

# Scene-centric 트레킹 기법을 이용한 CCTV 영상 분석

장동재<sup>1</sup>, 최민준<sup>1</sup>, Rakshitha Nagaraj<sup>1</sup>, 김현우<sup>2</sup>, 고석주<sup>3</sup>

<sup>1</sup> 경북대학교 컴퓨터학부 학부생

<sup>2</sup> 사로리스 주식회사 대표

<sup>3</sup> 경북대학교 컴퓨터학부 교수

qazw1227@knu.ac.kr, minjunic@knu.ac.kr, rakshitha03@knu.ac.kr, maker@saloris.world, sjkho@knu.ac.kr

## Scene-Centric Tracking for CCTV Video Analysis

Dong-Jae Jang<sup>1</sup>, Min-Jun Choi<sup>1</sup>, Rakshitha Nagaraj<sup>1</sup>, Hyum-Woo Kim<sup>2</sup>, Seok-Joo Koh<sup>3</sup>

<sup>1</sup>School of Computer Science and Engineering, Kyungpook National University

<sup>2</sup>Saloris Inc.

<sup>3</sup> School of Computer Science and Engineering, Kyungpook National University

### 요약

CCTV 를 활용한 교통 모니터링은 스마트시티의 핵심 요소로, 본 연구는 배경 기반(Scene-Centric) 접근법을 통해 차량 이동 경로를 추적하는 새로운 트래킹 기법을 제안한다. 제안 시스템은 Cityscapes 데이터셋으로 학습된 SegFormer 모델을 활용해 도로와 차량 영역을 세그멘테이션하고, 픽셀(Pixel) 기반 움직임 감지를 통해 추적 대상을 식별한다. 이를 시작 위치 기반 그리드 클러스터링(Grid Clustering)으로 그룹화하여 상행/하행 방향을 자동 식별하며, 역주행 및 과속과 같은 이상 행동을 실시간 탐지한다. 야간 환경에서 헤드라이트로 인한 오탐지를 줄이기 위해 HSV 색상 공간 기반 적응형 가중치를 적용해 보았으나, 탐지 성능 개선이 필요함을 확인했다. 실험 결과, 제안 시스템은 주간 환경에서 평균 방향 식별 정확도 67.37%, 야간 환경에서 58.11%를 달성했으며, 야간 영상에서 YOLO+ByteTrack 대비 프레임 처리 속도가 62.19% 빠르고 FPS 는 164.48% 향상되었다.

### 1. 서론

CCTV 를 활용한 교통 흐름 분석, 이상 행동 탐지, 사고 예방은 스마트시티에서 실시간 데이터 분석의 필요성을 높이고 있다. 기존 객체 기반(Object-Centric) 접근법은 복잡한 교통 환경에서 가림(Occlusion)이나 조명 변화로 인해 추적 정확도가 저하되는 한계가 있었다. 반면, 배경 기반(Scene-Centric) 접근법은 비디오 인스턴스 세그멘테이션에서 성능 향상을 보여주고 있다[1]. 그러나 기존 배경 기반 연구는 통계적 모델이나 단순 궤적 분석에 의존하여 차선 단위 분석과 직관적 시각화에 한계가 있었다.

본 연구는 Cityscapes 데이터셋[2]으로 학습된 SegFormer 모델[3]을 활용해 픽셀 기반 움직임 감지와 시작 위치 기반 그리드 클러스터링을 도입하여 차선 단위로 차량을 그룹화하고, 상행/하행 방향을 자동 식별한다. 이를 통해 차선별 이동 패턴을 시각화하며, 역주행 및 과속과 같은 이상 행동을 실시간 탐지한다. 또한, HSV 색상 공간 기반 적응형 가중치를 적용해 그림자와 헤드라이트로 인한 오탐지를 줄이고자 했다.

### 2. 본론

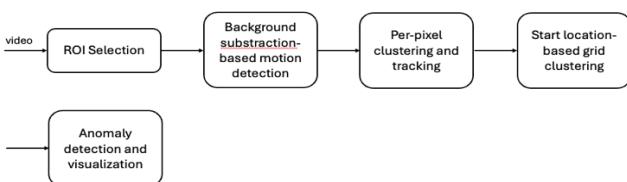
#### 2.1 기존 연구의 한계

기존 객체 기반 접근법은 개별 객체의 ID 를 추적하는 방식으로, 실시간 처리 시 연산량 증가로 성능이 저하되었다. 특히, 복잡한 교통 환경에서 차량 간 가림이나 조명 변화로 인해 객체 ID 가 자주 손실되어 추적의 연속성이 깨지는 문제가 있었다. 배경 기반 접근법[1]은 통계 기반 이상 행동 탐지에 주로 활용되었으나, 시멘틱(Semantic) 정보 기반의 차선 단위 시각적 분석은 제공하지 못했다.

#### 2.2 제안하는 시스템 구조

본 연구는 Cityscapes 데이터셋[2]으로 학습된 SegFormer[3] 기반의 시멘틱 세그멘테이션을 활용하여 도로와 차량 영역을 추출한다. 그럼 1 과 같이, 시스템은 다음의 단계를 따른다. 먼저, 관심 영역(Region of Interest, ROI)을 설정하여 분석 대상을 지정한다. 이어서, MOG2 기법을 활용한 배경 차분 기반 움직임 감지(Background Subtraction-based Motion Detection)를 통해 픽셀 단위로 움직임을 검출한다. 감지된 픽셀은 픽셀 단위 클러스터링 및 추적(per-pixel clustering and tracking)을 통해 그룹화되며, 이때 면적(500 픽셀 이상)과 종횡비(0.5~2.0)를 만족하는 객체만을 추적 대상으로 선정하고 HSV 색상 공간 기반 적응형 가중치를

적용해 그림자와 헤드라이트로 인한 오탐지를 최소화하였다. 이후 각 객체는 중심점(centroid)을 기준으로 궤적을 저장하고, 시작 위치 기반 그리드 클러스터링(start location-based grid clustering)을 통해 차량의 주행 방향을 상행/하행으로 자동 분류한다. 마지막으로, 이상 행동 탐지 및 시각화(anomaly detection and visualization)를 수행하여 역주행, 과속 등의 이상 행동을 실시간으로 탐지하며, 차선 단위 이동 패턴을 시각적으로 표현한다.



(그림 1) 시스템 아키텍처

### 2.3 실험 설계

실험은 주간/야간 CCTV 영상을 테스트 데이터로 사용하여, 두 가지 접근법을 비교했다. 첫째, Cityscapes 데이터셋[2]으로 학습된 SegFormer 모델[3]을 활용해 도로와 차량 영역을 세그멘테이션하고, MOG2 배경 차분과 그리드 클러스터링을 통해 차량의 이동 경로와 방향을 추적했다(SegFormer-Grid). 둘째, YOLOv8[4] 모델로 차량 객체를 검출하고, YOLOv8 내장 ByteTrack 트래커[3]를 활용해 객체의 이동 경로와 방향을 추적했다. 두 접근법 모두 연속 프레임 간 중심점 이동 벡터를 분석해 차량의 상행/하행 방향을 식별했으며, 방향 식별 정확도와 실시간 성능(프레임 처리 속도, FPS)을 기준으로 성능을 평가했다.

### 2.4 실험 결과

실험은 주간과 야간 환경에서 진행되었다. 주간 환경에서 총 21 개 차량 트랙을 추적하며 SegFormer-Grid 는 평균 방향 식별 정확도 67.37%, YOLO+ByteTrack[4]은 80.34%를 달성했다. 야간 환경에서는 11 개 트랙을 추적하며 SegFormer-Grid 는 평균 방향 식별 정확도 58.11%, YOLO+ByteTrack[4]은 88.68%를 기록했다. 야간 환경에서 SegFormer-Grid 는 헤드라이트 그림자로 인해 추적이 불안정해 정차 탐지와 같은 오탐지가 발생했다. 그러나 실시간 성능에서는 SegFormer-Grid 가 YOLO+ByteTrack[4] 대비 프레임 처리 속도가 62.19% 빠르고 FPS 는 164.48% 향상되었다.

### 2.5 결과 분석

제안 시스템(SegFormer-Grid)은 주간 및 야간 환경에서 방향 식별 성능을 평가했다. 주간 환경에서 SegFormer-Grid 의 평균 방향 식별 정확도는 67.37%, YOLO+ByteTrack[4]은 80.34%로, YOLO+ByteTrack[4] 이 더 높은 정확도를 보였다. 야간 환경에서는 SegFormer-Grid 가 58.11%, YOLO+ByteTrack[4]가 88.68%로, 야간 환경에서 성능 차이가 더 두드러졌다. 이는 HSV 색상 공간 기반 적응형 가중치를 적

용했으나 헤드라이트로 인한 오탐지 문제를 완전히 해결하지 못했기 때문이다. SegFormer-Grid 의 주요 장점은 차선 단위로 차량을 그룹화하여 이동 패턴을 시각화할 수 있다는 점과, 픽셀 기반 접근법으로 연산 효율성이 높아 실시간 성능에서 우수하다는 점이다. 실시간 성능 평가에서 SegFormer-Grid 는 YOLO+ByteTrack [4] 대비 프레임 처리 속도가 62.19% 빠르고 FPS 는 164.48% 향상되어 실시간 교통 모니터링에 유리한 성능을 보였다. 그러나 방향 식별 정확도가 YOLO+ByteTrack[4]보다 낮은 점은 배경 기반 접근법의 한계로, 특히 조명 변화에 민감한 MOG2 배경 차분 기법의 특성에서 기인한다. 향후 야간 환경에서의 객체 탐지 성능 개선과 이동 경로 곡률을 고려한 방향 식별 알고리즘 도입이 필요하다.

## 3. 결론

본 연구는 배경 기반(Scene-Centric) 접근법을 통해 차량 이동 경로를 추적하고, 차선 단위로 그룹화하며, 상행/하행 방향을 자동 식별하는 시스템(SegFormer-Grid)을 제안했다. Cityscapes 데이터셋[2]으로 학습된 SegFormer 모델[3]을 활용해 도로와 차량 영역을 세그멘테이션하고, 픽셀 기반 움직임 감지와 그리드 클러스터링을 통해 차량 추적 및 방향 식별을 수행했다. 실험 결과, 주간 환경에서 평균 방향 식별 정확도 67.37%, 야간 환경에서 58.11%를 달성했으며, YOLO+ByteTrack[4](주간 80.34%, 야간 88.68%) 대비 낮은 정확도를 보였다. 그러나 실시간 성능 평가에서 YOLO+ByteTrack[4] 대비 프레임 처리 속도가 62.19% 빠르고 FPS 는 164.48% 향상되어 실시간 교통 모니터링에 유리함을 입증했다. 향후 연구에서는 야간 데이터로 SegFormer 모델을 추가 학습시켜 객체 탐지 성능을 개선하고, 다양한 환경에서의 성능 평가와 더 복잡한 이상 행동 탐지 알고리즘을 개발하여 시스템의 실용성을 강화할 예정이다.

## 사사문구

“이 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW 중심대학사업 지원을 통해 수행되었음”(2021-0-01082)”

## 참고문헌

- [1] Hwang, S., Heo, M., Oh, S. W., & Kim, S. J. (2021). "Video Instance Segmentation with IFC." Advances in Neural Information Processing Systems (NeurIPS), 2021.
- [2] Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., ... & Schiele, B. (2016). "The Cityscapes Dataset for Semantic Urban Scene Understanding." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
- [3] Xie, E., Wang, W., Yu, Z., Anandkumar, A., Alvarez, J. M., & Luo, P. (2021). "SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers." Advances in Neural Information Processing Systems (NeurIPS), 2021.
- [4] Ultralytics (2023). "YOLOv8: A High-Performance Object Detection Model with ByteTrack." <https://docs.ultralytics.com>