Video Anomaly Detection by Estimating Likelihood of Representations

Yuqi Ouyang

Victor Sanchez

Signal and Information Processing (SIP) Lab

Department of Computer Science

The University of Warwick

ICPR 2020

Contact: yuqi.ouyang@warwick.ac.uk

Outline

1. What is Video Anomaly Detection?
2. Model Overview.
3. Training.
4. Anomaly Inference.
5. Experiments & Analysis.
6. Future Work.

What is Video Anomaly Detection?

Detect abnormal events in surveillance video data.

Infeasible to gather all anomaly video data.

Outlier detection – only train with normality, test to detect abnormality

GMM-DAE: Model Overview

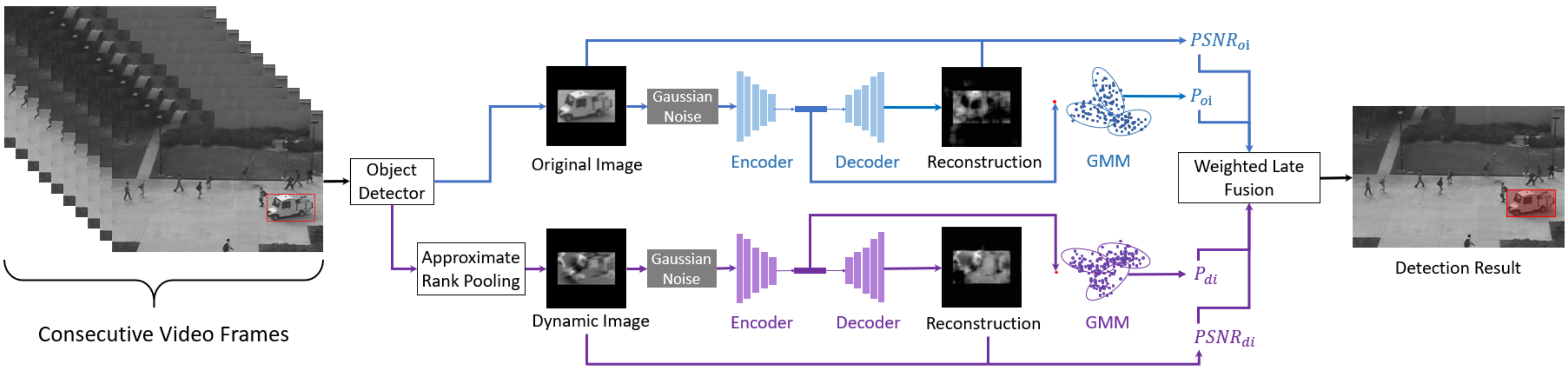


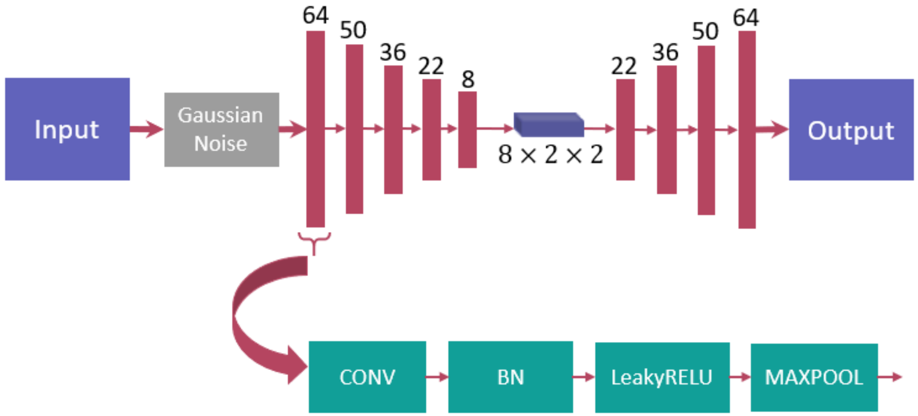
Figure: The proposed GMM-DAE model.

GMM-DAE: Training

Train Autoencoder (Eq. 2.):

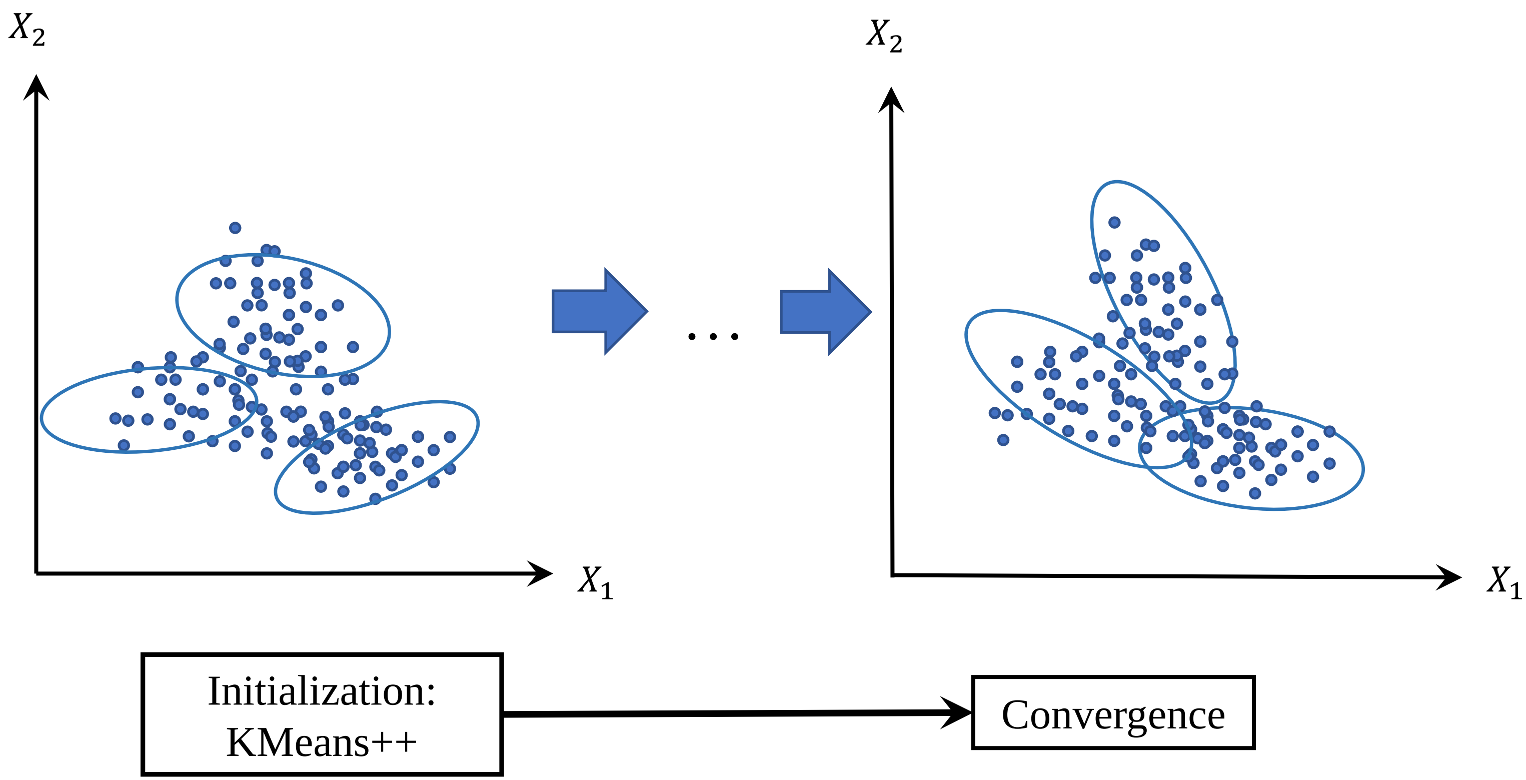
$$L\left(\boldsymbol{X}, \hat{\boldsymbol{X}}\right) = \frac{1}{n}\sum_{i=1}^{n} \|x_i - \hat{x}_i\|_2^2 + \beta\|\boldsymbol{W}\|_2^2,$$

$\boldsymbol{X}$: n uncorrupted images.
$\widehat{\boldsymbol{X}}$: n output images,
$\|W\|_2^2$: ℓ_2-regularization on the weights by a factor β.



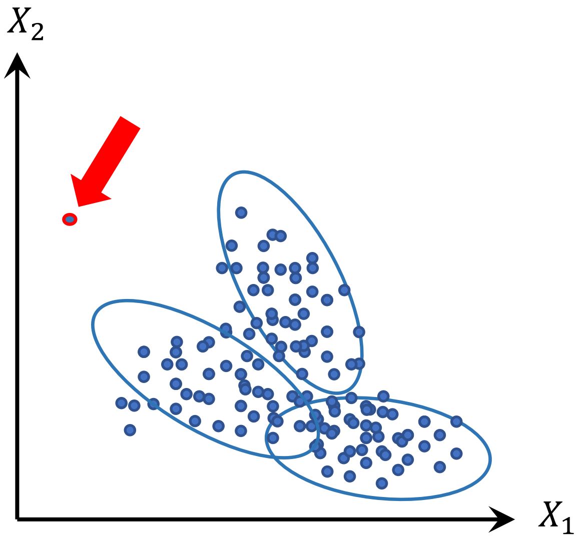
GMM-DAE: Training

Train GMM (Eq. 3-11): Expectation-Maximization.



GMM-DAE: Anomaly Inference

Anomaly Inference (Eq. 12-15):

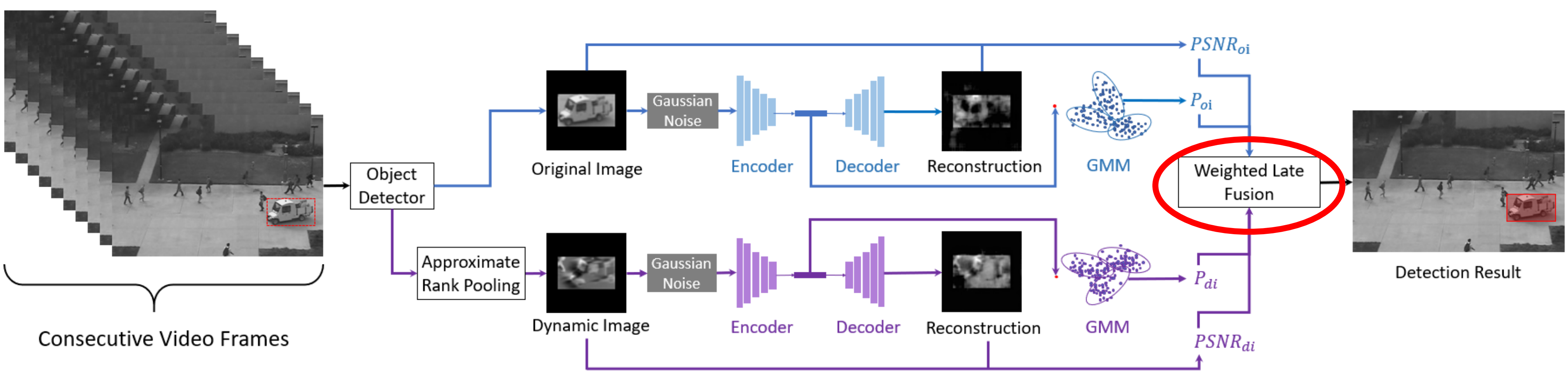


$$PSNR(x,\hat{x}) = 10\log_{10}\frac{max(x)}{MSE(x,\hat{x})} \tag{12}$$

$$P(z) = \log\sum_{j=1}^{K}\frac{\hat{\phi}_j \cdot \exp -\frac{1}{2}\left(z-\hat{\mu}_j\right)^T \hat{\Sigma}_j^{-1}\left(z-\hat{\mu}_j\right)}{\sqrt{|2\pi\hat{\Sigma}_j|}}, \tag{13}$$

$$A(x^t) = -\left[\lambda_1 P(z_{x^t}) + \lambda_2 \cdot PSNR(x^t,\hat{x}^t) + \lambda_3 P(z_{d^t}) + \lambda_4 \cdot PSNR(d^t,\hat{d}^t)\right], \tag{14}$$

$$A(I^t) = max\{A(x_1^t), A(x_2^t), ..., A(x_n^t)\}. \tag{15}$$



GMM-DAE: Experiments & Analysis

Model	UCSD Ped2	CUHK Avenue	ShanghaiTech
MPPCA [14]	69.3	-	-
MPPCA+SFA [25]	61.3	-	-
MDT [25]	82.9	-	-
Unmasking [13]	82.2	80.6	-
AMDN [41]	90.8	-	-
FRCN action [10]	92.2	**89.8**	-
Conv-AE [9]	90.0	70.2	60.9
STAE [42]	91.2	77.1	-
GANs [30]	93.5	-	-
FFP+MC [18]	95.4	85.1	72.8
LSA [1]	95.4	-	72.5
MLAD [40]	**99.21**	71.54	-
sRNN-AE [22]	92.21	83.48	69.63
MemAE [8]	94.1	83.3	71.2
MLEP-FP [19]	-	89.2	73.4
MemAE2020 [29]	97.0	88.5	70.5
SDOR [28]	83.2	-	-
SAGC [26]	-	-	**76.1**
GMM-DAE	**96.5**	**89.3**	**81.2**

Table: Frame-level AUROC curve (%) comparison with other baseline models, on three datasets (Higher is better).

Method	AUROC
DAE+OP	90.1
DAE+OP+DP	93.9
DAE+OP+OL	94.2
DAE+OP+OL+DP+DL (GMM-DAE)	**96.5**
AE+OP+OL+DP+DL	95.8

Table: Frame-level AUROC curve (%) on the UCSD Ped2 dataset using different components of the proposed GMM-DAE model.

GMM-DAE: Experiments & Analysis

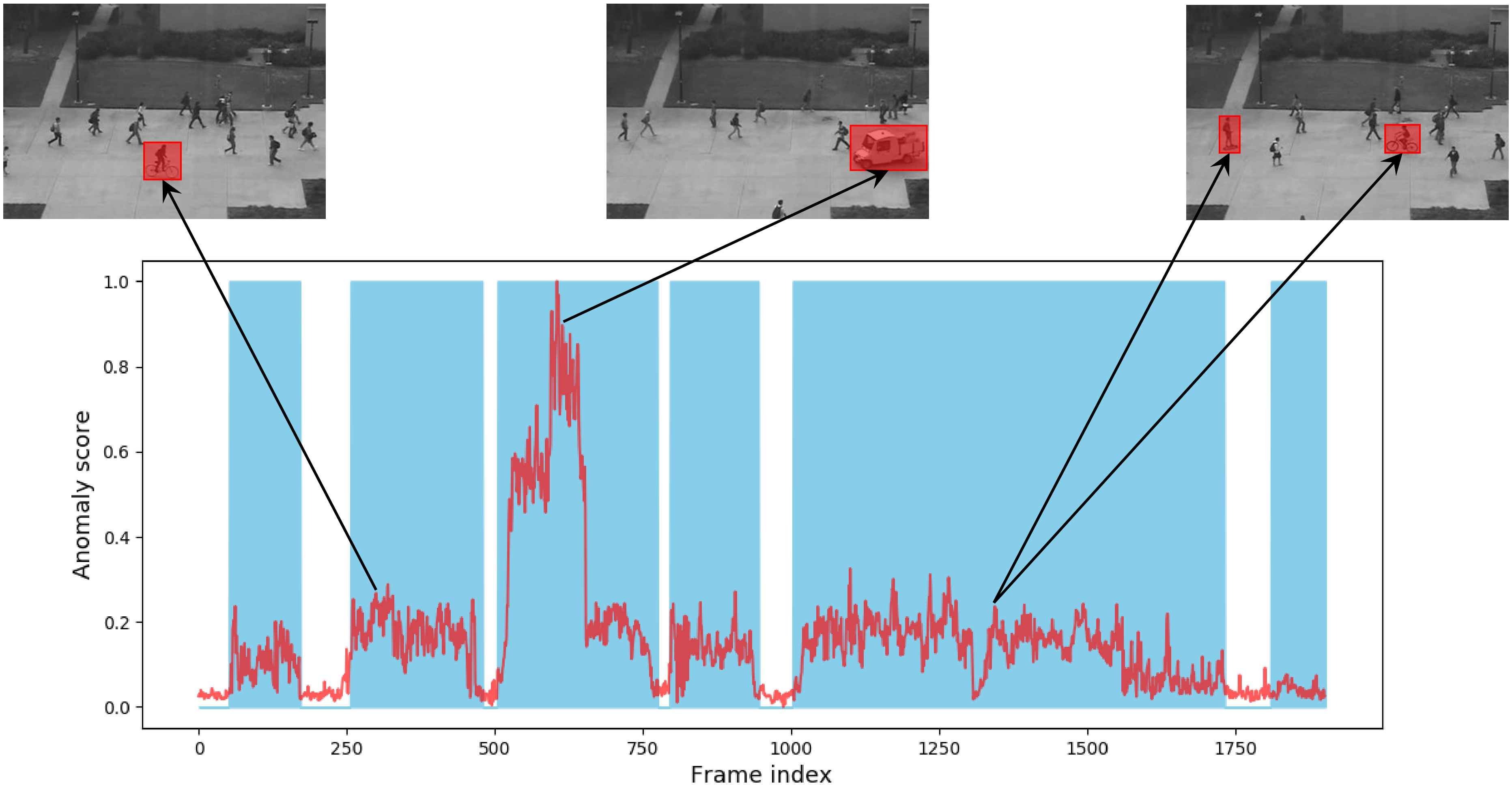


Figure: Performance analysis on the UCSD Ped2 dataset.

GMM-DAE: Experiments & Analysis

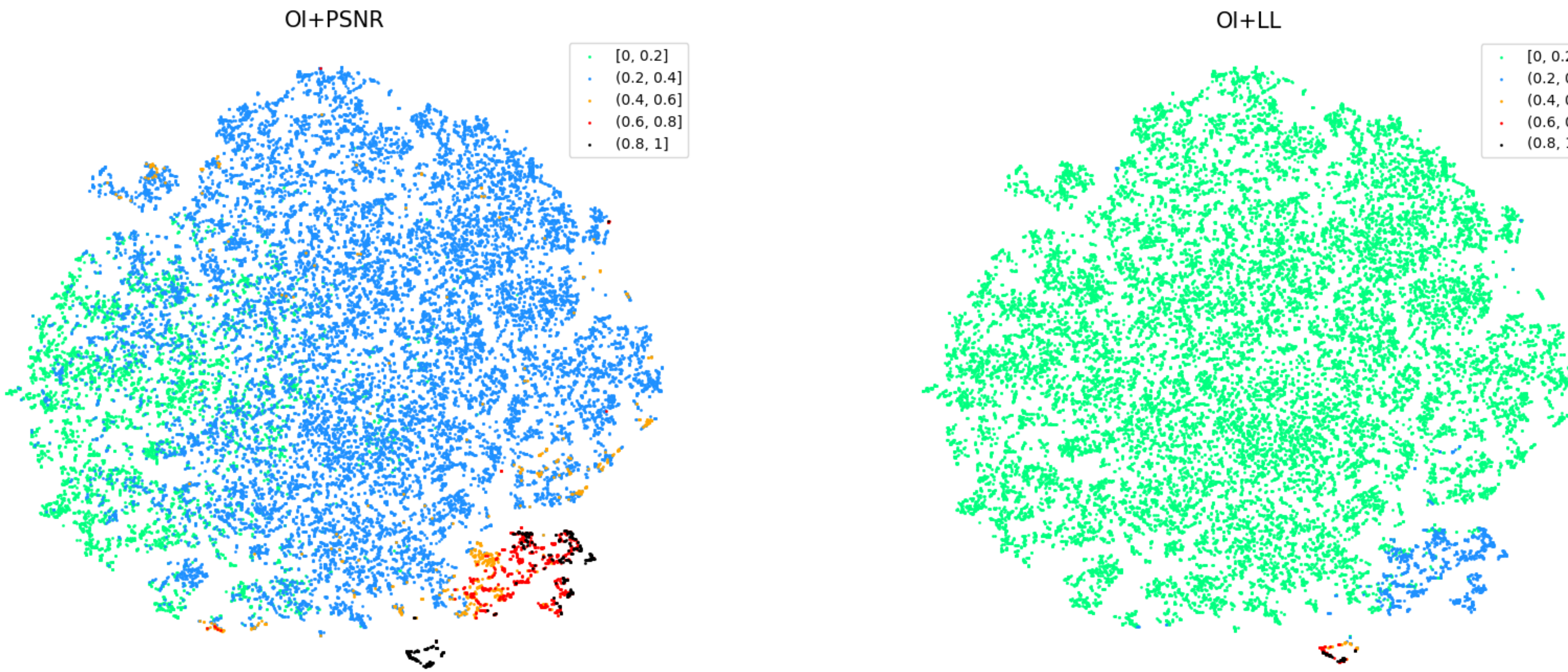


Figure: Distribution of the generated manifolds based on the normalized anomaly scores using PSNR values (left: OI+PSNR) and latent likelihood values (right: OI+LL) on the UCSD Ped2 test videos. Each color represents a range of anomaly scores.

Future Work

- We want to remove the object detector (because it causes miss-detection issue).
- Model training in an end-to-end way.
- Design a model to perform robust and accurate pixel-level detection.

Thank you for listening, Q?

Stay safe, be happy