

AI비전 기반 오프라인 공간의 행동 분석을 위한 모니터링 시스템 설계에 관한 연구

김민준, 김린희, 성정규, 윤민상, 최진혁, 윤수연*
국민대학교, *국민대학교

kkmj1374@kookmin.ac.kr, flsghl@kookmin.ac.kr, juku1133@kookmin.ac.kr,
nornen20@kookmin.ac.kr, skan9708@kookmin.ac.kr, *1104py@kookmin.ac.kr

Designing an AI Vision-Based Monitoring System for Behavioral Analysis in Offline Environments

Kim Min Jun, Kim Lin hoe, Seong Jeong Kyu, Yoon Min Sang, Choi Jin Hyeok, Yoon Soo Yeon*
Kookmin Univ., *Kookmin Univ.

요약

본 연구는 YOLOv8과 DeepSORT를 결합한 실시간 객체 인식 후 추적 시스템을 통해 오프라인 공간의 고객 행동을 시각적으로 분석하였다. 히트맵과 트래픽맵을 활용해 이동 경로와 체류 패턴을 직관적으로 확인할 수 있으며, MOTA 71.2%, ID Switch 116회의 성능을 기록하여 기존 기법 대비 우수한 정확도를 보였다. 이를 통해 마케팅 전략 수립과 공간 운영 효율성을 높이기 위하여 활용될 것으로 기대한다.

I. 서론

1.1 연구 배경 및 필요성

최근 팝업스토어, 플래그십 스토어와 같은 다양한 형태의 오프라인 행사가 증가하면서 소비자들의 행동 데이터를 정밀하게 수집하고 분석할 필요성이 높아지고 있다. 그러나 기존의 구매 데이터 중심의 분석 방식으로는 복합적이고 다양한 목적을 가진 오프라인 행사에서 소비자 행동의 세부적인 특성을 충분히 파악하기 어렵다. 특히 오프라인 공간은 상품 판매 외에도 브랜드 인지도 상승과 소비자 참여 및 마케팅 효과 평가 등 다양한 목적으로 운영되고 있어, 보다 정교한 행동 예측 모델의 필요성이 요구되고 있다. 최근 개인정보보호법의 특례 조항 도입과 YOLO v8과 같은 객체 인식 알고리즘 및 객체 추적 기술의 발전은 이러한 연구를 기술적으로 가능하게 하였다.

II. 관련 연구

2.1 오프라인 공간의 행동 예측을 위한 AI 알고리즘
객체 탐지와 추적은 오프라인 고객의 행동을 실시간으로 분석하는 데 핵심적인 역할을 하며, 이 분야에서는 YOLO와 DeepSORT 알고리즘이 널리 활용되고 있다. YOLO는 이미지 전역을 단일 패스만으로 분석함으로써 객체 탐지 속도와 정확도 면에서 우수한 성능을 제공하며, 특히 YOLOv8은 경량성과 정밀도를 동시에 달성한 모델로 평가받는다. 이와 더불어 DeepSORT는 기존 SORT의 약점을 보완하여 객체 재식별(Re-ID) 기능을 포함하고 있으며, 다중 객체 간 혼잡도가 높은 환경에서도 안정적인 객체 추적을 가능케 한다.

III. 실험설계

3.1 실험 환경, 데이터 구성

3.1.1 실험 환경 구성

다양한 환경의 영상 데이터 수집 본 연구에서는 실제 오프라인 매장에서 수집된 다양한 환경의 영상 데이터를

사용하였다. 이 데이터세트는 다양한 조명 조건, 카메라 각도, 인구 밀집도를 고려하여 구성하였다.

3.1.2 데이터세트 구성

객체 탐지 및 추적을 위한 레이블링 객체 탐지 및 추적 알고리즘 학습을 위해 개별 객체의 바운딩 박스와 군중 수에 대한 정확한 레이블링 작업을 진행하였다. 이러한 세부적인 데이터 구성은 행동 분석 모델의 정확한 학습과 평가에 핵심적인 요소로서 작용하고 있다.

3.2 오프라인 공간의 행동 분석 모델

3.2.1 YOLO v8 객체 인식

사람을 탐지하기 위하여 객체 검출 모델인 YOLO-v8을 활용하였다. YOLOv8은 anchor-free 방식의 decoupled head를 도입하여 검출 정확도를 높였으며, 연산 효율을 유지하도록 설계되었다. 본 연구에서는 YOLO-v8을 파인튜닝하여 각 프레임의 사람 수를 검출하고 이를 객체 추적 모듈의 입력으로 활용했다. 그 결과 mAP지표에서 높은 성능을 보였으며, 실제 장면에서 실시간(30 FPS 이상) 처리 가능한 속도를 확인하였다.

3.2.2 DeepSort 객체 추적

YOLO v8을 통해 검출된 (프레임별 바운딩 박스)를 입력으로 받아 동일한 객체가 시간 연속 프레임에서 어떤 경로로 이동하는지 ID를 부여하여 추적했다. DeepSort를 적용한 결과 복잡한 군중 속에서도 대부분 사람 개체에 대한 일관된 ID 추적이 이루어졌으며 ID Switch는 SORT만 사용했을 경우보다 현저히 감소했다. 추적 알고리즘의 주요 파라미터로는 IoU임계값, 최대 실종 프레임수 등을 조정하여 군중 환경에 최적화하였다.

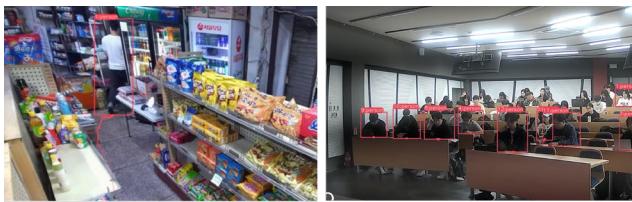


그림 1. 객체 인식 및 추적 학습
Datasets for AiHub 실내 구매 행동 데이터

3.2.3 히트맵 및 트래픽 맵 시각화

분석된 이동 경로와 체류 시간 데이터를 히트맵 및 트래픽 맵으로 시각화하여 사용자들이 직관적으로 공간 내에서의 행동 패턴을 파악할 수 있도록 제공하였다.

IV. 실험 결과, 성능 분석

4.1 히트맵 분석

YOLOv8과 DeepSort 기반으로 생성된 히트맵은 고객들이 주로 밀집하거나 많은 관심을 보인 공간을 시각적으로 나타내었다. 분석 결과, 특정 부스나 제품 전시 구역에서 높은 밀집도가 관찰되어 마케팅 전략의 초점을 효과적으로 맞출 수 있는 정보를 제공하였다.

4.2 트래픽 맵 분석

트래픽 맵은 방문객의 이동 경로와 이동 빈도를 시각화한 것으로, 주요 이동 경로를 파악할 수 있었다. DeepSort 알고리즘을 통해 보다 정확한 객체 추적이 가능하여 고객의 공간 내 이동 패턴을 명확히 분석하고 효율적인 공간 배치 및 동선 관리 전략 수립에 활용할 수 있었다.



그림 2. 좌측 트래픽맵
우측 히트맵

4.3 객체 탐지 및 추적 성능

YOLOv8 기반 객체 탐지 모델을 혼합한 도심 거리 영상(해상도 1920×1080, 약 5분 분량)에 적용한 결과, 조밀하게 밀집한 사람들 대부분이 정확히 탐지되었으며, 원근에 위치한 소형 객체도 안정적으로 인식되었다. 객체 누락은 주로 완전히 가려진 대상이거나 영상 경계에서 일부만 포착된 경우에 발생하였다.

추적 성능은 Multi-Object Tracking Accuracy (MOTA) 와 ID Switch 횟수로 평가되었다. 본 연구에서 YOLOv8과 DeepSORT를 결합하여 실험한 결과, MOTA는 71.2%, ID Switch는 116회로 나타났다. 이는 동일 영상에서 Tracking-by-Detection 기법이 기록한 MOTA 50.9%, ID Switch 2,474회와 비교할 때 현저히 향상된 수치로, 추적 대상 간 빈번한 겹침이나 재진입 상황에서도 안정적인 식별이 이루어졌음을 의미한다. 특히 DeepSORT의 Re-Identification (Re-ID) 모듈이 ID 변경 오류를 효과적으로 억제하여, 대부분의 객체가 일관된 ID로 추적되었다. 이는 군중 밀집 지역의 사전 대응, 공간 재배치, 상품 진열 최적화 등 다양한 응용 가능성을 시사한다.

표 1. 알고리즘 성능 평가표

알고리즘 조합	MOTA (%)	ID Switch 수	비고
Tracking-by-Detection	50.9	2,474	기본 베이스라인
YOLOv8 + DeepSort (본 연구)	71.2	116	실시간 성능과 성능의 균형 확보
YOLOv8 + Strong Sort	79.6	101	가장 높은 성능, 단연 자원 요구 높음

V. 결론 및 시사점

AI 기반 기술의 활용 가능성을 실험적으로 입증하였다. YOLOv8과 DeepSort 알고리즘을 결합하여 오프라인 공간에서의 실시간 행동 데이터를 정밀하게 분석할 수 있었다. 해당 행동 데이터는 히트맵과 트래픽 맵을 통해 직관적으로 파악이 가능하기 때문에, 이는 마케팅 전략 수립과 공간 운영의 효율성을 높일 수 있을 것이다. 다양한 오프라인 환경에서의 추가적인 데이터 수집과 분석, 모델의 지속적인 최적화를 통해 더욱 높은 정확도와 신뢰성을 확보할 필요가 있다. 또한, 다양한 산업 분야에서의 적용 사례 연구를 통해 본 연구의 확장 가능성을 탐색하는 것이 중요하다.

ACKNOWLEDGMENT

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW 중심 대학사업의 연구 결과로 수행되었음. (2022-0-00964)

This research was supported by the Ministry of Science and ICT (MSIT) and the Institute for Information & Communications Technology Planning & Evaluation (IITP) under the National Program of Excellence in Software (2022-0-00964).

참고 문헌

- [1] Kim, S. (2025). *A deep learning-based system for people counting and crowd density estimation using smart devices* (Master's thesis, Chosun University, Graduate School, Gwangju, South Korea).
- [2] Ahn, S. (2024). *Development of an AI-based video analysis system for urban crowd density safety*. Retrieved from <https://kr.mathworks.com/>
- [3] MATLAB Korea. (2022, February 7). *Understanding Kalman Filters with MATLAB*. Retrieved from <https://kr.mathworks.com/>
- [4] Avola, D., Cinque, L., Foresti, G. L., Pannone, D., & Smeragliuolo, R. (2021). MS-Faster R-CNN: Multi-stream backbone for improved Faster R-CNN object detection and aerial tracking from UAV images. *Remote Sensing*, 13(9), 1670. <https://doi.org/10.3390/rs13091670>
- [5] Chen, I.-H., Li, C.-P., Shen, Y.-P., Lee, C.-H., & Chuang, Y.-Y. (2024). Improving point-based crowd counting and localization based on auxiliary point guidance. In *European Conference on Computer Vision* (pp. 428–444). Cham: Springer Nature Switzerland.