것이 아니라 새로운 심층 신경망을 사용하는 것을 고안하였다. 기존의 알고리즘은 주로 CPU 자원을 활용한 연산을 사용하여 비교적 많은 연산 시간이 소요되었지만, 심층 신경망을 통한다면 GPU를 사용하기 때문에 연산 시간이 단축된다. 이 같은 심층 신경망은 관심 영역의 후보군을 제공하는 기능을 한다고 하여 Region Proposal Network (RPN)이라고 하는데, Backbone Network에 비해 상대적으로 적은 신경망 층을 가지고 있다. 이 신경망을 통해 객체 인식을 하는 필요한 하드웨어 자원의 대부분이 CPU에서 GPU로 이동하였다.

하지만 이 같은 구조로 완성된 모델을 만들기 위해서는 여러 단계를 거친 학습 수행이 필요하다. Faster R-CNN은 Backbone Network, Region Proposal Network, Classifier Network로 구성되며 이 3개의 심층 신경망을 학습하기 위해서는 여러 단계의 수행이 필요하므로 다중 단계 학습 기법으로 완성된 모델을 구하는데 시간이 연장되었다. 그럼에도 불구하고 인공지능 모델이 실제 적용되는 상황에서는 많은 속도 향상이 이루어졌고, 실시간 객체 인식 기술에 한층 더 가까워졌다.

R-CNN 계열 방법의 마지막인 Mask R-CNN의 가장 큰 특징은 객체 인식의 분할 (Segmentation)까지 가능하다는 점이다[12]. 그 전까지 객체 인식은 영상 내에서 사각 박스를 통해 객체의 위치를 추정하고 그 종류를 구분하였다. 하지만 Mask R-CNN에서는 객체 인식 기술에 객체의 위치를 화소 단위로 구분할 수 있도록 심층 신경망을 설계하였다. 이 심층 신경망의 역할은 각 화소가 객체에 해당되는 것인지 아닌지를 구분하는 것으로 Binary Mask라고 한다. Backbone Network로부터 얻은 특징지도 (Feature Map)에 Binary Mask를 씌워서 화소 단위로 구분을 짓는다. 하지만, 일반적인 특징 지도의 크기는 원본 영상의 크기도 4분의 1일에서 16분의 1 크기의 수준으로 작다. 이렇게 작아진 특징지도에서 화소 단위로 객체 영역을 결정하더라도 원본 영상에서는 그 오차가 크게 나타난다. 이 같은 오차를 감소시키기 위해

여 ROI Align이라는 기법을 사용하는데, 이는 ROI Pool 영역에서 생기는 소수점 오차를 2차 선형보간법으로 감소시키는 방법이다. 이를 통해 정확한 객체 인식과 분할이 가능해졌다.

지금까지 소개한 영상 처리 기술은 영상 내에서 특정 객체를 인식하여 정보를 제공하는 인공 지능 기술이다. 하지만 본 논문에서는 이와 유사한 영상 처리 기술을 이용하여 건설 기계로부터 작업자 안전사고를 예방하기 위한 활용 기술을 제안하고자 한다. 먼저 본 연구에서 수행한 실험에 필요한 가정에 대하여 소개하고, 다음으로 안전사고 예방을 위하여 적용한 기술에 대해 언급한다. 끝으로 실제 건설 현장에 적용한 결과에 대하여 분석하고 결론을 맺고자 한다.

2. 현장 실험과 알고리즘 구현

2.1 현장 실험을 위한 가정

건설기계와 작업자가 혼재된 환경에서 부주의에 의한 사고를 예방하기 위하여 가장 필요한 것은 정보의 공유이다. 이를 위하여 Fig. 3과 같은 자동화 시스템이 필요하다. 이는 건설기계의 운전자와 주변 작업자 간의 작업 상황 정보를 공유하고 인지하며 작업에 반영하여 안전사고를 예방할 수 있도록 하는 시스템이다.

본 연구에서는 시스템의 현실적인 구현을 위하여 건설 현장의 많은 공정 중 지반 다짐 작업을 대상으로 설정하였다. 이는 건설 분야에서는 건축 공사와 도로 공사에서 공통적이고 기본적으로 포함되는 작업이며, 다짐 장비는 기본적으로 전진 운행 뿐만 아니라 후진 운행이 많아 운전자의 비가시권 영역에 대한 작업이 대부분을 차지하기 때문이다. 이 같은 이유로 건설기계 주변 작업자의 인지 및 상태 판단이 필요하다. 또한 Fig. 3에 표시된 건설작업 안전 시스템의 정보 제공에 대해서는 현재 굴삭기 등에 적용된 머신가이던스 장비에 적용이 가능하다[15]. 대표적인 예로 Kim et al.의 소형 디스플레이 화면을 통한 운전자 정보 제공 채널에 관한 연구에 앞서 많은 선행 연구가 진행되었다. 따라서 본 연구에서는 단순한 영상 정보 제공 보다 환경 인식과 상황 판단을 위한 알고리즘을 제안하고자 한다.

본 연구에서는 카메라를 이용한 인식연구에 초점을 맞춘다. 영상을 이용한 측정 센서는 레이저와 UWB 같은 센서에 비해 알고리즘이 매우 복잡하지만, 유일하게 객체 인식 기술 적용이 가능하다. 이런 기술들 가운데 최근 들어 인공지능 기반의 영상 인식 기술에 활발하게 적용되

고 있다. 이 같은 점을 반영하여 본 연구에서는 인공지능 기반의 영상 처리 기법을 활용하고자 한다.

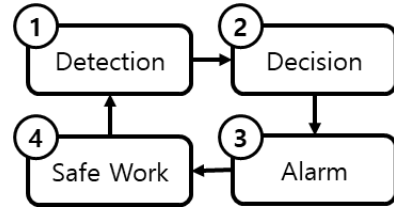


Fig. 3. Construction safety system flow

끝으로 본 논문에서는 현장에서 녹화된 영상을 활용하여 실시간으로 동작할 수 있는 알고리즘 개발에 초점을 맞춘다. 본 연구는 실시간 영상 처리를 바탕으로 환경 정보를 수집하고 안전 정보를 공유하는 시스템 제작의 선행 연구이다. 이 시스템을 위해서는 Fig. 3에서 나타난 것처럼 세부 기능이 구현되어야 하는데, 본 논문에서는 ①과 ②에 해당하는 기능 개발에 집중한다. 따라서 실시간으로 사용될 수 있는 알고리즘을 사전에 수집된 영상을 바탕으로 우선적으로 개발하고 검증하고자 한다.

2.2 YOLO-v3를 활용한 실험 환경 구축

앞에서 언급한 가정에 따라 Fig. 4에서와 같이 Detection과 Decision 단계에 해당하는 기능을 구현하고자 한다. 먼저 Detection 단계에서는 영상을 통해 얻은 정보 분석을 통하여 사람 또는 차량의 유무를 판단하는 것은 물론이고 영상 내에서 위치를 탐지한다. 다음으로 Decision 단계에서는 위치 정보를 활용하여 건설 장비의 주변에 위험 상황 여부를 판단하는 알고리즘을 수행한다. 본 논문에서는 2가지 단계를 수행할 수 있는 기술 개발을 위하여 실험 환경을 구축하였다.

본 실험은 세부 기능 개발이 목적이기 때문에 전체 시스템이 아닌 세부 기능과 관련된 사양 정의가 우선적으로 요구된다. 그 중에서도 가장 먼저 결정해야 하는 것은 영상 정보를 수집하기 위한 카메라이다. 건설기계 주변 환경 인식을 위하여 건설 현장에서 일반적으로 사용되고 있는 다짐 장비 후미에 Fig. 4와 같이 카메라를 설치하였다. 본 연구에서 일반적으로 사용되는 Gopro 카메라를 사용하였고 해상도는 FHD이며 초당 30 프레임으로 촬영하도록 설정하였다. 카메라의 설치는 우선 장비의 후방을 모두 볼 수 있도록 하고, 지면과 평행하게 하며 화면에서 지평선이 대략 3분의 2에 위치하도록 임의로 각도를 조정하였다. 그리고 다짐 장비가 작업을 하는 동안 계속해서 후방 영역을 녹화하였다.

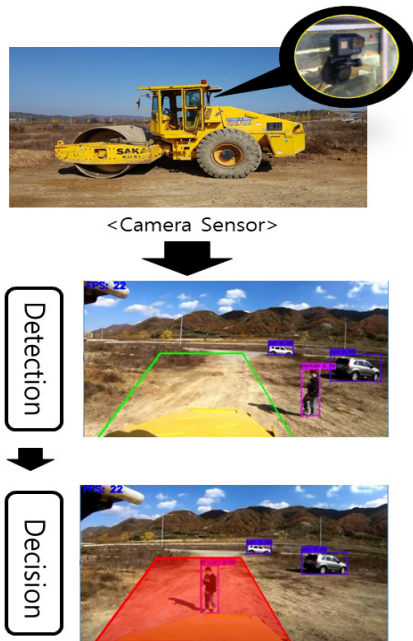


Fig. 4. Part of construction safety system for recognition of the surrounding objects and decision of the situation using a camera.

다음으로 대상 객체를 인식하기 위한 인공지능 알고리즘의 활용이다. 본 논문에서는 녹화된 동영상에 YOLO-v3을 적용하여 영상 처리를 수행하였다. 본 기술은 위험 상황을 빠르게 인식하는 것이 목표이므로 빠른 영상처리 속도를 가진 알고리즘이 필요하다. 이 같은 이유로 여러 객체 인식 알고리즘 중 YOLO-v3를 선택하였다. YOLO는 Fig. 5와 같이 하나의 합성곱 신경망이 동시에 여러 개의 경계 상자 (Bounding Box)를 예측하고 각 경계 상자에 대하여 분류 확률 (Class Probability)을 예측하는 알고리즘이다. 이때 사용되는 합성곱 신경망은 특징 지도를 생성하는 용도로 활용되는데 알고리즘의 중추를 담당한다고 하여 Backbone Network라고 부른다. 이 특징 지도는 여러 개의 Grid Cell로 구성되는데, 각 Cell 마다 Score 방식을 적용하여 대상 객체의 종류와 위치를 동시에 결정하는 알고리즘이다.

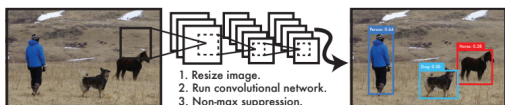


Fig. 5. The YOLO Detection System[16]

이 알고리즘은 조금씩 수정되고 발전해오면서 여러 가지

형태로 존재하지만 가장 최근에 발표된 것은 YOLO-v3이다. 이 알고리즘은 차량, 사람, 비행기, 동물 등을 포함하여 총 20가지 종류를 매 프레임마다 인식하도록 개발되었다. 하지만 본 논문에서는 건설현장에 있는 차량과 사람에 대한 인식 정보만을 사용하였다. 그리고 Backbone Network는 DarkNet-53로 선택하였고, 기 학습된 가중치를 사용하였다. 또한 본 모델의 입력 동영상 크기는 480×272로 설정하였다. Redmon et al.의 실험에 따르면 416×416일 때 29ms의 소요시간을 가지는 것으로 드러났는데, 이를 감안한 실시간 객체 인식을 위하여 이 같은 입력 영상의 크기를 설정하였다[17].

알고리즘 수행을 위해 Intel-i7, NVIDIA Geforce 1080 TI, 32GB Memory가 장착된 컴퓨터를 이용하였으며, 초당 약 21 프레임 이상을 처리하며 차량 및 작업자 인식이 가능한 것을 확인하였다. 대부분의 다짐 장비의 이동 속도는 최대 4m/s인 점을 감안하면, 0.27m 마다 발생할 수 있는 위험 상황을 감지할 수 있게 된다.

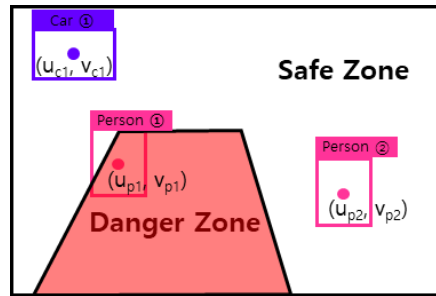


Fig. 6. Boundary between safe and dangerous region and positions of the detected objects in the image

2.3 건설 기계 주변 상황 판단 알고리즘

이 같은 인식 정보를 바탕으로 상황 판단을 할 수 있는 기준을 수립하였다. 건설 공사 작업 중 위험 상황 여부를 판단하기 위하여 다짐 장비 후방 영역을 2개의 영역으로 구분하였다. 일반적인 다짐 장비 작업의 경우 급격한 좌우 선회가 많지 않다. 그러므로 다짐 장비가 후진하는 경로에 장애물이 없을 경우 안전사고 발생이 크게 감소할 것으로 가정하였다. 또한 다짐 장비는 대부분 저속으로 운행하고 고속의 영상처리 속도를 고려했을 때 후방 10m 내에 있는 객체를 주의 깊게 볼 필요성이 있음을 가정하였다. 이 같은 가정을 바탕으로 Fig. 6과 같이 다짐 장비의 운행 방향을 고려하여 안전영역 (Safe Zone)과 위험영역 (Danger Zone)을 구분하였다.

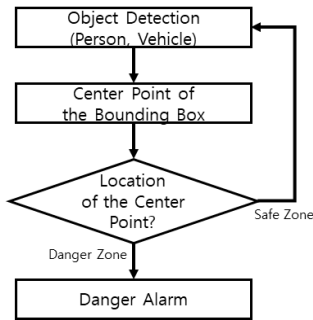


Fig. 7. Flowchart of detection algorithm for risk situation in construction sites

이 같은 기준으로 위험 상황을 판단하기 위하여 Fig. 7과 같은 규칙을 세웠다. 우선 인공지능 모델을 통하여 입력된 영상에서 차량 또는 사람을 인식한다. 인식한 결과는 경계 상자 (Bounding Box)의 형태로 표시한다. 다음으로 이 직사각형 모양의 경계 상자의 중심점 좌표를 구한다. 끝으로 이 중심점 좌표가 영상에서 설정한 위험 영역 내에 포함되는지 여부를 판단한다. 만일 포함될 경우 위험 상황을 알린다. Fig. 6에서 영상 내 1대의 차량과 2명의 사람이 있을 경우 인공지능은 3개의 객체를 인식한다. 그리고 3개의 객체에서 중심점 좌표를 직각 좌표계로 각각 계산한다. 그 결과 car①의 경우에는 (uc1, vc1), person①의 경우에는 (up1, vp1), person②의 경우에는 (up2, vp2)로 나타난다. 이 경우에는 person ①의 중심점 좌표는 현재 위험 영역 내에 있으므로 운전자에게 위험 알림을 주게 된다.



Fig. 8. Images for validation of detection algorithm

2.4 실험 결과 분석

이 알고리즘에 대한 성능을 정량적으로 검증하기 위해서 총 105장의 영상을 확보하였다. 이 영상들은 성토 다짐 장비 후방에 장착된 카메라를 통해 얻은 것으로 실제

다짐 작업을 하는 중에 촬영한 데이터이다. 총 4시간 이상 촬영한 동영상 중에서 사람과 차량이 포함된 영상을 추출한 데이터이다. 이 데이터들은 가운데 Fig. 8과 같이 위험 영역에서 사람 또는 차량을 포함하고 있는 것들로 선별하였다. 그 결과 84장의 사람을 포함하고 있는 영상과 46장의 차량을 포함하고 있는 영상을 확보하였다.

Table 2. Accuracy of detection algorithm for risk situation in construction sites

Number of Images	Person	Car	Images
Total Counts	84	46	105
Correct Counts	76	44	95
Accuracy(%)	90.48	95.65	90.48

이 데이터를 이용하여 알고리즘을 수행하였고, 인식 성능은 Table 2와 같다. Danger Zone에 사람이 포함된 84장의 영상들 가운데, 76장에 대해서 정확히 동작하였다. 그리고 차량이 포함된 46장의 영상들 중에서 44장에서 위험 알람을 주는 것으로 나타났다. 종합해보면, 총 105장의 영상 중에서 10장의 영상에서만 오류가 발생하는 것으로 확인되었다. 따라서 최종적으로 본 논문에서 제안하는 상황 판단 알고리즘은 90.48%의 정확도를 갖는다. 아울러 105장의 영상처리를 하는데 6.97초의 시간이 소요되어 초당 15.06 장의 영상을 처리하는 것으로 나타났다. 이는 대상 객체만을 인식하는 경우보다 소요 시간이 증가하는데, 그 원인은 판단 알고리즘 때문이다.

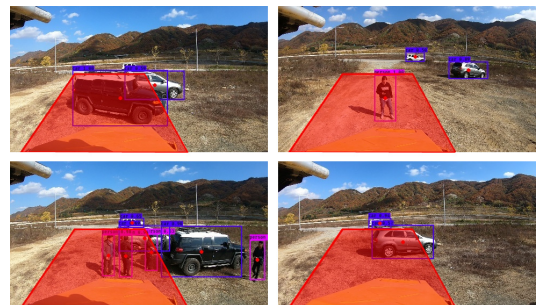


Fig. 9. Results of detection algorithm

이 같은 알고리즘을 적용한 결과는 Fig. 9와 같다. 사람 또는 차량이 위험 영역에 존재할 경우 해당 영역을 붉은색으로 표시하고 그렇지 않은 경우 녹색 테두리만 표시하도록 하였다. 그 결과 다짐 장비 주변에 있는 사람의

위치에 따라 다르게 표시되는 것을 확인하였는데, 특히 위험 영역에 객체가 있는 경우는 운전자에게 알림을 줄 수 있는 알고리즘이 정상 동작하는 점 또한 확인하였다.

다만 10장의 영상에서는 제대로 인식하지 못하는 현상이 나타났는데, 2가지의 원인으로 요약할 수 있다. 첫째는 다짐 장비로부터 측정하고자 하는 대상 객체가 영상에서 차지하는 면적이 작기 때문이다. YOLO-v3의 경우 여러 가지 mean AP를 기준으로 성능을 평가하였는데 그 중에서도 3가지의 면적 기준을 사용하였다 [17]. 그 기준은 첫째 면적이 32^2 화소 미만인 경우를 AP_S 로, 둘째 면적이 32^2 화소 이상 96^2 화소 미만인 경우를 AP_M 으로, 셋째 면적이 96^2 화소 이상인 경우를 AP_L 로 설정하였다. Redmon et al.의 실험에 따르면 입력 영상의 크기가 608×608 일 때, Darknet-53을 사용할 경우 AP_S 에서 18.3%의 성능을 보였다. 이 알고리즘을 동일하게 사용하고 입력 영상이 480×272 인 점을 감안하면 본 실험 결과의 인식 성능은 더욱 낮을 것으로 예상된다. 둘째는 2개의 객체가 서로 겹쳐 있는 경우이다. 이럴 경우 카메라와 가까이 있는 객체는 인식을 하지만, 카메라와 멀리 있는 객체의 경우 형태가 가려지기 때문에 정확하게 인식되지 않는다.

3. 결론

본 논문에서는 건설 기계와 주변 작업자 간 발생할 수 있는 안전사고를 예방하기 위한 방법에 대하여 살펴보았다. 특히 카메라와 인공 지능 객체 인식 기술을 활용한 위험 상황 판단 알고리즘에 대해 언급하였다. 그리하여 다짐 장비 주변에 사람과 차량을 인식하고 구분할 수 있는 객체 인식 기술을 바탕으로 다짐 작업 시 위험 상황을 알려줄 수 있는 방법을 실험하였다. 그 결과 초당 15.06 프레임을 처리하면서 90.48%의 정확도를 갖는 알고리즘을 확보하였다.

하지만 본 연구에서의 한계는 우선 해상도다. 본 실험에서 녹화 시 설정한 해상도는 1920×1080 이지만, 알고리즘에 사용한 해상도는 $1/16$ 정도로 축소되었다. 이로 인하여 영상에서 사람과 차량에 해당하는 면적이 축소되는 현상이 나타났다. 이 같은 이유로 녹화된 원본 영상에서는 보이지만, 축소된 영상에서는 보이지 않게 되어 인식이 불가능하게 되었다. 향후 이러한 점을 개선하기 위한 대안으로 고해상도 영상을 그대로 사용하여 연산이 가능한 고성능 연산장치가 필요할 것으로 예상된다.

다음으로 객체 인식을 위한 다양한 연구를 진행하고자 한다. 우선 카메라의 설치 조건에 따른 인식 성능에 대한 연구를 통해 동작 조건에 대한 명확한 규명이 필요할 것이다. 그리고 건설 현장에서 흔히 볼 수 있는 굴삭기, 트럭, 다짐 장비와 같은 객체를 인식하기 위하여 연구가 필수적일 것이다. 이를 위하여 인공지능 기반 학습을 위한 라벨 영상(Labeled Image) 마련이 필요하다. 또한 운전자에게 정확한 정보를 필요한 시간에 제공하기 위해서 실시간으로 구동될 수 있는 시스템 구축이 필요할 것으로 예상된다. 이를 위하여 인공지능 연산용 임베디드 보드를 통한 현장 적용과 운전자 정보 제공을 위한 증강 현실 기반의 인터페이스가 필수적인 기술이 될 것이라 전망된다.

References

- [1] Ministry of Land, Infrastructure and Transport. (Oct., 31, 2018). Smart Construction Technology Road Map. Available From: http://www.molit.go.kr/USR/NEWS/m_71/dtl.jsp?id=95081506 (accessed Feb., 8, 2019)
- [2] B. Jo, Y. Lee, D. Kim, J. Kim, P. Choi, "Image-based proximity warning system for excavator of construction sites", *Journal of the Korea Contents Association*, vol. 16, no. 10, pp. 588-597, 2016. DOI: <http://dx.doi.org/10.5392/JKCA.2016.16.10.588>
- [3] J. Y. Soh, J. Lee, C. H. Han, "Development of Omnidirectional Object Detecting Technology for a safer excavator", *Journal of the Korea Institute of Building Construction*, vol. 10, no. 4, pp. 105-112, 2010. DOI: <https://doi.org/10.5345/JKIC.2010.10.4.105>
- [4] J. Seo, S. Han, S. Lee, H. Kim, "Computer vision techniques for construction safety and health monitoring", *Advanced Engineering Informatics*, vol. 29, pp. 239-251, 2015. DOI: <http://dx.doi.org/10.1016/j.aei.2015.02.001>
- [5] J. Na, S. Lee, C. Kim, H. Son, C. Kim. "Real-time vision-based proximity detection for improved worker safety in construction equipment operation", *Proc. of Architectural institute of Korea*, vol. 35, no. 2, pp. 31-32, 2015.
- [6] S. Han, S. Lee, "A vision-based motion capture and recognition framework for behavior-based safety management", *Automation in Construction*, vol. 35, pp. 131-141, 2013. DOI: <http://dx.doi.org/10.1016/j.autcon.2013.05.001>
- [7] H. Kim, H. Kim, Y. W. Hong, H. Byun, "Detecting construction equipment using a region-based fully convolutional network and transfer learning", *Journal*

of computing in Civil Engineering, vol. 32, no. 2, 04017082, 2018.

DOI: [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000731](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000731)

- [8] S. Ren, K. He, R. Girshick, J. Sun, "Faster R-CNN: towards real-time object detection with region proposal networks", *Proc. of Advances in Neural Information Processing Systems 28 (NIPS 2015)*, Montreal, Canada, pp. 91-99, 2015.
- [9] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, L. Fei-Fei, "ImageNet large scale visual recognition challenge", *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211-252, 2015. DOI: <https://doi.org/10.1007/s11263-015-0816-y>
- [10] R. Girshick, J. Donahue, T. Darrell, J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation", *Proc. of the IEEE conference on computer vision and pattern recognition*, Columbus, Ohio, pp. 580-587, 2014.
- [11] R. Girshick, "Fast r-cnn", *Proc. of The IEEE International Conference on Computer Vision (ICCV)*, Santiago, Chile, pp. 1440-1448, 2015.
- [12] K. He, G. Gkioxari, P. Dollár, R. Girshick, "Mask R-CNN", *Proc. of The IEEE International Conference on Computer Vision (ICCV)*, Venezia, Italy, pp. 2980-2988, 2017.
- [13] X. Glorot, Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks", *Proc. of the thirteenth international conference on artificial intelligence and statistics*, Sardinia, Italy, pp. 249-256, 2010.
- [14] J. R. Uijlings, K. E. Van De Sande, T. Gevers, A. W. Smeulders, "Selective Search for Object Recognition", *International journal of computer vision*, vol. 104, no. 2, pp. 154-171, 2013. DOI: <https://doi.org/10.1007/s11263-013-0620-5>
- [15] W. Kim, S. Park, R. Lee, J. Seo, "A case study on the application of machine guidance in construction field", *Journal of the Korean Society of Civil Engineers*, vol. 38, no. 5, pp. 721-731, 2018. DOI: <https://doi.org/10.12652/Ksce.2018.38.5.0721>
- [16] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, "You only look once: Unified, real-time object detection", *Proc. the IEEE conference on computer vision and pattern recognition*, pp. 779-788, 2016.
- [17] J. Redmon, A. Farhadi, "Yolov3: An incremental improvement", arXiv preprint arXiv:1804.02767, 2018.

심 승 보(Seungbo Shim)

[정회원]



- 2011년 2월 : 포항공과대학교 전자전기공학과 (공학석사)
- 2017년 1월 : 삼성중공업 중앙연구소 (선임연구원)
- 2017년 1월 ~ 현재 : 한국건설기술연구원 차세대 인프라연구센터 전임연구원

<관심분야>

인공지능, 영상처리, 머신비전

최 상 일(Sang-Il Choi)

[정회원]



- 2017년 2월 : 경북대학교 IT대학 컴퓨터학과 (공학박사)
- 2017년 1월 ~ 현재 : 한국건설기술연구원 차세대 인프라연구센터 전임연구원

<관심분야>

통신네트워크, 사물인터넷, 블록체인