Journal of KIIT. Vol. 24, No. 8, pp. 57-67, Aug. 31, 2026. pISSN 1598-8619, eISSN 2093-7571
http://dx.doi.org/10.14801/jkiit.2026.24.8.57

AI 기반 부정승차 실시간 감지 시스템

김민지*[1], 윤효정*[2], 이동근*[3], 이지현*[4], 정경용**

AI-based Real-Time Fare Evasion Detection System

Minji Kim*[1], Hyojeong Yoon*[2], Donggeun Lee*[3], Jihyun Lee*[4], and Kyungyong Chung**

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학지원사업의 연구결과로 수행되었음
(2021-0-01393)

요 약

최근 대중교통내 물리적 부정승차가 증가하고 있으나, 기존 육안 감시와 카드 데이터 분석은 개찰구 뛰어넘기, 꼬리물기 등 물리적 무단 진입을 탐지하는 데 구조적 한계가 있다. 본 논문은 이를 해결하기 위해 YOLOv11 객체 탐지와 ByteTrack 추적, LineZone 라인 크로싱 감지를 결합한 실시간 부정승차 감지 시스템 GateGuard를 제안한다. 제안 시스템은 얼굴 비식별화, 실시간 알림 및 증거 영상 자동 저장 기능을 통합하여 개인정보 보호 요건과 현장 적용성을 동시에 충족하고 역무원의 오탐 피드백을 모델 재학습에 반영하는 자가 개선 구조를 통해 운용 단계에서의 점진적 성능 향상을 도모한다. AI Hub 공개 데이터셋과 자체 촬영 영상을 결합한 실험 결과, 기존 모델 대비 mAP@0.5:0.95 8.8%p, MOTA 10.4%p 향상된 성능을 달성하였다.

Abstract

Recently, physical fare evasion in public transit is rising, yet conventional surveillance and card analysis often miss unauthorized entries like gate jumping and tailgating. This paper proposes GateGuard, a real-time detection system combining YOLOv11 detection, ByteTrack tracking, and LineZone line-crossing analysis. It adds face de-identification, real-time alerts, and automatic evidence storage for privacy and practical use. A self-improving framework also feeds staff feedback on false detections into retraining. On the AI Hub dataset and self-collected videos, it improves mAP@0.5:0.95 by 8.8% and MOTA by 10.4% over the baseline.

Keywords
face de-identification, fare evasion detection, object detection, YOLOv11, ByteTrack

* 경기대학교 AI컴퓨터공학부 학사과정
- ORCID[1]: https://orcid.org/0009-0000-4437-6171
- ORCID[2]: https://orcid.org/0009-0001-8483-1711
- ORCID[3]: https://orcid.org/0009-0004-2286-806X
- ORCID[4]: https://orcid.org/0009-0002-5150-7584
** 경기대학교 AI컴퓨터공학부 교수(교신저자)
- ORCID: https://orcid.org/0000-0002-6439-9992

· Received: May 15, 2026, Revised: Jun. 21, 2026, Accepted: Jun. 24, 2026
· Corresponding Author: Kyungyong Chung
Dept. of AI Computer Science and Enginerring Kyonggi University, Korea
Tel.: +82-32-249-1382, Email: dragonhci@gmail.com

Ⅰ. 서 론

개찰구를 뛰어넘거나, 결제한 사람 뒤를 바짝 따라 들어오거나, 비상문을 슬쩍 열고 통과하는 물리적 부정승차는 교통카드 기록 자체가 남지 않아 기존 탐지 체계로는 사실상 추적이 불가능하다. 서울교통공사 공식 발표에 따르면 2020년부터 2025년 10월까지 누적 적발 건수는 31만 3,217건, 징수금은 약 142억 9,000만 원에 달하며, 적발되지 않은 건수까지 고려하면 실제 피해 규모는 이를 훨씬 상회할 것으로 추정된다[1]. 현재 대부분의 역사에서는 역무원의 육안 감시와 사후 CCTV 확인에 의존하고 있으나, 이는 24시간 실시간 대응이 구조적으로 불가능하며 인력 비용 부담도 크다.

이러한 문제를 해결하기 위한 기술적 접근이 다양하게 시도되어 왔으나 기존 연구들은 실제 운영 환경 적용에 있어 뚜렷한 한계를 드러내고 있다. 스켈레톤 기반 행동 탐지 방식은 계산 복잡도가 높고, 혼잡 환경에서의 가림 현상으로 인해 탐지 신뢰도가 급격히 저하된다. 딥 뉴럴 네트워크 기반 연구는 통제된 실험 환경에 머물러 있어 실제 성능이 검증되지 않았으며, 프레임 간 동일 인물의 연속 추적을 위해 별도 Re-ID 모델을 요구하는 한계를 지닌다[2]. 나아가 기존 영상 기반 이상행동 탐지 시스템은 개인 식별 정보를 수집하여 법적·윤리적 제약으로 실제 배포가 제한된다는 문제도 제기된다[3].

이에 본 논문은 실시간 지하철 부정승차 감지 시스템의 설계 및 구현 방안을 제안한다. 제안 시스템은 YOLOv11 기반의 실시간 객체 탐지와 다중 객체 추적 ByteTrack(MOT, Multi-Object Tracking) 파이프라인을 결합하여 승객의 동선을 추적한다. Supervision 라이브러리의 LineZone을 활용하여 개찰구 라인 크로싱을 정밀하게 감지하며, OpenCV Gaussian Blur 기반 얼굴 비식별화 처리를 통해 개인정보 보호 법적 요건을 충족한다. FastAPI와 WebSocket을 통해 관리자 대시보드에 실시간 알림을 전송하며, AWS S3와 Celery 비동기 처리를 활용해 증거 영상을 자동으로 저장하는 통합 환경을 제공한다. 이러한 시스템의 전체 구조를 그림 1에 나타내었다.

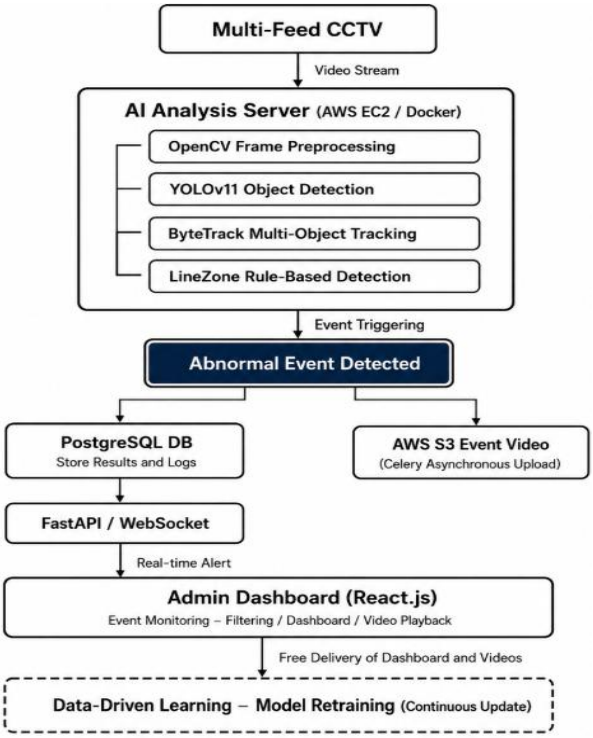

그림 1. 전체 시스템 구조
Fig. 1. Overall system architecture

본 연구의 핵심 기여는 세 가지이다. 첫째, 부정승차 탐지 도메인에서 얼굴 비식별화를 파이프라인 전단 입력 전처리 단계에 통합한 Privacy-by-Design 구현 사례를 최초로 제시하며, 비식별화 적용 시 탐지 성능 저하가 mAP@0.5 기준 −0.2%p에 불과함을 정량적으로 확인하였다. 둘째, 별도 Re-ID 모델 없이 YOLOv11과 ByteTrack 단일 추적기만으로 혼잡 환경에서 실시간 다중 객체 추적을 실현하여, 기존 선행연구의 연산 부담 및 가림 현상 취약성을 동시에 해소하였다. 셋째, 역무원 오탐 신고 데이터가 자동으로 재학습 파이프라인에 투입되는 Human-in-the-Loop 자가개선 폐루프를 End-to-End 운영 아키텍처에 내재화하였다.

Ⅱ. 관련 연구

2.1 스켈레톤 및 자세 분석 기반 부정승차 탐지

비전 기반 행동 인식 기술을 부정승차 탐지에 직접 적용하려는 시도는 꾸준히 이어져 왔다. Huang et al.[4]은 지하철 환경에서 승객의 골격 좌표를 시

간축으로 추적하여 부정승차 행동 유형을 분류하는 기법을 제안하였다. 스켈레톤 시퀀스와 시계열 분석을 결합함으로써 행동의 구조적 패턴을 포착한다는 점에서 의의가 있으나, 다단계 파이프라인은 연산 부담이 크고 혼잡한 실제 환경에서 가림 현상이 발생하면 탐지 신뢰도가 급격히 저하된다는 한계가 있다.

Nicodème[5]은 자동 개찰기 주변에서 Mask-RCNN으로 게이트 영역을 분할한 뒤 자세 분석을 수행하는 프레임워크를 제안하였으나, 단일 게이트 형태와 정적 배경에 최적화된 실험 설계로 인해 다양한 역사 환경으로의 일반화에 제약이 따른다. Vyver[2]는 LSTM 키포인트 추출과 모델을 결합하여 실시간 영상에서 요금 납부 여부를 예측하는 방식을 제안하였으나, 동일 인물을 프레임 간 지속 추적하기 위한 Re-ID 모델이 별도로 요구되는 구조적 한계를 안고 있다. 반면, 본 연구는 골격 추출 없이 YOLOv11 기반 바운딩 박스 탐지와 ByteTrack 추적만으로 혼잡 환경에서의 실시간 감지를 구현하여, 가림 현상에 따른 신뢰도 저하 문제와 높은 연산 부담을 동시에 해소한다.

2.2 교통카드 데이터 기반 부정승차 패턴 감지

교통카드 거래 내역과 머신러닝을 결합하여 요금 사기를 탐지하는 연구도 축적되어 있다. Claiborne와 Gupta[6]는 SVM, 랜덤 포레스트, 인공신경망으로 요금 매체 사기를 이진 분류하는 모델들을 비교하여, 랜덤 포레스트와 신경망이 SVM 대비 우월한 성능을 보임을 확인하였다. Delfau et al.[7]은 머신러닝 기반 단속 요원 배치 최적화를 통해 부정승차가 집중되는 역사와 시간대를 특정하는 접근을 제안하였다. 그러나 이러한 데이터 기반 방식은 카드 거래 기록이 존재하는 경우에만 분석이 가능하다는 근본적 한계를 지닌다. 개찰구를 뛰어넘거나 결제자 뒤를 바짝 따라 진입하는 테일게이팅, 비상문을 이용한 무단 통과와 같은 물리적 부정승차는 카드 태그 자체가 이루어지지 않아 데이터가 생성되지 않으며, 결과적으로 탐지 사각지대가 구조적으로 발생한다.

2.3 개인정보 보호 요건을 고려한 영상 이상 탐지 연구 동향

영상 기반 이상 행동 탐지(VAD, Video Anomaly Detection) 분야는 딥러닝의 발전과 함께 급속히 성장하였으나, 실제 배포에는 법적·윤리적 장벽이 존재한다. Liu et al.[8]은 기존 VAD 시스템이 개인 식별 정보를 포함한 원본 영상을 그대로 수집·처리함으로써 개인정보 침해 우려를 야기하며, 이것이 실세계 적용을 제한하는 핵심 요인임을 지적하였다. 국내에서는 개인정보 보호법 및 개인정보보호위원회 가이드라인에 따라 공공장소 CCTV 영상의 수집·활용 시 얼굴 등 식별 가능 정보의 비식별 처리 또는 열람 제한 조치가 요구된다. 이에 골격 키포인트만을 입력으로 활용하거나 얼굴 등 식별 영역을 마스킹하는 프라이버시 보존형 접근이 주목받고 있으나, 부정승차 탐지 도메인에서 이를 실제 시스템 설계에 통합한 사례는 아직 보고되지 않고 있다. 본 연구는 OpenCV Gaussian Blur 기반 얼굴 비식별화 모듈을 탐지 파이프라인 내 입력 전처리 단계에 통합함으로써, 개인정보보호법 제25조 요건을 충족하면서도 탐지 성능을 유지하는 운영 가능한 시스템 구조를 제시한다[9]. 또한 본 연구는 비식별화 모듈을 독립 컴포넌트로 분리·설계하여, 향후 규제 변화에 따라 k-익명성, 차분 프라이버시, 연합 학습(Federated learning) 기반 방식 등 대안적 기법으로 교체 가능한 확장 구조를 채택하였다. 이를 통해 본 시스템은 현행 법적 요건을 충족하면서도 미래 규제 환경 변화에 유연하게 대응할 수 있는 프라이버시 보존형 운영 모델을 제공한다.

III. 실시간 부정승차 감지 시스템 설계 및 구현

3.1 시스템 개요 및 설계 목표

본 시스템의 개발 목표는 세 가지로 나뉜다. 첫째, AI 기반 자동 감지 및 기록을 통해 부정승차 행위가 빠짐없이 적발된다는 인식을 형성하여 부정승차를 사전에 억제한다. 둘째, AI가 사전 필터링한

이상 이벤트만을 역무원에게 전달함으로써 수십 개의 CCTV를 직접 모니터링하는 부담을 줄이고 업무 효율을 극대화한다. 셋째, 시스템 도입을 통해 확보한 미징수 요금과 절감된 인건비를 시민 서비스 개선에 재투자하는 선순환 구조를 구축한다.

본 시스템은 CCTV 영상 수집, AI 분석 서버, 이벤트 관리, 관리자 인터페이스의 네 단계로 구성된 처리 파이프라인을 제공한다. 다채널 CCTV에서 수집된 영상은 AWS EC2 기반 AI 분석 서버에서 YOLOv11 객체 탐지, ByteTrack 추적, Supervision LineZone 라인 크로싱 감지 순으로 처리된다. 판정된 이벤트는 PostgreSQL 데이터베이스에 기록되는 동시에 FastAPI·WebSocket을 통해 관리자 대시보드에 실시간 전달되며, 증거 영상 클립은 Celery 비동기 처리를 통해 AWS S3에 자동 저장된다. 역무원은 이벤트를 확인하고 처리완료 또는 오탐신고로 피드백을 전송하며, 오탐 데이터를 모델 재학습에 활용하는 Human-in-the-Loop 기반 자가 개선 구조를 설계에 반영하였으며, 실제 역사 환경 배포 이후 오탐률 감소 효과를 정량적으로 검증할 계획이다. 그림 2는 이러한 시스템 전체 처리 흐름을 나타낸 구성도이다.

3.2 AI 기반 부정승차 탐지 기법

본 연구에서는 부정승차 탐지 및 실시간 대응을 위해 YOLOv11 기반의 객체 탐지 및 추적 시스템을 설계하였다[10]. CCTV로부터 수신된 실시간 영상 스트림은 모델 입력 규격에 맞추어 정규화 전처리를 수행한다. 이와 동시에 OpenCV 기반의 얼굴 검출 및 Gaussian Blur 처리가 수행되어 개인 식별 정보가 제거된다. 이를 통해 탐지에 필요한 신체 윤곽 및 이동 궤적 정보는 보존하면서도 개인정보 보호 법적 요건을 충족한다. 비식별화 적용 전후 탐지 성능을 동일한 테스트셋 조건에서 비교한 결과, mAP@0.5 기준 성능 저하는 −0.2%에 불과하였다. 이는 얼굴 비식별화가 신체 윤곽 및 이동 궤적 정보를 보존하기 때문으로, 프라이버시 보호와 탐지 성능이 실질적으로 양립 가능함을 실험적으로 입증한 결과이다.

객체 탐지 단계에서는 YOLOv11 모델로 프레임 내 사람 객체를 검출하고 바운딩 박스 좌표와 신뢰도 점수를 산출한다. 검출된 객체에는 ByteTrack 알고리즘을 적용하여 고유 ID를 부여하고 추적한다[11]. ByteTrack은 높은 신뢰도와 낮은 신뢰도의 탐지 결과를 모두 연관 과정에 활용하여 혼잡 환경에서의 ID 소실을 억제한다.

탐지 및 추적이 완료된 객체는 Supervision 라이브러리의 라인크로싱 및 구역 통과 모듈을 통해 분석된다[12]. 개찰구 영역에 가상 라인을 정의하고 탐지된 객체가 해당 라인을 통과하는 시점에 부여된 행동 클래스(Jump, Tailgating, Crawling)가 부정승차 유형에 해당하면 의심 이벤트로 판정한다.

역무원이 오탐(False positive)으로 판정한 이벤트는 오탐 사유와 함께 학습용 데이터셋으로 분류되어 기존 데이터에 병합되고, 일정 주기마다 미세 조정 방식의 점진적 학습을 수행한다. 이러한 Human-in-the-Loop 방식을 설계에 반영하였으며, 실제 역사 환경 배포 이후 오탐률 감소 효과를 정량적으로 검증할 계획이다.

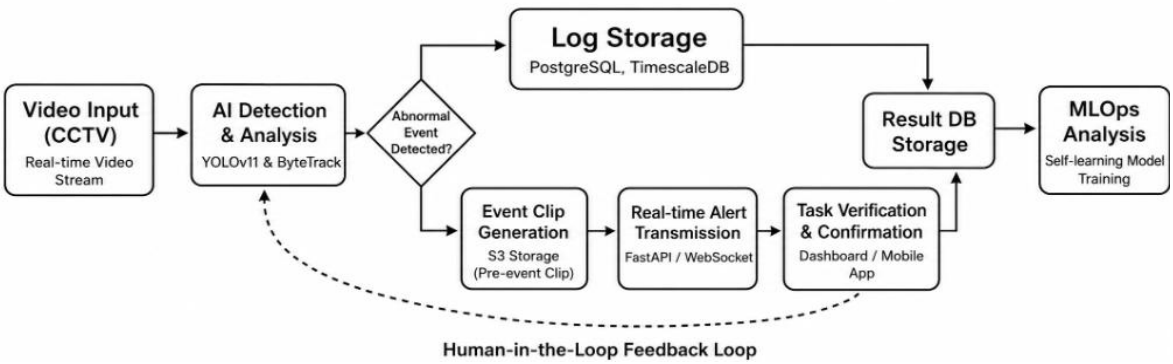

그림 2. 시스템 전체 처리 흐름
Fig. 2. Overall system processing flow

3.3 관리자 대시보드 및 이벤트 처리

본 시스템의 관리자 인터페이스는 React.js 기반의 웹 대시보드로 구현되며 역무원이 감지된 부정승차 이벤트를 실시간으로 수신하고 확인 및 처리할 수 있도록 설계하였다. 그림 3은 대시보드 메인 화면을 나타낸다.

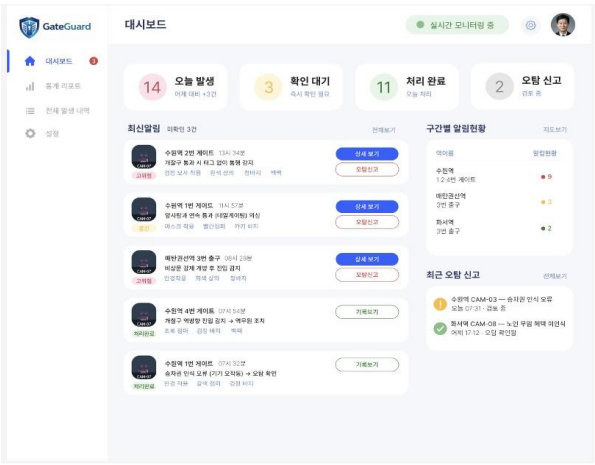

그림 3. 대시보드 메인 화면
Fig. 3. Dashboard main screen

그림 3의 대시보드 상단에는 오늘 발생, 확인 대기, 처리 완료, 오탐 신고 건수 현황이 표시된다. 우측 구간별 알림 현황 패널에서는 역사별 알림 건수를 한눈에 볼 수 있다. 대시보드 화면에는 최신 이벤트 목록이 리스트 형식으로 표시되며, 새로운 부정승차 의심 이벤트를 생성할 때마다 WebSocket 프로토콜을 통해 해당 이벤트가 즉시 목록에 추가된다. 역무원이 목록에서 특정 이벤트를 선택하면 상세보기 화면으로 전환된다. 그림 4는 이벤트 상세 정보 화면을 나타낸다.

그림 4와 같이 상세 정보 화면 좌측에는 실시간 알림 목록이, 우측에는 선택한 이벤트의 감지 영상 클립과 함께 기록 시각, 발생 위치, AI가 분석한 인상착의 정보가 표시된다. 역무원은 영상과 분석 결과를 검토한 후 해당 이벤트에 대해 현장 대응하거나 오탐신고 혹은 처리완료로 상태를 변경할 수 있다. 처리완료는 이벤트가 실제 부정승차에 해당함을 확인하고 대응을 마친 경우에 선택하고, 오탐신고는 AI가 정상 통행을 부정승차로 잘못 판정한 경우에 선택해 오탐 사유를 함께 입력하여 전송한다.

그림 4. 이벤트 상세 정보 화면
Fig. 4. Event details screen

Ⅳ. 결과 및 성능 평가

4.1 시스템 구현 결과

본 시스템은 YOLOv11 기반 실시간 객체 탐지, ByteTrack 다중 객체 추적, Supervision LineZone 라인 크로싱 감지, OpenCV Gaussian Blur 기반 얼굴 비식별화 파이프라인을 구현하여 FastAPI 서버와 통합하였다. 감지된 이벤트는 WebSocket을 통해 실시간으로 관리자 대시보드에 전송되며, 증거 영상 클립은 Celery 비동기 처리를 통해 AWS S3에 자동 저장된다. 본 시스템은 연산 특성이 다른 두 계층(2-tier)으로 구성된다. YOLOv11·ByteTrack 추론 계층은 GPU 가속을 전제로 하며(표2의 FPS는 본 계층 기준), FastAPI·WebSocket·Celery·PostgreSQL·S3 업로드 등 경량 CPU 워크로드를 담당하는 애플리케이션 계층은 Docker 서비스 통합 및 HTTPS SSL 적용 등 실제 운영 환경에 준하는 배포 구성을 검증하였다 [13]. 다중 채널 실시간 추론은 AWS g4dn 인스턴스(NVIDIA T4 GPU)에서 수행한다.

4.2 성능 평가 및 기대 효과

4.2.1 실험 데이터셋 구성 및 학습 환경 설정

본 연구에서는 개인정보보호법에 따른 실제 역사 내 무단 영상 수집의 제약을 극복하기 위해, AI Hub에서 제공하는 '지하철 이상행동 영상' 공개 데

이터셋과 자체 수집 영상을 결합하여 평가 데이터셋을 구축하였다[14]. AI Hub 데이터셋은 Crawling, Jump, Normal, Tailgating 클래스를 포함하나, Crawling과 Tailgating 클래스는 샘플 수가 상대적으로 부족하여 클래스 불균형 문제가 존재한다. 이를 보완하기 위해 대학 도서관 출입 게이트를 테스트베드로 활용하여 해당 클래스의 영상을 자체 촬영하여 추가하였다. 다만 자체 촬영 데이터는 통제된 실내 환경에서 수집되어 실제 혼잡 시간대의 다중 인원 밀집·가림 상황을 충분히 반영하지 못한다는 한계가 있다. 향후 운영 기관과의 협력을 통해 혼잡 시간대 실측 데이터를 추가 확보할 계획이다. 표 1에서는 영상 데이터셋의 구성 현황을 나타낸다.

표 1. 영상 데이터셋 구성 현황
Table 1. Composition of the video dataset

Source	Recording environment	Usage
AI hub subway anomalous behavior dataset [14]	Real subway stations, various lighting conditions and crowd densities	Training / validation
Self-collected dataset (University library gate)	Indoor fixed camera, controlled lighting conditions	Crawling·tailgating class balancing
Integrated dataset (with augmentation)	Mixed environments	Training / validation / test

표 1과 같이 평가 데이터셋은 AI Hub 공개 데이터셋, 자체 촬영 데이터셋, 두 소스를 결합한 통합 데이터셋의 세 가지로 구성된다. 그림 5는 데이터셋 이미지를 부정승차 유형별로 나타낸다.

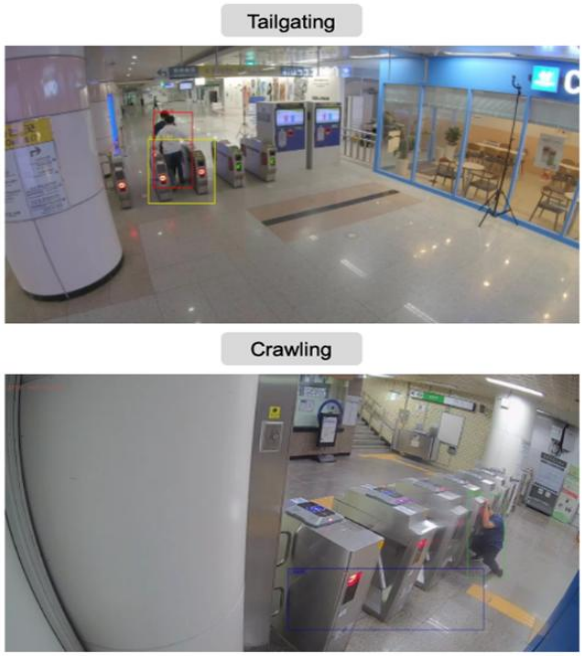

그림 5. 부정승차 유형별 데이터셋 이미지
Fig. 5. Images of fare evasion types in dataset

그림 5에서 Jump는 승객이 개찰구를 뛰어넘는 행위, Tailgating은 정상 결제자 뒤를 바짝 따라 진입하는 행위, Crawling은 개찰구 하단을 기어서 통과하는 행위로, 각 유형은 외형적 특징이 뚜렷하게 구분된다. 그림 6은 자체 촬영 데이터셋과 AI Hub 지하철 데이터셋의 촬영 환경 차이를 나타낸다.

그림 6. 도서관 촬영 영상과 지하철 영상 비교
Fig. 6. Comparison of library and subway videos

지하철 영상은 실제 역사의 다양한 조도 및 배경 복잡도를 포함하는 반면, 도서관 영상은 고정 카메라와 통제된 조명 환경에서 촬영되어 배경이 단순하고 혼잡도가 낮다. 이러한 도메인 차이로 인해 자체 촬영 데이터만으로는 실제 지하철 환경을 완전히 재현하기 어렵다는 한계를 지닌다. 앞서 기술한 자체 촬영 데이터의 보강을 통해 클래스 불균형을 완화한 최종 데이터셋의 클래스별 프레임 분포는 그림 7에 나타내었다.

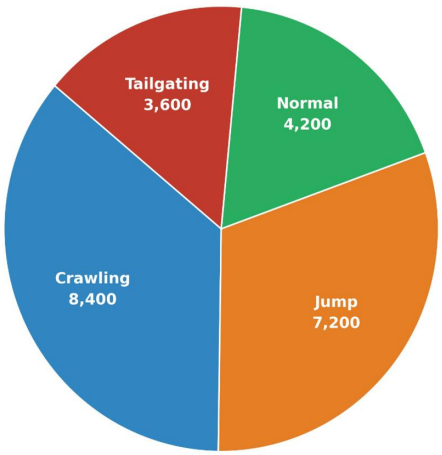

그림 7. 데이터셋 클래스 프레임 분포
Fig. 7. Frame distribution by class in the integrated dataset

각 클립에서 10프레임 간격으로 프레임을 추출하고 데이터 증강을 적용하여, 최종 데이터셋은 총 23,400프레임(Normal 4,200프레임, Jump 7,200프레임, Crawling 8,400프레임, Tailgating 3,600프레임)으로 구성된다. 학습·검증·테스트 데이터는 7:1.5:1.5 비율로 분할하였다. 모델 학습에는 input resolution 640×640, batch size16, epoch 100, learning rate 0.01, confidence threshold 0.5의 하이퍼파라미터를 공통 적용하였다. 데이터 증강으로 좌우 반전, 밝기 조정(±30%), 가우시안 노이즈를 적용하여 조도 변화에 대한 강건성을 확보하였다. 실험은 NVIDIA RTX 3060(12GB) GPU 환경에서 진행되었다.

4.2.2 파이프라인 성능 비교

본 연구는 동일 데이터셋(23,400프레임)·동일 하이퍼파라미터·동일 분할 비율(7:1.5:1.5) 조건 하에 세 가지 파이프라인을 비교함으로써, 통제된 조건에서의 공정한 성능 비교로서 학술적 타당성을 확보하였다. 객체 탐지 성능은 평균 정밀도(mAP)로 평가하였으며, 다중 객체 추적 성능은 혼잡 환경에서의 추적 안정성을 대변하는 MOTA와 IDF1(ID F1 Score) 지표를 적용하였다. 특히 MOTA는 식 (1)과 같이 오탐지(FP), 미탐지(FN), ID 전환(IDSW)을 모두 고려하여 종합적인 신뢰성을 평가한다.

$$MOTA = 1 - \frac{\sum_t FN_t + FP_t + IDSW_t}{\sum_t GT_t} \quad (1)$$

여기서 GT는 Ground Truth 객체 수, FP는 오탐지 건수, FN은 미탐지 건수, IDSW는 ID 전환 횟수를 의미하며, 값이 높을수록 추적 정확도가 우수함을 나타낸다. 표 2는 동일한 데이터셋(23,400프레임, 7:1.5:1.5 분할)과 동일한 하이퍼파라미터 조건에서 YOLOv8m+DeepSORT, YOLOv10m+BoT-SORT, YOLOv11m+ByteTrack(제안 파이프라인)을 비교한 결과이다.

표 2. 데이터셋 기반 AI 파이프라인 성능 비교 평가 결과
Table 2. Performance comparison of AI pipelines on dataset

Pipeline	mAP@0.5	mAP@0.5:0.95	MOTA	IDF1	FPS
YOLOv8m + DeepSORT	88.2	62.4	72.1	71.3	24.3
YOLOv10m + BoT-SORT	91.5	66.8	76.4	75.8	28.7
YOLOv11m + ByteTrack (Proposed pipeline)	94.3	71.2	82.5	80.1	34.8

AI Hub 지하철 이상행동 데이터셋을 동일한 조건으로 활용한 외부 선행 연구가 현재까지 공개되어 있지 않아 외부 연구와의 직접 수치 비교는 본 연구 범위에서 제외하였다. 표 3은 동일 목적의 선행연구들과 방법론적 특성을 비교한 것이다.

표 3. 선행연구와의 방법론 비교
Table 3. Methodological comparison with prior studies

Study	Method	Dataset	Limitations
Huang et al. [4]	Skeleton + time-series	Self-collected	High computational cost; reliability drops under occlusion
Nicodème [5]	Mask-RCNN + posture analysis	Single station	Limited to single gate; poor generalizability
van der Vyver [2]	LSTM + keypoint	Self-collected	Requires separate Re-ID model for inter-frame tracking
Proposed	YOLOv11 + ByteTrack	AI Hub + self-collected	Re-ID-free; de-identification integrated; multi-station applicable

표 3에서 알 수 있듯이 Huang et al.[4]의 스켈레톤 기반 접근은 다단계 파이프라인의 연산 부담으로 실시간 처리에 제약이 있으며, Nicodème[5]은 단일 게이트 환경에 최적화되어 다양한 역사 환경으로의 일반화에 한계가 있다. van der Vyver[2]는 프레임 간 동일 인물 추적을 위해 별도 Re-ID 모델을 요구하는 구조적 제약이 있다. 본 연구는 이러한 한계를 Re-ID 없는 ByteTrack 단일 추적기와 얼굴 비식별화 통합 설계로 해소하여 mAP@0.5 94.3%, MOTA 82.5%를 달성하였으며, 부정승차 탐지 도메인에서 비식별화를 파이프라인 전단에 통합한 유일한 구현 사례라는 점에서 차별성을 갖는다.

제안하는 파이프라인은 이전 세대 모델인 YOLOv8m+DeepSORT 대비 mAP@0.5:0.95 기준 8.8%, MOTA 기준 10.4% 향상된 정확도를 기록하였다. 특히 혼잡한 개찰구 통과 시 발생하는 겹침 상황에서도 ByteTrack의 유연한 박스 연관 기법이 작용하여 IDF1 수치가 크게 상승(80.1%)하였다. FPS는 NVIDIA RTX 3060 GPU 기준 제안 모델 34.8 FPS로, ByteTrack이 Re-ID 연산을 생략해 DeepSORT 대비 높은 처리 속도를 달성하였다. 다만 이 FPS는 GPU 추론 계층의 성능 지표이며, 본 시스템은 연산 특성이 다른 두 계층으로 분리 운용된다. YOLOv11·ByteTrack 추론은 GPU 가속을 전제로 하고(표2의 FPS는 본 계층 기준), AWS EC2 t3.micro는 FastAPI·WebSocket·Celery·PostgreSQL·S3 업로드 등 경량 애플리케이션·오케스트레이션 계층만을 담당한다. GPU가 없는 t3.micro(2 vCPU)는 YOLOv11 실시간 추론에 부적합하므로, 실배포 시 다중 채널 실시간 추론은 AWS g4dn 인스턴스(NVIDIA T4 GPU) 또는 GPU 기반 엣지 디바이스에서 수행한다. 즉 34.8 FPS는 GPU 추론 계층 기준 실시간 처리 성능이며, t3.micro는 추론을 담당하지 않는 배포·서비스 계층임을 명확히 한다. 탐지 오류 유형별 세부 분석을 위해 파이프라인별 FP, FN, IDSW 절대 수치를 표 4에 제시한다.

표 4. 파이프라인별 FP/FN/IDSW 세부 오류 지표
Table 4. Detailed error metrics FP/FN/IDSW by pipeline

Pipeline	FP	FN	IDSW	MOTA(%)	FNR(%)
YOLOv8m + DeepSORT	312	187	94	72.1	9.2
YOLOv10m + BoT-SORT	271	154	61	76.4	7.6
YOLOv11m + ByteTrack (Proposed pipeline)	198	97	38	82.5	4.8

표 4의 세부 오류 지표를 기반으로 MOTA 향상폭(10.4%)의 기여 요인을 분석하면, YOLOv8m+DeepSORT 대비 FP는 312건에서 198건으로 114건, FN은 187건에서 97건으로 90건, IDSW는 94건에서 38건으로 56건 감소하였으며, 전체 오류 감소량(260건) 대비 기여 비율은 FP 감소가 약 43.8%로 가장 크고, FN 감소가 약 34.6%, IDSW 감소가 약 21.5%를 차지한다. 이로부터 오탐지(FP) 감소가 MOTA 향상의 핵심 기여 요인임을 확인하였다. 이는 ByteTrack의 유연한 박스 연관 기법이 혼잡 환경에서의 잘못된 탐지 억제에 효과적임을 시사한다. 수치를 FP·FN 기반으로 정규화하기 위해 테스트셋 내 Ground Truth 어노테이션 총 수를 집계한 결과 2,029건(Jump 381건, Tailgating 897건, Crawling 438건, Normal 313건)으로 확인되었다. 여기서 미탐률(FNR)은 시스템 수준의 미탐을 정규화하기 위해 전 클래스 어노테이션을 분모로 하며,

FNR = ΣFN ÷ Σ(전체 GT 어노테이션) = 97 ÷ 2,029 = 4.8%로 산출된다. 이는 베이스라인(YOLOv8m+DeepSORT)의 9.2% 대비 4.4% 개선된 값이다. MOTA의 ΣGT는 표준 MOT 평가 방식에 따라 프레임 단위 Ground Truth 총합으로 산출되며, 이는 FNR 정규화에 사용한 어노테이션 건수(2,029)와 모집단이 다르다. 한편 표 5의 클래스별 FNR은 각 클래스의 GT를 분모로 하며, 정상(Normal) 클래스는 부정승차 이벤트가 아니므로 클래스별 FNR을 정의하지 않고 "−"로 표기하되 전체 미탐률 산출에는 포함된다.

표 5. 유형별 FP·FN 분석
Table 5. Class-level FP/FN analysis of proposed pipeline

Class	FP	FN	FNR(%)
Jump	21	8	2.1
Tailgating	98	52	5.8
Crawling	18	14	3.2
Normal	61	23	-
Total	198	97	4.8

표 5와 같이 Tailgating 클래스에서 FP 98건, FN 52건으로 가장 높은 오류율이 관찰되었다. 이는 정상 결제자 뒤를 따라 진입하는 동선 패턴이 일반 통행과 외형적으로 유사하여 판별 경계가 모호하기 때문으로 분석된다. Normal 클래스에서도 FP 61건이 관찰되었는데, 이는 정상 통행자의 빠른 보행 또는 복수 인원 동시 통과 시 부정승차 유형과 외형이 유사해지는 경우에 기인하는 것으로 분석된다. 반면 Jump와 Crawling은 신체 동작의 외형적 특징이 뚜렷하여 상대적으로 낮은 오류율을 보였다.

실제 도입 효과를 간접 추정하면, 서울교통공사 2024년 공식 통계(연간 적발 60,719건) 기반 물리적 부정승차 추산 규모(약 16,597건/년, 현행 탐지율 30% 가정 기준)에 제안 파이프라인의 FNR 개선분 4.4%p를 적용할 경우 연간 약 730건의 추가 탐지 가능성이 산출된다. 단, 이는 공개·자체 데이터셋 기반 AI 파이프라인 간 성능 개선분을 공식 통계에 적용한 보수적 추정치이며, 실제 역사 운영 로그 기반 실측값이 아님을 명시한다. 실제 도입 효과는 향후 파일럿 역사에서 도입 전후 적발 건수를 직접 비교하여 실증할 계획이다.

Ⅴ. 결론 및 향후 과제

본 논문의 핵심 기여는 다음 세 가지이다. 첫째, 부정승차 탐지 도메인에서 얼굴 비식별화를 파이프라인 전단 입력 전처리 단계에 통합한 Privacy-by-Design 구현 사례를 최초로 제시하였으며, 비식별화 적용 시 mAP@0.5 기준 성능 저하가 −0.2%p에 불과함을 정량적으로 확인하였다. 이는 Huang et al.[4], Nicodème[5], van der Vyver[2] 어느 선행연구에서도 구현된 바 없는 차별적 기여로서, 공개 데이터셋 기반 실험에서 베이스라인(YOLOv8m+DeepSORT) 대비 mAP@0.5:0.95 8.8%, MOTA 10.4% 향상을 달성하였다. 둘째, 별도 Re-ID 모델 없이 YOLOv11과 ByteTrack 단일 추적기만으로 혼잡 환경에서 실시간 다중 객체 추적을 실현하여, 기존 선행연구의 연산 부담 및 가림 현상 취약성을 동시에 해소하였다. 셋째, 역무원 오탐 신고 데이터가 자동으로 재학습 파이프라인에 투입되는 Human-in-the-Loop 자가개선 폐루프를 FastAPI·WebSocket 기반 관리자 대시보드, AWS S3 증거 영상 저장, PostgreSQL 이벤트 로그와 통합하는 End-to-End 운영 아키텍처에 내재화하여, 실제 역사 배포 이후 운영 데이터 축적에 따른 점진적 성능 향상이 가능한 구조를 제시하였다.

본 시스템은 물리적 수단으로 개찰구를 무단 통과하는 부정승차 문제를 해결하기 위해 실시간 감지 시스템의 설계 및 구현 방안을 제안하였다. 제안 시스템은 YOLOv11 기반 객체 탐지와 ByteTrack 다중 객체 추적을 결합하고, Supervision LineZone으로 개찰구 통과를 정밀 판별한다. OpenCV Gaussian Blur 기반 얼굴 비식별화로 개인정보 보호 법적 요건을 충족하였으며, FastAPI·WebSocket 기반 실시간 알림과 AWS S3 증거 영상 자동 저장을 통합하여 현장 적용성을 높였다.

본 연구의 한계는 다음과 같다. 첫째, 규제로 인한 실측 데이터 수집 제약으로 AI Hub 공개 데이터셋과 자체 수집 영상을 결합한 총 23,400 프레임 규모의 데이터셋에 의존하였으며, 실제 역사의 다양한 조도·혼잡도 조건을 완전히 재현하지 못하였다. 둘째, 동일 가중치·입력 해상도로 t3.micro에서 단일 스트림 측정 시 3.2 FPS에 그쳐 다수의 CCTV 채널

을 실시간 처리하기 어려우며, 다중 채널 실시간 추론에는 AWS g4dn 인스턴스(NVIDIA T4 GPU)가 요구된다. 셋째, 테일게이팅과 같이 정상 통과와 동선 패턴이 유사한 유형은 오탐률이 높아 강건성 확보가 요구된다. 표 4에 제시된 바와 같이 제안 파이프라인의 FP는 198건으로 FN(97건) 대비 약 2배 수준이며, 이는 오탐이 미탐보다 상대적으로 높은 비중을 차지함을 나타낸다. 표 5에 제시된 유형별 분석은 현재 데이터셋 기반의 결과이며, 실제 역사 배포 이후 실측 데이터를 활용한 추가 검증을 수행할 계획이다. 넷째, 실제 지하철 역사 환경에서의 시스템 테스트가 이루어지지 못하였다. 개인정보보호법 제25조에 따른 영상 수집 제한, 운영 기관의 협조 확보 어려움, 실제 개찰구 장비와의 연동 인터페이스 부재 등의 제약으로 인해 실제 역사 환경에서의 성능 측정 및 현장 실증이 불가능하였다. 이로 인해 실제 운영 조건(다양한 카메라 앵글, 역사별 조명 편차, 출퇴근 혼잡 군중 등)에서의 탐지 성능이 본 실험 결과와 상이할 수 있다. 이러한 제약의 해소를 위해 향후 세 가지 방안을 추진할 계획이다. 첫째, 정부 규제 샌드박스 제도를 활용하여 역사 내 실측 영상 수집의 법적 근거를 확보한다. 둘째, 서울교통공사 등 운영 기관과의 MOU 체결을 통해 협조 체계를 구축한다. 셋째, 실제 개찰구 AFC 시스템과 연동을 위한 표준 인터페이스를 사전 설계하여 현장 배포 준비도를 높인다.

향후 연구에서는 운영 기관과의 협력 및 규제 샌드박스를 활용하여 현장 실측 데이터셋을 구축하고 모델 파인튜닝으로 탐지 정확도를 향상시킬 예정이다. 또한 비식별화 모듈의 독립 컴포넌트 구조를 활용하여, 향후 규제 변화에 따라 k-익명성, 차분 프라이버시, 연합 학습(Federated learning) 기반 방식 등 대안적 기법으로 교체하는 확장 연구를 수행할 계획이다. 더욱이 다중 채널 실시간 처리에는 GPU 추론 계층이 요구되므로, GPU 기반 엣지 컴퓨팅 및 AWS g4dn 인스턴스(NVIDIA T4 GPU) 도입을 통해 다중 카메라 병렬 실시간 처리를 실증하고, Human-in-the-Loop 재학습 파이프라인을 실제 운영 환경에 적용하여 정량적으로 검증할 계획이다. 구체적으로는 1~2개 파일럿 역사에서 시스템 도입 전후 부정승차 적발 건수, 역무원 실제 대응 시간(알림 수신~처리 완료), 오탐 신고 건수 변화를 정량 비교하여 시스템의 현장 적용성을 실증할 예정이다.

References

[1] Seoul Economic Daily, "Surge in fare evasion among young subway riders", Dec. 2025. https://www.sedaily.com/NewsView/2H1L461L4M. [accessed: Feb. 10, 2026].

[2] J. van der Vyver, "A Deep Neural Network Approach to Fare Evasion", arXiv preprint, arXiv:2405.17855, May 2024. https://doi.org/10.48550/arXiv.2405.17855.

[3] Y. Liu, et al., "Privacy-Preserving Video Anomaly Detection: A Survey", arXiv preprint, arXiv:2411.14565, Nov. 2024. https://doi.org/10.48550/arXiv.2411.14565.

[4] S. Huang, X. Liu, W. Chen, G. Song, Z. Zhang, L. Yang, and B. Zhang, "A detection method of individual fare evasion behaviours on metros based on skeleton sequence and time series", Information Sciences, Vol. 589, pp. 62-79, Apr. 2022. https://doi.org/10.1016/j.ins.2021.12.088.

[5] C. Nicodème, "Fare-evasion detection at ticket gates using posture analysis", Proc. of IEEE 25th Int. Conf. on Intelligent Transportation Systems (ITSC), Macau, China, pp. 1128-1132, Oct. 2022. https://doi.org/10.1109/ITSC55140.2022.9922039.

[6] J. Claiborne and A. Gupta, "Machine Learning Classifiers for Predicting Transit Fraud", Proc. of 24th Americas Conf. on Information Systems (AMCIS), New Orleans, LA, USA, Paper 37, Aug. 2018. https://aisel.aisnet.org/amcis2018/DataScience/Presentations/37. [accessed: Feb. 10, 2026]

[7] J. B. Delfau, D. Pertsekos, and M. Chouiten, "Optimization of Control Agents Shifts in Public Transportation: Tackling Fare Evasion with Machine-Learning", Proc. of IEEE 30th Int. Conf. on Tools with Artificial Intelligence (ICTAI), Volos, Greece, pp. 409-413, Nov. 2018. https://doi.org/10.1109/ICTAI.2018.00070.

[8] Personal Information Protection Commission

(PIPC), "Guideline for Establishment of Video Surveillance Device Operation and Management Policy", 2023. https://www.pipc.go.kr. [accessed: Feb. 10, 2026]

[9] Korea Ministry of Government Legislation, Restriction on Installation and Operation of Fixed Visual Data Processing Devices, Personal Information Protection Act (PIPA), Art. 25, Act. no. 19234, amended Mar. 14, 2023, effective Sep. 15, 2023. https://www.law.go.kr/LSW/lsInfoP.do?lsId=011357. [accessed: Feb. 10, 2026]

[10] G. Jocher, J. Qiu, and A. Chaurasia, "Ultralytics YOLO (v8.3.0)", Zenodo, Sep. 2024. https://doi.org/10.5281/zenodo.13858602.

[11] Y. Zhang, P. Sun, Y. Jiang, D. Yu, F. Weng, Z. Yuan, P. Luo, W. Liu, and X. Wang, "ByteTrack: Multi-Object Tracking by Associating Every Detection Box", Proc. of European Conf. on Computer Vision (ECCV), Tel Aviv, Israel, pp. 1-21, Oct. 2022. https://doi.org/10.1007/978-3-031-20047-2_1.

[12] Roboflow, "Supervision", version 0.27.0, Jun. 2026. https://github.com/roboflow/supervision. [accessed: Feb. 10, 2026]

[13] Seoul Metro, "Smart Station Operation Status", 2026. https://www.seoulmetro.co.kr. [accessed: Feb. 10, 2026]

[14] AI Hub, "Subway Abnormal Behavior Video Dataset", National Information Society Agency (NIA), 2022. https://aihub.or.kr/aihubdata/data/view.do?dataSetSn=174. [accessed: Feb. 10, 2026]

저자소개

김 민 지 (Minji Kim)

2024년 2월 ~ 현재 : 경기대학교 컴퓨터공학과 학사과정
관심분야 : 데이터마이닝, 컴퓨터비전, 인공지능, 빅데이터

윤 효 정 (Hyojeong Yoon)

2022년 2월 ~ 현재 : 경기대학교 컴퓨터공학과 학사과정
관심분야 : 컴퓨터비전, 인공지능, 객체 탐지, 딥러닝

이 동 근 (Donggeun Lee)

2022년 2월 ~ 현재 : 경기대학교 컴퓨터공학과 학사과정
관심분야 : 컴퓨터 그래픽스, 게임 엔진, 인터랙티브 시스템

이 지 현 (Jihyun Lee)

2023년 2월 ~ 현재 : 경기대학교 컴퓨터공학과 학사과정
관심분야 : 사이버 보안, 접근 제어 시스템, 개인정보 보호

정 경 용 (Kyungyong Chung)

2000년 2월 : 인하대학교 전자계산공학과(공학사)
2002년 2월 : 인하대학교 전자계산공학과(공학석사)
2005년 8월 : 인하대학교 전자계산공학과(공학박사)
2006년 3월 ~ 2017년 2월 : 상지대학교 컴퓨터정보공학부 교수
2017년 3월 ~ 현재 : 경기대학교 AL컴퓨터공학부 교수
관심분야 : 데이터 마이닝, 헬스케어, 빅데이터, HCI, 지능시스템, 인공지능