Open DMQA Seminar

Anomaly Detection in Surveillance Video

김경수

Korea University

Data **M**ining & **Q**uality **A**nalytics Lab.

23.07.07

발표자 소개

- 김경수 (Kyungsoo Kim)
  - ✓ 고려대학교 산업경영공학과
  - ✓ Data Mining & Quality Analytics Lab. (김성범 교수님)
  - ✓ M.S Student (2022.03 ~ Present)

- Research Interest
  - ✓ Machine Learning & Deep Learning for tabular data
  - ✓ Anomaly Detection

- Contact
  - ✓ E-mail : kyungsoo@korea.ac.kr

Contents

- **Introduction**
- **Anomaly Detection in Surveillance Video**
- **결론**

Introduction

Introduction

❖ Surveillance Video(감시카메라)?

- Surveillance Video는 상황의 모니터링 및 기록을 위한 CCTV영상
- 단순히 영상을 촬영하여 보여주거나 저장을 하는 용도로 사용

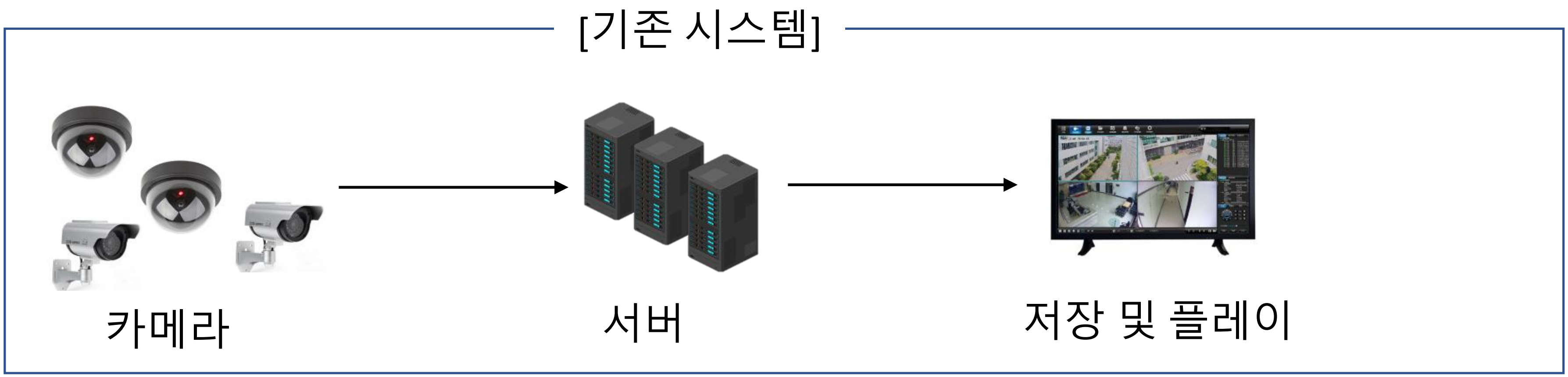
https://www.hani.co.kr/arti/economy/it/1093078.html

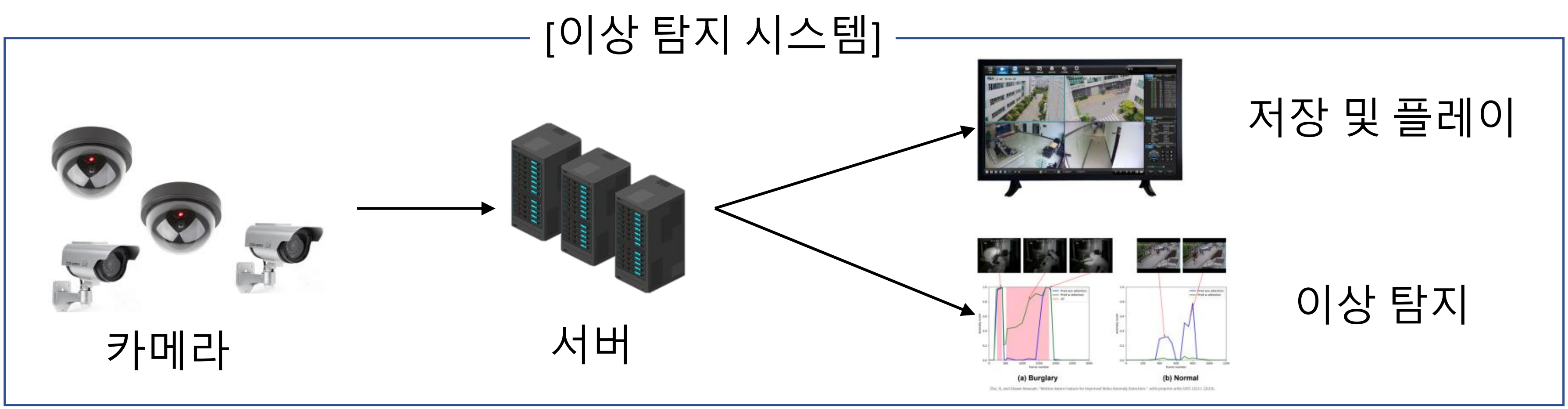
- 서울의 감시카메라 수는 14만 4513대로 단위면적당 카메라 수는 세계 2위 (2023)
- 지속적인 감시카메라의 증가

Introduction

❖ Video anomaly detection의 필요성

- 감시카메라의 수는 증가하고 다수의 카메라 모니터링이 필요
- 실시간 모니터링은 많은 인력과 비용이 필요 → 이상 탐지에 대한 자동화 필요





세미나 주제

Anomaly Detection in Surveillance Video

Data Set

❖ 대표적인 Surveillance Video Dataset

Dataset	Data 내용
1. UCSD	보행자 감지 및 추척을 위해 수집된 비디오 데이터
2. Subway Entrance/Exit	지하철 출입구에서 수집된 비디오 데이터
3. CUHK Avenue	도로와 도시 환경에서 수집된 비디오
4. ShanghaiTech	인구밀도가 높은 환경에서 수집된 비디오 데이터
5. UCF-Crime (Weakly Supervised)	CCTV에서 발생한 범죄관련 데이터
6. Traffic-Train	교통 관련 CCTV에서 수집된 데이터
7. Belleview	도로와 주차장에서 수집된 비디오
8. Street Scene (WACV 2020)	도로 및 도시환경에서 수집된 비디오 데이터
9. IITB-Corridor (WACV 2020)	건물 복도를 촬영한 비디오 데이터
10. XD-Violence (ECCV 2020)	폭력적인 행동 감지를 위한 데이터
11. ADOC (ACCV 2020)	도로에서의 차량들을 촬영한 데이터
12. UBnormal (CVPR 2022)	도로와 교차로에서 발생하는 정상적 비정상적인 교통상황 데이터

Data Set

❖ 대표적인 Video Dataset (UCF-Crime)

- 13가지 범죄 유형을 포함한 데이터 (사고, 절도, 싸움, 폭발, 강도, 총격 등)
- 950개 정상 데이터 950개 이상 데이터

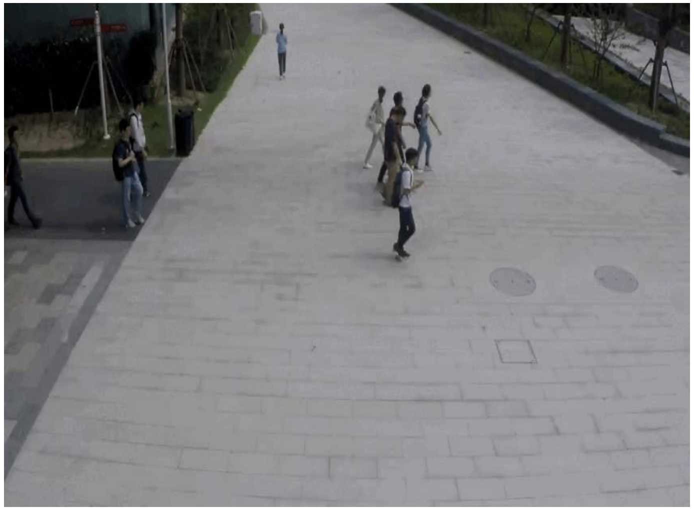
[차사고]

[폭발]

Data Set

❖ **대표적인 Surveillance Video Data set2 (Shanghai tech)**

- 다양한 이상 유형을 포함한 데이터 (추적, 싸움, 이상움직임)
- Campus의 CCTV영상으로 856X480 해상도의 데이터



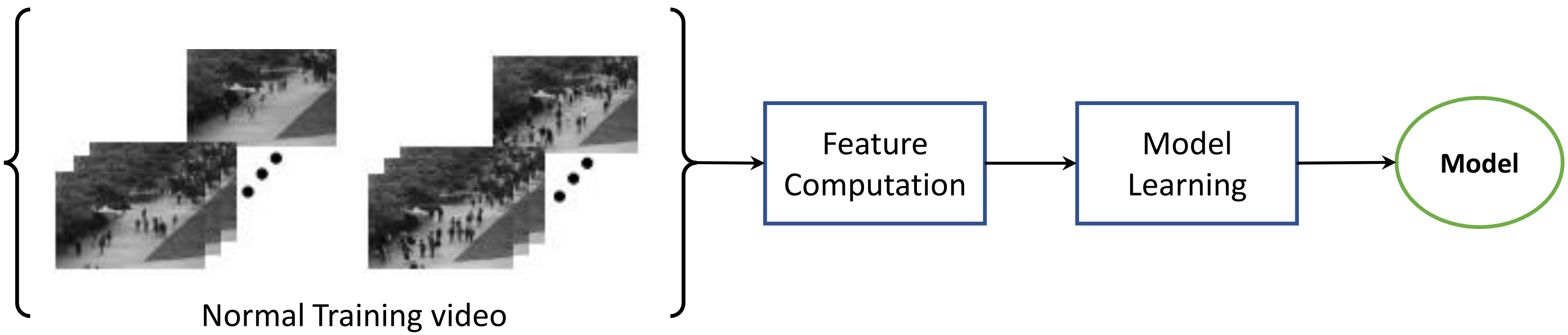
[보행로 자전거1]

[보행로 자전거2]

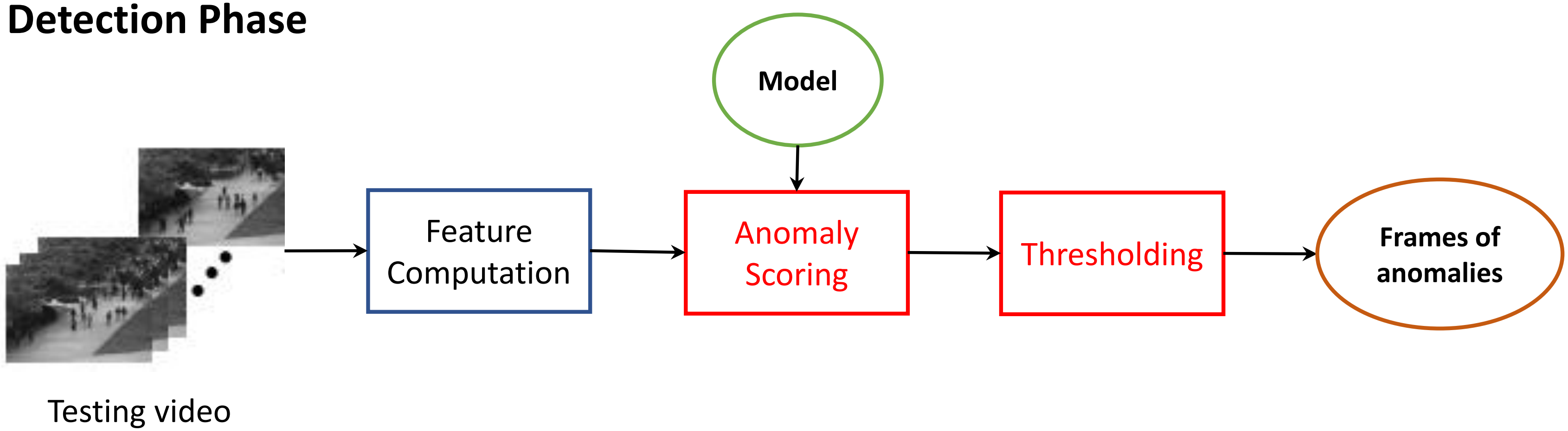
Video Anomaly Detection

❖ Video Anomaly Detection ?

1. Training Phase



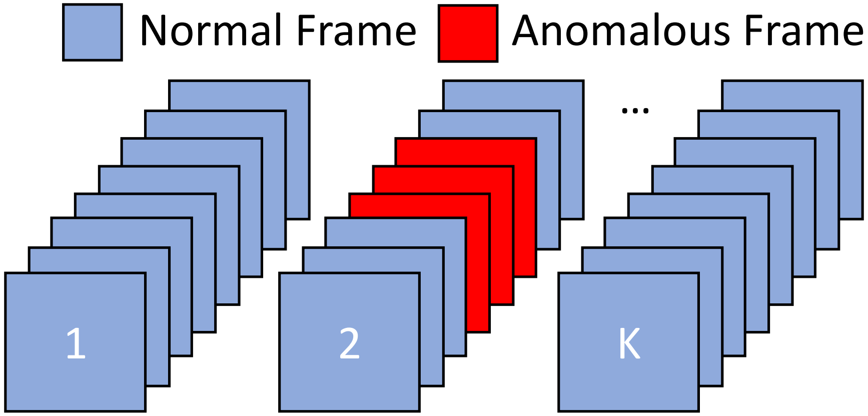
2. Detection Phase



Video Anomaly Detection

❖ Video Anomaly Detection 종류 (Fully Supervised)

- 1개 영상을 1개의 Dataset으로 설정하고 Frame별 Label 설정
- 아래 그림에서 2번째 데이터의 4,5,6 Frame은 Anomalous Frame
- Labeling을 위한 시간이 많이 소요

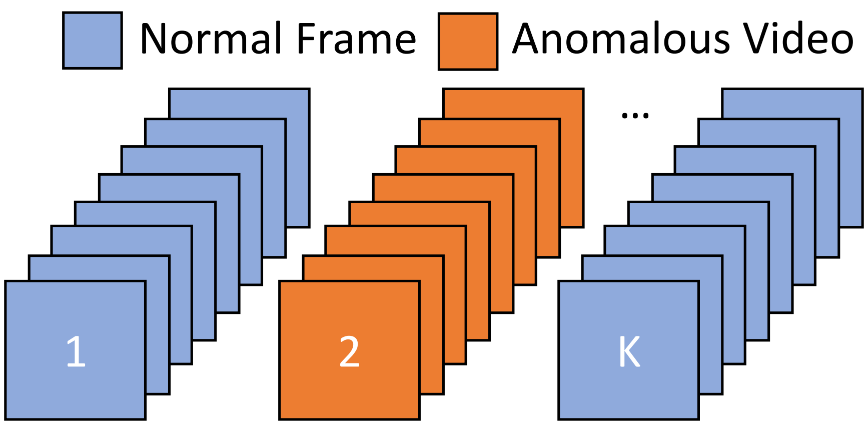


[Fully Supervised]

Video Anomaly Detection

❖ Video Anomaly Detection 종류 (Weakly Supervised)

- 1개 영상을 1개의 Dataset으로 설정
- 해당 영상 중에 Anomalous Data 포함여부 Labeling
- 아래 그림에서 2번째 데이터의 Anomalous Frame을 포함한 영상
- Labeling시 Supervised 방법보다 적은 시간 소요

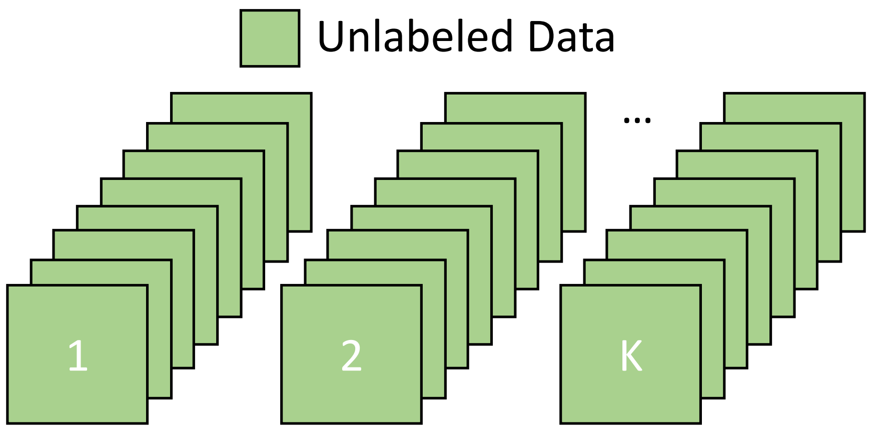


[Weakly Supervised]

Video Anomaly Detection

❖ Video Anomaly Detection 종류 (Unsupervised)

- 1개 영상을 1개의 Dataset으로 설정
- 이상 데이터가 학습에 포함된다는 가정하에 이상 Frame분류
- 정상 적인 Frame이 더 많고 이상적인 현상이 더 적게 발생한다는 가정하에 학습
- Labeling 없이 학습



[Unsupervised]

Video Anomaly Detection

❖ Real-world Anomaly Detection in Surveillance Videos

- Sultani, W., Chen, C., & Shah, M. (2018). Real-world anomaly detection in surveillance videos. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 6479-6488).(1199회 인용)

Real-world Anomaly Detection in Surveillance Videos

Waqas Sultani[1]
[1]Department of Computer Science
Information Technology University, Pakistan

Chen Chen[2], Mubarak Shah[2]
[2]Center for Research in Computer Vision
University of Central Florida, Orlando, FL,USA

waqas5163@gmail.com, chenchen870713@gmail.com, shah@crcv.ucf.edu

Abstract

Surveillance videos are able to capture a variety of realistic anomalies. In this paper, we propose to learn anomalies by exploiting both normal and anomalous videos. To avoid annotating the anomalous segments or clips in training videos, which is very time consuming, we propose to learn anomaly through the deep multiple instance ranking framework by leveraging weakly labeled training videos, i.e. *the training labels (anomalous or normal) are at video-level instead of clip-level. In our approach, we consider normal and anomalous videos as bags and video segments as instances in multiple instance learning (MIL), and automatically learn a deep anomaly ranking model that predicts high anomaly scores for anomalous video segments. Furthermore, we introduce sparsity and temporal smoothness constraints in the ranking loss function to better localize anomaly during training.*

We also introduce a new large-scale first of its kind dataset of 128 hours of videos. It consists of 1900 long and untrimmed real-world surveillance videos, with 13 realistic anomalies such as fighting, road accident, burglary, robbery, etc. as well as normal activities. This dataset can be used for two tasks. First, general anomaly detection considering all anomalies in one group and all normal activities in

etc. to increase public safety. However, the monitoring capability of law enforcement agencies has not kept pace. The result is that there is a glaring deficiency in the utilization of surveillance cameras and an unworkable ratio of cameras to human monitors. One critical task in video surveillance is detecting anomalous events such as traffic accidents, crimes or illegal activities. Generally, anomalous events rarely occur as compared to normal activities. Therefore, to alleviate the waste of labor and time, developing intelligent computer vision algorithms for automatic video anomaly detection is a pressing need. The goal of a practical anomaly detection system is to timely signal an activity that deviates normal patterns and identify the time window of the occurring anomaly. Therefore, anomaly detection can be considered as coarse level video understanding, which filters out anomalies from normal patterns. Once an anomaly is detected, it can further be categorized into one of the specific activities using classification techniques.

A small step towards addressing anomaly detection is to develop algorithms to detect a specific anomalous event, for example violence detector [30] and traffic accident detector [23, 35]. However, it is obvious that such solutions cannot be generalized to detect other anomalous events, therefore they render a limited use in practice.

Video Anomaly Detection

❖ Real-world Anomaly Detection in Surveillance Videos

- 문제점

- 감시카메라를 활용한 이상 탐지에서 모든 이상현상들을 나열할 수 없다.

- 정상과 비정상의 경계가 모호하다.

- 측정된 환경에 따라서 시간대별로 급격하게 변화 → High false alarm rate for normal

- 해결을 위한 제안 방안

- Supervision을 최소화 → weakly labeled training video 사용

Frame1:0 Frame2:1 Frame3:1 Frame4:1

[선행 연구(Frame별 Label)]

Video1:1 Video2:0

[논문1(Video별 Label)]

Video Anomaly Detection

❖ Real-world Anomaly Detection in Surveillance Videos

- 기여점 4가지

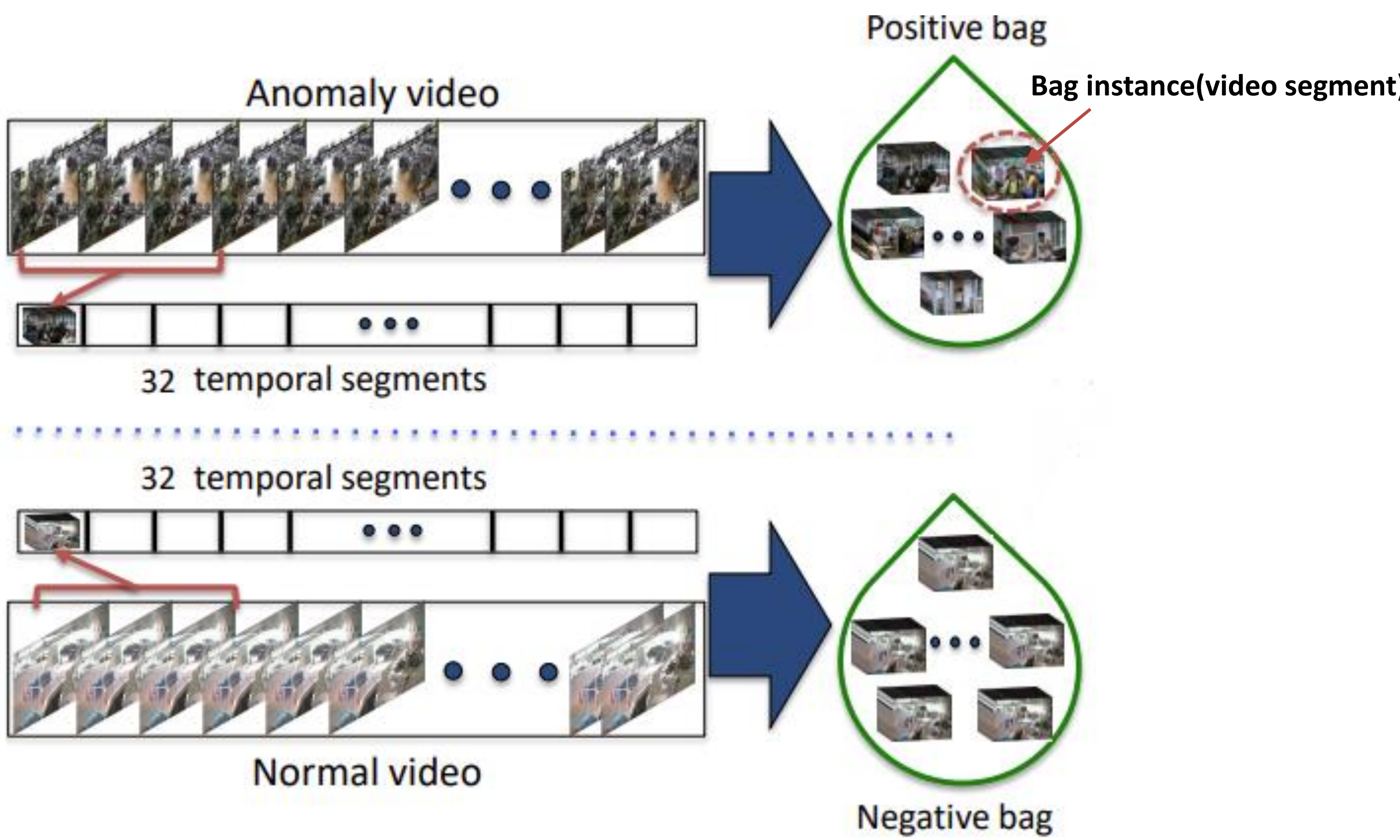
  - MIL(Multiple instance learning)을 사용하여 weakly labeled training video로 이상탐지
  - 1900개의 실제 감시카메라 비디오를 활용하여 13개의 다른 이상 현상 감지
  - SOTA 대비 좋은 성능
  - 가공되지 않은 Video 데이터를 사용한 연구

Video Anomaly Detection

❖ Real-world Anomaly Detection in Surveillance Videos

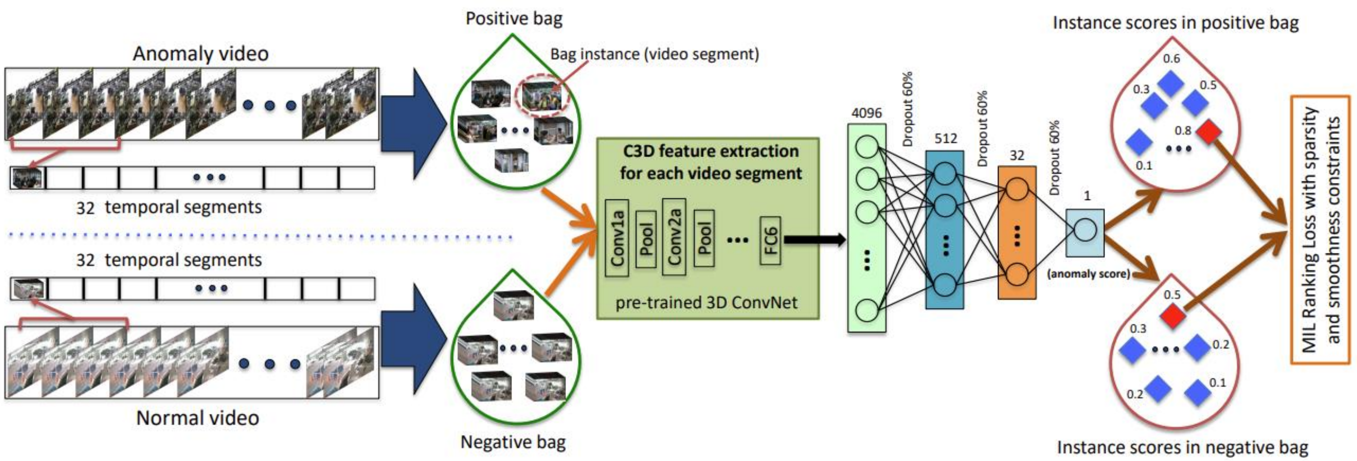
- 제안 방법론 (Multiple instance learning)

- Anomaly video는 Positive bag / Normal video는 negative bag



Video Anomaly Detection

❖ Real-world Anomaly Detection in Surveillance Videos



- 제안 방법론 (Deep MIL Ranking Model)

  segment를 C3D feature extractor에 넣어서 4096개의 feature를 추출
  → 추출한 데이터를 활용하여 segment별 anomaly score를 계산

Positive bag에서 나온 Anomaly score의 Max가 negative bag의 Max보다 높다

$$f(\mathcal{V}_a) > f(\mathcal{V}_n),$$

Frame 단위의 anomaly score 비교

$$\max_{i \in \mathcal{B}_a} f(\mathcal{V}_a^i) > \max_{i \in \mathcal{B}_n} f(\mathcal{V}_n^i),$$

Segment(frame 여러 개 뭉치)의 Score 비교

Video Anomaly Detection

❖ Real-world Anomaly Detection in Surveillance Videos

- 제안 방법론 (Deep MIL Ranking Model)

SVM Loss

$$l(\mathcal{B}_a, \mathcal{B}_n) = \boxed{\max(0, 1 - \max_{i \in \mathcal{B}_a} f(\mathcal{V}_a^i) + \max_{i \in \mathcal{B}_n} f(\mathcal{V}_n^i))}$$

$$+\lambda_1 \overbrace{\sum_i^{(n-1)} (f(\mathcal{V}_a^i) - f(\mathcal{V}_a^{i+1}))^2}^{①} + \lambda_2 \overbrace{\sum_i^{n} f(\mathcal{V}_a^i)}^{②},$$

① Term : Loss function에 Smoothness constrain 추가
→ 이전 Frame과 현재 Frame간 anomaly score의 차이가 작아야 한다.

② Term : Sparsly constrain 추가
→ Anomaly가 normal에 비해서 적게 발생
anomaly가 많이 나오지 않도록 조정하는 constrain (Score 자체를 제약)

Video Anomaly Detection

❖ Real-world Anomaly Detection in Surveillance Videos

- Experiment Result

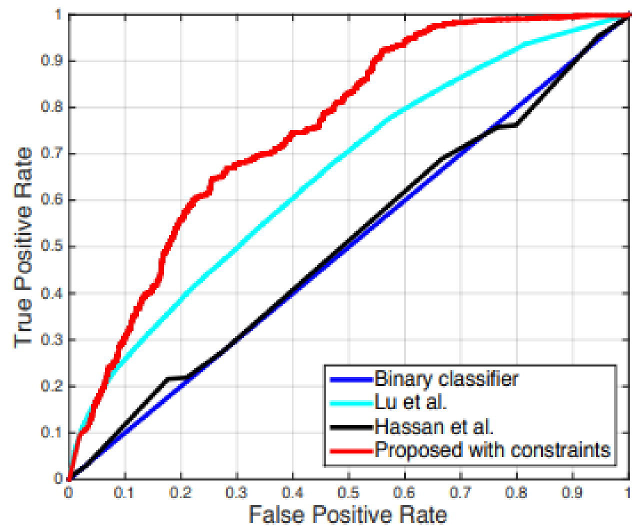


Figure 6. ROC comparison of binary classifier (blue), Lu *et al.* [28] (cyan), Hasan *et al.* [18] (black), proposed method without constraints (magenta) and with constraints (red).

Lu et al 제안 방법 : Dictionary 기반 접근 방식으로 정상동작 학습하고 재구축 오차로 감지

Hassan et al 제안 방법 : Deep Auto Encoder를 활용하여 이상 감지

Video Anomaly Detection

❖ Real-world Anomaly Detection in Surveillance Videos

- Experiment Result

- 제안된 방법론이 다른 방법론 대비 더 좋은 성능
- 첫번째 방법(Lu et al.)은 비정상 비디오에 대한 재구축 오차가 낮은 경우가 있음.
- 두번째 방법(Hasanet al.) 정상 패턴을 잘 학습하지만 새로운 정상 패턴에 대해서 높은 이상 점수 생성

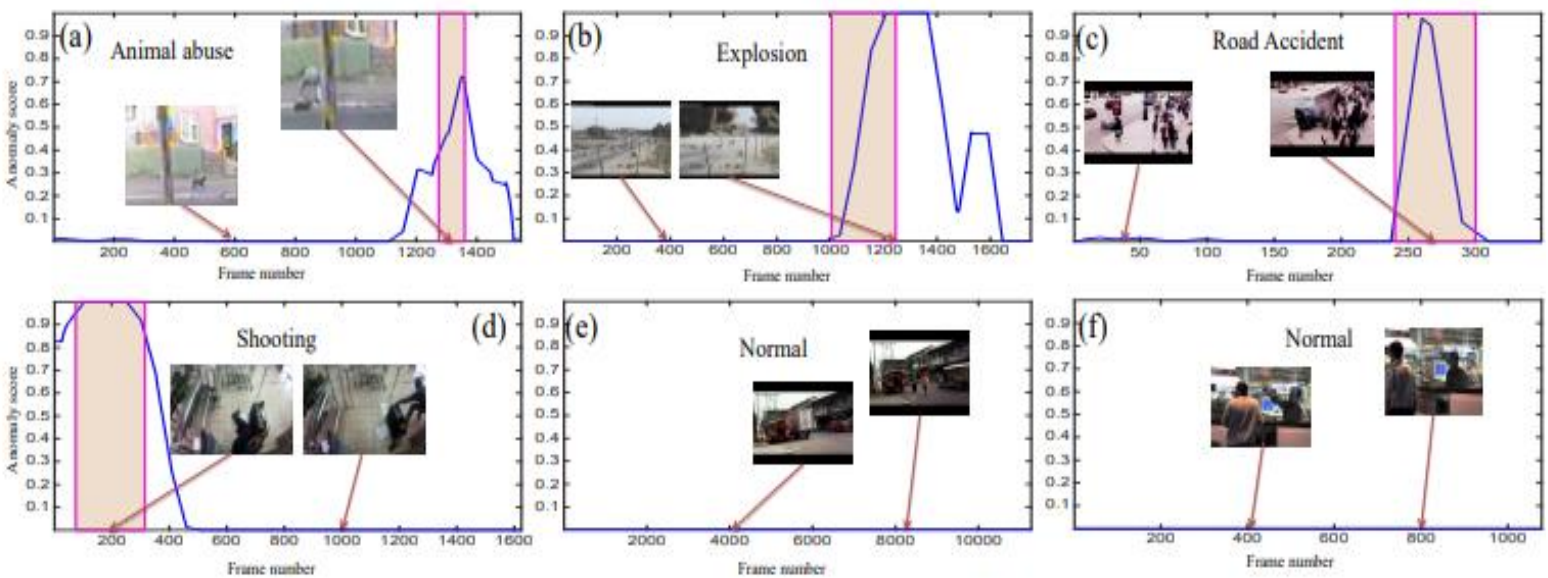
Method	AUC
Binary classifier	50.0
Hasan *et al.* [18]	50.6
Lu *et al.* [28]	65.51
Proposed w/o constraints	74.44
Proposed w constraints	**75.41**

Table 3. AUC comparison of various approaches on our dataset.

Video Anomaly Detection

❖ Real-world Anomaly Detection in Surveillance Videos

- Experiment Result



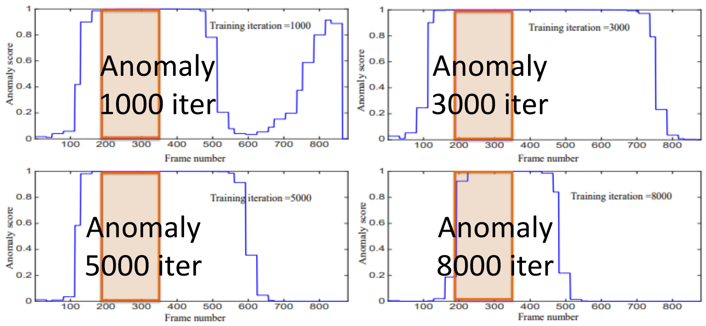
(a) Animal abuse, (b) Explosion, (c) Car accident, (d) Shooting, (e),(f) normal

Video Anomaly Detection

❖ Real-world Anomaly Detection in Surveillance Videos

- Experiment Result

Model training



- 표시된 영역은 Anomaly Frame

- 학습이 진행됨에 따라 이상 판정 정도 향상

- Robust한 이상 감지를 위해서는 이상 및 정상 비디오 모두를 사용 필요

False Alarm

Method	[18]	[28]	**Proposed**
False alarm rate	27.2	3.1	**1.9**

Table 4. False alarm rate comparison on normal testing videos.

- 실제 감시카메라는 대부분 정상적인 상황

- 좋은 모델은 False Alarm rate가 낮아야함

- 정상 Data만 사용한 모델들 대비 좋은 성능

- Loss Function에 제약식 활용

Video Anomaly Detection

❖ Self-Training Multi-Sequence Learning with Transformer for Weakly Supervised Video Anomaly Detection

- Li, S., Liu, F., & Jiao, L. (2022, June). Self-training multi-sequence learning with transformer for weakly supervised video anomaly detection. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 36, No. 2, pp. 1395-1403). (38회 인용)

Self-Training Multi-Sequence Learning with Transformer for Weakly Supervised Video Anomaly Detection

Shuo Li, Fang Liu*, Licheng Jiao

Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education,
International Research Center for Intelligent Perception and Computation,
Joint International Research Laboratory of Intelligent Perception and Computation,
School of Artificial Intelligent, Xidian University, Xi'an, 710071, P.R. China,
alisure@stu.xidian.edu.cn, f63liu@163.com, lchjiao@mail.xidian.edu.cn

Abstract

Weakly supervised Video Anomaly Detection (VAD) using Multi-Instance Learning (MIL) is usually based on the fact that the anomaly score of an abnormal snippet is higher than that of a normal snippet. In the beginning of training, due to the limited accuracy of the model, it is easy to select the wrong abnormal snippet. In order to reduce the probability of selection errors, we first propose a Multi-Sequence Learning (MSL) method and a hinge-based MSL ranking loss that uses a sequence composed of multiple snippets as an optimization unit. We then design a Transformer-based MSL network to learn both video-level anomaly probability and snippet-level anomaly scores. In the inference stage, we propose to use the video-level anomaly probability to suppress the fluctuation of snippet-level anomaly scores. Finally, since VAD needs to predict the snippet-level anomaly scores, by gradually reducing the length of selected sequence, we propose a self-training strategy to gradually refine the anomaly scores. Experimental results show that our method achieves significant improvements on ShanghaiTech, UCF-Crime, and XD-Violence.

Introduction

Hong, and Zheng 2021). Recently, many researchers have focused on weakly supervised VAD (Zhong et al. 2019).

Most weakly supervised VADs are based on Multi-Instance Learning (MIL) (Sultani, Chen, and Shah 2018; Zhu and Newsam 2019; Wan et al. 2020; Tian et al. 2021). MIL-based methods treat a video as a bag, which contains multiple instances. Each instance is a snippet. The bag generated from an abnormal video is called a positive bag, and the bag generated from a normal video is called a negative bag. Since the video-level label indicates whether the video contains anomalies, the positive bag contains at least one abnormal snippet and the negative bag contains no abnormal snippet. MIL-based methods learn instance-level anomaly scores through the bag-level labels (Zhong et al. 2019).

In MIL-based methods, at least one instance of the positive bag contains the anomaly, and any instance of the negative bag does not contain the anomaly (Sultani, Chen, and Shah 2018). Generally, MIL-based methods assume that the instance with the highest anomaly score in the positive bag should rank higher than the instance with the highest anomaly score in the negative bag (Zhu and Newsam 2019). Therefore, the important thing for MIL-based methods is to cor-

Video Anomaly Detection

❖ **Self-Training Multi-Sequence Learning with Transformer for Weakly Supervised Video Anomaly Detection**

- 기존 MIL(Multi-Instance Learning) 방법론의 문제점

  - 정상 Instance가 비정상 Instance로 예측되면 후속 Instance 선택 시 오류 증가

  - 비정상 이벤트는 여러 개의 Instance의 집합이지만 MIL은 고려하지 않음.

- 단점을 보완하기 위해 본 논문에서는 MSL(Multi-Sequence Learning)을 제안

  - Instance 1개를 Loss Function을 만들어서 최적화를 하지 않음.

  - 여러 Instance로 구성된 Sequence를 최적화의 단위로 사용

  - MSL에서는 이상 점수의 합계가 가장 높은 Sequence를 선택 및 학습에 활용

Video Anomaly Detection

❖ Multi-Sequence Learning

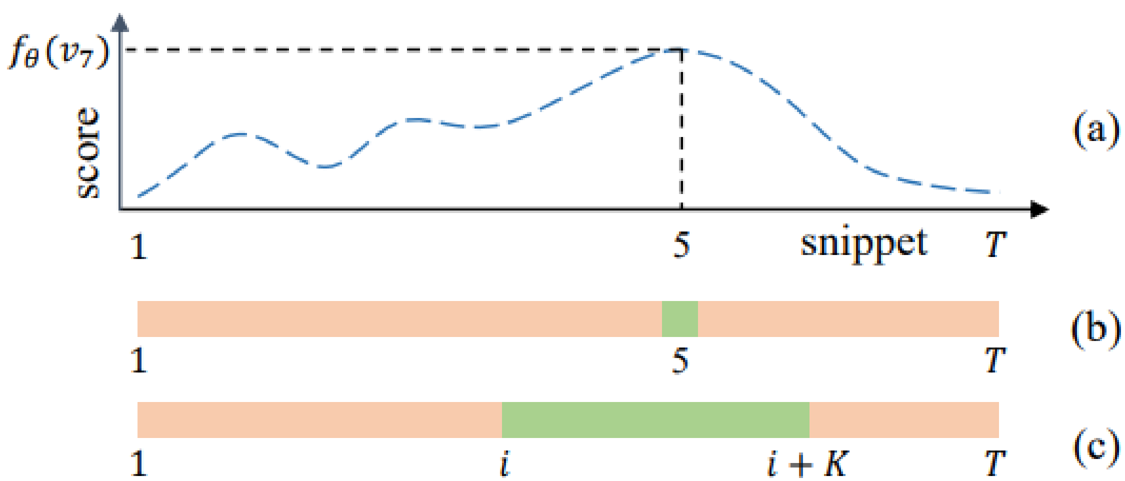
- MIL과 MSL간에 Instance 선택 방법 비교

  - (a) T개의 Snippet(Instance)을 갖는 비디오의 Anomaly Score

  - 5번째 Instance가 가장 큰 Anomaly Score

  - 5번째 Instance를 선택하는 MIL 방법론

  - MSL은 i번째 Instance부터 K개의 연속 Instance로 구성된 Sequence를 선택



Video Anomaly Detection

❖ Multi-Sequence Learning

- MSL 방법론
  - K(Hyperparameter)를 갖고 K개의 연속 Instance를 선택
  - K개 모두 Anomaly Score 계산 → Sequence별 평균 계산
  - Anomaly Score의 평균이 가장 큰 Sequence 선택

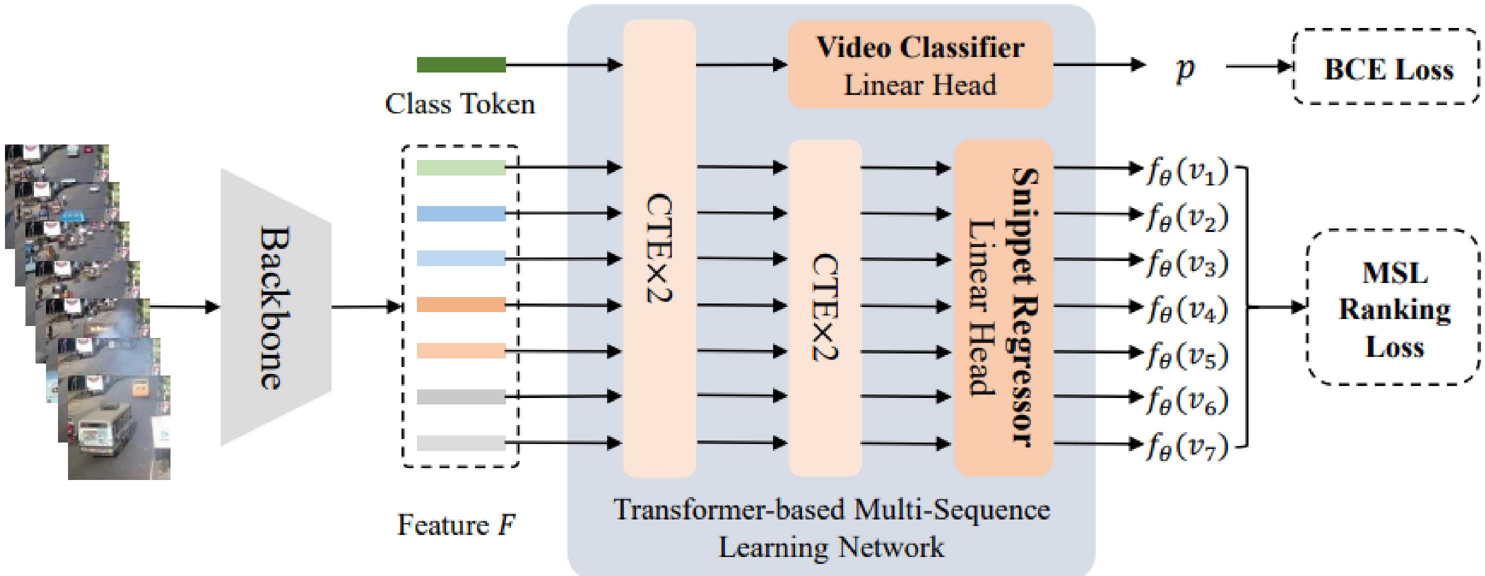
$$S = \{s_i\}_{i=1}^{T-K}, \quad s_i = \frac{1}{K}\sum_{k=0}^{K-1} f_\theta(v_{i+k}), \tag{4}$$

$$\begin{gathered} \max_{s_{a,i} \in S_a} s_{a,i} > \max_{s_{n,i} \in S_n} s_{n,i}, \\ s_{a,i} = \frac{1}{K}\sum_{k=0}^{K-1} f_\theta(a_{i+k}), \quad s_{n,i} = \frac{1}{K}\sum_{k=0}^{K-1} f_\theta(n_{i+k}). \end{gathered} \tag{5}$$

$$\mathcal{L}(\mathcal{B}_a, \mathcal{B}_n) = \max(0, 1 - \max_{s_{a,i} \in S_a} s_{a,i} + \max_{s_{n,i} \in S_n} s_{n,i}). \tag{6}$$

Video Anomaly Detection

❖ Transformer-based MSL Network (Overview)

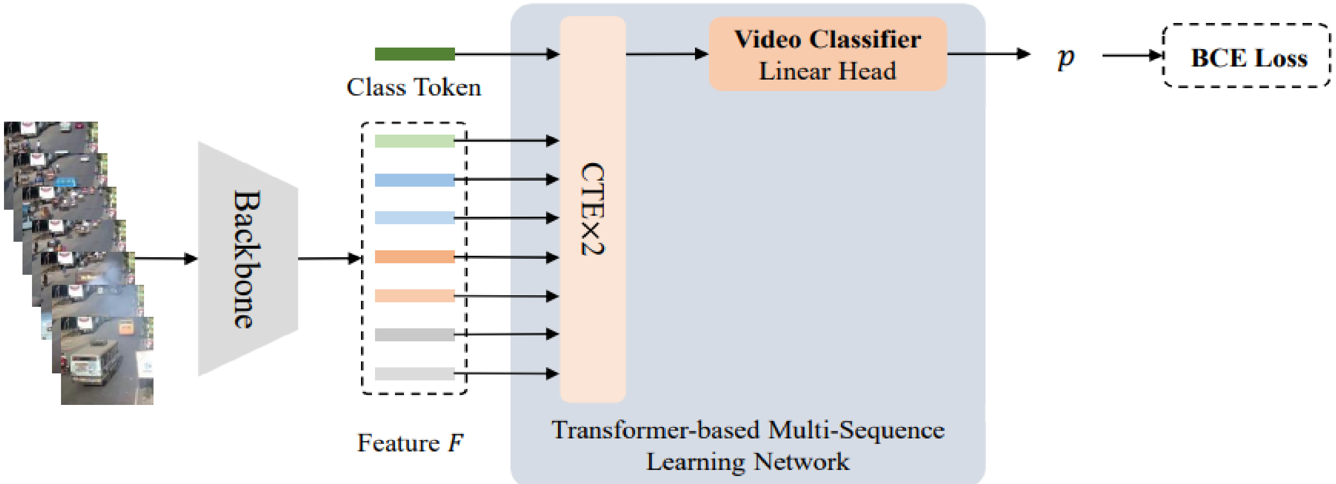


(a) Multi-Sequence Learning Architecture

- Instance들을 Backbone에서 특징을 추출 (C3D, I3D, VideoSwin 사용 + Pre-train W)
- 추출된 Feature와 Class Token을 CTE1,2(convolutional transformer encoder) 입력
- Feature는 Regressor로 입력되어 Instance별 Anomaly Score를 산출 → MSL Ranking L
- Class token은 Classifier를 통과하여 BCE(Binary cross entropy p and Y) Loss 계산

Video Anomaly Detection

❖ Transformer-based MSL Network (Video Classifier)



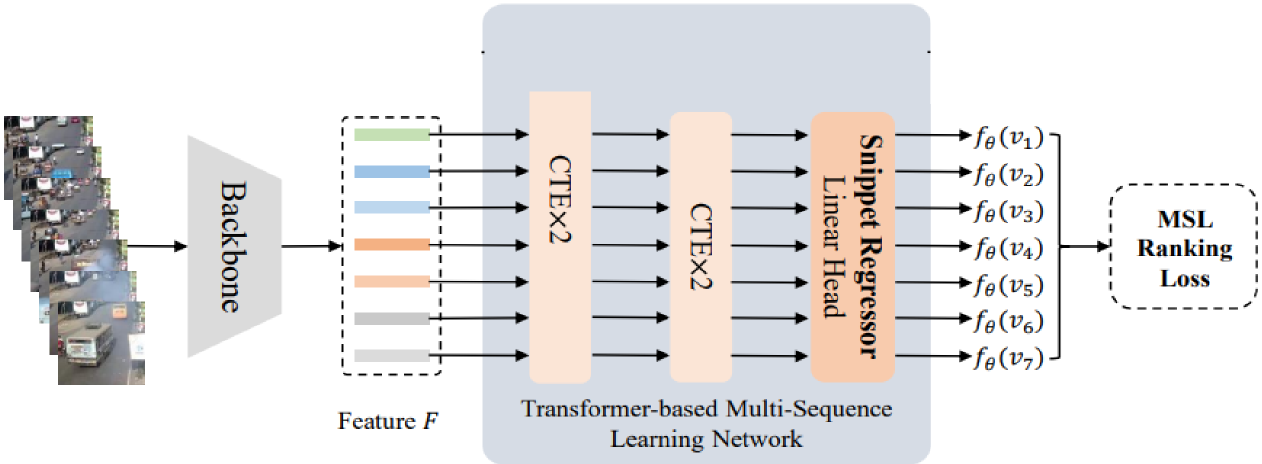
(a) Multi-Sequence Learning Architecture

$$p = \sigma(\mathcal{W}^c \cdot E^c[0]), \quad E^c = CTE_{\times 2}(class\ token||F), \ (7)$$

- Feature과 Class Token은 CTE로 입력되고 새로운 Feature(Ec) 생성
- W는 Video Classifier의 Parameter, p는 video에 Anomalies가 포함될 확률
- Class token은 CTE에서 생성된 특징들을 집계하여 비디오의 전체적인 특징을 만듬(E)

Video Anomaly Detection

❖ Transformer-based MSL Network (Snippet Regressor)



(a) Multi-Sequence Learning Architecture

$$f_\theta(v_i) = \sigma(\mathcal{W}^r \cdot E^r[i]), \quad E^r = CTE_{\times 2}(E^c), \tag{8}$$

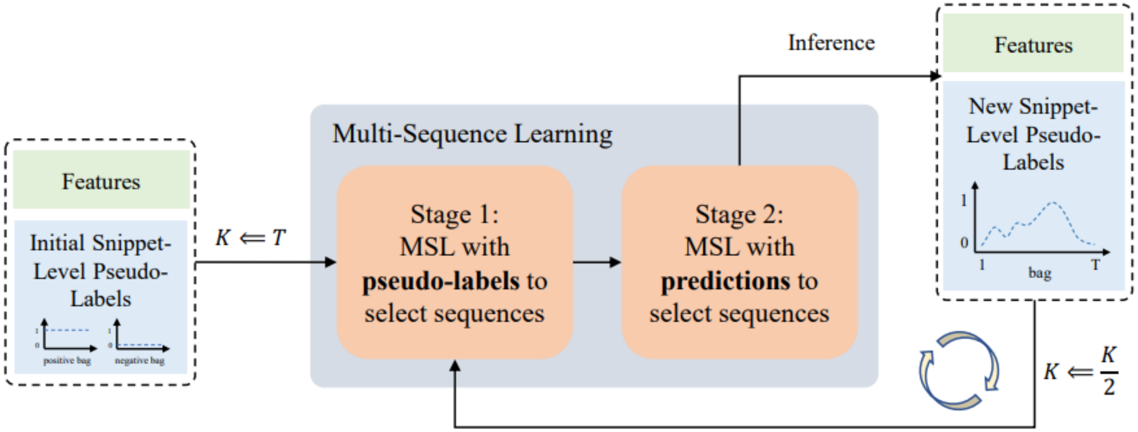
$$\mathcal{L} = \mathcal{L}(\mathcal{B}_a, \mathcal{B}_n) + BCE(p, Y), \tag{9}$$

- MSL Ranking Loss에 BCE(binary cross entropy)
- Anomaly score의 변동을 줄이기 위해 Score correction 방법을 사용

$$\hat{f}_\theta(v_i) = f_\theta(v_i) \times p. \tag{10}$$

Video Anomaly Detection

❖ Self-Training MSL



(b) Self-Training Multi-Sequence Learning Pipeline

- 초기 학습 시 Model의 Anomaly Score 예측 성능은 너무 낮음(잘못된 Sequence 선택)
  → Weakly Supervised 이기때문에 동영상에 부여된 Label을 Pseudo Label로 활용
  → 1. Pseudo Label로 Sequence를 선택 후 학습 $$\mathcal{L}(\mathcal{B}_a, \mathcal{B}_n) = \max(0, 1 - s_{a,i} + s_{n,i}), \quad (11)$$
- 2. 예측한 값으로 시퀀스를 선택하여 학습 + pseudo label update
- 32(T) → 16 → 8 → 4 → 2 → 1 K개수를 작게 가져가면서 미세 튜닝
- 1이된 이후 학습에서도 Self Training 반복하면서 Pseudo Label의 학습

Video Anomaly Detection

❖ Experiment Result

Method	Feature	Crop	AUC(%) ↑
MIL-Rank†	I3D RGB	one	85.33
GCN	C3D-RGB	ten	76.44
GCN	TSN-Flow	ten	84.13
GCN	TSN-RGB	ten	84.44
IBL	I3D-RGB	one	82.50
AR-Net†	C3D RGB	one	85.01
AR-Net	I3D Flow	one	82.32
AR-Net	I3D RGB	one	85.38
AR-Net	I3D-RGB+Flow	one	91.24
CLAWS	C3D-RGB	one	89.67
MIST	C3D-RGB	one	93.13
MIST	I3D-RGB	one	94.83
RTFM	C3D-RGB	ten	91.51
RTFM	I3D-RGB	ten	**97.21**
RTFM*	VideoSwin-RGB	ten	96.76
Ours	C3D-RGB	one	**94.23**
Ours	I3D-RGB	one	**95.45**
Ours	VideoSwin-RGB	one	**96.93**
Ours	C3D-RGB	ten	**94.81**
Ours	I3D-RGB	ten	96.08
Ours	VideoSwin-RGB	ten	**97.32**

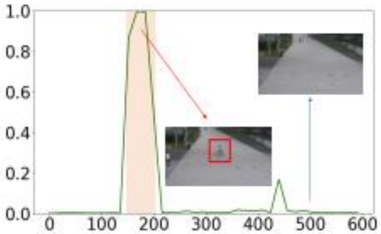
Table 1: Compared with related methods on ShanghaiTech. The methods with † are reported by (Feng, Hong, and Zheng 2021) or (Tian et al. 2021). * indicates we re-train the method. Under the same feature, the highest result is bolded.

Method	Feature	Crop	AUC(%) ↑
MIL-Rank	C3D RGB	one	75.41
MIL-Rank†	I3D RGB	one	77.92
Motion-Aware	PWC-Flow	one	79.00
GCN	C3D-RGB	ten	81.08
GCN	TSN-Flow	ten	78.08
GCN	TSN-RGB	ten	82.12
IBL	C3D-RGB	one	78.66
CLAWS	C3D-RGB	ten	83.03
MIST	C3D-RGB	one	81.40
MIST	I3D-RGB	one	82.30
RTFM	C3D-RGB	ten	**83.28**
RTFM	I3D-RGB	ten	84.03
RTFM*	VideoSwin-RGB	one	83.31
Ours	C3D-RGB	one	82.85
Ours	I3D-RGB	one	**85.30**
Ours	VideoSwin-RGB	one	**85.62**

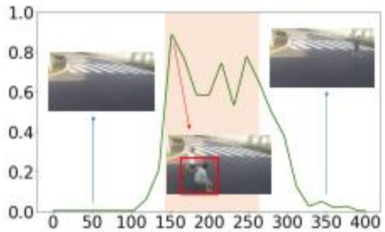
Table 2: Compared with other methods on UCF-Crime. The method with † is reported by (Tian et al. 2021). * indicates we re-train the method. Bold represents the best results.

Video Anomaly Detection

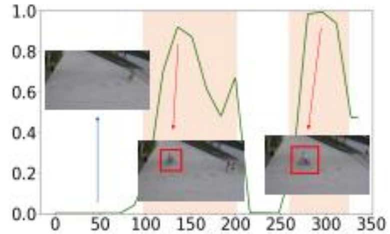
❖ Experiment Result



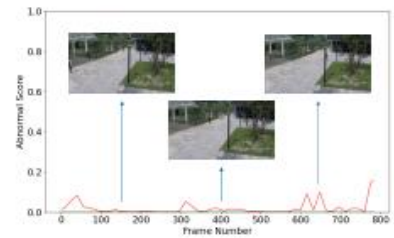
(a) Abnormal Video *01_0025*



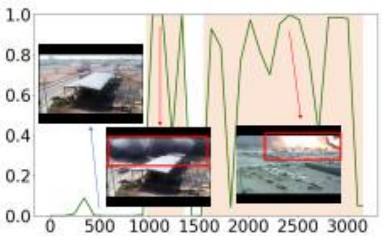
(b) Abnormal Video *03_0032*



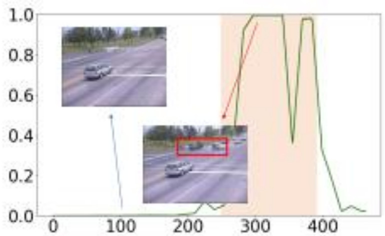
(c) Abnormal Video *01_0051*



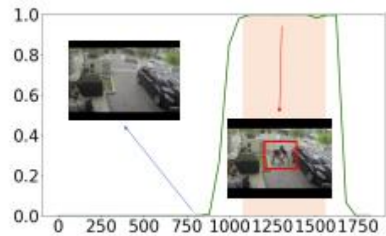
(d) Normal Video *08_045*



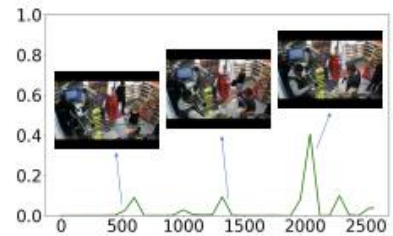
(e) Explosion *033_x264*



(f) RoadAccidents *012_x264*



(g) Robbery *102_x264*



(h) Normal Video *894_x264*

Basic Layer	ShanghaiTech	UCF-Crime
Transformer	96.51	85.41
CTE	**96.93** (**+0.42**)	**85.62** (**+0.21**)

Table 4: Compared with Transformer (Dosovitskiy et al. 2021), AUC(%) improvement brought by CTE on the ShanghaiTech and UCF-Crime datasets.

Score correction	ShanghaiTech	UCF-Crime
×	95.98	84.94
✓	**96.93** (**+0.95**)	**85.62** (**+0.68**)

Table 5: Performance improvement brought by the score correction method in the inference stage measured by AUC(%) on the ShanghaiTech and UCF-Crime datasets.

Conclusion

❖ 결론

- 이번 세미나의 주제는 Anomaly Detection in surveillance video
- Anomaly Detection에 대한 필요성
- Surveillance Video가 무엇이고 어떤 Dataset이 있는지 설명
- Video Anomaly detection 분류
- 논문 리뷰1

  - Real-world Anomaly Detection in Surveillance Videos 2018
- 논문 리뷰2

  - Self-Training Multi-Sequence Learning with Transformer for Weakly Supervised Video Anomaly Detection 2022

참고자료

Ramachandra, B., Jones, M. J., & Vatsavai, R. R. (2020). A survey of single-scene video anomaly detection. *IEEE transactions on pattern analysis and machine intelligence*, *44*(5), 2293-2312.

Suarez, J. J. P., & Naval Jr, P. C. (2020). A survey on deep learning techniques for video anomaly detection. arXiv preprint arXiv:2009.14146.

Zaheer, M. Z., Mahmood, A., Khan, M. H., Segu, M., Yu, F., & Lee, S. I. (2022). Generative cooperative learning for unsupervised video anomaly detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 14744-14754).

Sultani, W., Chen, C., & Shah, M. (2018). Real-world anomaly detection in surveillance videos. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 6479-6488).

Li, S., Liu, F., & Jiao, L. (2022, June). Self-training multi-sequence learning with transformer for weakly supervised video anomaly detection. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 36, No. 2, pp. 1395-1403).

감사합니다