

# 视频监控下的异常检测

西安电子科技大学

杨文昊、吴婕、李永苏、朱世健

课程教师：王颖

# 楔子——图像处理VS计算机视觉



- 图像处理：图像→图像，例如 异质人脸图像（人脸图像素描合成）
- 计算机视觉：图像→知识，例如 目标检测，行人重识别
- 在大部分视觉课题中，图像处理与计算机视觉的知识都是相互融合的，异常检测的大部分方法就是通过图像处理的方法使计算机理解。

# 目录



01

异常检测的  
定义

02

背景与研究  
意义

03

思路与算法  
框架

04

结果展示  
demo

05

未来发展与  
挑战

06

参考文献



# 异常检测的定义



# 异常检测的定义

- 异常也被称为数据挖掘和统计文献中的异常，偏差或异常值。简而言之，异常检测就是一种识别不正常情况与挖掘非逻辑数据的技术。
- 例如在计算机视觉应用中，有人在抖音发表的一个视频，在边骑车边打电话，那这就是一个不符合规范的视频，我们可以用异常检测的方法将其检测出来。



# 异常检测的定义

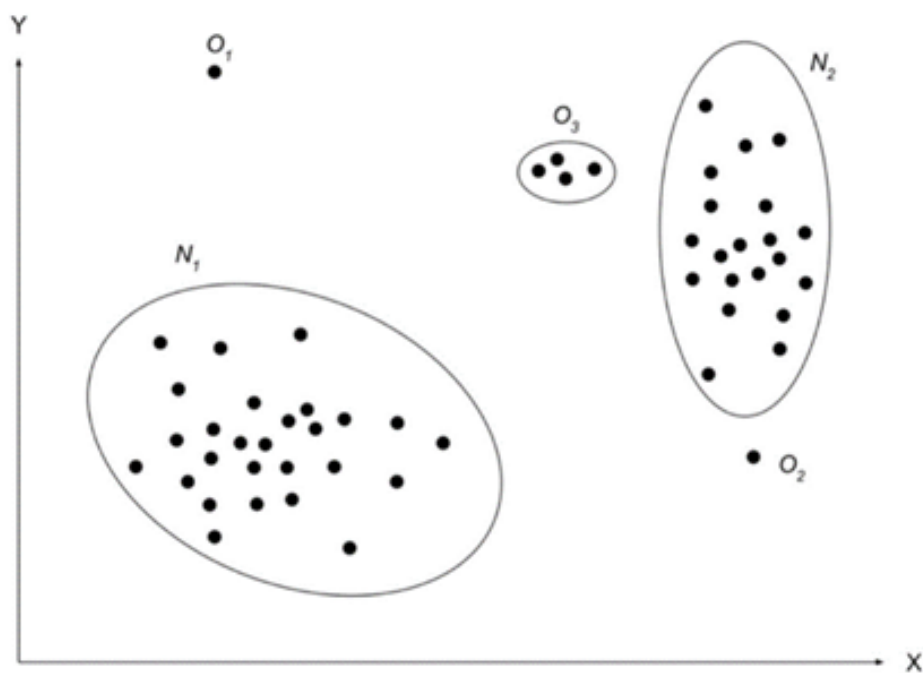


Figure 3: Illustration of anomalies in two-dimensional data set.

- 如图所示， $N_1$ 和 $N_2$ 是由大多数观察组成的区域，因此被认为是正常数据实例区域，而区域 $O_3$ 和数据点 $O_1$ 和 $O_2$ 是少于远离大部分的数据点。数据点并因此被视为异常。



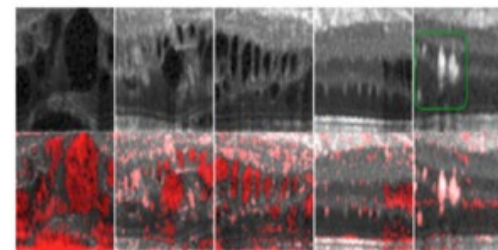
# 异常检测的常见应用



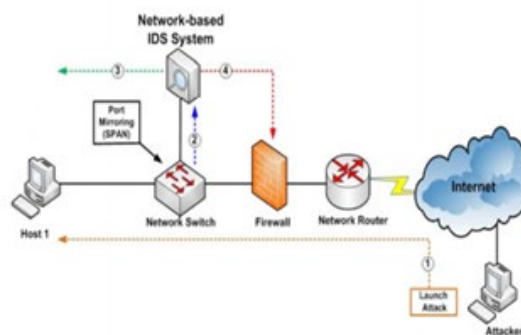
- 诈骗检测
- 网络侵入检测
- 医学异常检测
- 传感器网络异常检测视屏监督
- 物联网大数据异常检测
- 日志异常检测
- 工业危害检测



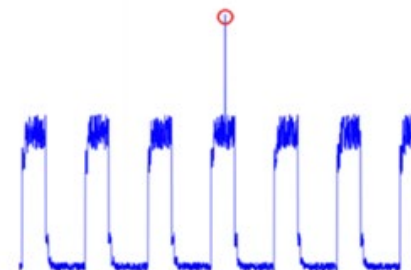
(a) Illegal Traffic Flow detection



(b) Detecting Retinal Damage



(c) Cyber-Network Intrusion detection



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



2

# 背景与研究意义



# 背景与研究意义

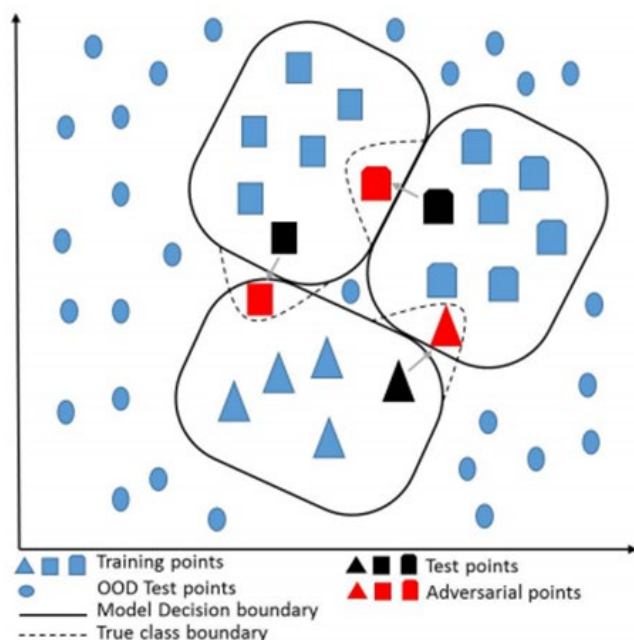


- 在生物信息学、医疗保健、网络安全和自动驾驶等领域越来越多地应用深度学习等技术，但深度学习在遇到与训练样本不同的数据，或者受噪声影响的数据时往往会得到异常的结果。因此对于数据的异常检测就显得尤为重要。
- 在图像分类中，通常将异常点分为无意的和有意的两种。无意的异常点包括：
  - • Anomaly Detection: 指在不属于该分类的数据集
  - • Novelty: 可能属于该分类但却没有出现过的数据集
  - • out-of-distribution: 多分类中不同目标的分布

# 背景与研究意义



- 有意的异常点通常指对抗性样本（Adversarial Examples），这是攻击者故意设计的数据输入，来迫使模型得到错误结果的数据集。



- 在目前研究中，异常检测的模型可依据数据标签的内容分为监督学习、半监督学习和无监督学习三种。



# 3

## 思路与算法框架

由于电院并没有系统地学习深度学习的内容，所以以下的算法将侧重于思路。

# 传统方法



- 异常检测的课题是在上个世纪就有很多学者进行了研究，主要面向于数据挖掘方向的课题。采用的算法例如混合高斯模型，半监督学习等。
- 深度学习大热之后，人们也开始研究将深度学习应用于各种异常任务中(也就是 Deep Anomaly Detection, 简称DAD), 并取得了很大的成功。

# DEEP LEARNING BASED



- Learning Temporal Regularity in Video Sequences, Hasan etal, CVPR 2016
- 网络结构设计、分类器设计、特征表达等方面
- 针对大部分深度学习异常检测的课题，AutoEncoder思路都是普遍的方法
- 而针对视频序列，主要用encoder去完成 reconstruction **重构** 和 prediction **预测** 这两个任务，从而判断图像/视频序列是否为异常。

# RECONSTRUCTION 重构



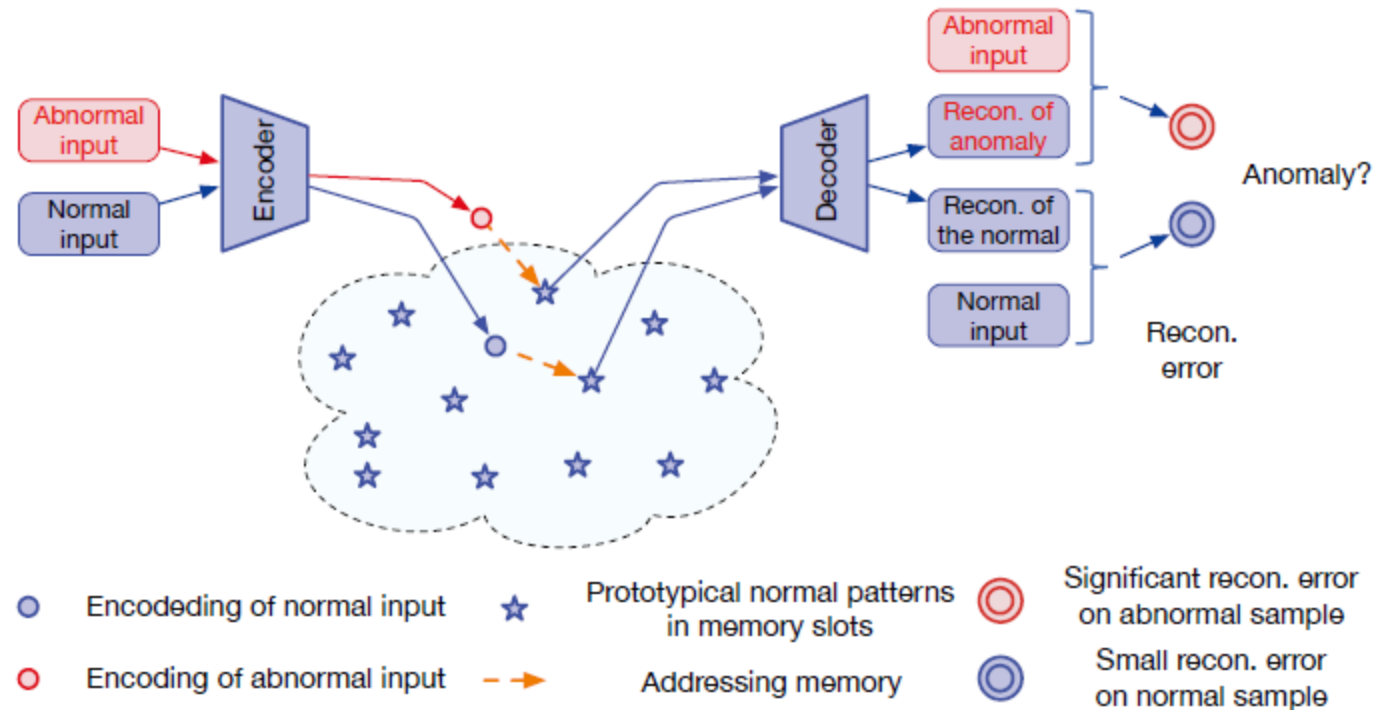
- Gong D, Liu L, Le V, et al. Memorizing normality to detect anomaly: Memory-augmented deep autoencoder for unsupervised anomaly detection[C]//Proceedings of the IEEE International Conference on Computer Vision. 2019: 1705-1714.
- 本人最喜欢的文章，不需要什么tricks就能得到非常稳定的模型。非常具有启发意义的文章，将重构任务完成的很好。



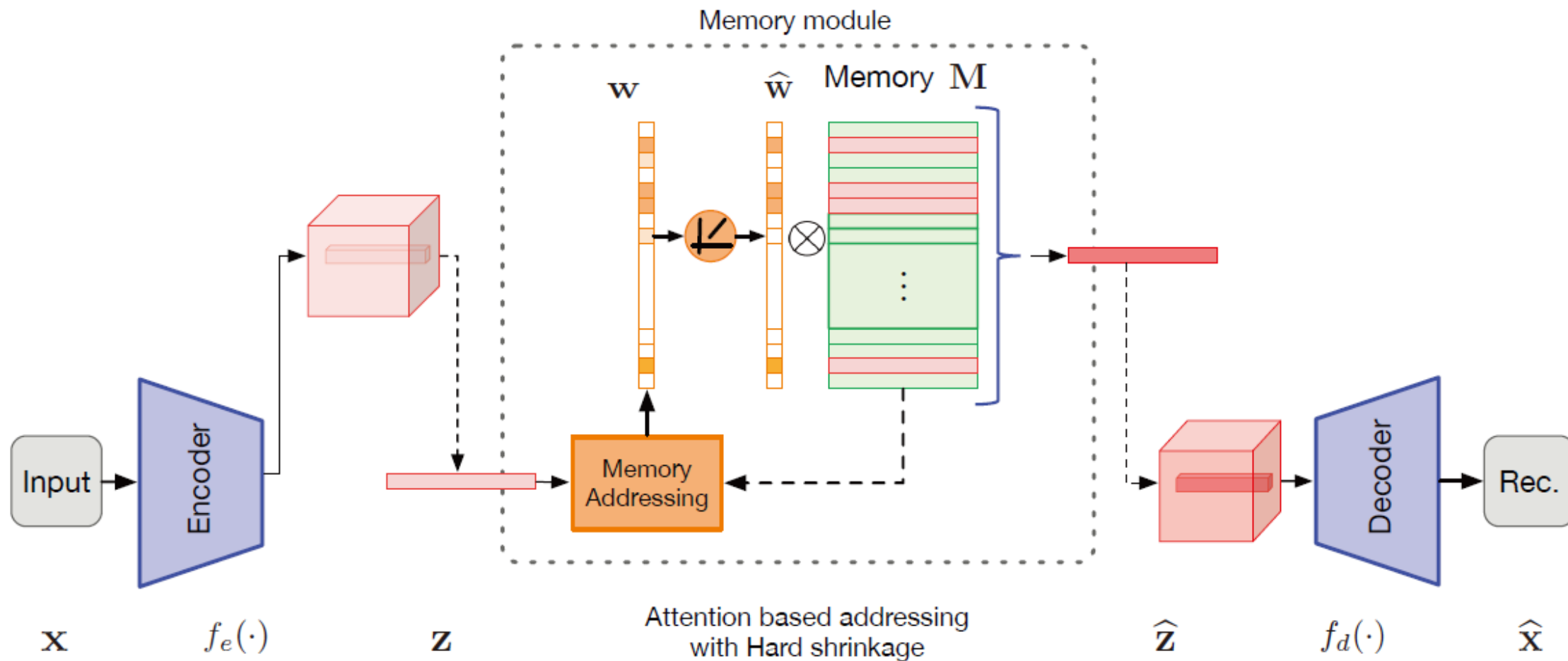
# RECONSTRUCTION 重构



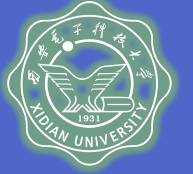
- How reconstruction error work ?
- Why unsupervised learning method ?
- Problems of reconstruction ?



# FRAMEWORK



# LOSS FUNCTION & TRAINING



■ Reconstruction Loss :  $R(\mathbf{x}^t, \hat{\mathbf{x}}^t) = \|\mathbf{x}^t - \hat{\mathbf{x}}^t\|_2^2$

■ Entropy Loss :  $E(\hat{\mathbf{w}}^t) = \sum_{i=1}^T -\hat{w}_i \cdot \log(\hat{w}_i)$

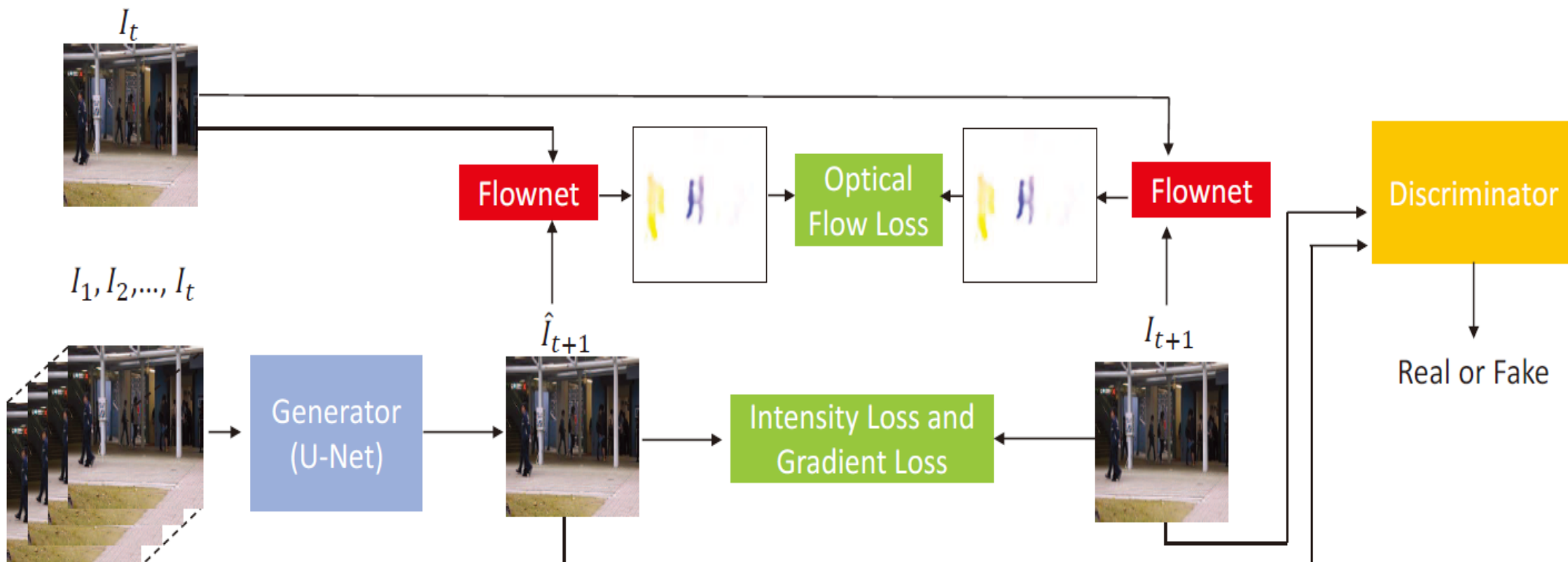
■ Combined Loss :  $L(\theta_e, \theta_d, \mathbf{M}) = \frac{1}{T} \sum_{t=1}^T \left( R(\mathbf{x}^t, \hat{\mathbf{x}}^t) + \alpha E(\hat{\mathbf{w}}^t) \right)$

# PREDICTION 预测



- Liu W, Luo W, Lian D, et al. Future frame prediction for anomaly detection—a new baseline[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 6536-6545.
- 上科大高盛华老师团队的论文，motivation很足而且整个研究工作非常完备充分，在encoding-decoding的基础上，摒弃之前的重构误差，使用预测误差，强调预测行为的不可预测性。使用GAN保证重构图像逼近真实图像，首次使用Flownet来产生光流。

# FRAMEWORK





- 针对视频序列，异常检测算法除了要输出当前图像或视频序列的标签（即是否为异常情况），还需要对当前图像进行打分。
- 当出现异常情况时，会出现低响应。
- 大部分文章采用MSE，本文采用了PSNR（Peak Signal to Noise Ratio）

$$S(t) = \frac{PSNR(I_t, \hat{I}_t) - \min_t PSNR(I_t, \hat{I}_t)}{\max_t PSNR(I_t, \hat{I}_t) - \min_t PSNR(I_t, \hat{I}_t)}$$



# OTHER WORK



- Large Margin Video Anomaly Detection with A Few Anomalies, Liu et al, IJCAI2019
- Zhao Y, Deng B, Shen C, et al. Spatio-temporal autoencoder for video anomaly detection[C]//Proceedings of the 25th ACM international conference on Multimedia. 2017: 1933-1941.



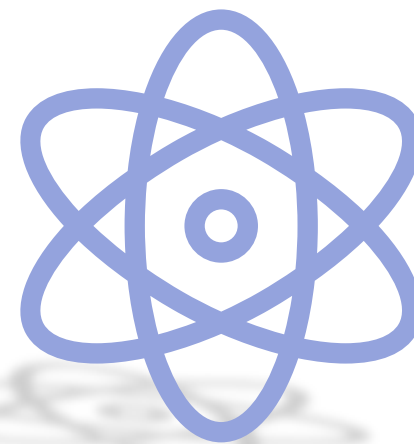
4

# 结果展示DEMO

# 实验结果



- Memorizing Normality to Detect Anomaly (MNDA)
- 小组工作：由于开源的源代码中无training code，所以首先对文章进行了复现工作，并经过了finetune得到了较好的结果。利用Adam优化器与初始learning rate得到了下列结果。



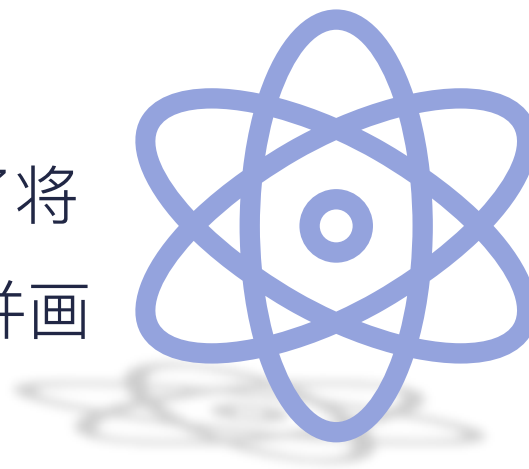
|           | This Paper | My Experiment |
|-----------|------------|---------------|
| USCD-ped2 | 94.1       | 92.7          |
| CUHK      | 83.3       | 81.79         |

- V100 48G
- Training: 8h

# 实验结果



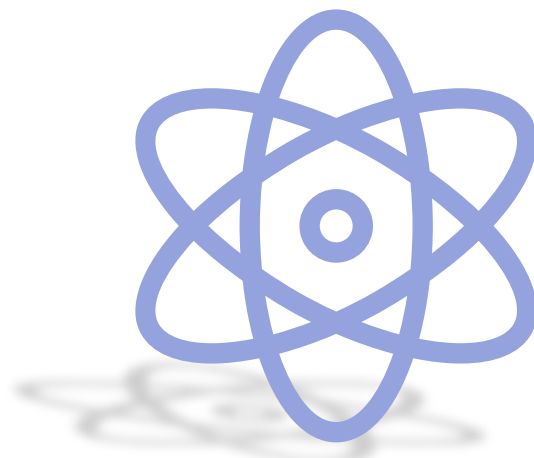
- Future Frame Prediction for Anomaly Detection (FFP)
- 小组工作：利用开源代码对文章进行了复现，并自主编写了将结果可视化的代码，测试了Avenue数据集的21个video，并画出psnr振荡图。（以video4为例）
- 模型训练时间：双卡RTX8000,48G显存，训练12h，单卡20h
- 21 Testing Videos: 3h



# 实验结果



| Method\Dataset |                 | UCSD-Ped2    | CUHK         | SH.Tech      |
|----------------|-----------------|--------------|--------------|--------------|
| Non-Recon.     | MPPCA [13]      | 0.693        | -            | -            |
|                | MPPCA+SFA [25]  | 0.613        | -            | -            |
|                | MDT [25]        | 0.829        | -            | -            |
|                | AMDN [40]       | 0.908        | -            | -            |
|                | Unmasking [36]  | 0.822        | 0.806        | -            |
|                | MT-FRCN [10]    | 0.922        | -            | -            |
|                | Frame-Pred [24] | <b>0.954</b> | <b>0.849</b> | <b>0.728</b> |
| Recon.         | AE-Conv2D [9]   | 0.850        | 0.800        | 0.609        |
|                | AE-Conv3D [44]  | 0.912        | 0.771        | -            |
|                | TSC [24]        | 0.910        | 0.806        | 0.679        |
|                | StackRNN [24]   | 0.922        | 0.817        | 0.680        |
|                | AE              | 0.917        | 0.810        | 0.697        |
|                | MemAE-nonSpar   | 0.929        | 0.821        | 0.688        |
|                | MemAE           | <b>0.941</b> | <b>0.833</b> | <b>0.712</b> |





5

# 未来挑战与思考



# 未来挑战与思考



- 1. 训练过程必须采用mini-batch，针对大规模的数据集，非常消耗资源与时间
- 2. 科研与工业脱节严重，工业上的异常检测问题通常是定位目标且需要判断类别
- 3. 数据集本身问题不符合实际情况，例如Shanghai Tech融合多角度视频
- 4. 深度学习不是万能的，即使有大量数据，无监督学习往往会发生过拟合的问题
- 5. 在视频分析与理解的领域，例如异常检测、行人重识别、行为定位与识别，总是会存在tricks>>model的情况，落地效果不稳定。



# 6

## 参考文献

# 我的博客 & 代表性论文



- 1. 阿尔法杨XDU, 异常检测 | Anomaly Detection综述
- 2. 阿法尔杨XDU, 异常检测 | 视频监控方向8篇论文模型浅析
- 3. 阿尔法杨XDU, 异常检测 | MemAE模型复现与思考
- 4. Future frame prediction for anomaly detection—a new baseline, CVPR2018
- 5. Memorizing normality to detect anomaly: Memory-augmented deep autoencoder for unsupervised anomaly detection, ICCV2019

谢谢观看！

Q&A