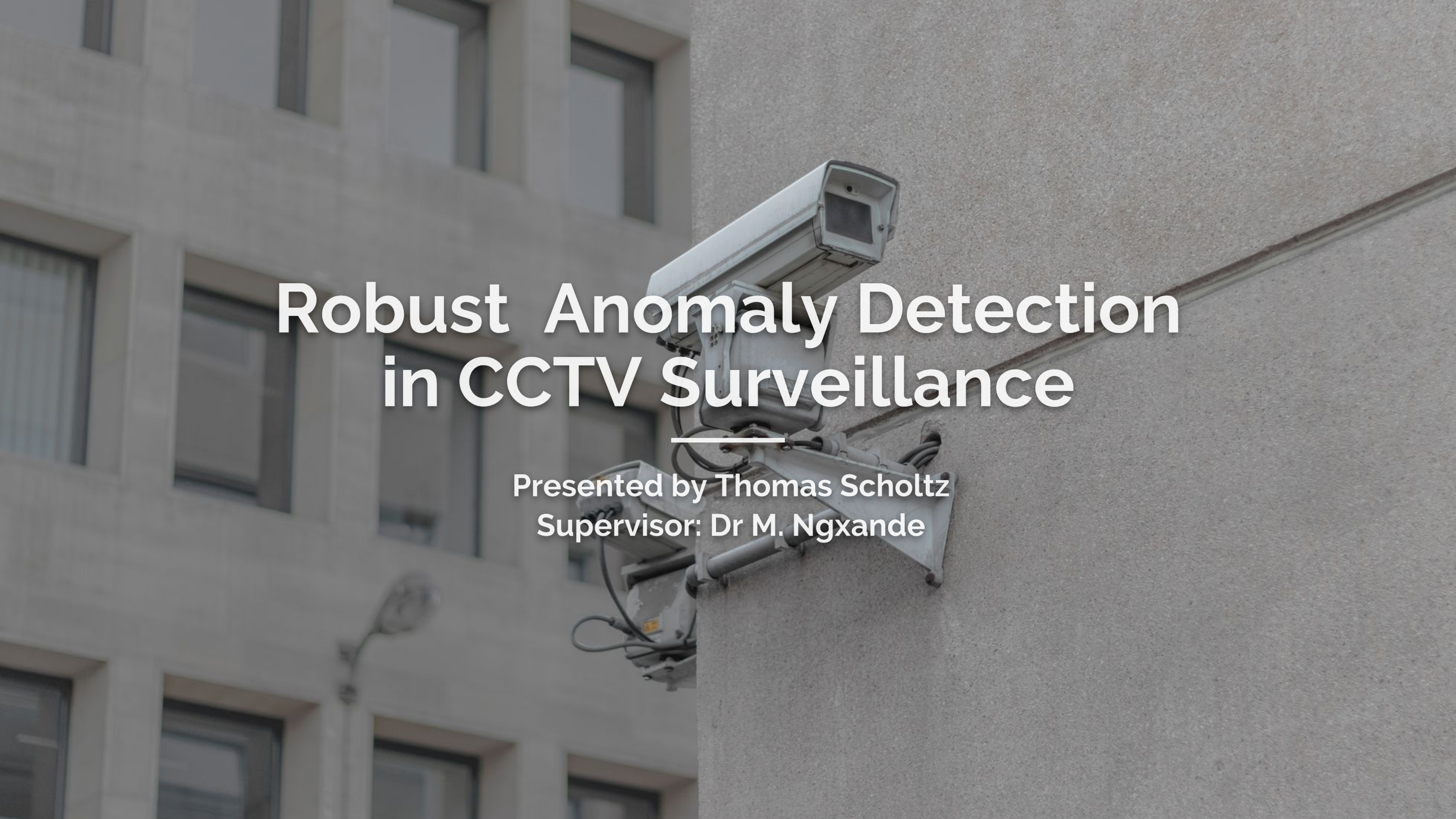


Robust Anomaly Detection in CCTV Surveillance

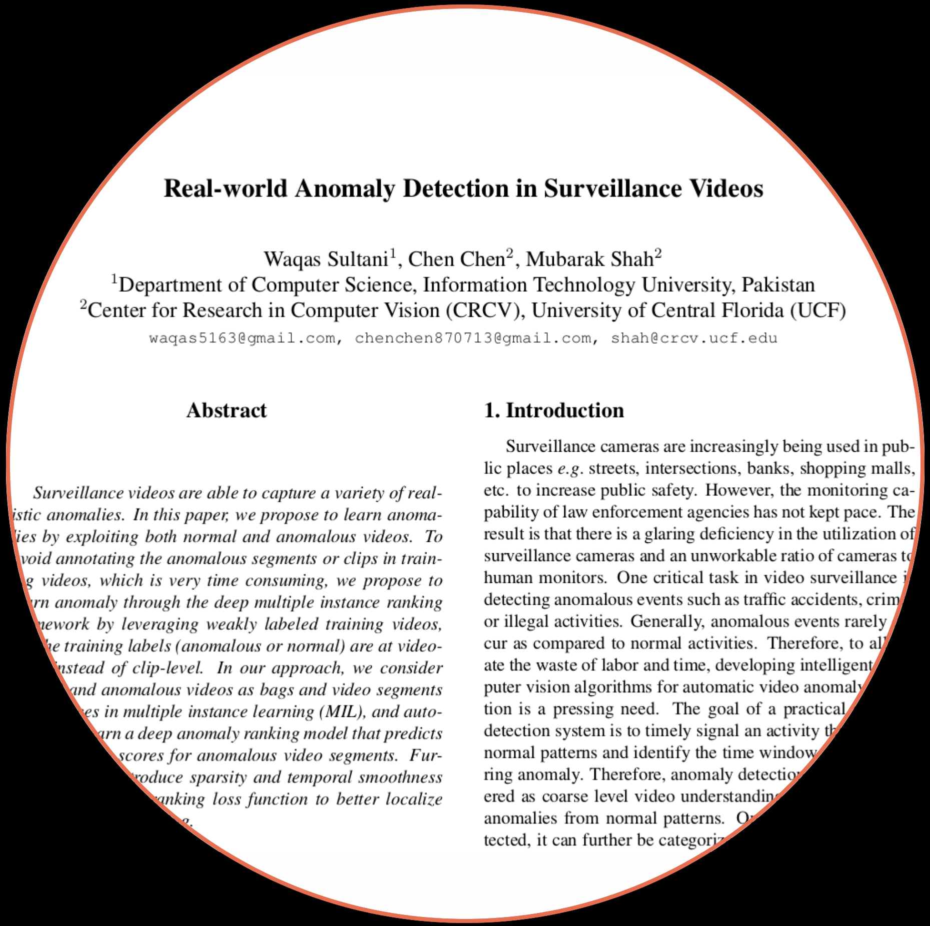
Presented by Thomas Scholtz
Supervisor: Dr M. Ngxande

Outline

Real-world Anomaly Detection

- Perform computations on CCTV surveillance
- **Automate** the role of a human monitor
- **Objective**: Maximize overall anomaly detection ability
- **Sub-objective**: Investigate a consensus framework

Outline



Real-world Anomaly Detection in Surveillance Videos

Waqas Sultani[1], Chen Chen[2], Mubarak Shah[2]
[1]Department of Computer Science, Information Technology University, Pakistan
[2]Center for Research in Computer Vision (CRCV), University of Central Florida (UCF)
waqas5163@gmail.com, chenchen870713@gmail.com, shah@crcv.ucf.edu

Abstract

*Surveillance videos are able to capture a variety of real-
istic anomalies. In this paper, we propose to learn anoma-
lies by exploiting both normal and anomalous videos. To
void annotating the anomalous segments or clips in train-
g videos, which is very time consuming, we propose to
rn anomaly through the deep multiple instance ranking
nework by leveraging weakly labeled training videos,
he training labels (anomalous or normal) are at video-
nstead of clip-level. In our approach, we consider
and anomalous videos as bags and video segments
es in multiple instance learning (MIL), and auto-
arn a deep anomaly ranking model that predicts
scores for anomalous video segments. Fur-
roduce sparsity and temporal smoothness
anking loss function to better localize*

1. Introduction

Surveillance cameras are increasingly being used in pub-
lic places *e.g.* streets, intersections, banks, shopping malls,
etc. to increase public safety. However, the monitoring ca-
pability 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 i
detecting anomalous events such as traffic accidents, crim
or illegal activities. Generally, anomalous events rarely
cur as compared to normal activities. Therefore, to al
ate the waste of labor and time, developing intelligent
puter vision algorithms for automatic video anomaly
tion is a pressing need. The goal of a practical
detection system is to timely signal an activity th
normal patterns and identify the time window
ring anomaly. Therefore, anomaly detectio
ered as coarse level video understandin
anomalies from normal patterns. O
tected, it can further be categoriz

Base Model

- Learns high-level context
- Focus on deterrence of false positives
- Multiple improvements implemented

Outline

Consensus Framework

- **Three** score profiles mapped to **one**
- **Base model (1/3)**
- **CRAFT (2/3)** - Consecutive frame construction with RAFT optical flow estimation
- **LKKM (3/3)** - Lukas Kanade K-Means pattern deviation

Outline

Findings

- Each of the three approaches are useful in isolation
- Deep learning approaches are indispensable to the development of state-of-the-art anomaly detection
- Score combination is surprisingly non-trivial

Problem Statement

Quantify the degree of abnormality in video

- **input:** untrimmed CCTV video
- *<assign anomaly scores along temporal axis>*
- **output**: score profile
- emphasis on a **highly generalizable solution**

Problem statement is inherently vague

Is one situation more anomalous than the other?

Video

Video

An occasional bicycle passes through a sidewalk of pedestrians

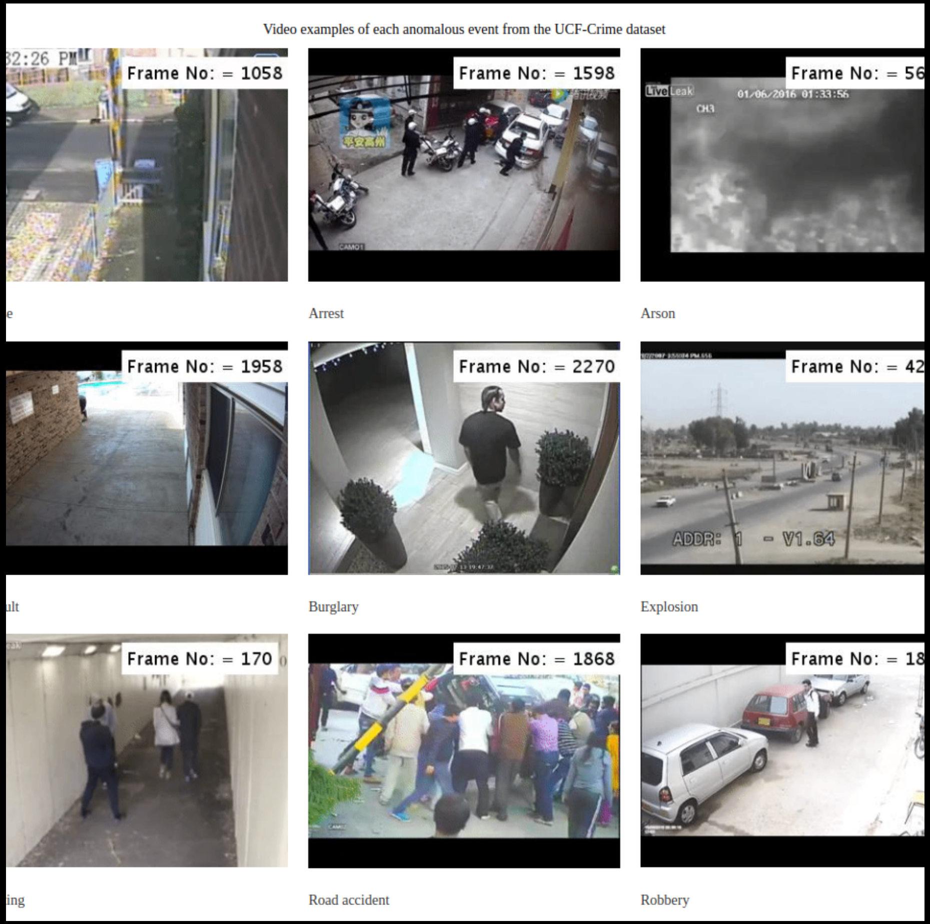
With human context: normal

Without: anomalous appearance and motion

A road accident

Appearance, motion and context indicative of anomaly

Dataset



UCF Crime Dataset [2]

- Abuse
- Arrest
- Arson
- Assault
- Burglary
- Explosion
- Fighting
- Road Accidents
- Robbery
- Shooting
- Shoplifting
- Stealing
- Vandalism

Existing Approaches

- **Abundant supply of papers on the topic**

- **Common approach:**
  - learn an idea of normal activity
  - does unseen footage conform to this idea?

- **Assumption:**
  - anomaly = rare event
  - therefore not learned by a framework

Existing Approaches: Shortfalls

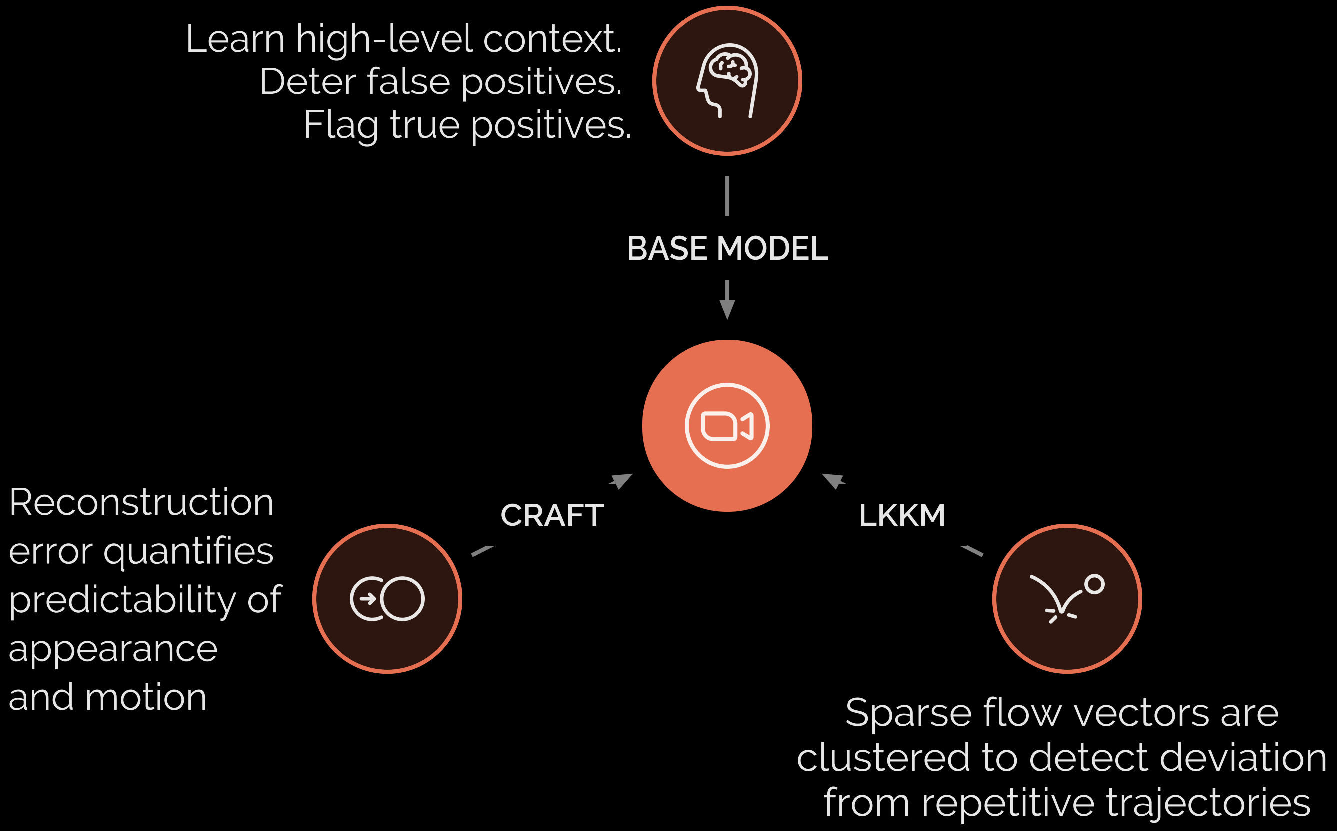
Frameworks struggle to cover the spectrum of anomalies

Difficulty in dealing with false positives

Sub-standard datasets

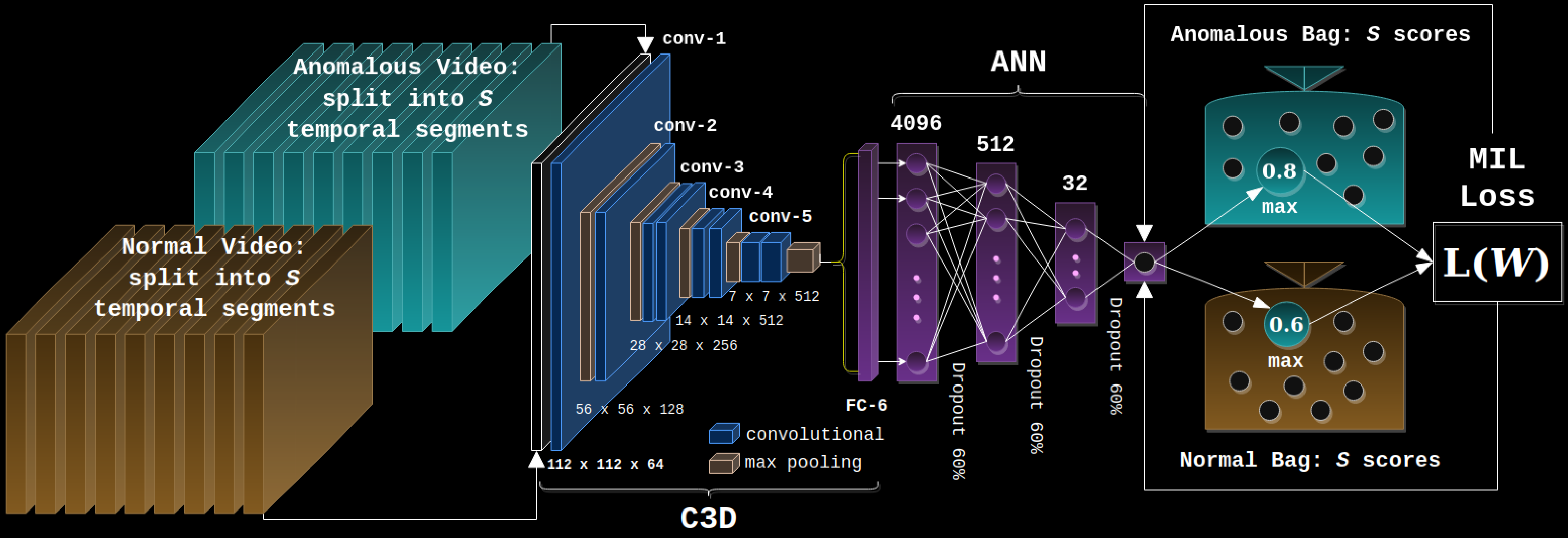
Proposed Framework

Each approach is developed independently



Implementation: Base model

- *Real-world Anomaly Detection in Surveillance Videos* by W. Sultani [1]
- **Weakly-supervised** regression problem
- **C3D** to **ANN**, wrapped in **MIL** ranking model
- Prioritizes a **low false alarm rate**



Implementation: CRAFT

Optical flow refers to estimation of **displacement of pixels** from one frame to the next

CRAFT uses **RAFT** [3] to **CONSTRUCT** future frames

$$f_i = RAFT(x_i, x_{i+1})$$

$$x_i + f_i = \hat{x}_{i+1}$$

$$error = ||\hat{x}_{i+1} - x_{i+1}||$$

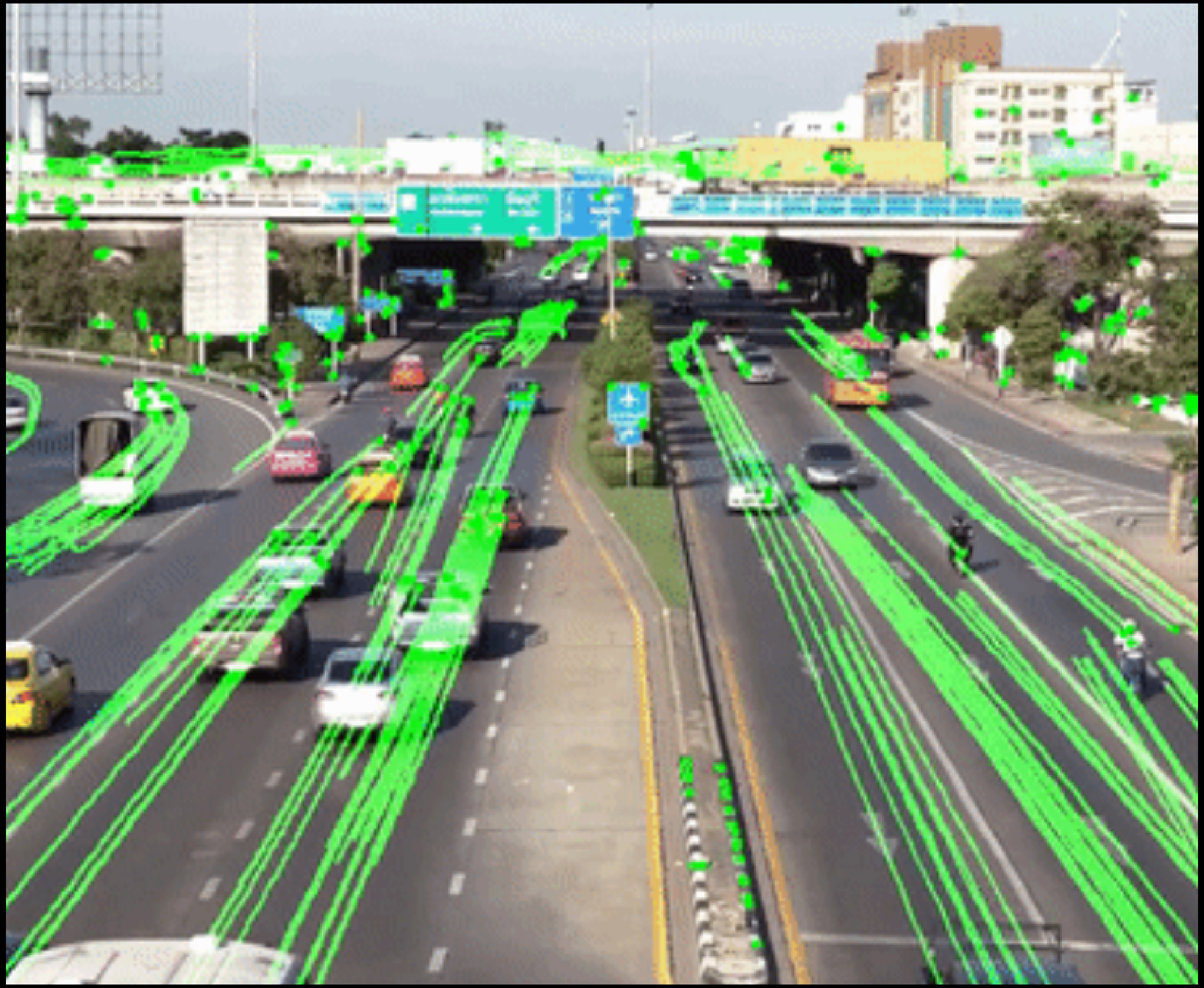
Exploit RAFT's **inaccurate predictions,** resulting from anomalous activity, to obtain **inaccurate constructions**

true frame

predicted frame

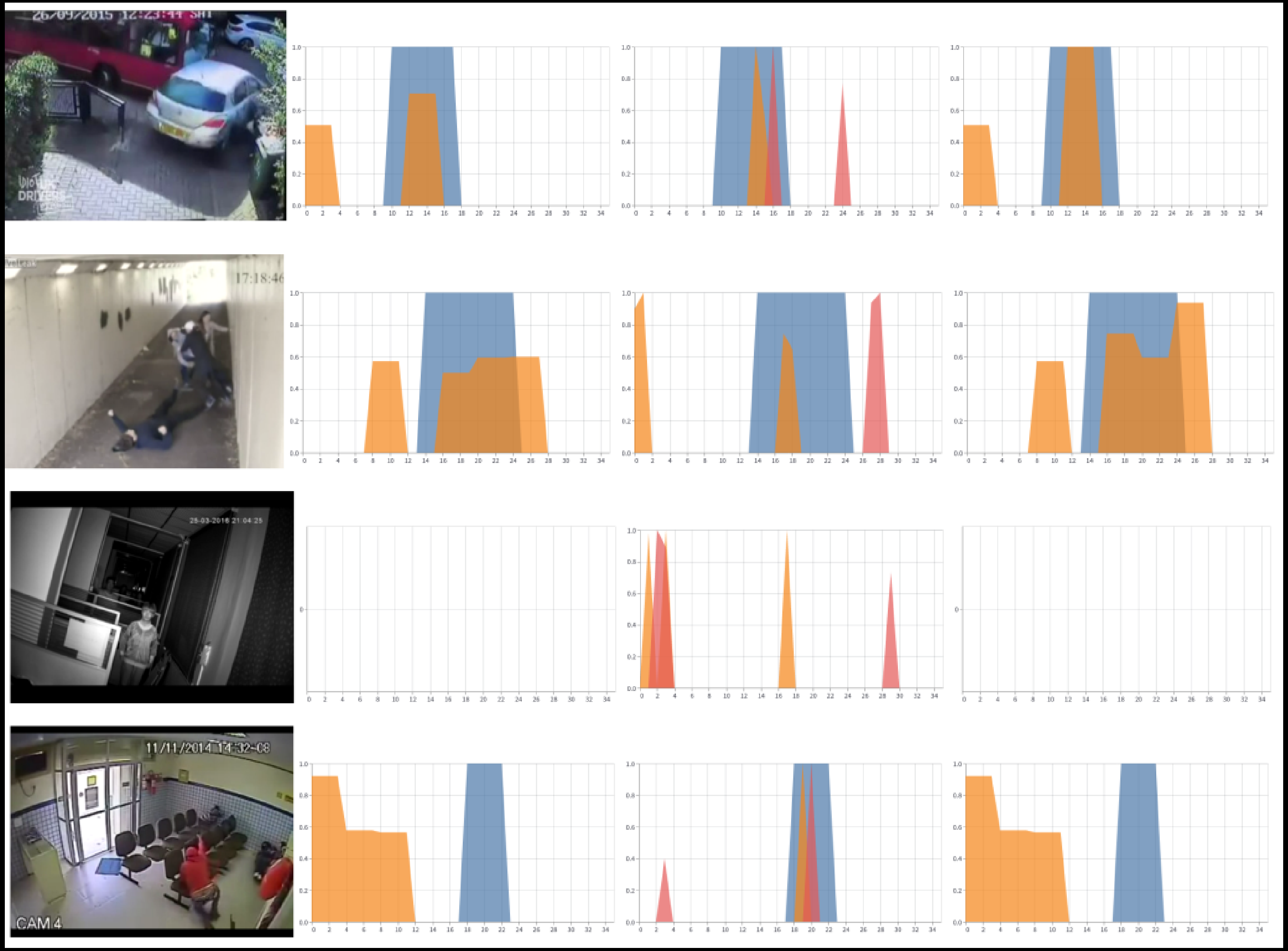
Implementation: LKKM

- **Lukas-Kanade** [4]: computes sparse optical flow
- **K-Means** [5]**:** data clustering algorithm
- **LKKM** forms a heuristic by clustering flow vectors
- **Quantify deviation from repetitive patterns**
  - Extract features
  - Compute sparse flow
  - Form cumulative clusters
  - Measure maximum distances between frame's flow vectors and nearest clusters



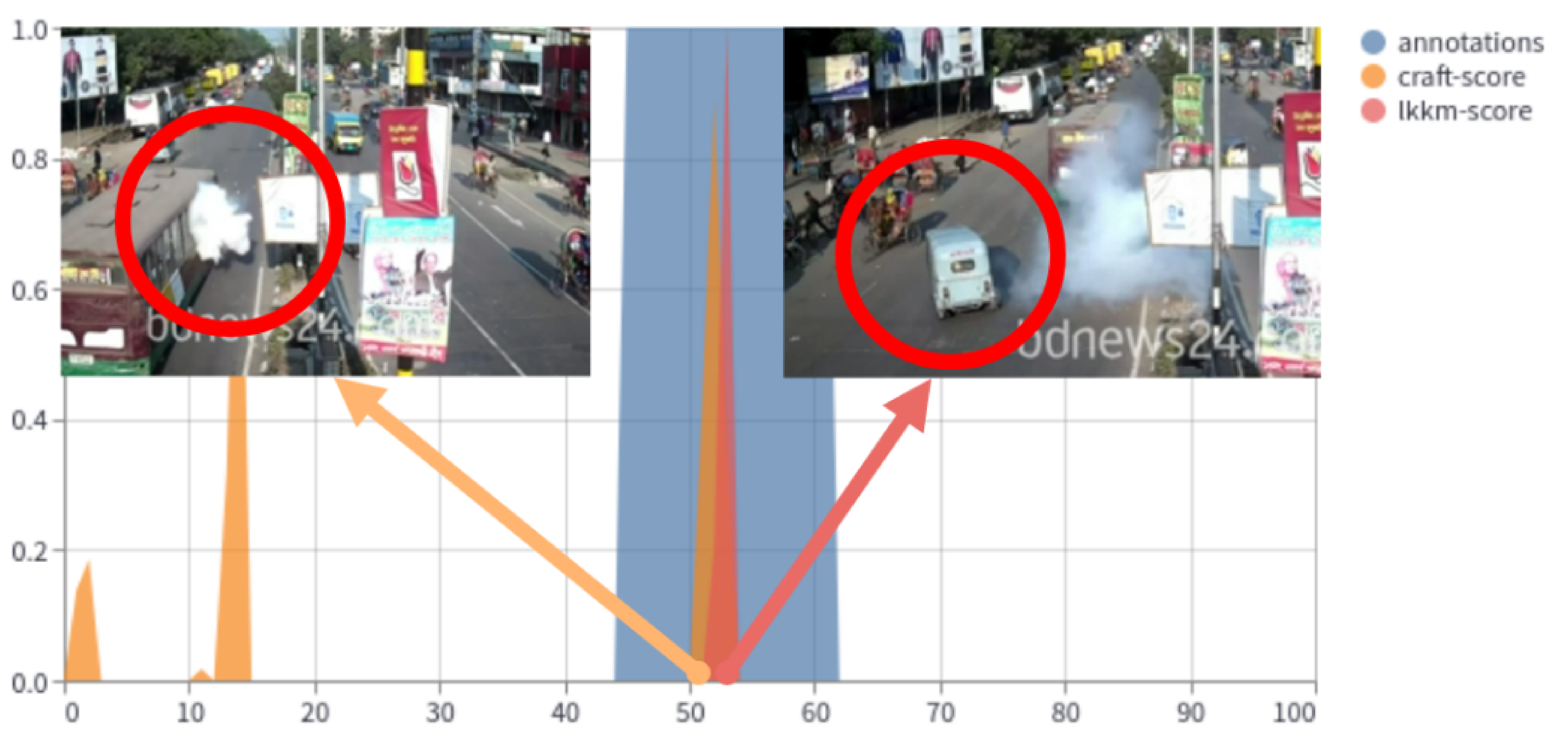
Results

https://share.streamlit.io/tomschdev/cctvanomalydetection/demo/src/pred_evaluation.py



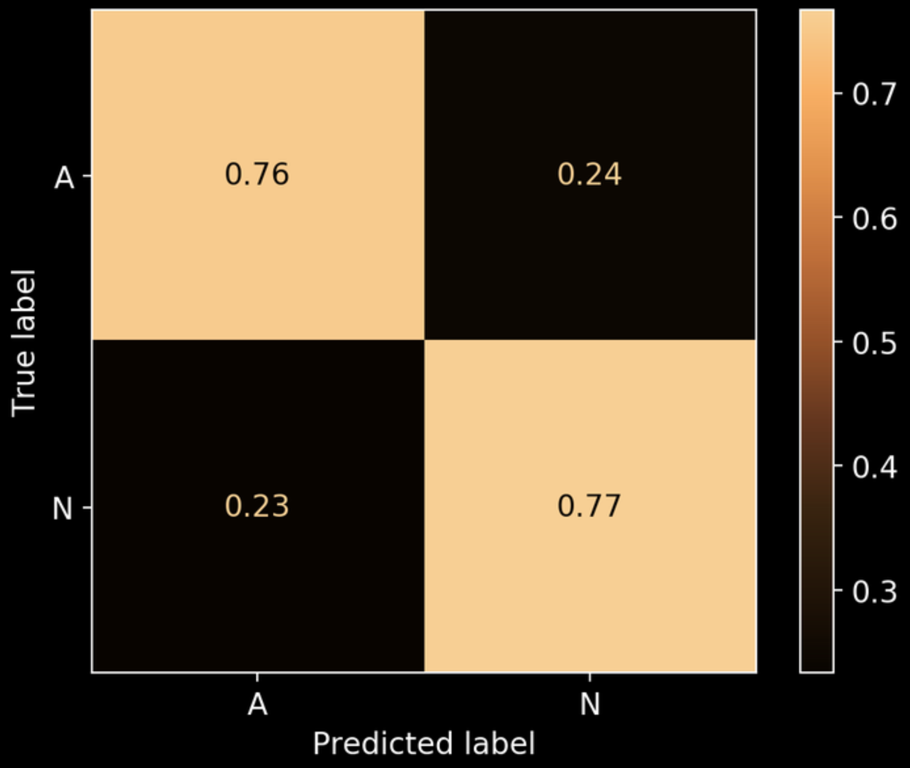
- Left: video instance.
- Second from left: base model score profile (orange) with annotation (blue)
- Second from right: CRAFT (orange) and LKKM (red) score profiles with annotation (blue).
- Right: consensus framework score profile (orange) with annotation (blue).

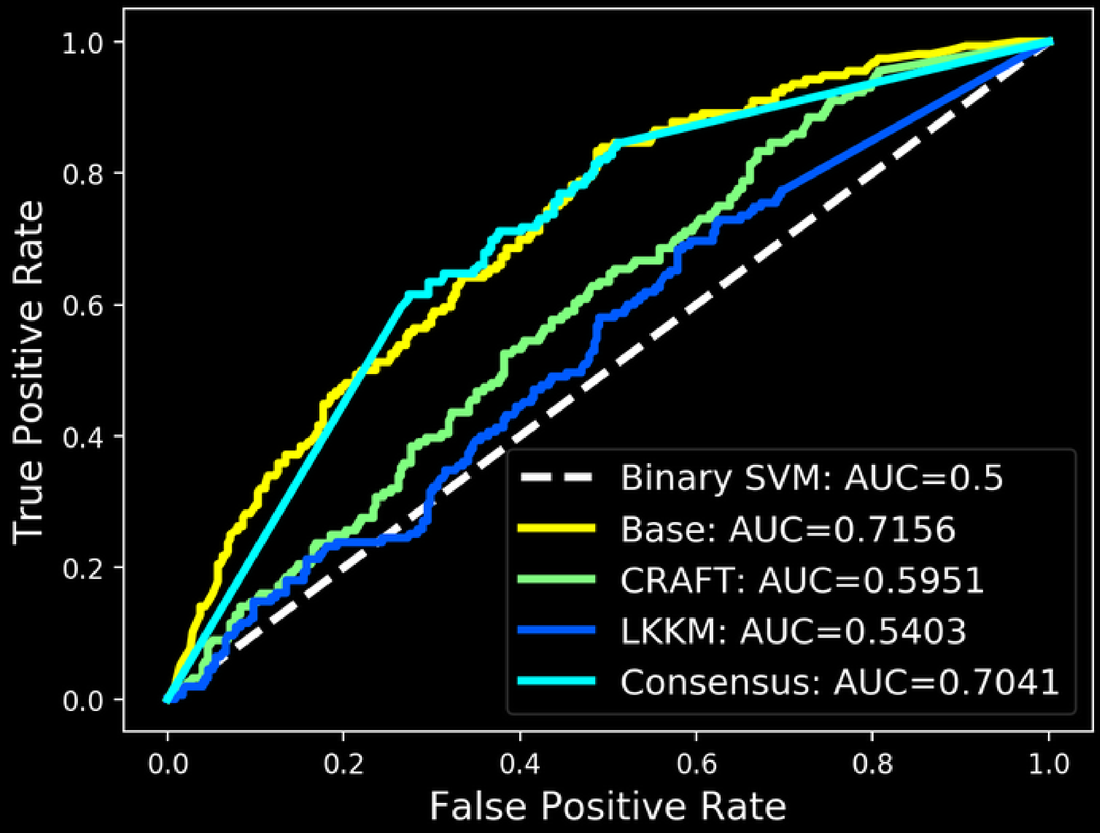
Results


1.0
0.8
0.6
0.4
0.2
0.0
0
10
20
30
40
50
60
70
80
90
100
annotations
craft-score
lkkm-score
bdnews24

Results

- High-level context of base model is necessary to achieve required levels of sophistication
- CRAFT and LKKM heuristics perform as intended
- Score combination is point of failure for consensus





- Optimal performance not attributable to consensus framework
- Instead, the source was a modified version of base model
    - deeper network

      further specification to training incentive

  TP and TN cases are well balanced

Conclusion

- The base model presents a realistic solution to real world anomaly detection
- CRAFT and LKKM are proven to be of useful but high false alarm rate is guaranteed
- Consensus framework ultimately fails to add value to the overall anomaly detection ability of the base model.

Future Work

- Cleaning of dataset
- End-to-end training instead of heuristic integration
- Extend event detection to activity classification
- Active learning in ANN component of base model

References

[1] Waqas Sultani, Chen Chen, Mubarak Shah, **Real-World Anomaly Detection in Surveillance Videos,** Cornell University Library, [v1] Fri, 12 Jan 2018.

[2] UCF-Crime Data set, https://www.crcv.ucf.edu/projects/real-world/

[3] RAFT: Teed, Z. and Deng, J., 2020, August. Raft: Recurrent all-pairs field transforms for optical flow. In *European conference on computer vision* (pp. 402-419). Springer, Cham.

[4] Lukas-Kanade: Bouguet, J.Y., 2001. Pyramidal implementation of the affine Lucas Kanade feature tracker description of the algorithm. *Intel corporation*, *5*(1-10), p.4.

[5] K-Means Clustering Burkardt, J., 2009. K-means clustering. *Virginia Tech, Advanced Research Computing, Interdisciplinary Center for Applied Mathematics*.

Questions?