

# 딥러닝 기반 객체 인식을 통한 군중밀집도 및 동선 추적 알고리즘에 관한 연구

김민준, 김린회, 성경규, 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

## A Study on Crowd Density and Trajectory Tracking Algorithms Using Deep Learning-Based Object Detection

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

### 요 약

본 논문에서는 딥러닝 기반 객체 인식 및 추적 기술을 결합하여 군중 밀집도 및 이동 동선을 실시간으로 분석하는 통합 시스템을 제안한다. YOLOv8 모델을 통해 영상 내 사람 객체를 검출하고, DeepSORT 알고리즘을 통해 실시간으로 객체 추적을 수행하며, APGCC 모델을 활용해 정지 이미지 상의 군중 수를 정밀하게 추정한다. 제안된 파이프라인은 고밀도 군중 환경에서도 정확한 인원 계수와 이동 경로 파악이 가능하며, 실시간 처리 성능을 갖춘 것이 특징이다. 본 논문에서는 객체 탐지 및 추정 시스템의 구조와 핵심 기능을 기술하고 군중이 밀집된 환경에서의 실험 결과를 바탕으로 안전 관리 및 공간 운영 효율화에의 활용 가능성을 제시한다.

### I. 서 론

#### 1.1 연구 배경 및 필요성

군중 밀집도 분석은 다양한 분야에서 중요한 역할을 한다. 특히 대규모 행사나 복잡한 도심 공간에서는 실시간 인구 밀집도 파악을 통해 사고를 예방하고 안전을 확보 해야한다. 예를 들어, 군중 밀집도가 지나치게 높으면 압사 사고 등의 위험이 증가하므로 사전에 이를 감지하고 대응할 수 있어야 한다. 또한, 군중의 이동 동선 분석은 공간 설계나 운영 효율화에 활용될 수 있다. 사람들이 많이 모이는 구역이나 이동 경로를 파악하면, 시설 배치나 안내 동선을 최적화하여 오프라인 공간의 운영 효율을 높일 수 있다. 이처럼 군중 밀집도와 이동 흐름에 대한 정확한 분석은 안전 관리에 활용될 수 있다.

### II. 실험 설계

#### 2.1 데이터셋 구성

본 연구에서는 두 가지 주요 데이터셋을 사용하였다. 첫째는 ShanghaiTech Part A 데이터셋으로 군중밀집도 및 인원계수를 예측하는 모델을 학습시켰다. ShanghaiTech Part A는 주로 번화가나 공공행사에 촬영된 것으로 밀집도가 높은 경향을 띄고있다.

둘째, 객체 추적 알고리즘을 평가하기 위해 다양한 환경에서 수집한 AI Hub 영상 데이터셋을 구축하였다. 또한 객체 인식 및 추적 알고리즘 학습을 위한 사람의 바운딩 박스와 군중 수에 대한 레이블을 부여하여 학습 및 평가에 사용하였다.

표 1. 데이터셋 구분

데이터셋 구분	학습 모델	개요 (특징)
ShanghaiTech Part A	APGCC 모델	고밀도 군중 이미지 머리 위치에 점 주석
AI Hub 영상 데이터셋	YOLO v8 모델	프레임데이터 바운딩 박스 및 Id 라벨

#### 2.2 실험 내용

본 실험에서는 딥러닝 기반 객체 인식 알고리즘을 활용해 군중 밀집도를 분석하고 실시간 영상에서 사람 탐지 및 동선 추적 데이터를 적재할 수 있는 전체 파이프라인을 구축했다.

#### 2.3 실험 모델

본 실험에서는 2 가지 주요 딥러닝 모델 및 알고리즘을 설계 및 구현하였다: 객체 인식을 위한 YOLO v8, 군중 밀집도 추정을 위한 APGCC 모델, 객체 추적 및 재식별을 위한 DeepSort기법을 파이프라인에 포함하였다. 각 구성 요소의 세부 내용은 다음과 같다.

##### 2.3.1 YOLO v8 - 객체 인식

사람 등을 탐지하기 위해 최신 객체 검출 모델인 YOLO v8을 활용하였다. YOLO v8은 anchor-free 방식의 decoupled head를 도입하여 검출 정확도를 높이면서도 연산 효율을 유지하도록 설계되었다.YOLO v8을 파인튜닝하여 각 프레임의 사람 수를 1차적으로 검출하고, 검출 결과는 군중 밀집도 분석과 추적 모듈의 입력으로 활용된다.

##### 2.3.2 APGCC모델 - 군중 밀집도 분석

APGCC모델을 사용하여 영상 내 사람 수를 추정하였다. APGCC는 P2PNet과 유사한 포인트 예측 기반 구조에 Auxiliary Point Guidance(APG) 및 Implicit Feature Interpolation(IFI) 모듈이 추가된 형태이다 VGG-16모델 활용하여 ImageNet 사전학습 가중치를 초기값으로 사용하고, 이후 군중 데이터에 맞춰 미세 조정하였다. GPU 메모리한계를 고려하여 하이퍼 파라미터로 Batch size는 8로 설정했다. Epoch 수는 3500까지 학습을 진행하여 정확성을 확보하고자 하였다.

옵티마이저는 Adam을 사용하였고, 초반 500 epochs 동안은 워밍업(warm-up)으로 학습률을 점진적으로 상승시킨 후 이후에는 plateau 스케줄러를 통해 검증셋 오차 감소가 정체되면 학습률을 줄이는 전략을 취하였다.

손실 함수로는 포인트 회귀의 L1 손실과 분류의 이진 교차 엔트로피(BCE) 손실을 사용하고, 두 손실을 균등 가중합하여 최종 손실을 구성하였다. 학습 과정은 300 에폭 단위로 중간 결과(예측 출력 시각화, MAE/MSE)를 모니터링하며 진행했고, 최종 에폭에서의 모델을 평가에 사용하였다.

##### 2.3.3 DeepSort - 객체 추적

개별 사람들의 실시간 동선을 추적하기 위해 DeepSort 알고리즘을 적용하였다. DeepSort는 앞서 사용한 YOLO 검출 결과(프레임별 바운딩 박스)를 입력으로 받아, 동일한

객체(사람)가 시간 연속 프레임에서 어느 경로로 이동하는지 ID를 부여하며 추적했다. DeepSort를 통해 얻은 데이터를 활용해 특정 시간 동안 사람들이 가장 많이 지나간 경로를 파악하여 동선 히트맵으로 표현할 수 있다.

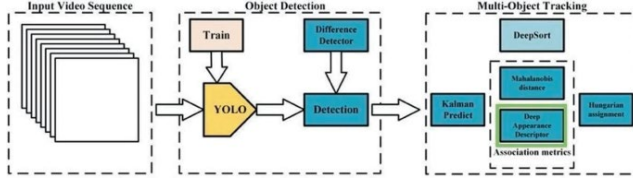


그림 1. YOLO v8 + DeepSort 구조도 설계

### III. 실험 결과 및 성능 분석

#### 3.1 군중 계수 모델(APGCC)의 성능 평가

구현된 APGCC 기반 군중 계수 모델을 ShanghaiTech Part A 데이터셋으로 평가한 결과 MAE (Mean Absolute Error)는 48.73명, MSE (Mean Squared Error)는 76.74명이 기록됐다. 또한 해당 모델 성능을 평가하기 위해 밀도 맵 기반 모델 CSRNet과 포인트 기반 모델 P2PNet모델의 성능을 비교하면 다음과 같다.

표 2.군중밀집도 모델 성능 비교

Model	MAE	MSE
APGCC	48명	76명
CSRNet	68명	115명
P2PNet	52명	85명

이는 한 이미지당 사람 수를 약  $\pm 48.7$ 명의 오차 범위로 예측함을 의미하는데, 해당 데이터셋의 인원수가 수천 명까지 분포하는 것을 감안하면 상당히 낮은 오차다.

특히 극도로 밀집된 이미지(인원 1000명 이상)에서도 APGCC는 예측값이 실제값 대비  $\pm 5\%$  이내로 근접하는 경향을 보였다. 이는 APGCC의 APG 전략 덕분에 학습 시 혼잡한 영역에서도 누락이나 중복 예측이 줄어든 결과로 해석된다.

#### 3.2 객체 탐지(YOLO v8) 및 추적(DeepSort) 성능

YOLOv8 객체 탐지 모델을 혼잡한 거리 영상에 적용한 결과는 다음과 같다. 추적 성능은 Multi-Object Tracking Accuracy (MOTA)와 ID Switch 수를 기준으로 평가하였다.

표 3. 객체 탐지 및 추적 성능 비교

Model	MOTA	ID Switch
기존 Tracking 방식	50.9%	2474
Deep Sort	71.2%	116
Strong Sort	79.6%	101

기존 방식에 비해 DeepSORT의 Re-Identification 모듈을 통해 이러한 오류가 효과적으로 감소되었다 ID Switch 수만 보면 다소 많아 보일 수 있으나, 테스트 영상이 수많은 인원이 동시에 이동하는 약 5분 길이의 복잡한 거리 영상이라는 점을 감안하면 이는 매우 우수한 수치다. 또한 누락되는 경우는 주로 매우 심한 가림이나 이미지 경계 부분에 얼굴 일부만 나오는 경우 등이었다. 참고로 최신 모델인 Strong SORT는 MOTA 79.6%, ID Switch 101회를 기록하지만, 연산량이 많아 본 연구의 실시간 환경에는 적합하지 않았다.

전반적으로 YOLOv8 + DeepSORT 모델은 각 인물의 이동 경로를 일관된 ID로 안정적으로 추적하며, 개인 단위의 이동 동선을 높은 정확도로 획득할 수 있음을 확인하였다.

### 4. 결론

본 연구에서는 딥러닝을 활용한 객체 탐지 기술을 활용하여 실시간 군중 분석 및 동선 추적 시스템을 구현하였다. YOLOv8 객체 탐지와 DeepSORT 추적을 통해 복잡한 영상 환경에서도 다수 개인의 위치와 이동 경로를 정확히 파악할 수 있었고, APGCC 기반 군중 계수 모델을 적용하여 정지 영상에서의 인원 추정 정확도를 높였다. 이 연구가 가지는 사회적 시사점은 크게 두 가지로 정리된다.

첫째, 군중 밀집도 실시간 분석을 통해 오프라인 공간의 안전성 강화에 기여할 수 있다. 예를 들어, 대규모 행사장에서 본 시스템을 도입하면, 실시간 인파 밀집 모니터링으로 위험 수준을 조기에 감지하고 관리자가 즉각 대처함으로써 사고를 예방할 수 있다.

둘째, 방문객의 동선 데이터와 밀집도 정보를 활용하여 운영 효율화를 달성할 수 있다. 상업 공간이나 박물관, 공항 등에서 사람들이 몰리는 구역과 시간대를 파악하면, 인력 배치나 시설 운영을 수요에 맞게 조정할 수 있어 자원 활용의 최적화가 가능하다.

향후 연구로 더욱 복잡한 시나리오(예: 카메라 다중 설치 환경에서의 영역 간 이동 추적등)에 대한 확장 연구, 그리고 모델 경량화 및 최적화를 통한 임베디드 시스템 실시간 처리 구현 등이 있다. 그럼에도 불구하고, 본 연구가 제시한 군중 밀집도 중심의 객체 인식 기술 프레임워크는 다양한 분야에서 안전하고 지능적인 공간 관리를 가능케 하는 초석이 될 것으로 기대한다.

### ACKNOWLEDGEMENT

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 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).

### 참 고 문 헌

- [1] WOJKE, Nicolai; BEWLEY, Alex; PAULUS, Dietrich. Simple online and realtime tracking with a deep association metric. In: 2017 IEEE international conference on image processing (ICIP). IEEE, 2017. p. 3645-3649.
- [2] DU, Yunhao, et al. Strongsort: Make deepsort great again. IEEE Transactions on Multimedia, 2023, 25: 8725-8737.
- [3] Understanding Kalman Filters with MATLAB, MATLAB Korea, 2022-02-07
- [4] AVOLA, Danilo, et al. MS-Faster R-CNN: Multi-stream backbone for improved Faster R-CNN object detection and aerial tracking from UAV images. Remote Sensing, 2021, 13.9: 1670.
- [5] CHEN, I.-Hsiang, et al. Improving point-based crowd counting and localization based on auxiliary point guidance. In: European Conference on Computer Vision. Cham: Springer Nature Switzerland, 2024. p. 428-444.
- [6] 안시현, et al. 도심 군중밀집 안전을 위한 인공지능 기반의 영상분석 시스템 개발. 한국정보기술학회논문지, 2024, 22.12: 55-64.