

# CCTV 영상 기반 건물 공간 실시간 재실 인원 추론 시스템 연구

차영민, 권동우, 권현정, 지영민

한국전자기술연구원

{cym8026, dwkwon, hyunjoy1221, ym.ji}.keti.re.kr

## A Study on the Real-Time Inference System of Building Space Based on CCTV

Youngmin Cha, Dongwoo Kwon, Hyunjung Kwon and Youngmin Ji

Korea Electronics Technology Institute (KETI).

### 요약

본 논문은 CCTV 데이터를 이용해 객체 탐지 및 객체 추적을 이용하여 인원의 위치를 파악하고 익명화된 데이터로 변환한다. 이후 이를 이용해 건물 내 공간의 재실 인원 수를 추론하는 시스템이다. 이를 통해 CCTV가 특정 공간 전체를 비추지 않더라도 해당 공간의 출입구를 CCTV로 확인할 수 있다면, 건물 내 특정 공간의 실시간 재실 인원 수를 실시간으로 추론할 수 있다.

### I. 서론

건물 자율제어를 통해 에너지 절감을 위해서 실내 환경 센서 측정도 중요한 부분이지만, 건물 내에 재실 인원의 유무, 재실 인원 수 등을 파악하는 것 역시 에너지 절감을 위해 사용할 수 있는 중요한 지표 중 하나가 된다.[1] 재실 관련 연구들 중 PIR, LiDAR 등 센서들을 이용한 다양한 연구들이 존재하지만, 이 중 성능이 우수하면서도 건물 내에 보편적으로 설치되어 있어 추가적인 센서들을 설치할 필요 없는 CCTV를 활용하는 방식을 이용하게 되었다.

본 논문에서는 CCTV를 이용해 건물 내 공간 실시간 재실 인원을 추론할 수 있는 시스템에 대해 설명한다. 해당 기술은 Vision AI 기술 중 객체 탐지(Object Detection), 객체 추적(Object Tracking) 기술을 적용하였고, 이후 이미지에서 검출한 bounding box와 id 등의 익명화된 정보로 변환되어 개인정보 보호법을 준수하며, 변환 데이터를 이용해 공간의 재실 인원 수를 추론한다.

### II. 본론

CCTV는 촬영을 담당하고, 영상을 저장은 DVR 장비를 이용해 저장된다. CCTV와 연결된 DVR 장비에서 RTSP(Real-time streaming protocol) 통신 방식을 이용해 실시간으로 CCTV에서 촬영중인 영상을 실시간으로 받아올 수 있다. RTSP가 연결된 후, 각 프레임별로 받아오는 이미지를 현재 시간과 함께 객체 탐지를 수행한다.

객체 탐지는 YOLO[2]를 선택하였는데, YOLO는 You Only Look Once의 약자로 객체 탐지 모델 중 하나이다. YOLO의 특징으로는 실시간 처리 능력이 매우 뛰어나다는 것과, 성능 및 정확도 역시 타 모델들에 비해 우수하다는 점이 있다. 이러한 특성으로 인해 실시간 처리를 진행해야 되기 때문에 YOLO를 선택하였으며, 여러 YOLO 시리즈 중 안정성과 정확성 측면을 고려하여 YOLOv7[3]을 사용하였으며, 환경에 맞게  $conf\_threshold = 0.2$ ,  $iou\_threshold = 0.35$  등 하이퍼파라미터를 조정하였다.  $conf\_threshold$ 는 신뢰도 임계값으로써 모델이 해당 객체를 얼마나 확신하는지 나타내는 매개변수이다.  $iou\_threshold$ 는 IOU 임계값으로, 한

객체에 대해 여러개의 예측이 나타나는 경우가 있는데, 이때  $iou\_threshold$ 를 통해 중복되는 bounding box를 제거하는 역할을 한다. 객체 탐지를 진행한 후에는 해당 영상은 사용하지 않게 되며, 객체 탐지로 얻은 데이터들을 사용하게 된다.

객체 탐지를 거친 후, 객체 추적을 통해 bounding box의 id를 생성하게 된다. 객체 추적은 비디오나 연속된 이미지 프레임에서 특징적인 정보간의 유사도를 이용하여 특정 객체를 지속적으로 식별하고, 그 위치와 움직임을 추적하는 과정이다. 재실 감지에 있어서 객체 추적은 상당히 중요한 역할을 수행하게 되는데, 객체 탐지 이후, 이미지는 사용하지 않기 때문에 인원이 건물의 공간으로 들어갈때와 나갈때를 알 수 있는 유일한 방법이다. 영상은 여러장의 이미지들을 FPS단위로 연속해서 재생하는데, 이를 객체 추적을 통해 id로 같은 객체를 연결하게 된다면, 연속적인 경로 추적을 진행함으로써 건물 내 특정 공간의 출입 여부를 알 수 있다. 현재 객체 추적은 BoT-SORT[4]를 사용하고 있으며, 환경에 맞게  $match\_thresh = 0.6$ ,  $track\_high\_thresh = 0.4$  등 하이퍼파라미터로 조정 한 후에 시스템에 적용하였다.  $match\_thresh$ 는 새로운 객체 검출 결과와 기존 추적하고 있는 추적 객체를 연결할 때 사용하는 임계값으로, 높을수록 객체간 매칭이 엄격해지지만, 추적 대상 객체를 놓칠 위험이 존재한다.  $track\_high\_thresh$ 는 해당 점수 이상의 객체들만 추적 객체로 간주하여 추적을 진행하는 매개변수이다.

객체 탐지 및 객체 추적을 진행하고 나면 이를 이용해 건물 내 공간 실시간 인원 수 추론을 진행하게 된다. 건물 내 공간의 출입구 부분에 출입 여부를 판단하는 경계선을 그은 후 객체 탐지 및 추적을 진행한 이후인 bounding box의 밑변의 중앙에 해당하는 부분, 발에 해당하는 좌표를 구하고, 좌표가 경계선을 넘어가게 된다면 해당 id의 상태를 공간에 들어왔는지 나갔는지에 대해 알 수 있게 된다. 이를 누적함으로써 실시간으로 재실 인원 수를 알 수 있게 된다.

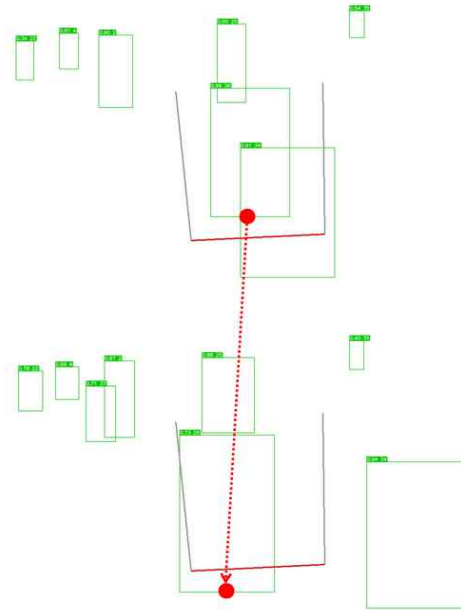
#### Algorithm 1 Count Algorithm

```
1: Initialize past state as undefined
2: Get line points  $(x_1, y_1), (x_2, y_2)$ 
3: Get bounding box corners  $(lt_x, lt_y), (rb_x, rb_y)$ 
4: Calculate foot position  $(foot_x, foot_y)$ 
5: Set  $foot_x = (lt_x + rb_x)/2$ 
6: Set  $foot_y = rb_y$ 
7: Compute slope and y-intercept
8: Set  $slope = (y_1 - y_2)/(x_1 - x_2)$ 
9: Set  $y\_intercept = y_1 - slope * x_1$ 
10: Calculate  $line_y = slope * (foot_x - x_1) + y_1$ 
11: if  $foot_x > x_1$  AND  $foot_x < x_2$  then
12:   if  $foot_y < line_y$  then
13:     Set current_state = "in"
14:   else
15:     Set current_state = "out"
16:   end if
17: else
18:   Set current_state = None
19: end if
20: if past_state  $\neq$  current_state then
21:   if past_state = "in" AND current_state = "out" then
22:     Set event = {detection_id: "out"}
23:   else
24:     if past_state = "out" AND current_state = "in" then
25:       Set event = {detection_id: "in"}
26:     else
27:       Set event = {detection_id: None}
28:     end if
29:   end if
30: end if
31: Set past_state = current_state
32: return event
```

(그림1) 재실 알고리즘 의사코드

(그림1)은 재실 알고리즘을 간단하게 표현하기 위해 작성한 의사코드이다. inout을 판단하는 경계선을 지나는 두 점  $(x_1, y_1), (x_2, y_2)$ 을 이용해 일차함수  $y = slope * (x - x_1) + y_1$ 을 구한다. 이후 bounding box  $lt_x, lt_y, rb_x, rb_y$ 에서 발 위치에 해당하는 좌표  $foot_x = (lt_x + rb_x)/2$ ,  $foot_y = rb_y$ 를 구한 후,  $foot_x$ 의 범위가  $x_1$  초과  $x_2$  미만일 때, 위에서 구한 일차함수 x,y에  $foot_x, foot_y$ 를 넣어 나온 값  $line_y$ 과  $foot_y$ 을 비교하여 현재 상태를 확인한다. 현재 상태가 확인되었다면, 해당 id의 과거 상태와 비교한다. id의 과거 상태가 없는 경우 None이며, 해당 id 과거 상태는 재실 알고리즘이 끝난 후 별도로 저장된다. 인원이 한명 들어갔다고 판단하는 경우, 해당 인원이 밖에서 안으로 들어간다는 것이며, 또한 밖에서 확인한 인원과 안에서 확인한 인원이 서로 같은 id를 가지고 있어야 한다. 이런 이유로 과거 id 상태와 현재 id 상태를 비교하여 재실의 출입 여부를 판단하며, 알고리즘이 끝난 후에는 과거 id상태를 현재값으로 업데이트하고 출입 여부를 반환한다. 반환한 출입 여부 이벤트를 누적시켜 재실자 인원 수를 알 수 있다.

(그림2)는 실제 실증 사이트에서 실제 이미지를 제외한 bounding box 등의 모습을 시각화로 나타낸 모습이다. 위 그림에서 bounding box는 초록색, 회색 선은 출입구 부분이며, 빨간 선이 in/out 재실 인원 수를 측정할 수 있는 경계선이다. 만약 예측하고자 하는 공간이 위쪽인 경우, 위에 위치한 이미지의 빨간 점을 가진 bounding box가 아래 이미지처럼 출입구의 경계선을 위에서 아래로 통과했기 때문에 재실 인원이 한명 나갔다고 판단한다.



(그림 2) 재실 인원 시스템 동작 예시

### III. 결론

본 논문에서는 CCTV 영상을 이용한 건물 내 공간 실시간 재실 인원 추론 시스템을 적용해 보았다. 객체 탐지 및 객체 추적을 이용해 공간 내 실시간 재실 인원 수를 좋은 성능으로 추론할 수 있을 것으로 예상하였다. 하지만, 객체 추적에서 ID switching 현상 및 새로운 ID로 할당되는 경우가 종종 발생해서 in/out을 판단하는 경계선을 넘을 때 제대로 카운트가 되지 않는 경우들이 종종 발생한다. 이로 인해 인원 수가 맞지 않는 경우가 발생한다는 단점이 존재한다. 따라서 향후에는 ID를 기반하지 않고 재실 인원 수를 추론하는 방법 혹은 더 성능이 좋은 객체 추적 모델을 이용하여 개선할 여지가 있다.

### ACKNOWLEDGMENT

본 연구는 산업통상자원부(MOTIE)와 한국에너지기술평가원(KETEP)의 지원을 받아 수행한 연구 과제입니다. (No. 20212020800120)

### 참고 문헌

- [1] M. Kim, and S. Son, "Study on the relationship between resident number and energy consumption using AI human body detection technology," Proceedings of Symposium of the Korean Institute of communications and Information Sciences, pp. 1008-1009, Feb 2022
- [2] Redmon, Joseph, et al. "You only look once: Unified, real-time object detection." Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 799-788, May 2016.
- [3] Wang, Chien-Yao, Alexey Bochkovskiy, and Hong-Yuan Mark Liao. "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 7464-7465, Jul 2023.
- [4] Aharon, Nir, Roy Orfaig, and Ben-Zion Bobrovsky. "BoT-SORT: Robust associations multi-pedestrian tracking." arXiv preprint arXiv:2206.14651, 2022.