대한산업공학회/한국경영과학회/한국시뮬레이션학회
2025년 춘계공동학술대회

Towards Global Eminence

TrackLens: 객체 추적 모델의 설명가능 기법

TrackLens: Explanation on Object Tracking Models

제주 해비치 호텔&리조트
2025.06.20(금)

경희대학교 빅데이터응용학과[1], 산업경영공학과[2]
산업지능연구실
정 솔[1], 한영석[1], 정재윤[1,2,*]
bohomi1995j@khu.ac.kr

- 본 연구는 산업통상자원부의 재원으로 한국산업기술기획평가원[전자부품산업기술개발사업, No. RS-2024-00442354]와 방위사업청의 재원으로 국방기술진흥연구소(V220023, 자율운항 성능 고도화를 위한 해상환경 학습 데이터셋 구축)의 지원을 받아 수행되었음.

목차

1. 연구 동기 및 목적

2. 배경 지식
   - SORT (Simple Online and Realtime Tracking)
   - 개선된 추적 기법들

3. TrackLens: Tracking을 위한 XAI
   - Appearance 기반 설명
   - Movement 기반 설명
   - Matching 기반 설명

4. 결론 및 추후 연구

참고문헌

연구 동기 및 목적

- **다중 객체 추적(Multi-Object Tracking, MOT)**
  - 감시, 자율 주행, 군중 분석 등 다양한 컴퓨터 비전 응용 분야에 있어 중요한 기술
  - Tracking-by-Detection 패러다임(Bewley, Alex, et al., 2016)에서 Deep learning(Sun, Yifan, et al., 2018)과 결합하여 안정적이고 빠른 추적이 가능한 알고리즘 등장 (DeepSORT, StrongSORT 등)

- **기존 기법의 한계**
  - 복잡한 알고리즘 (블랙박스) → 성능 ↑ → 사용자의 이해 ↓
  - Detection이 어떻게 특정 track에 할당되었는지, 어떤 시각적 이유로 추적이 유지되거나 변경되었는지에 대한 명확한 설명을 제공하지 못함

    → 안전성, 투명성, 신뢰성이 중요한 실제 응용 환경에서 한계로 작용

- **연구의 필요성**
  - 복잡한 MOT 모형의 작동 방법과 추적 결과에 대하여 확인할 수 있는 설명가능한 AI(XAI) 기법의 개발 필요

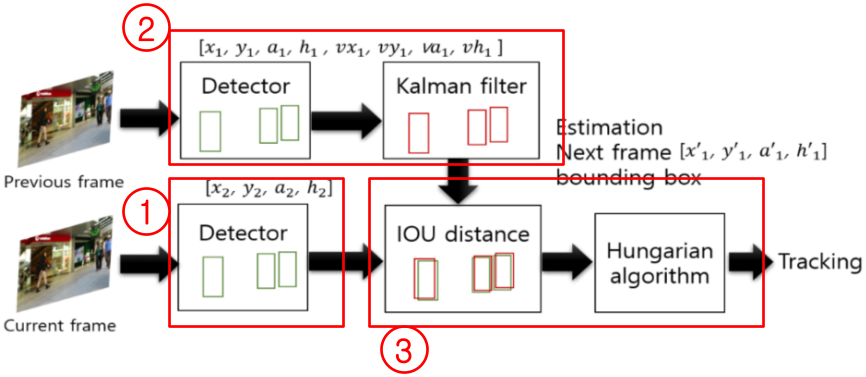
- **제안 기법: TrackLens (XAI for StrongSORT)**
  - Tracker의 객체 추적 과정을 체계적으로 확인하고 해석할 수 있는 방안을 **세 가지 측면에서 제안**

1) **Appearance 측면**: Re-ID feature map을 활용하여 Tracker가 지속적으로 활용 중인 영역을 시각화
2) **Movement 측면**: 칼만 필터 기반의 예측과 지나온 경로를 시각화
3) **Matching 측면:** StrongSORT의 2가지 matching 단계를 시각화

→ 단순한 객체 추적에서 나아가, ID 할당 오류나 drift 등 오류 원인을 분석

- **SORT (Simple Online and Realtime Tracking)**
  - 다중 객체 추적 문제에 대한 추적 감지 프레임워크
  - 3가지 단계로 수행

    (1. Detection → 2. Predict → 3. Association)

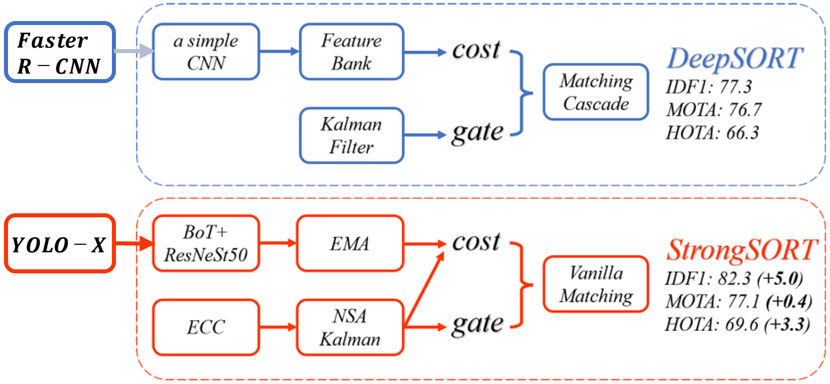


SORT 객체 추적 예시 (김경훈 외., 2020)

- **Re-Identification (Re-ID) Model**
  - 객체를 다른 시점에서 다시 식별하기 위한 모델
  - 고유한 특징(appearance) 을 학습해 동일한 객체를 재식별할 수 있도록 훈련

Re-ID 학습데이터 예시 (Sun, Yifan, et al., 2018)

- **Advanced Tracker (DeepSORT, StrongSORT 등)**
  - SORT + Deep Learning
  - Re-ID(CNN)(Sun, Yifan, et al., 2018)을 이용해 appearance extracting (embedding)
  - 추출한 appearance 정보를 movement 정보를 결합해 추적에 활용

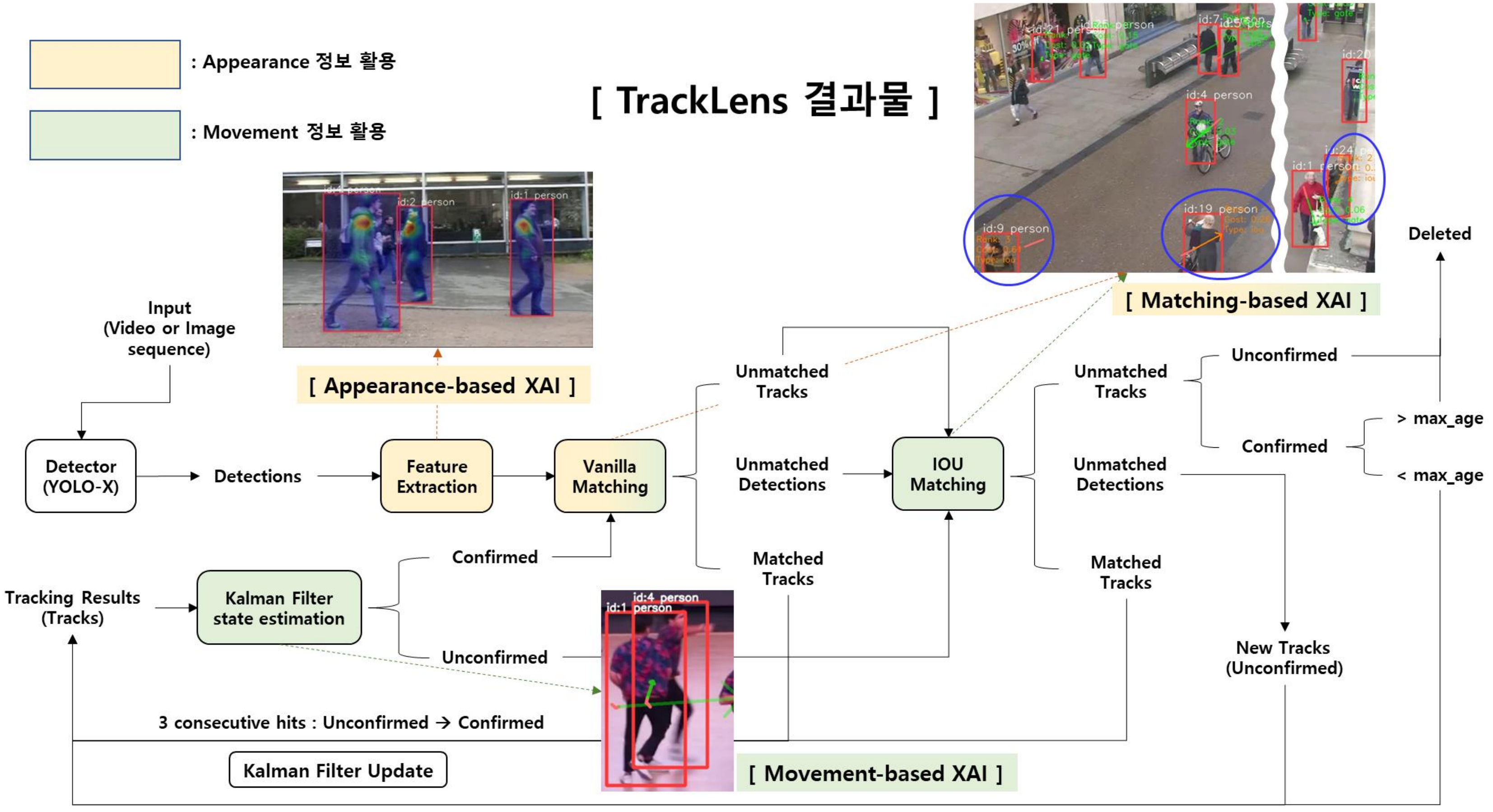


DeepSORT와 StrongSORT의 프레임워크 비교(Du, Yunhao, et al., 2023)

TrackLens 프레임워크

Towards Global Eminence

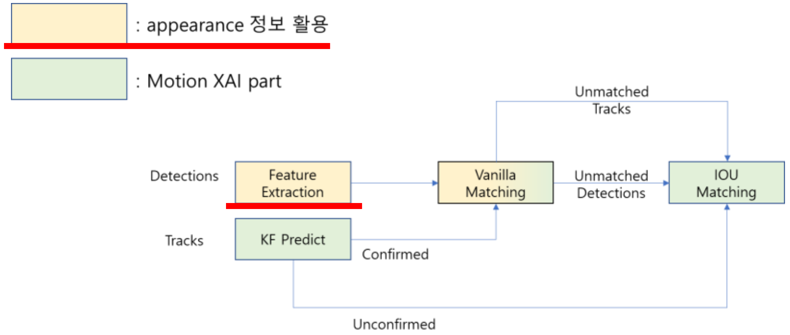
- **TrackLens**
  - StrongSORT를 위한 설명가능AI(XAI) 기법
  - Appearance와 Movement, Matching를 기반으로 객체 추적의 설명 방법 제공



TrackLens 프레임워크

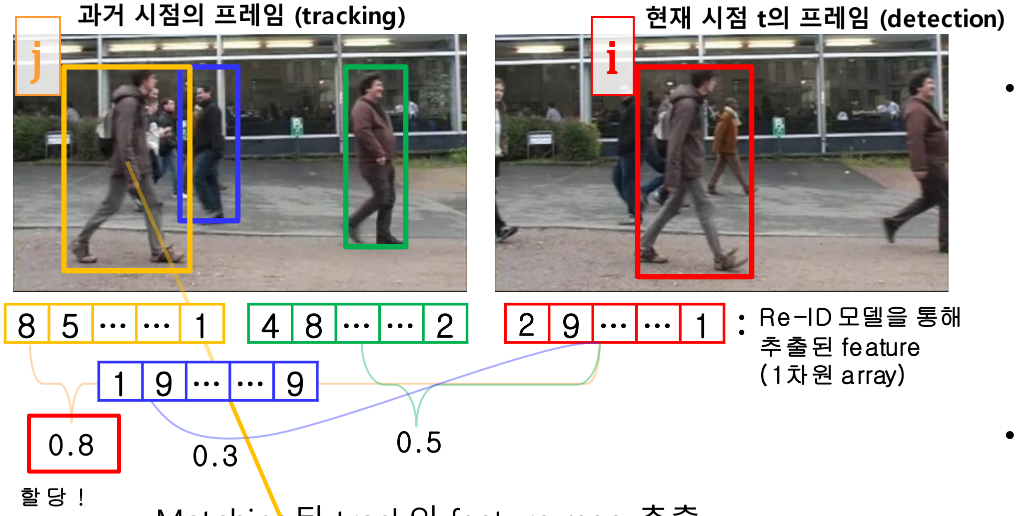
TrackLens: Appearance 기반 설명

- **Appearance 기반 설명**
  - track과 detection으로부터 Re-ID모델을 이용하여 feature 추출
  - Vanilla matching 단계에서 추출된 feature간의 cosine 유사도 계산
  - Matching된 track의 feature map를 통해 saliency map 생성
  - Tracker가 객체의 어떤 부분에 주의를 기울이는지 시각화

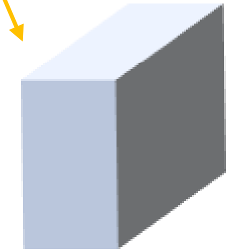


TrackLens의 프레임워크 (간략화)

  - Feature vector 추출 및 코사인 유사도 계산



  - Matching된 track의 feature map 추출 (3차원 tensor)

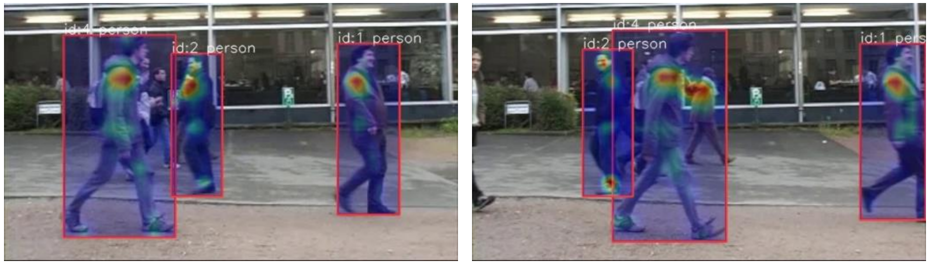


- Feature map은 일시적인 detection의 노이즈를 방지하기 위한 Exponential Moving Average (EMA)가 적용된 형태

$$E_i^t = \alpha E_i^{t-1} + (1-\alpha) F_i^t$$

  - F_i^t : i-번째 detection의 t 프레임에서의 feature map
  - α : momentum term (0.9) (Du, Yunhao, et al., 2023)

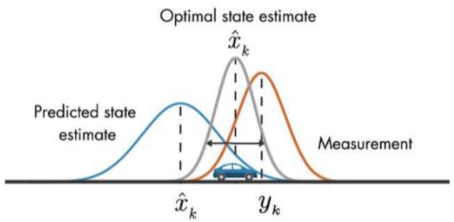
- Track이 특정 객체를 추적하는데 중점을 두는 위치를 확인하기 위한 activation map을 시각화 (Du, Yunhao, et al., 2019)



feature map 시각화 결과 (왼쪽-matching된 track, 오른쪽-detections)

TrackLens: Movement 기반 설명

- **Movement 기반 설명**
  - 칼만 필터(Kalman, Rudolph Emil., 1960) 의 예측값과 객체 추적 결과를 비교하는 **경로(Movement)** 설명 기법 개발
  - 칼만 필터의 예측값과 객체 추적 결과를 이동의 결과값을 사용하여 객체 이동방향을 어떻게 추적하는지를 화살표로 시각화하여 제시



칼만 필터의 예시

- 칼만 필터를 통해 직전 프레임에서 예측한 결과를 화살표로 표시 (직선을 통해 이동 궤적 표현)

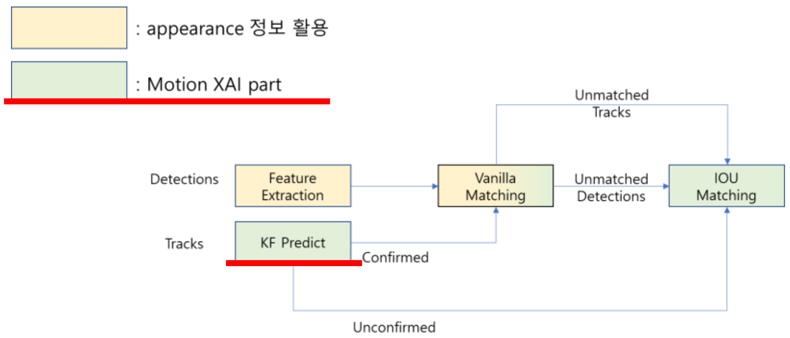
$$\hat{x}'_k = F_k \hat{x}_{k-1}$$

$$P'_k = F_k P_{k-1} F_k^T + Q_k$$

- $\hat{x}_{k-1}, P_{k-1}$: k-1 시점에서의 평균과 분산
- $\hat{x}$: 칼만 필터의 평균, $(x, y, w, h, \dot{x}, \dot{y}, \dot{w}, \dot{h})$
  - (x, y, w, h) : bounding box의 좌측 상단 좌표 (x, y)와 width, height
  - $(\dot{x}, \dot{y}, \dot{w}, \dot{h})$: 각각의 속도

- **화살표의 크기가 클수록 이동 속도 ↑**
  (프레임의 간격에 따라 화살표의 크기 조정)

- 투명도를 통해 matching의 신뢰도 표현
  (진할 수록 신뢰도 ↑)

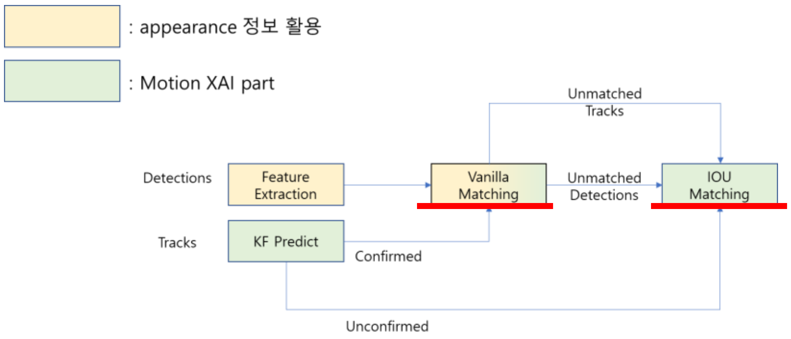


TrackLens의 프레임워크 (간략화)

track의 이동 경로와 화살표로 표현된 직전 프레임에서의 칼만필터의 예측 방향

TrackLens: Matching 기반 설명

- **Matching(Assignment) 기반 설명**
  - StrongSORT는 2가지 matching 단계 존재
    - 1단계: appearance를 통한 vanilla matching
    - 2단계: 칼만 필터를 통한 IoU matching
  - 색깔을 통해 matching이 이루어진 단계 구분
    - 초록색: vanilla matching
    - 주황색: IoU matching
  - Matching (assignment)(Kuhn and Harold W., 1955)이 될 때의 cost에 따라서 화살표의 투명도 조정



TrackLens의 프레임워크 (간략화)

TrackLens를 통한 StrongSORT의 matching 단계 시각화

할당 문제 (Assignment Problem)

- 두 개의 이산 집합(예: 작업자 집합과 작업 집합) 사이에서 **일대일 매칭**을 수행하여, 주어진 **비용(cost)** 행렬에 따라 **전체 비용을 최소화**하는 **조합 최적화 문제**

	Detection 1	Detection 2	Detection 3
Track 1	4	7	3
Track 2	5	6	1
Track 3	2	5	3

Assignment problem solution의 예시

결론 및 추후 연구

- **연구 요약**
  - Multi-Object Tracking을 위해 고안된, Deep Learning과 SORT를 융합한 StrongSORT를 대상으로 XAI 진행
  - StrongSORT를 크게 Appearance, Movement, Matching의 세 가지 측면에서 설명 제공
    - **Appearance:** CNN 계열의 Re-ID 모델을 통해 추출된 feature map을 통해 추적기가 track의 어느 부분에 집중하고 있는지를 시각화
    - **Movement:** 직전 프레임에서 Kalman Filter를 통해 다음 위치를 예측하는 작동방식을 활용해, 화살표를 통해 track이 나아갈 방향을 표시 (지나온 궤적은 line으로 표현)
    - **Matching:** 또한 화살표 및 글자의 색깔을 통해 어떤 단계에서 matching이 이루어졌는지 직관적으로 알아볼 수 있도록 구현

- **연구 공헌점**
  - 객체 추적(MOT)을 위한 설명가능한 AI의 개념을 정의하고 기법을 최초로 제안
  - 이를 통해 복잡한 tracking 알고리즘의 작동방식을 이해하고 오류 발생 원인 및 개선을 위한 통찰 제공

- **추후 연구**
  - 다양한 동영상에 대한 TrackLens의 적용 및 효과 검증
  - 최신 MOT 알고리즘(예. Transformer 기반)에 대하여 TrackLens의 확장 적용

참고문헌

김경훈 외, "CPU 환경에서의 실시간 동작을 위한 딥러닝 기반 다중 객체추적시스템", 2020.

Bewley, Alex, et al. "Simple online and realtime tracking." 2016 IEEE international conference on image processing (ICIP), 2016.

Chu, Peng, and Haibin Ling. "Famnet: Joint learning of feature, affinity and multi-dimensional assignment for online multiple object tracking." Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019.

Du, Yunhao, et al. "StrongSort: Make deepsort great again." IEEE Transactions on Multimedia 25 (2023): 8725-8737.

Kalman, Rudolph Emil. "A new approach to linear filtering and prediction problems." (1960): 35-45.

Kuhn, Harold W. "The Hungarian method for the assignment problem." Naval research logistics quarterly 2.1-2 (1955): 83-97.

Sun, Yifan, et al. "Beyond part models: Person retrieval with refined part pooling (and a strong convolutional baseline)." Proceedings of the European conference on computer vision (ECCV). 2018.

Wojke, Nicolai, Alex Bewley, and Dietrich Paulus. "Simple online and realtime tracking with a deep association metric." 2017 IEEE international conference on image processing (ICIP). IEEE, 2017.

Zhou, Kaiyang, et al. "Omni-scale feature learning for person re-identification." Proceedings of the IEEE/CVF international conference on computer vision. 2019.