

# **Object-centric Auto-encoders and Dummy Anomalies for Abnormal Event Detection in Video**

CVPR 2019

Radu Tudor Ionescu, Fahad Shahbaz Khan, Mariana-Iuliana Georgescu, Ling Shao

SeulGi Park

February 06, 2020

# Contents

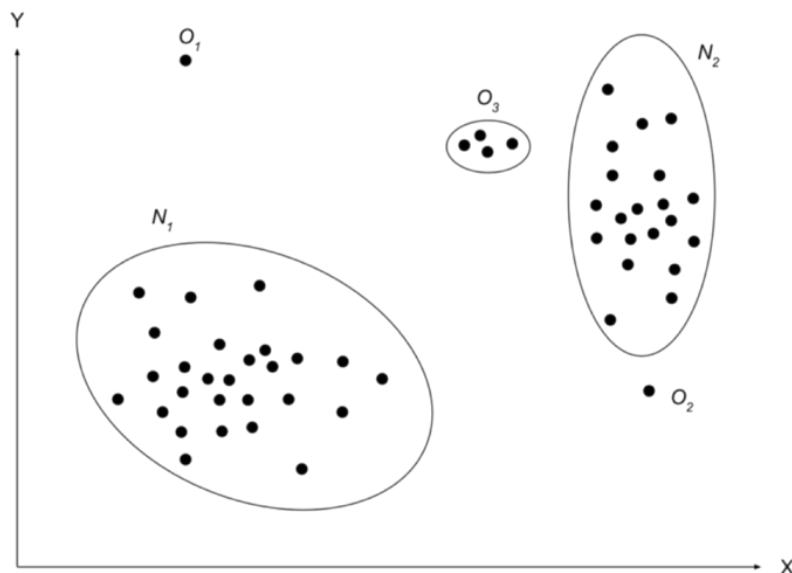
---

1. Introduction
2. Related Work
3. Method
4. Experiments
5. Conclusion

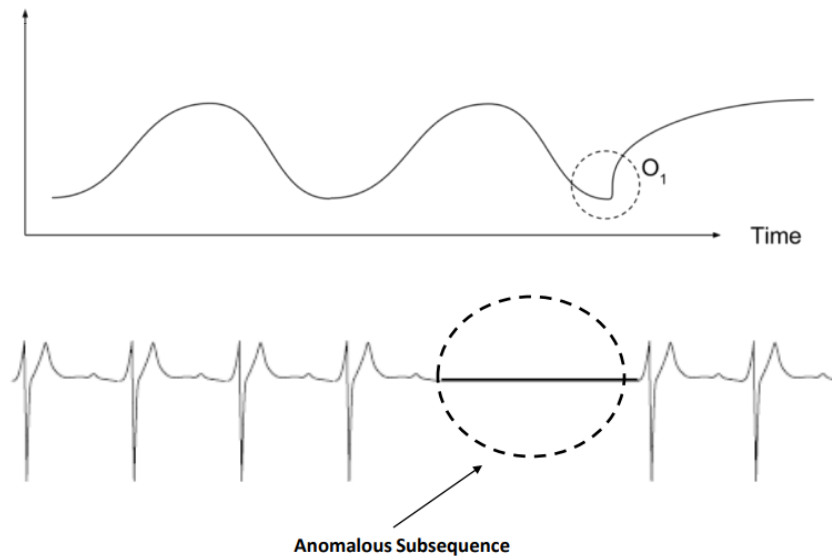
# 1. Introduction

- 이상 탐지(Anomaly Detection)?

- Anomalies = Outliers
- 일상생활에서 예상과는 다른 패턴을 보이는 개체나 데이터를 찾는 것을 말함



군집형 데이터 이상 탐지



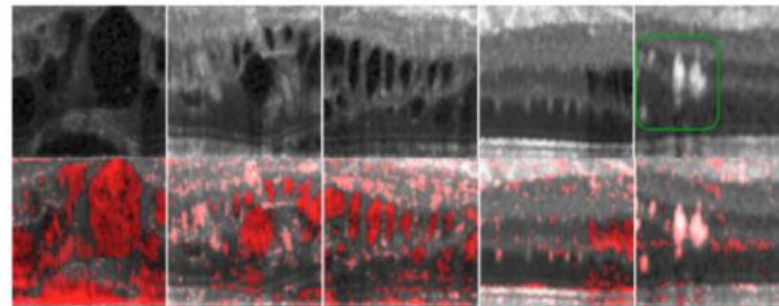
시계열 데이터 이상 탐지

# 1. Introduction

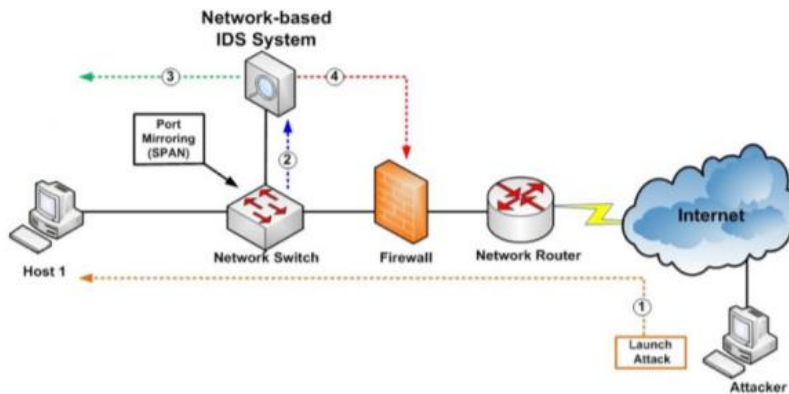
- 최근 이상 탐지는 불법 트래픽 탐지와 같은 다양한 분야에서 기존의 정확도를 높이거나 새로운 문제를 해결하는 방법으로 활발하게 사용되고 있음



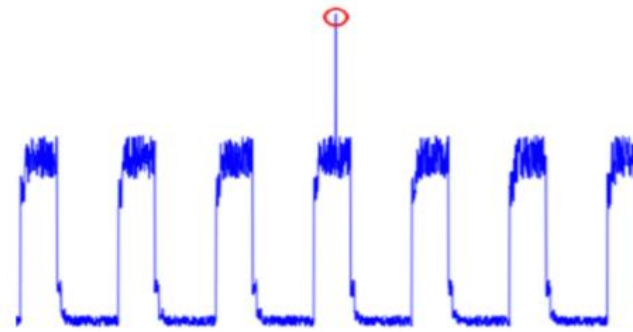
(a) 비디오 감시 : 불법 트래픽 탐지



(b) 건강 관리 : 망막 손상 감지



(c) 네트워크 : 사이버 침입 탐지



(d) 센서 네트워크 : IoT 빅 데이터 이상 탐지

# 1. Introduction

---

- 이상 탐지(Anomaly Detection)?
  - 비디오를 이용한 이상 탐지 연구는 지속적으로 연구되어 왔음
  - 정상 이벤트와 비정상 이벤트의 구분이 문제에 따라 다양하기 때문에, 문제에 의존하는 경우가 많음
    - 정상: 거리에서 트럭을 운전하는 시나리오 / 비정상: 보행자 도보에서 트럭을 운전하는 시나리오
  - 비정상 이벤트보다 정상 이벤트가 상대적으로 많이 발생함
  - 다양한 문제에 따른 규칙을 만들 수 없기 때문에 전통적인 supervised 방법(rule based)을 사용하기 어려움
  - 기존 이상탐지 방법은 정상 이벤트만 포함된 학습 비디오에서 정규 모델을 학습
  - 정규 모델에서 벗어나는 이벤트의 경우 비정상으로 판단

# 1. Introduction

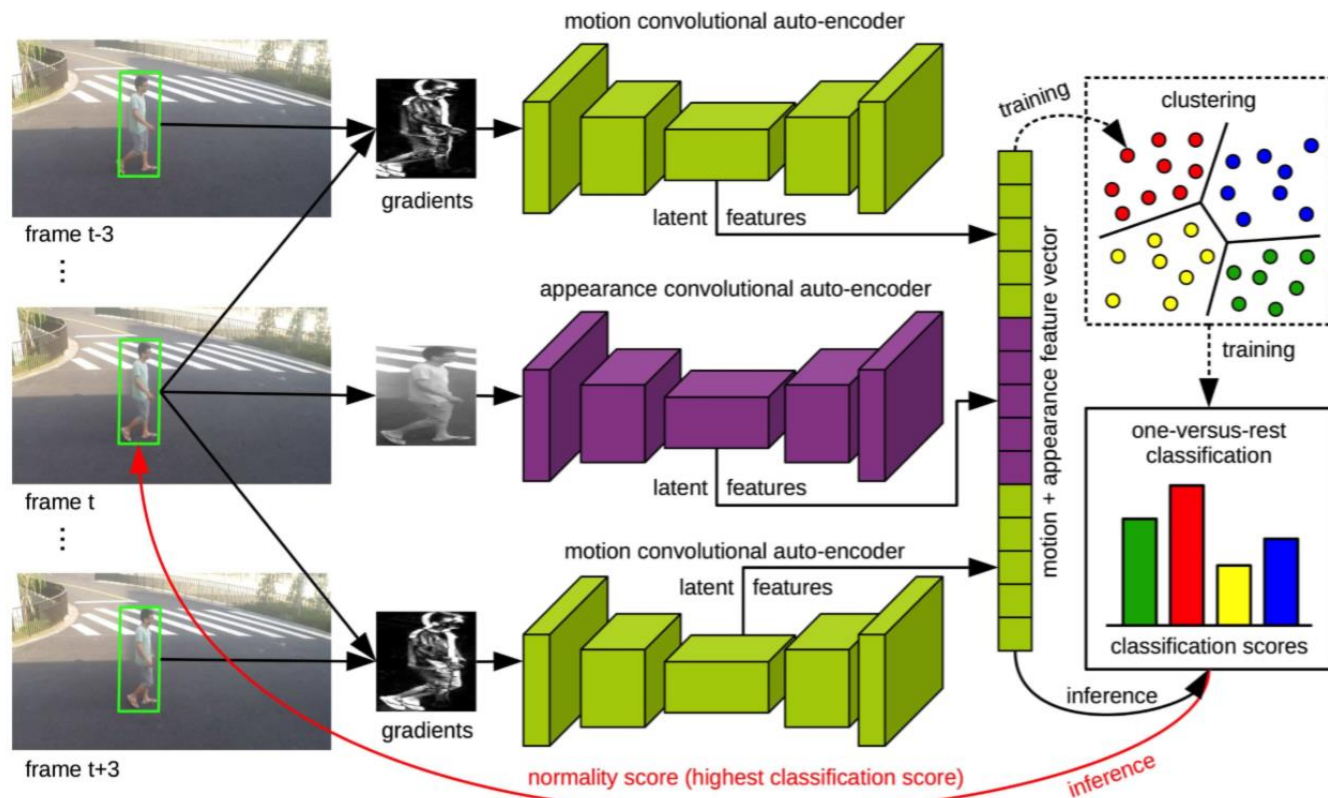
---

- 제안하는 방법?
  - 본 논문에서는 다중 클래스 분류 문제로 비정상 이벤트를 감지
  - 학습 데이터에서는 정상 이벤트만 포함, 다양한 유형을 나타내는 군집을 찾기 위해 k-means clustering 사용
  - 비정상 이벤트 감지를 OvR 바이너리 분류 문제로 지정
  - 첫번째, 모션 중심의 모양 정보를 인코딩하기 위해 객체 중심 컨볼루션 자동 인코더를 기반으로 하는 unsupervised 학습 프레임 워크를 제안
  - 두번째, 학습 데이터를 클러스터링 하기 위해 supervised classification 제안
  - Inference 동안에 test data에 대응하는 최고 classification 점수는 각 data의 정상 점수를 나타냄
  - 점수가 음수인 경우 비정상으로 판단

\* One-versus-rest(OvR = One-vs-the-Rest)?  
K개의 클래스가 존재하는 경우, 각각의 클래스에 대해 표본이 속하는지( $y=1$ ) 속하지 않는지( $y=0$ )의 이진 문제를 푸는 것

# 1. Introduction

- 제안하는 방법?
  - 감지된 객체 위에 CAE(Convolutional Auto-Encoder)를 사용하여 각 프레임에서의 unsupervised 학습
  - 이를 통해 각 장면에 존재하는 오브젝트에만 초점을 맞출 수 있으며, 각 프레임에서 이상을 정확하게 localize 할 수 있음
  - 추론 단계에서 가장 높은 점수가 음수이면, 어떤 클래스에도 귀속되지 않기 때문에 비정상으로 판단함



\* One-versus-rest(OvR = One-vs-the-Rest)?  
K개의 클래스가 존재하는 경우, 각각의 클래스에 대해 표본이 속하는지( $y=1$ ) 속하지 않는지( $y=0$ )의 이진 문제를 푸는 것



## 2. Related Work

- Avenue



- ShanghaiTech

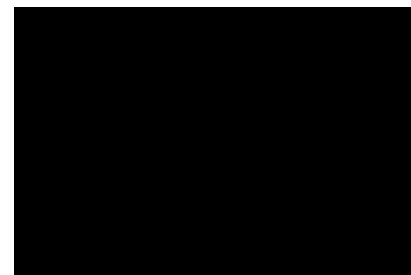
- (a) 원본 군중 이미지 / (b) 얼굴이 블러 처리된 이미지





## 2. Related Work

- UCSD(Ped1 / Ped2)
  - 1) 공원 도로에서 사람이 아닌 자전거, 오토바이 등 이동
  - 2) 사람이 도로가 아닌 잔디에 있는 것



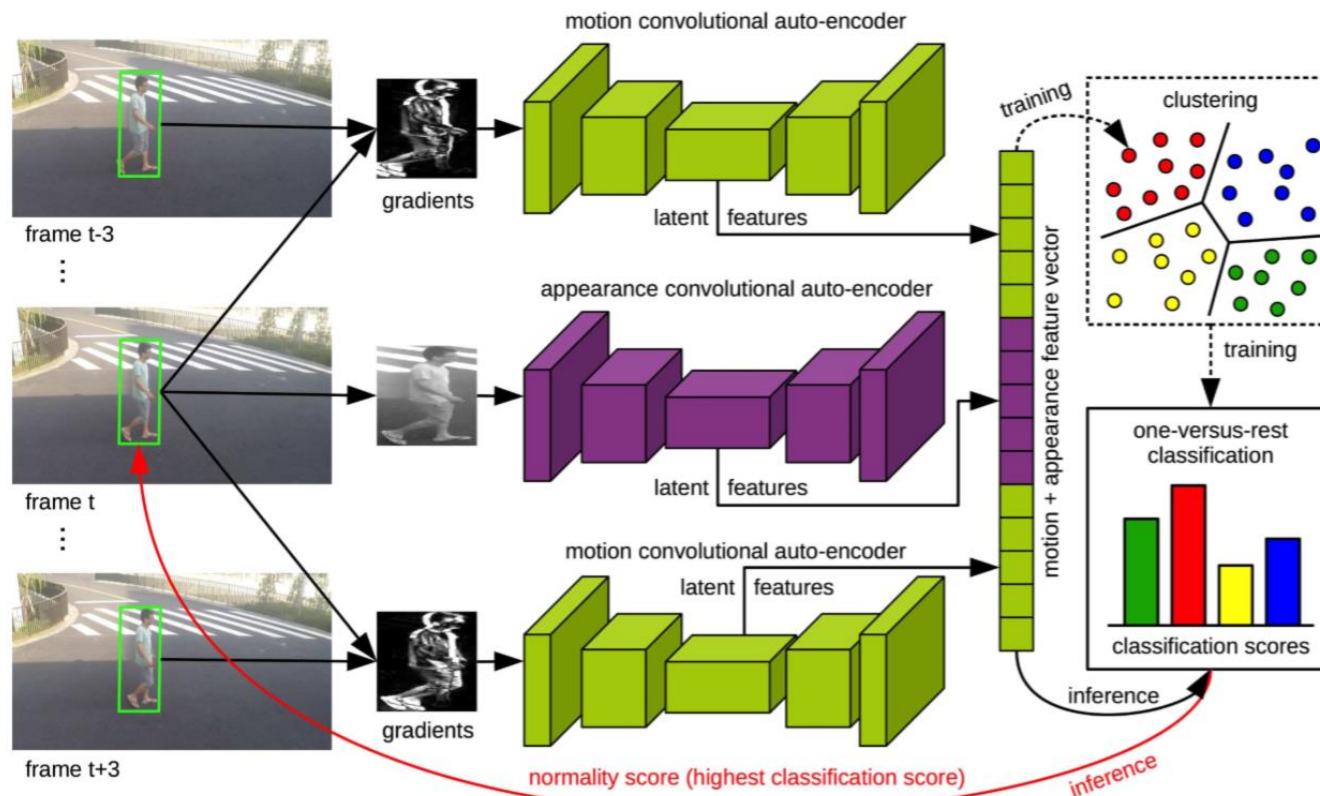
- UMN(Subway Entrance / Subway Exit)
  - 1) 지하철 입구인데, 나오는 사람
  - 2) 비용을 지불하지 않고 들어가는 사람

- 1) 지하철 출구인데, 들어가는 사람



# 3. Method

- Motivation?
  - 기존 이상 탐지를 위한 데이터셋이 다양하지 않음
  - 이상 탐지에서 이상을 정의하는 것이 어려움
  - 학습 데이터에는 정상 이벤트만 포함
  - 우수한 성능을 추출하기 위해, 객체의 다양한 정보를 이용하는 것이 좋음



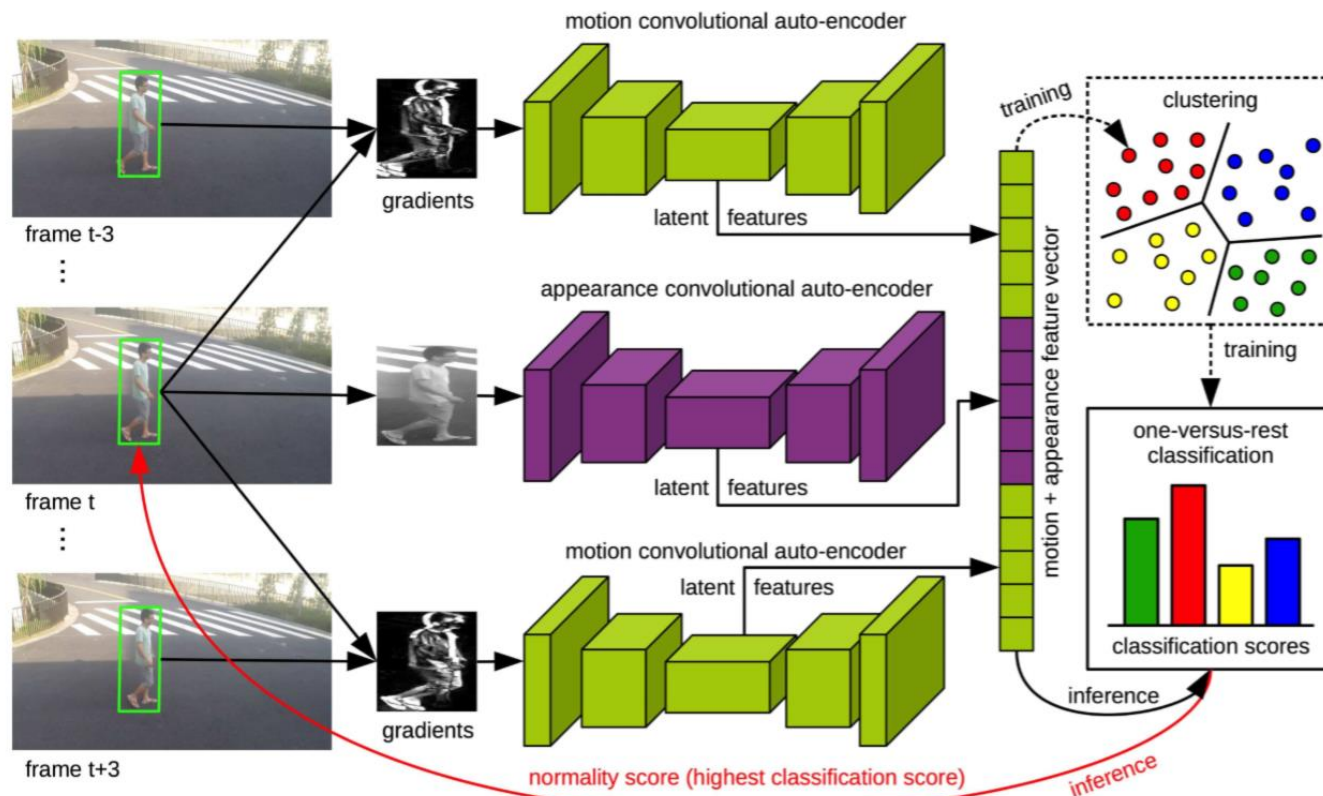
# 3. Method

---

- Object detection?
  - Single-shot object detection(FPN=Feature Pyramid Network)
    - 정확도와 속도 사이에 최적 균형
    - 작은 객체를 정확하게 감지할 수 있으며, GPU에서 초당 13 프레임 처리 가능
  - 각 프레임당 object detection, object 마다 bounding box로 crop
  - Appearance and Motion features
    - 자른 객체 이미지를 기준으로  $t-3$ ,  $t$ ,  $t+3$  프레임으로 gradients를 계산

# 3. Method

- Feature learning?
  - 각 물체에 대한 특징 벡터를 얻기 위해, 3개의 Convolution Auto-Encoder(CAE)를 학습
  - frame 1, Auto-Encoder는 객체를 포함하는 입력 이미지를 객체 기준으로 자르고, appearance features를 학습
  - frame t-3 and frame t+3, 객체가 어떻게 이동하였는지를 확인하는 gradients 입력으로 사용, motion 학습
  - frame t-3 and frame t+3 의 특징 dummy를 이용하여 학습 데이터 부족을 해결

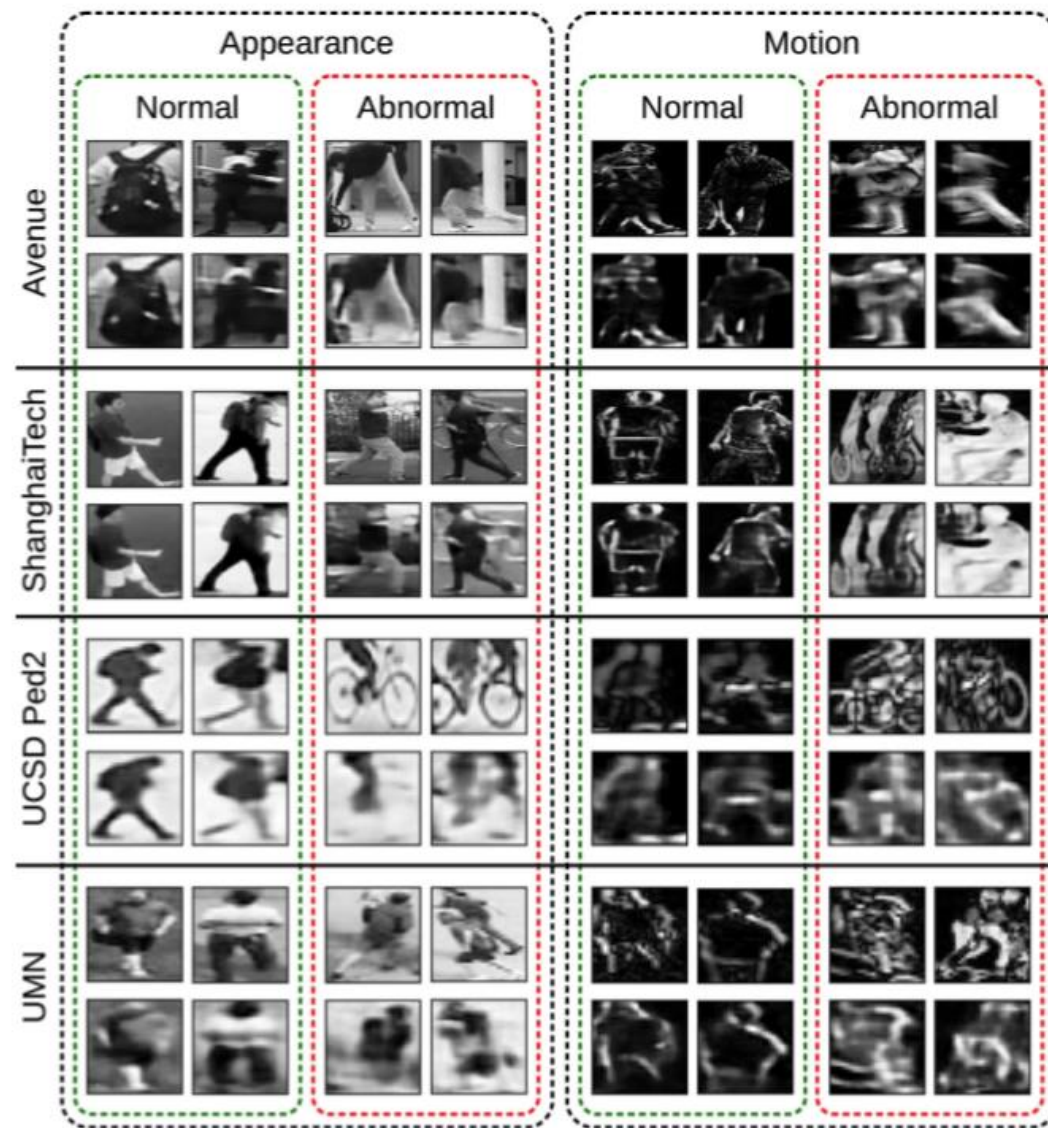




# 3. Method

- Object detection?

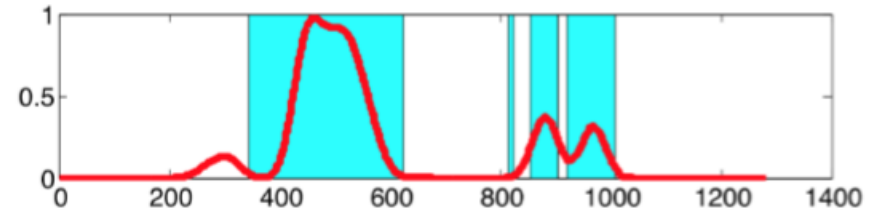
Figure 2. Normal and abnormal objects (left) and gradients (right) with reconstructions provided by the appearance (left) and the motion (right) convolutional auto-encoders. The samples are selected from the Avenue [23], the ShanghaiTech [24], the UCSD Ped2 [25] and the UMN [26] test videos, and are not seen during training the auto-encoders.



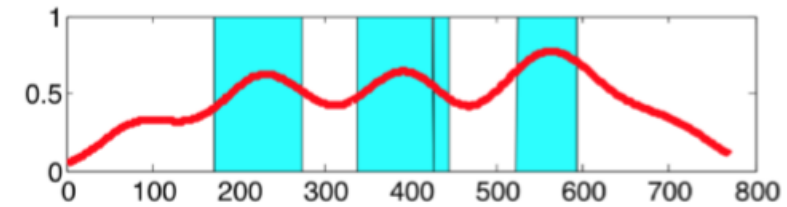
# 4. Experiments

- Results

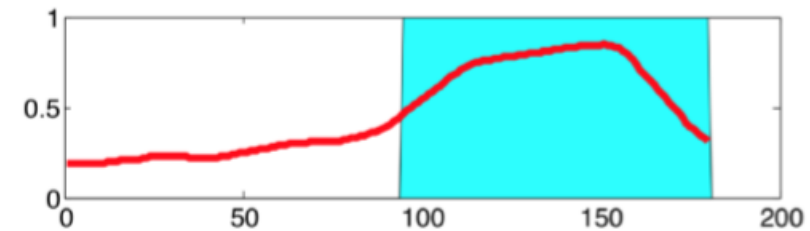
Method	Avenue	Shanghai Tech	UCSD Ped2	UMN
Kim et al. [15]	-	-	69.3	-
Mehran et al. [26]	-	-	55.6	96.0
Mahadevan et al. [25]	-	-	82.9	-
Cong et al. [6]	-	-	-	97.8
Saligrama et al. [32]	-	-	-	98.5
Lu et al. [23]	80.9	-	-	-
Dutta et al. [9]	-	-	-	99.5
Xu et al. [36, 37]	-	-	90.8	-
Hasan et al. [11]	70.2	60.9	90.0	-
Del Giorno et al. [7]	78.3	-	-	91.0
Zhang et al. [38]	-	-	91.0	98.7
Smeureanu et al. [33]	84.6	-	-	97.1
Ionescu et al. [13]	80.6	-	82.2	95.1
Luo et al. [24]	81.7	68.0	92.2	-
Hinami et al. [12]	-	-	92.2	-
Ravanbakhsh et al. [28]	-	-	93.5	99.0
Sabokrou et al. [31]	-	-	-	<b>99.6</b>
Ravanbakhsh et al. [27]	-	-	88.4	98.8
Liu et al. [21]	85.1	72.8	95.4	-
Liu et al. [22]	84.4	-	87.5	96.1
Sultani et al. [34]	-	76.5	-	-
Ionescu et al. [14]	88.9	-	-	99.3
<b>Ours</b>	<b>90.4</b>	<b>84.9</b>	<b>97.8</b>	<b>99.6</b>



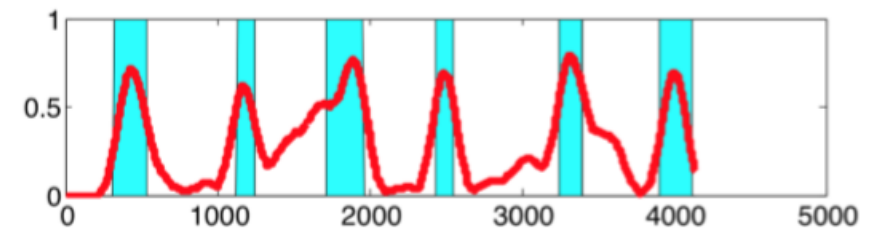
a) Anomaly scores on Avenue (test video 06)



b) Anomaly scores on ShanghaiTech (test video 05\_0023)



c) Anomaly scores on UCSD Ped2 (test video 02)

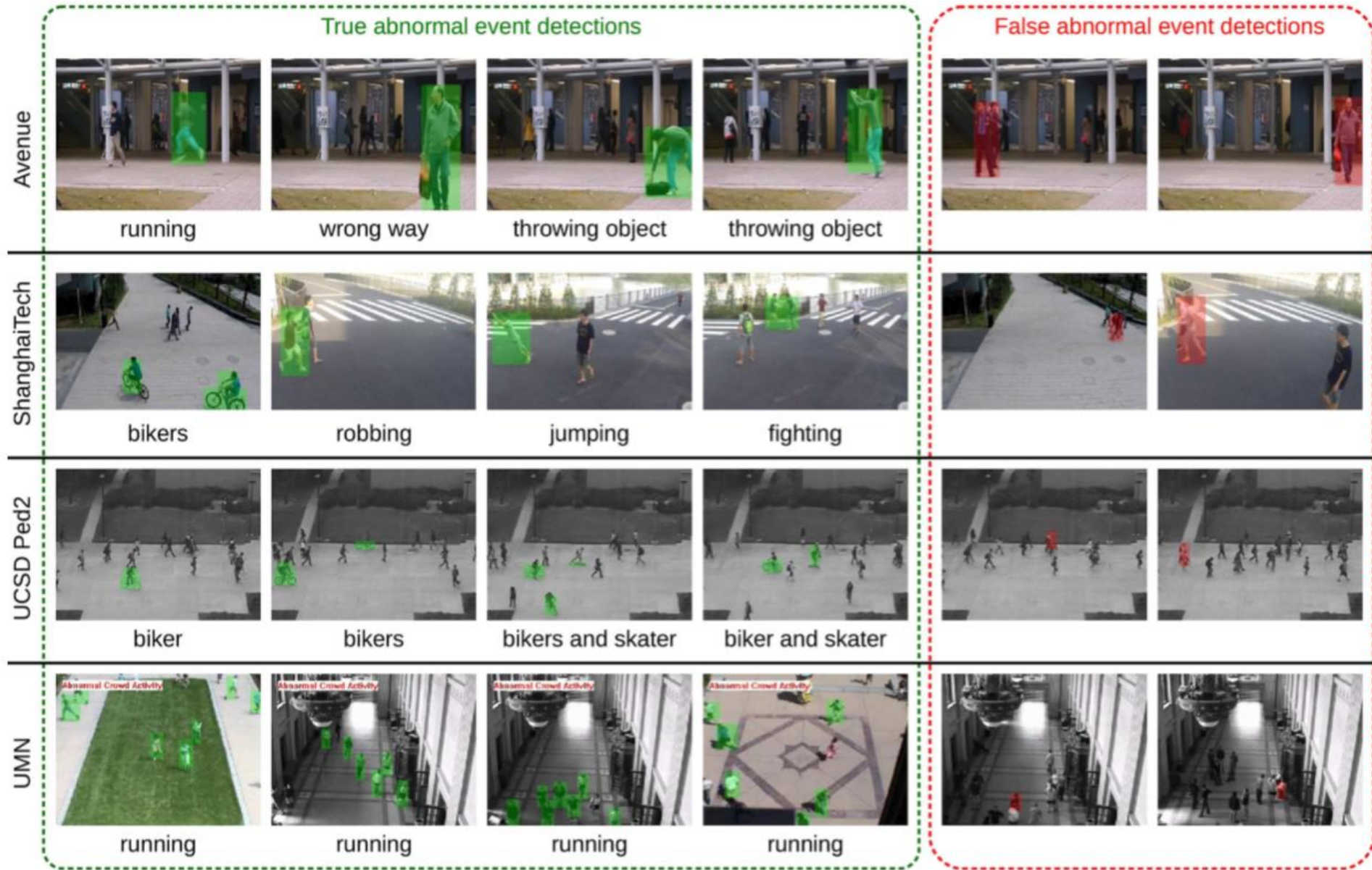


d) Anomaly scores on UMN (scene 2)



# 4. Experiments

- Results



## 5. Conclusion

---

- 이상 탐지를 위한 새로운 모델 제안
  - 1) 객체 중심의 컨볼루션 자동 인코더 학습
  - 2) 이상 탐지를 다중 클래스 문제로 변환하여 해결
- 4가지 데이터셋(Avenue, ShanghaiTech, UCSD, UMN)에서 우수한 결과를 도출함

**Thank you!**