Anomaly Detection in Videos

Elaborated by :

Hassen MNEJJA – Third Year Student – SISY

Supervised by :

Mr. Nizar BOUGUILA – Professor

Ms. Fatma NAJJAR – Research Assistant

Within :

Academic Year : 2021/2022

PLAN

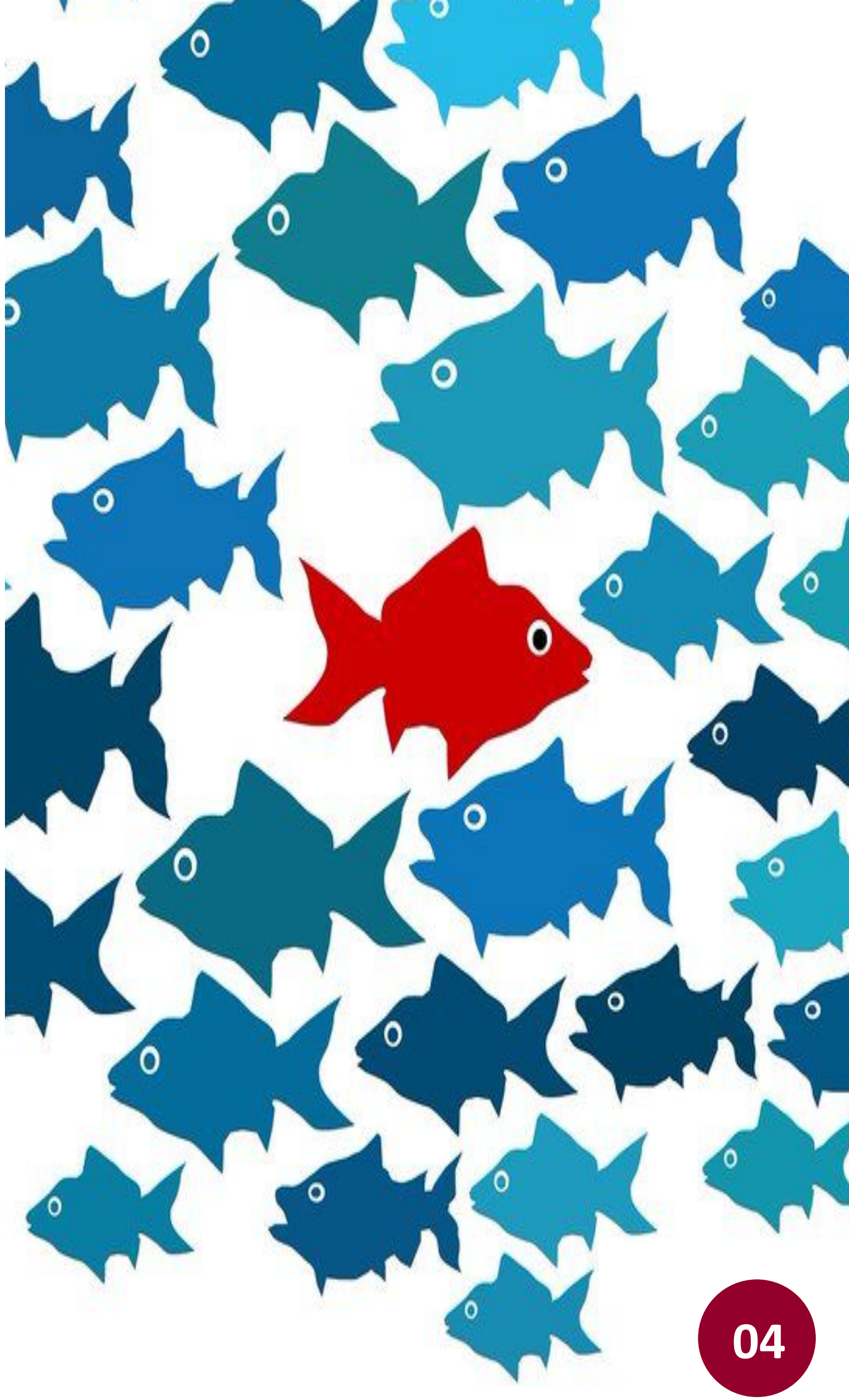
1. Introduction
2. General Framework of the Project
3. Problem Modeling
4. Exploring Experimental Results
5. Conclusion

INTRODUCTION

INTRODUCTION

Anomaly detection is the process of identifying unexpected items or events in data sets, which differ from the norm. And anomaly detection is often applied on unlabeled data which is known as unsupervised anomaly detection. Anomaly detection has two basic assumptions:

- Anomalies only occur very rarely in the data.
- Their features differ from the normal instances significantly.



Université Concordia University

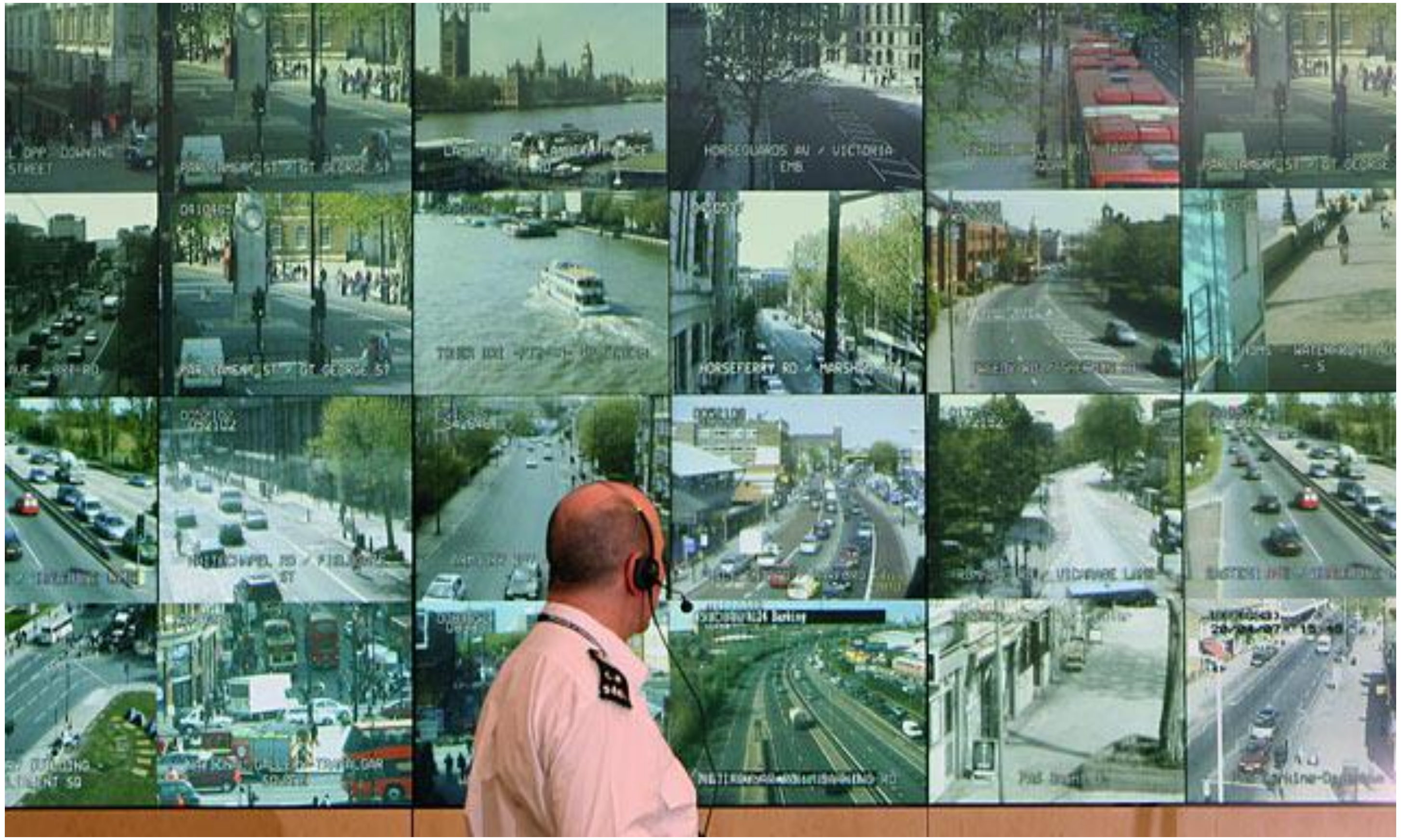
GENERAL FRAMEWORK OF THE PROJECT

HOSTING ORGANIZATION

- Merger of Loyola College and Sir George Williams University
- Gina Cody School of Engineering and Computer Science
- Concordia Institute for Information Systems Engineering
- Cybersecurity, Internet of Things, Artificial Intelligence, …

PROBLEM STATEMENT

PROBLEM STATEMENT



UNIVERSITÉ Concordia UNIVERSITY

ANOMALY DETECTION IN VIDEOS

PROBLEM MODELING

UNIVERSITÉ Concordia UNIVERSITY

PYTHON PACKAGES

NumPy

K

OpenCV

matplotlib

DATASETS USED

The UCSD pedestrian Dataset

The ShanghaiTech dataset

The CUHK Avenue dataset

Université Concordia University

The UCSD pedestrian Dataset

- A stationary video camera with a resolution of 238 X 158 pixels.
- Recorded at a frame rate of 26 frames per second.
- The Ped 1 dataset includes 34 train and 36 test video samples.

DATASETS USED

The UCSD pedestrian Dataset

The ShanghaiTech dataset

The CUHK Avenue dataset

The UCSD pedestrian Dataset

- A stationary video camera with a resolution of 238 X 158 pixels.
- Recorded at a frame rate of 26 frames per second.
- The Ped 2 dataset contains 16 train and 12 test video samples.

DATASETS USED

The UCSD pedestrian Dataset

The ShanghaiTech dataset

The CUHK Avenue dataset

The ShanghaiTech dataset

- It includes footage from 13 various cameras located throughout the ShanghaiTech University campus.
- It contains about 270, 000 training frames and 130 anomalous events.

DATASETS USED

The UCSD pedestrian Dataset

The ShanghaiTech dataset

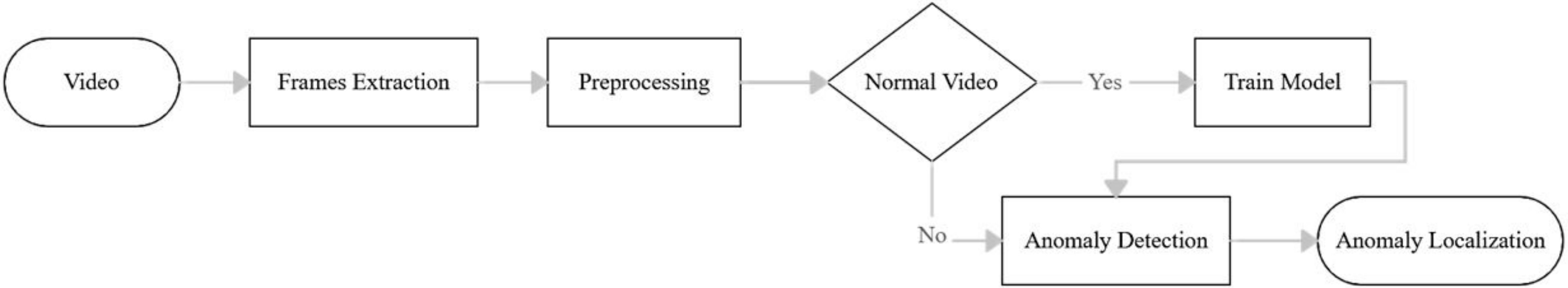
The CUHK Avenue dataset

The CUHK Avenue dataset

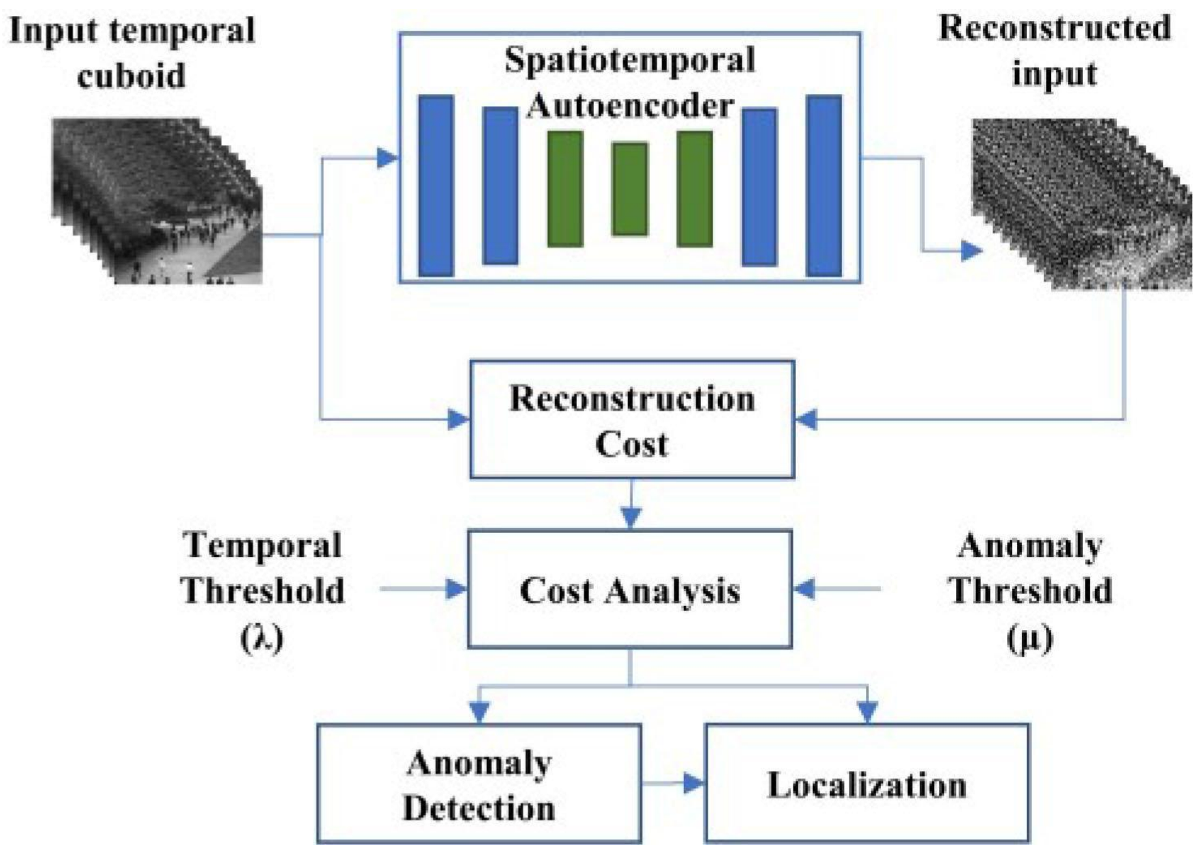
- created by filming street activities at the Chinese University of Hong Kong
- A stationary camera with a resolution of 640 X 360 pixels.
- Contains 16 train video examples and 21 test video samples.

UNIVERSITÉ Concordia UNIVERSITY

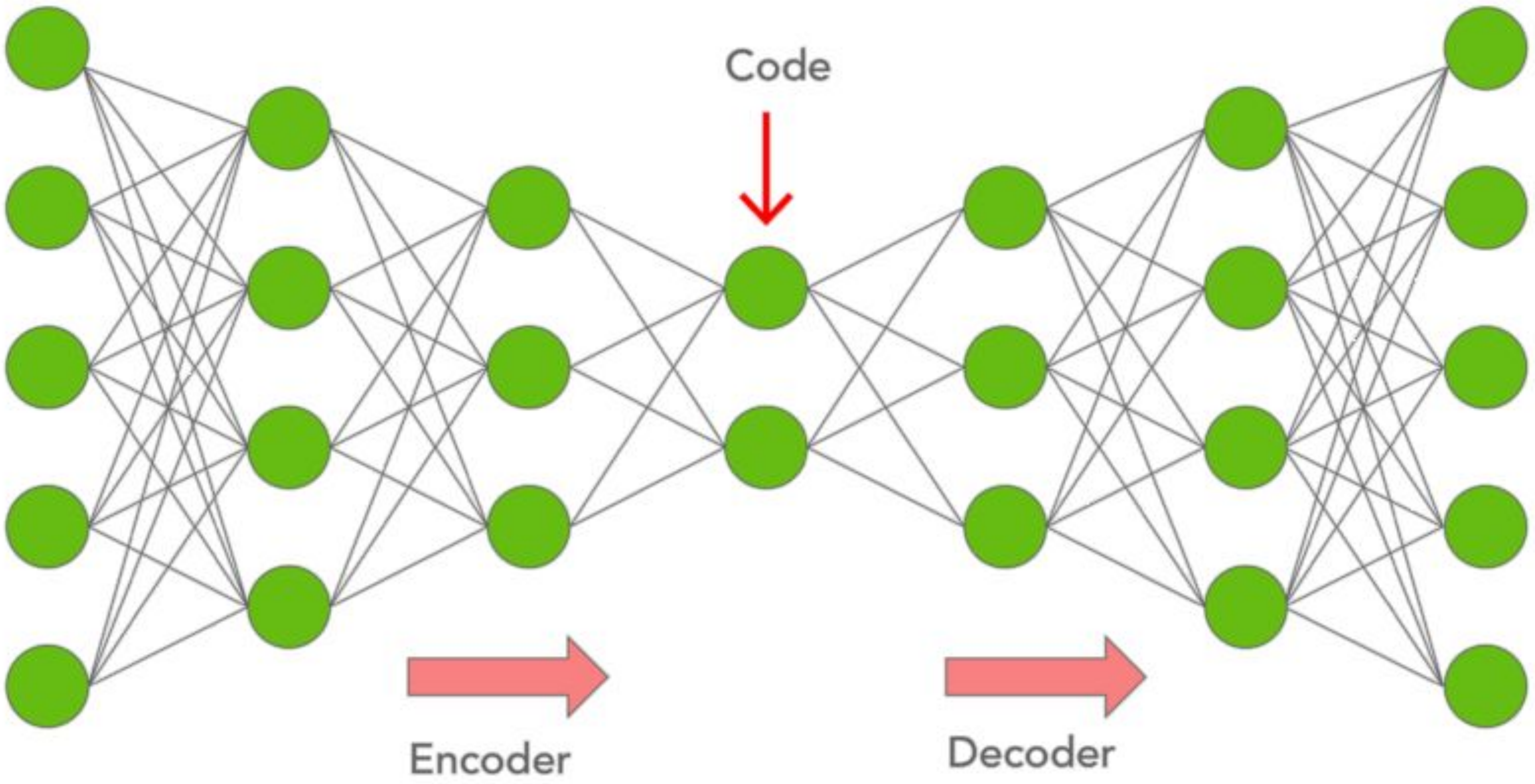
PROJECT’S PIPELINE



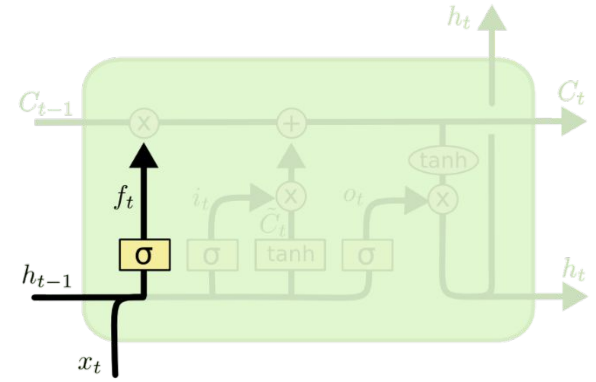
OUR SOLUTION



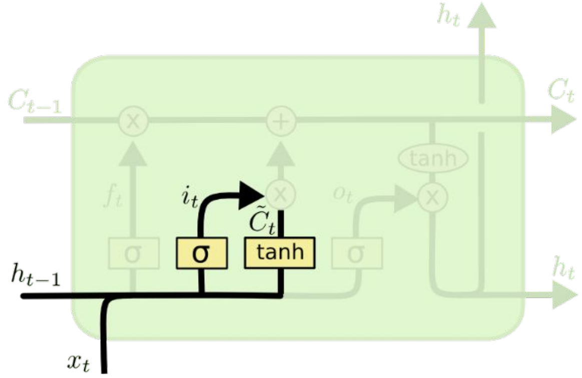
AUTO-ENCODER



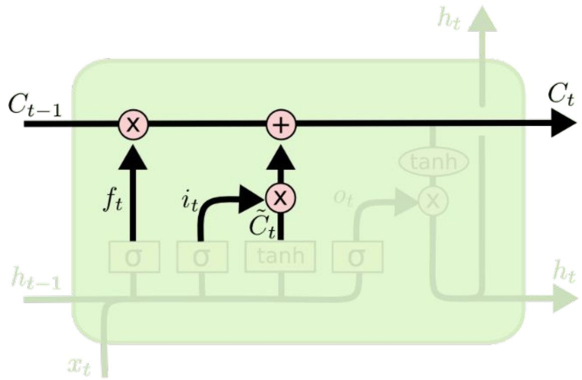
LONG AND SHORT TERM MEMORY



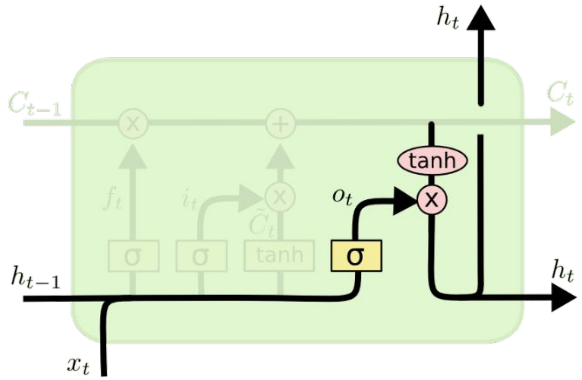
$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] \ + \ b_f\right)$$



$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] \ + \ b_i\right)$$
$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] \ + \ b_C)$$



$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$



$$o_t = \sigma\left(W_o \ [h_{t-1}, x_t] \ + \ b_o\right)$$
$$h_t = o_t * \tanh\left(C_t\right)$$

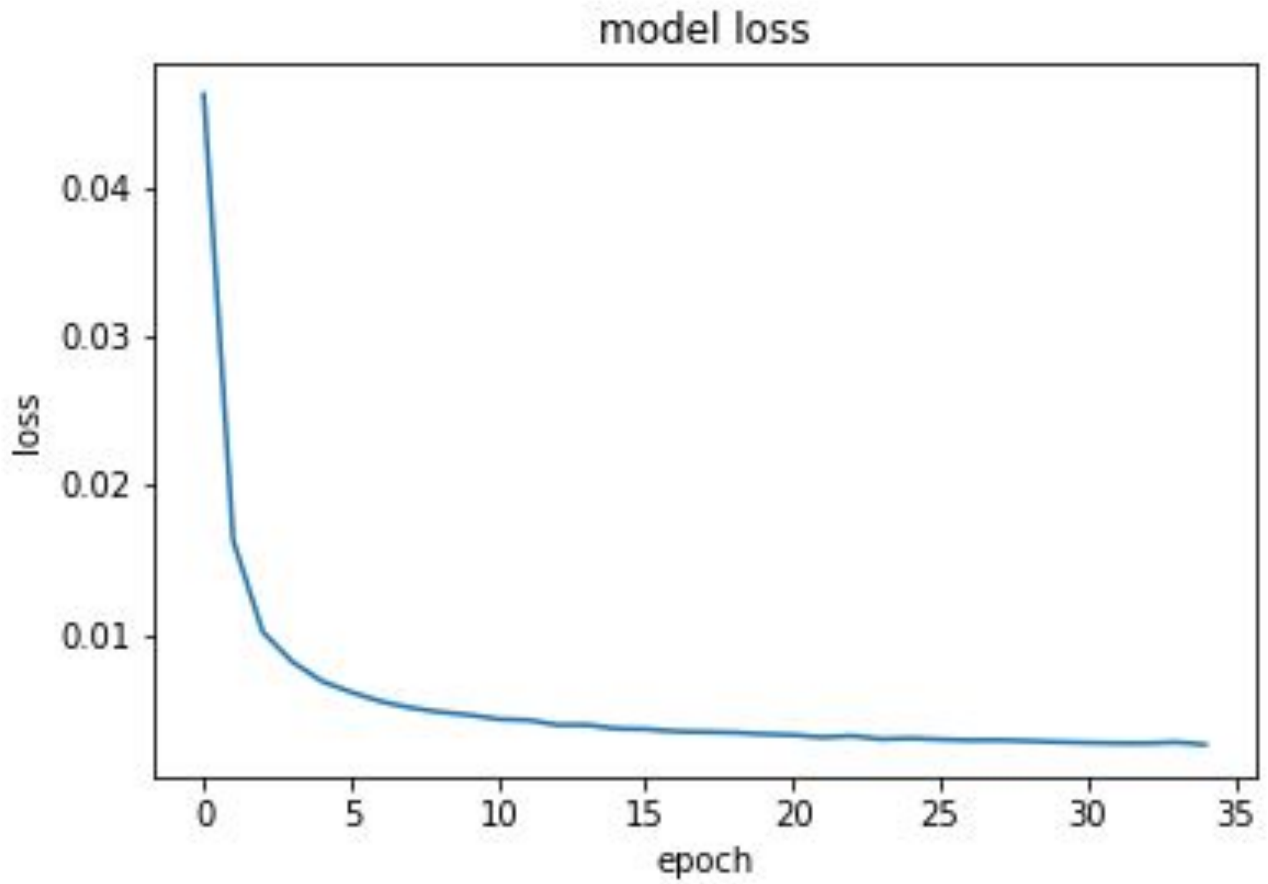
MODEL ARCHITECTURE

Type of layer	Output Shape	Weights (N)	Weights (%)
Input	(8, 224, 224, 1)		
TimeDistributed	(8, 56, 56, 128)	93440	1.1%
LayerNormalization	(8, 56, 56, 128)	256	0.0%
TimeDistributed	(8, 28, 28, 64)	1384512	15.6%
LayerNormalization	(8, 28, 28, 64)	128	0.0%
ConvLSTM2D	(8, 28, 28, 64)	295168	3.3%
LayerNormalization	(8, 28, 28, 64)	128	0.0%
ConvLSTM2D	(8, 28, 28, 32)	110720	1.2%
LayerNormalization	(8, 28, 28, 32)	64	0.0%
ConvLSTM2D	(8, 28, 28, 64)	221440	2.5%
LayerNormalization	(8, 28, 28, 64)	64	0.0%
TimeDistributed	(8, 56, 56, 64)	692288	7.8%
LayerNormalization	(8, 56, 56, 64)	128	0.0%
TimeDistributed	(8, 224, 224, 128)	5972096	67.4%
LayerNormalization	(8, 224, 224, 128)	256	0.0%
TimeDistributed	(8, 224, 224, 1)	93313	1.1%

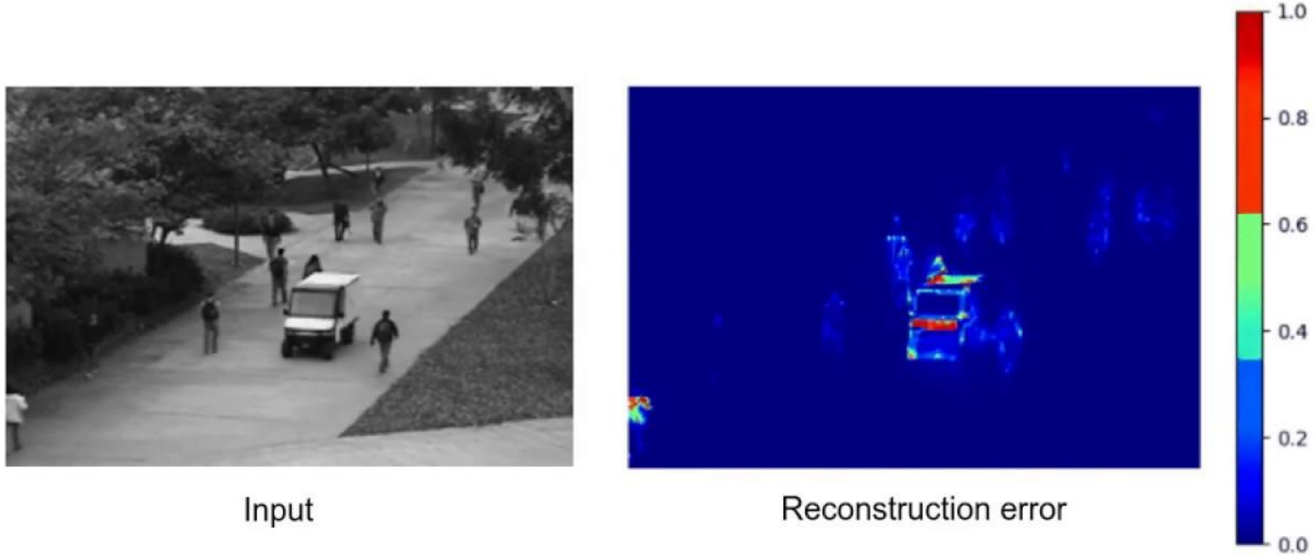
EXPLORING EXPERIMENTAL RESULTS

Université Concordia University

MODEL TRAINING



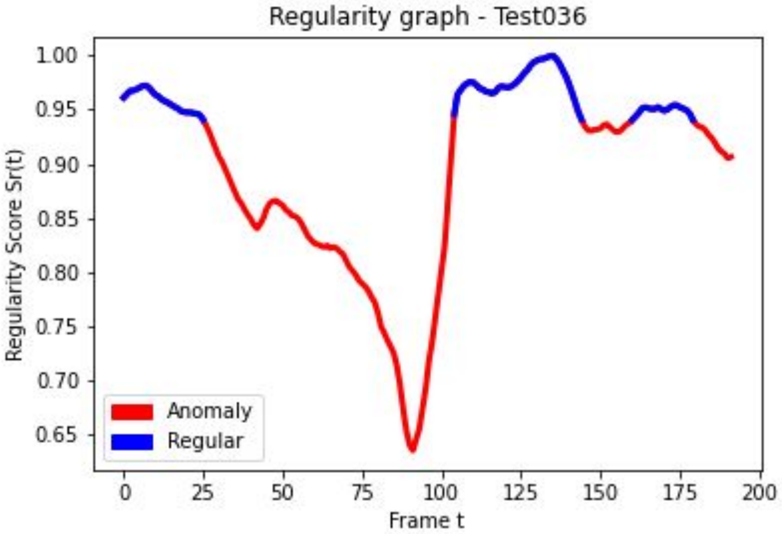
TEST OUR MODEL



Input

Reconstruction error

TEST OUR MODEL



COMPARISON

Model	Ped 1	Ped 2	Avenue
Conv-AE	81.0	90.0	70.2
S-RBM	70.3	86.4	78.8
Unmasking	68.4	82.2	80.6
ISTL	75.2	91.1	76.8
Our algorithm	70.4	84.2	74.6

Comparison of AUC

Université Concordia University

CONCLUSION

Université Concordia University

CONCLUSION

- Our model is based on the spatiotemporal auto-encoder.
- Our model does not only detect anomalies but also can localize them in the spatial reference.
- Despite the low available computational resources, our model achieved good results.

PERSPECTIVES

- Minimize the prediction serving latency
- Enhance quality of results
- Use of active learning techniques
- Real-time anomaly detection

THANK YOU

Université Concordia University