

광각 카메라를 활용한 합성곱 신경망 기반 재실감지

김상현*, 권동우*, 지영민*

*전자부품연구원

email: {kei97103, dwkwon, ym.ji}@keti.re.kr

CNN based Occupancy Detection with Wideangle Camera

Sanghun Kim*, Dongwoo Kwon*, Youngmin Ji*

*Korea Electronics Technology Institute

요 약

효과적인 에너지 절약 시스템을 구현하기 위해서는 실내의 재실자 존재 여부를 판별할 수 있는 근거가 필요하다. 본 연구에서는 천장에 광각렌즈 카메라를 설치하여 재실자 수를 탐지하는 방법을 소개한다. 인식 기술은 합성곱 신경망의 한 종류인 Yolo v2를 사용하였으며, 건물 내부의 다양한 장소의 천장에 카메라 센서를 설치하여 필요한 데이터를 수집하였다. 수집한 2,200장의 데이터를 기준으로 학습을 수행하였고, 인식의 정확도를 측정해본 결과 96.15%의 정확도와 91.72%의 재현율을 얻어냈다.

1. 서론

에너지 소비량이 점차 증가함에 따라 에너지 절약을 위한 연구가 활발히 진행되고 있다. 에너지 절약에 있어서 가장 핵심적인 것은 실내 냉, 난방 기구의 낭비를 줄이는 것이며, 이러한 냉, 난방 기구는 실내의 사람을 중심으로 사용된다. 따라서 재실자가 있는지를 판별하여 실내에 사람이 없을 때 이러한 냉난방 장치를 제어할 수 있다면 에너지 절감에 크게 기여할 수 있을 것이다.

기존 연구에서는 라즈베리파이에 이산화탄소 센서를 부착하여 재실자 수를 판단하고자 하였다[1]. 하지만 밀폐된 환경에서는 사람이 밖으로 나가도 공기가 순환되지 않았기에 정확한 재실자 수를 판단할 수 없었다. 본 논문에서는 천장에서 찍은 사진을 토대로 재실자 수를 확인하기 위한 방법과 그 결과를 소개하고자 한다.

2. 이론적 배경

합성곱 신경망(Convolutional Neural Network)이란 이미지 인식에 주로 사용되는 Neural Network로써 Convolution Layer와 Pooling Layer를 반복적으로 사용하여 이미지에서 객체를 도출해낸다. 본 연구에서는 널리 알려진 합성곱 신경망 모델들 중 하나인 Yolo v2(You Only Look Once v2)를 사용한다[2]. 라즈베리파이가 720x480 크기의 사진을 Yolo v2의 framework인 Darknet에 넣어주면, Darknet은 해당 이미지를 608x608 크기로 변경(resize)한 뒤 19x19개의 셀로 나누어 해당 셀에서 객체가 존재하는지, 해당 객체가 어떠한 종류로 분류되는지에 대한 확률을 계산하여, 결과를 사용자에게 보여준다.

3. 실험 방법

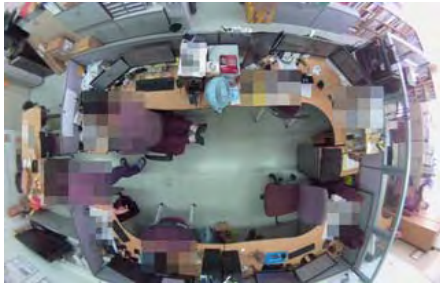
그림 1의 센서는 천장에 설치하여 재실자를 탐지하기 위해 설계된 디바이스이다. 라즈베리파이3에 카메라만을 연결한 단순한 구조이며, 적은 수의 센서로도 넓은 범위를 커버하기 위해 일반 카메라 렌즈가 아닌 광각렌즈(Camera Module v2 Wide Angle Lens)로 설치하였다. 이 센서는 사진을 찍은 뒤 인식을 수행하고, 그 결과를 데이터베이스로 전송하는 역할을 한다.



(그림 1) 천장설치용 카메라 센서

본 연구에서는 하나의 연구실에 그림 1의 센서 두 대를 설치하여 연구실 전체 재실자 수를 측정할 수 있는 환경을 구성하였다.

그림 2에서 나타나듯이 광각렌즈에서 찍히는 사진들은 일반렌즈에서 찍히는 사진들과 달리 사진이 굴곡지게 변형되기에, Yolo v2에서 제공하는 기본 가중치로는 제대로 된 결과를 얻을 수 없었다. 따라서 천장센서를 위한 고유한 가중치를 얻어야 할 필요가 있었고, 이를 위해 다양한 위치에 센서를 추가로 설치하였다. 이 센서들에서부터 지속적으로 수집되는 데이터를 기반으로 Yolo v2의 학습을 수행하였다.



(그림 2) 광각 카메라로 찍은 사진

Yolo v2의 학습은 GPU가 세팅된 컴퓨터에서 수행하였다. YOLO annotation tool을 활용하여 사진에서 보이는 사람의 위치와 크기를 표시하여 라벨파일로 저장하였다. 또한 수집되는 데이터가 한정적인 상황으로 인해 과적합(overfitting)이 발생하는 것을 방지하고자 사진을 좌우 반전, 상하 반전, 180° 회전의 3가지 방식을 사용하여 변형을 수행하였다. 3가지 형태의 변형 데이터에 대한 라벨파일은 원본 라벨파일에서 단순한 뉘셈 연산을 수행하는 것으로 도출해낼 수 있다. 이렇게 해서 550장의 원본데이터와 1,650장의 변형데이터를 합쳐 총 2,200장의 학습 데이터를 토대로 학습을 수행하였다.

4. 분석 결과

검증 위치에서 수집한 1,000장의 사진을 테스트 데이터로 선정해 인식률을 측정했다. 인식의 임계치(threshold)는 35%로 설정하였으며, 인식률 측정 기준은 Confusion Matrix를 사용하였다. 본 연구에서는 Yolo v2의 클래스를 사람 클래스만으로 제한하여 인식을 수행했기에 True Negative(사진에서 사물로 탐색된 사람 수)를 반영할 수 없었다. 따라서 True Negative를 반영하는 지표인 정확도(Accuracy)은 제외하고, 정밀도(Precision)와 재현율(Recall)만을 지표로 사용하였다. 각 지표의 계산방법은 식 1, 식 2와 같다.

정밀도는 인식된 객체가 사람일 확률을 의미한다. 정확도가 높을수록 Yolo v2가 사진에서 사람을 인식해낼 확률이 높다는 것을 의미한다.

재현율은 사람 수가 정확하게 측정된 정도를 나타낸다. 재현율이 높을수록 센서가 도출해내는 실내 재실자의 수의 정확도가 높아진다.

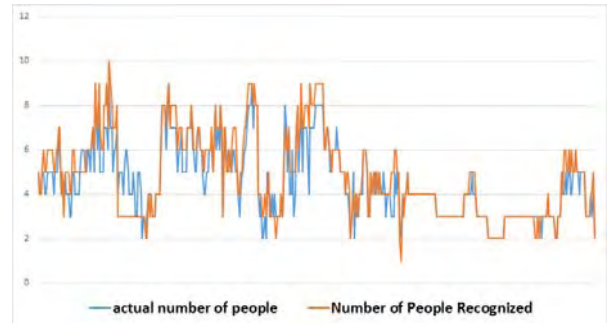
1,000장의 사진을 직접 비교해 본 결과 정밀도는 96.15%, 재현율은 91.72%로 높게 측정되었다.

$$\text{precision} = \frac{TP}{(TP + FP)} \quad (1)$$

$$\text{Recall} = \frac{TP}{(TP + FN)} \quad (2)$$

그림 3은 실제 사람의 숫자와 인식된 사람 수의 차이를 보여준다. 두 그래프 간의 차이가 발생하는 부분의 원인을 분석하면 다음과 같다. 우선 검증 위치에 설치된 두대의 광각렌즈 카메라 간에 겹치는 부분이 존재한다. 광각

렌즈는 일반렌즈에 비해 초점 거리가 짧아, 사진의 외곽부가 굴곡지게 나타난다. 이로 인해 두 카메라 간 겹치는 부분이 존재하여 이 부분에서 중복으로 사람이 인식되는 현상이 발생하였다. 또한 한정적인 장소에서만 데이터를 수집하다보니 데이터의 다양성이 부족했다. 이로 인해 특징적인 사물이 사람으로 인식되어 사람 수가 높게 측정되거나, 반대로 사람이 겹쳐서 찍힐 경우 사람 수가 낮게 측정되는 상황이 발생하였다.



(그림 3) 센서로 인식한 사람 수와 실제 사람 수간의 비교

카메라의 영역이 겹치는 문제에 대해서는 카메라의 위치를 재조정하여 겹치는 부분을 최소화하고, 사진의 해당 부분을 보이지 않게 만드는 방법을 활용하여 중복인식을 처리할 계획이다. 데이터의 다양성 측면에서 발생하는 문제는 데이터 수를 더 늘려가면서 발생하는 에러들을 제거해나갈 계획이다.

5. 결론

본 논문에서는 건물 내 에너지 절약을 위해 천장에 광각렌즈 카메라 센서를 부착한 뒤 찍은 사진을 바탕으로 재실자의 존재여부를 판단하였다. 학습에 사용되는 데이터가 2,200장으로 많지 않은 수준이었기에 에러가 발생하였지만, 지속적으로 데이터를 수집하여 사람에 대한 다양한 형태들을 학습시킨다면 훨씬 더 높은 인식률을 보일 수 있을 것으로 기대된다.

ACKNOWLEDGMENT

본 연구는 산업통상자원부(MOTIE)와 한국에너지기술연구원(KETEP)의 지원을 받아 수행한 연구 과제입니다. (No. 20182010106460)

참고문헌

[1] 지영민, 유준재, “지능형 IoT 기반 건물 자동제어 시스템”, 설비저널 제46권 제7호, pp.32-40, 2017.7.
 [2] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection,” The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.779-788, 2016.