

360도 카메라를 이용한 딥러닝 기반 탑승객 탐지 시스템

이혜영, 신수용*

금오공과대학교

a4125a@kumoh.ac.kr, *wdragon@kumoh.ac.kr

Deep Learning Based Passenger Detection System Using 360 Degree Camera

Hye Yeong Lee, Soo Young Shin*

Kumoh National Institute of Technology

요 약

본 논문에서는 360도 카메라를 이용한 딥러닝 기반 탑승객 탐지 방법을 제안한다. 360도 카메라를 차량 내부의 천장에 장착하여 딥러닝 기반 잔여 인원 파악을 목표로 한다. 360도 카메라는 기존 카메라 대비 하나의 카메라로 넓은 영역을 촬영할 수 있다. 다만 360도 카메라로 촬영한 이미지는 왜곡이 있어 직접 촬영한 이미지를 활용해 학습 데이터를 만들었다. 또한, 인원 카운팅의 정확도를 높이기 위해 추적 기법을 활용해 움직이는 물체가 프레임이 변했을 때 같은 물체인지 확인하는 기법을 적용하였다.

I. 서 론

최근 어린이 보호 차량의 질식사 사고가 빈번히 발생하고 있다. 현재는 관리자인 보호 교사, 운전자가 직접 차량 내부 어린이들의 안전에 관여해야 하는 시스템으로 관리자의 작은 실수가 큰 사고로 이어질 수 있다. 이를 해결하기 위해 어린이 보호 차량의 등하교를 보조하기 위한 보조교사가 제안되었으나, 인건비를 충족할 수 있는 여건이 부족한 실정이다. 또한, 안전띠 사용 여부에 따라 차량 내부의 인원 파악을 자동으로 할 수 있는 안전띠에 장착하는 장치가 제안되었으나 안전띠 제거 후 사고가 발생할 경우를 대비하기가 어렵다.

본 논문에서는 360도 카메라를 이용한 딥러닝 기반 탑승객 탐지 방법을 제안한다. 360도 카메라를 차량 내부의 천장에 장착하여 실시간 스트리밍을 통해 딥러닝 기반 잔여 인원 파악을 목표로 한다. 360도 카메라는 기존 카메라 대비 시야각이 넓으므로 단일 카메라로 넓은 영역을 촬영할 수 있다. 하지만 360도 카메라로 촬영한 이미지는 왜곡이 있으므로, 기존 공개된 데이터셋의 활용이 어렵다. 따라서 본 논문에서는 이를 해결하고자 학습 데이터셋의 수집 방법을 제안한다. 또한, 탐지 정확도를 높이기 위해 이미지 프로세싱의 방법의 하나인 배경 제거(Background subtraction)와 딥러닝을 활용한 잔여 인원 탐지 방법을 제안한다.

II. 시스템 모델

본 논문에서 제안하는 360도 카메라를 이용한 딥러닝 기반 탑승객 탐지 시스템은 그림 1과 같다. 먼저, 360도 카메라는 일반적으로 구형대의 fisheye 뷰와 fisheye 뷰를 펼친 파노라마 뷰가 있다. 제안하는 시스템은 구형대의 fisheye를 기준으로 한다. 360도 카메라로 촬영한 실시간 비디오 스트림은 입력 데이터이다. 받은 입력 데이터는 각 프레임 당 이미지 프로세싱 과정인 배경 제거(Background subtraction) 방식을 거친다. 배경 제거는 배경 이미지와 다른 프레임의 이미지 비교나 필터를 이용해 배경만 제거하는 방식으로, 움직이는 객체 탐지(Object detection)에 많이 사용되고 있다[1]. 따라서 효과적인 배경 제거를 위해 일반적으로 배경이 변하지 않는 고정된 카메라로 촬영한 동영상상을 사용한다. 본 논문에서도 차량 내

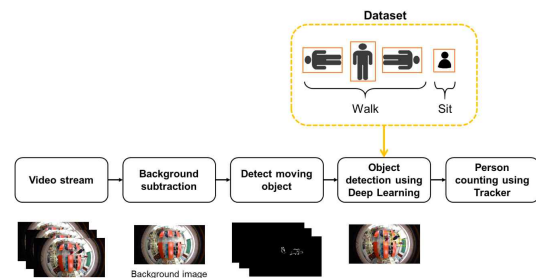


그림 1 시스템 모델

부의 천장에 카메라를 거치하여 사람이 없는 상황을 촬영하였고, 해당 이미지를 참조 이미지로 사용하여 움직이는 객체인 탑승객을 탐지하였다. 딥러닝으로 객체 탐지가 가능하나 배경 제거 방식을 사용 시 360도 카메라의 왜곡된 이미지 내에서 움직이는 객체를 먼저 탐지하여 최종 객체 탐지 정확도를 높일 수 있다. 사람이 없는 공간의 배경 이미지를 기준으로 threshold 기준을 넘을 때 움직이는 물체를 검출한다. 배경을 제거한 이미지에서 움직이는 객체만 찾을 수 있으며, 이는 딥러닝의 입력 데이터가 된다. 탑승객 인원 파악을 위해서는 360도 카메라의 왜곡된 사람 형태를 탐지해야 한다. 따라서 기존 데이터셋과 다른 학습 데이터가 필요하다. 카메라를 차량 내부의 천장에 거치하여 촬영하였을 때, 이미지에서 확인할 수 있는 사람의 형태는 크게 사람이 걷는 형태(Walk)와 앉은 형태(Sit)로 나눌 수 있다. 그림 1의 데이터셋 부분에서 나타났듯이 걷는 형태는 사람의 머리 방향에 따라 3가지 형태로 나뉘며, 앉은 형태는 하나로 고려하였다. 촬영한 동영상상을 기준으로 학습 데이터를 만들었으며, 회전, 반전, 확대, 잠음 등을 이용한 data augmentation 방식을 활용하여 부족한 학습 데이터 수를 증가시켰다[2]. 이렇게 얻은 학습 데이터를 기반으로 딥러닝을 이용한 객체 인식을 통해 사람을 탐지한다. 사람 탐지를 위한 딥러닝 기법 중 yolov3-tiny를 이용하였다[3]. yolov3-tiny는 yolov3에 비해 낮은 정확도를 가지나 빠른 속도를 가져 목표로 하는 실시간 시스템을 임베디드 보드에 적용하기 적합하여 tiny 모델을 딥러닝 모델로 결정하였다.

잔여 인원 파악을 위해 객체 탐지 결과를 바탕으로 카운팅 기능이 필요

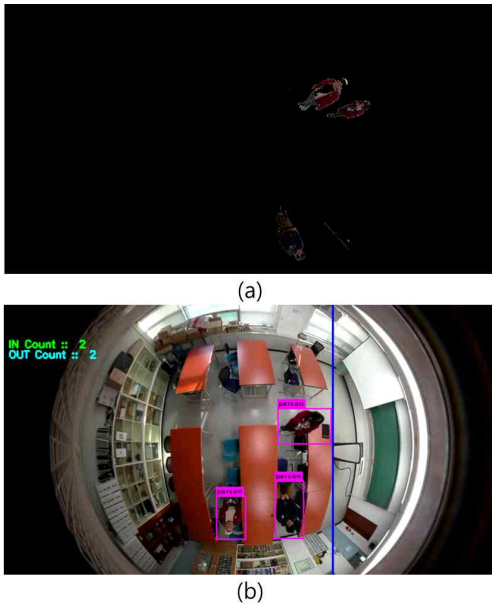


그림 2 결과 이미지 (a)배경 제거 결과 (b)모의실험 결과

하다. 입구를 기준으로 차량 내부로 들어오는 사람과 나가는 사람 수를 카운팅할 수 있다. 입구를 기준선으로 두고 객체 탐지 결과인 바운딩 박스(Bounding box)의 중앙점 좌표값을 이용해 기준선의 x 좌표값과 비교해 사람이 출입을 확인하여 카운팅하였다. 하지만 객체 탐지 기법인 yolo는 각 프레임마다 작동하여 같은 객체를 다른 바운딩 박스로 인식할 수 있다. 따라서 인원 카운팅 기능에서 같은 객체를 다른 바운딩 박스로 카운팅할 경우 정확하게 인원을 세릴 수 없다. 이를 해결하는 데 필요한 기법이 추적(tracker)이다. 추적 기법을 통해 같은 객체의 움직임을 파악하여 하나의 객체라고 인식하면 인원 카운팅의 정확도를 높일 수 있다. 본 논문에서는 sort(Simple Online and Realtime Tracking)를 활용하여 이전 프레임과 현재 프레임의 객체 추적을 통해 같은 객체임을 파악하였다[4].

III. 모의실험 결과

실험 환경으로 연구실 천장에 360도 카메라를 장착하여 촬영한 동영상을 이용해 객체 탐지 및 잔여 인원 카운팅 기능을 구현하고 성능 분석을 하였다. 360도 카메라는 kodak사의 pixpro sp360 4K를 이용하였다. 동영상 해상도는 1072×1072 30fps로 설정하였다. 직접 촬영한 동영상으로 학습 데이터 이미지 200장을 얻었다. data augmentation 기법 중 수직 반전, 수평 반전, zoom 추가, 회전 4가지 기능을 이용해 기존 이미지를 변형시켜 800장의 데이터셋을 얻었다. 얻은 데이터셋을 기준으로 라벨링 작업을 진행하였다. 객체 탐지 기법인 yolo3-tiny의 필터 수는 식(1)에 따라 사람(class=1)만 인식하기 위해 18로 설정하였다.

$$\text{filter} = (\text{classes} + 5) \times 3 \quad (1)$$

학습에 따른 테스트 결과는 그림 2와 같다. 그림 2.(a)는 배경 제거 후 이미지로 이미지에서 움직이는 객체만 검출할 수 있다. 그림 2.(b)는 전체 시스템 결과로 오른쪽 파란선을 입구인 기준선으로 가정하였다. 배경 제거를 기준으로 움직이는 물체를 검출하고, 학습된 weight 파일을 이용해 최종적으로 객체를 탐지할 수 있다. 학습은 mAP(mean average precision) 비교를 통해 학습을 진행하였으며, 2100번째일 때 가장 높은 mAP 성능인 98.89%로 확인할 수 있었다. mAP 결과는 반복 학습이 1000번째일때는

86.45%를 가졌고, 2000번째일 경우 97.48%를 가졌다. 2200번째일때는 97.7%로 성능이 더 좋아졌으며, 그 이후 현재 평균 손실값이 더 이상 줄어들지 않아 학습을 마치고 최적 학습을 통한 2100번째 weight 파일을 사용하였다.

IV. 결론

본 논문에서는 360도 카메라를 이용한 딥러닝 기반 탐승자 탐지 방법을 제안한다. 이미지 프로세싱 기법인 배경 제거와 객체 탐지 기법인 yolo3-tiny를 사용하여 움직이는 물체를 먼저 검출하고, 사람인지 판별한다. 그리고 추적 기법인 sort를 활용하여 같은 물체를 추적하여 인원 카운팅의 정확도를 높였다. 추후 실시간 스트리밍을 위한 전체 시스템 구현을 위해 Jetson Xavier와 서버를 통한 시스템 구축 및 성능 평가를 목표로 한다.

ACKNOWLEDGMENT

이 논문은 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2019R1A2C1089542).

참 고 문 헌

- [1] Piccardi, Massimo. "Background subtraction techniques: a review." 2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No. 04CH37583). Vol. 4. IEEE, 2004.
- [2] Shorten, Connor, and Taghi M. Khoshgoufar. "A survey on image data augmentation for deep learning." *Journal of Big Data* 6.1 (2019): 60.
- [3] Redmon, Joseph, et al. "You only look once: Unified, real-time object detection." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
- [4] Wojke, Nicolai, Alex Bewley, and Dietrich Paulus. "Simple online and realtime tracking with a deep association metric." 2017 IEEE international conference on image processing (ICIP). IEEE, 2017.