# Real-world Anomaly Detection in Surveillance Videos

刘洋

工程与应用技术研究院

计算机应用技术 2020 级

20110860015@fudan.edu.cn

2020 年 10 月 16 日

# Contents

1. Introduction to Anomaly Detection
   - Definition
   - Applications
   - Classification
2. Real-world Anomaly Detection in Surveillance Videos
   - Intoduction
   - Multiple Instance Learning
   - Deep MIL Ranking Model
   - Overall Framework
   - Results
3. References

# **Section 1:** Introduction to Anomaly Detection[1]

[1] R. Chalapathy and S. Chawla, "Deep learning for anomaly detection: A survey," *arXivpreprint arXiv:1901.03407*, 2019.

## Anomaly Dedection and Novelty Detection

**Anomaly Detection:** Determining instances stand out as being dissimilar to all others.

**Novelty Detection:** Identification of a novel or unobserved patterns in the data [1].

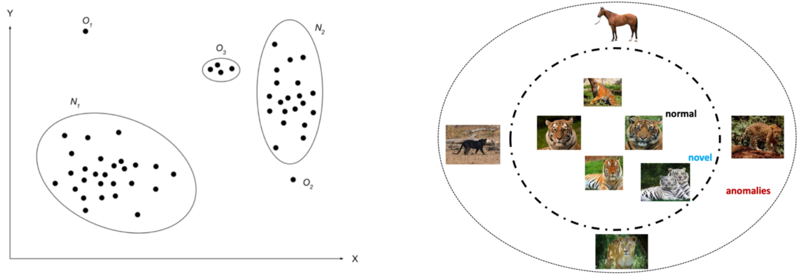
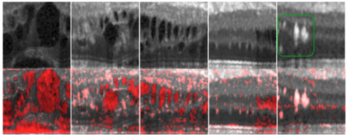


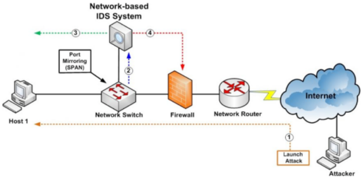

图 1: Illustration of anomaly detection and novelty detection

# Applications of Anomaly Detection

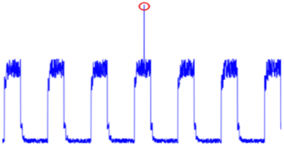
(a) Illegal Traffic Flow detection



(b) Detecting Retinal Damage




(c) Cyber-Network Intrusion detection



(d) Internet Of Things (IoT) Big-Data Anomaly detection

图 2: Applications of anomaly detection technique

# Classification of Models

1. **Unspervised**
   One-class NN/SVM
   Auto-encoder
   Generative Network
2. **Semi-supervised**
   Reinforcement Learning
3. **Weakly-supervised**
   Multi-instance learning
4. **Hybrid**
   Extractor + classifier

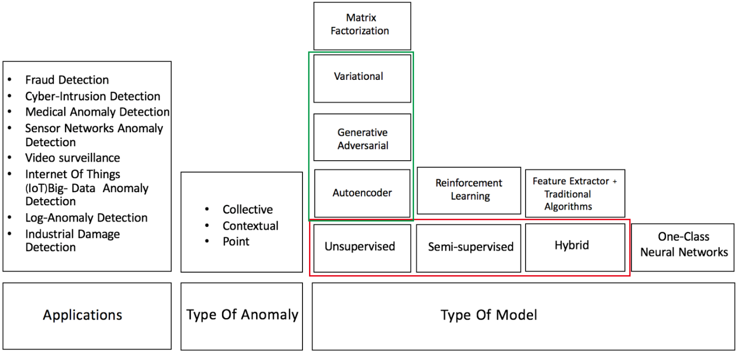



图 3: Key components in DL-based AD models

# **Section 2:** Real-world Anomaly Detection in Surveillance Videos[2]

[2] W. Sultani, C. Chen, and M. Shah, "Real-world anomaly detection in surveillance videos," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018,pp. 6479–6488.

# Introduction

Assumption of other approaches: *Any pattern that deviates from the learned normal patterns would be considered as an anomaly* [2].

- Difficult to define a normal event.
- Ambiguous boundary.
- Behaviour nature changed with condition.


图 4: Examples of different anomalies in real-world

**Video-level labels:** A video is normal or contains anomaly somewhere, but do not know where.

**Clip-level lables**: Uneasy to acquire.

*Considering normal and anomalous videos as bags and video segments as instances.*

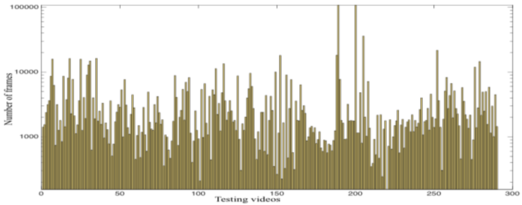



图 5: Distribution of video frames

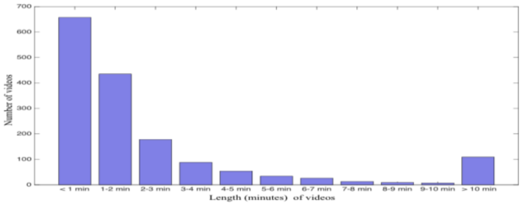



图 6: Distribution of videos length

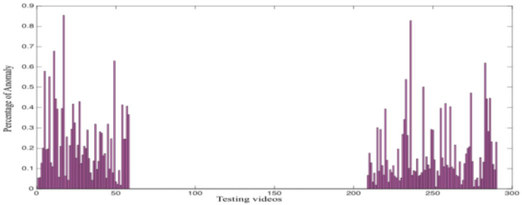



图 7: Percentage of anomaly in each video

# Multiple Instance Learning

**Optimization function of SVM:**

$$\min_{\mathbf{w}} \frac{1}{k} \sum_{i=1}^{k} \overbrace{\max\left(0, 1 - y_i(\mathbf{w} \cdot \phi(x) - b)\right)}^{1} + \frac{1}{2}\|\mathbf{w}\|^2 \qquad (1)$$

*where $\phi x$ denotes feature representation of an image patch or a video segment, $\mathbf{w}$ is the classifier to be learned.*

**Multiple Instance Learning**:

Positive bag$\mathcal{B}_a$: a video with anomalies containing temporal segments (instances) $\left(p^1, p^2, \ldots, p^m\right)$.

Negative bag$\mathcal{B}_n$: a video without anomalies, $\left(n^1, n^2, \ldots, n^m\right)$.

$$\min_{\mathbf{w}} \frac{1}{z} \sum_{j=1}^{z} \max\left(0, 1 - Y_{\mathcal{B}_j}\left(\max_{i \in \mathcal{B}_j}\left(\mathbf{w}.\phi\left(x_i\right)\right) - b\right)\right) + \frac{1}{2}\|\mathbf{w}\|^2 \qquad (2)$$

# Deep MIL Ranking Model

Encouraging high scores for anomalous video segments as compared to normal segments:

$$f(\mathcal{V}_a) > f(\mathcal{V}_n) \tag{3}$$

*where $\mathcal{V}_a$, $\mathcal{V}_b$ represent anomalous and normal video segments, $f(\mathcal{V}_a)$, $f(\mathcal{V}_n)$ represent the corresponding predicted anomaly scores ranging from 0 to 1.*

**Multiple instance ranking objective function:**

$$\max_{i \in \mathcal{B}_a} f(\mathcal{V}_a^i) > \max_{i \in \mathcal{B}_n} f(\mathcal{V}_n^i) \tag{4}$$

**Ranking loss:**

$$l(\mathcal{B}_a, \mathcal{B}_n) = \max\left(0, 1 - \max_{i \in \mathcal{B}_a} f(\mathcal{V}_a^i) + \max_{i \in \mathcal{B}_n} f(\mathcal{V}_n^i)\right) \tag{5}$$

# Deep MIL Ranking Mode

**Loss function with parsity and smoothness constraints:**

$$l(\mathcal{B}_a, \mathcal{B}_n) = \max\left(0, 1 - \max_{i \in \mathcal{B}_a} f(\mathcal{V}_a^i) + \max_{i \in \mathcal{B}_n} f(\mathcal{V}_n^i)\right) + \lambda_1 \sum_{1}^{n-1} \left(f(\mathcal{V}_a^i) - f(\mathcal{V}_a^{i+1})\right)^2 + \lambda_2 \sum_{i}^{n} f(\mathcal{V}_a^i) \tag{6}$$

**Final loss function:**

$$\mathcal{L}(\mathcal{W}) = l(\mathcal{B}_a, \mathcal{B}_n) + \lambda_3 \|\mathcal{W}\|_F \tag{7}$$

- In real-world scenarios, anomaly often occurs only for a short time.only a few segments may contain the anomaly.
- Since the video is a sequence of segments, the anomaly score should vary smoothly between video segments.

# Model Framework

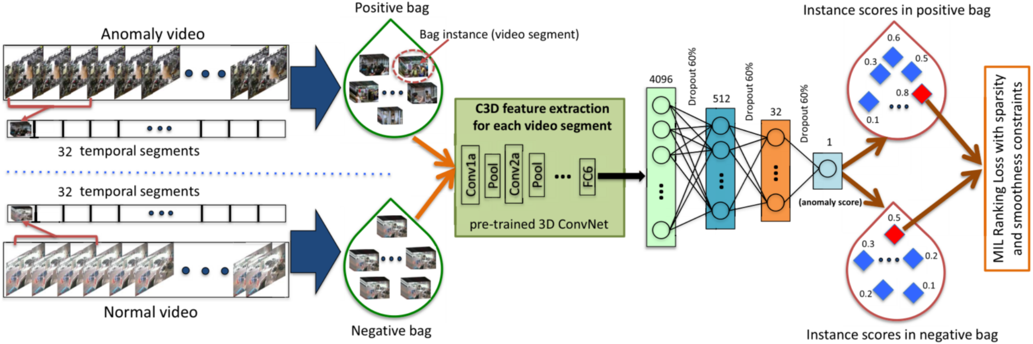



图 8: Flow diagram of the proposed anomaly detection approach

# Model Framework

1. re-size each video frame to 240 × 320 pixels and fix the frame rate to 30 fps.
2. compute C3D features for every 16-frame video clip followed by $L_2$ normalization.
3. take the average of all 16-frame clip features as features of a segment.
4. input these features (4096D) to a 3-layer FC neural network.
   - 4096 units, 60% dropout, relu
   - 512 units, 60% dropout, relu
   - 32 units, 60% dropout, sigmoid

Using adagrad optimizer with learning rate $\alpha = 0.001$, set $\lambda_1 = \lambda_2 = 8 \times 10^{-5}, \lambda_3 = 0.01$.

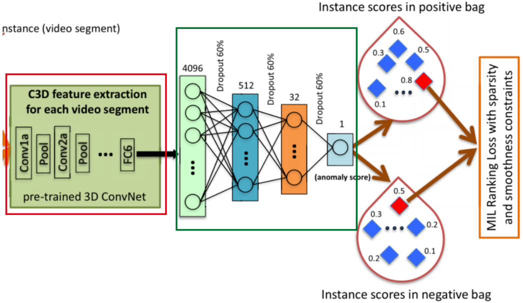



图 9: Deep MIL ranking model

# Results

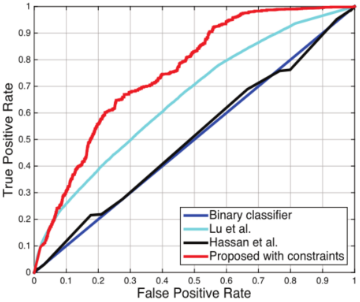



图 10: ROC comparison

| 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** |

图 11: AUC comparison

# Results

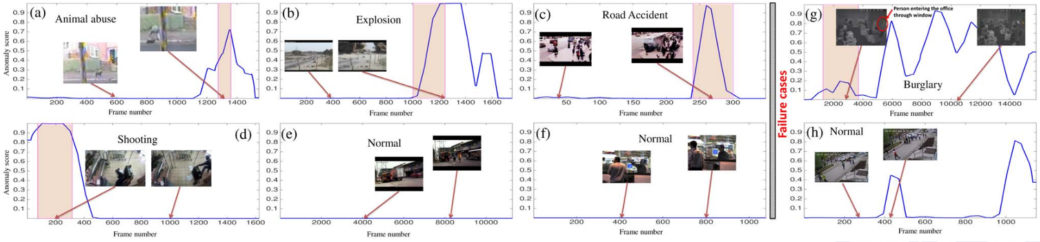



图 12: Qualitative results on testing videos

- (a), (b), (c) and (d) show videos containing animal abuse (beating a dog), explosion, road accident and shooting, respectively.
- (e) and (f) show normal videos with no anomaly.
- (g) and (h) present two failure cases of our anomaly detection method.

# References

[1] R. Chalapathy and S. Chawla, "Deep learning for anomaly detection: A survey," *arXiv preprint arXiv:1901.03407*, 2019.

[2] W. Sultani, C. Chen, and M. Shah, "Real-world anomaly detection in surveillance videos," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 6479–6488.

# Thank you for listening

## Please feel free to ask questions