DyAnNet: A Scene Dynamicity Guided Self-Trained Video Anomaly Detection Network (Supplementary Document)

This supplementary document shows a few more experimental and qualitative results. In Section 1, we have shown the performance variation with 3 different values of (τ). In the next Section 2, we have shown the performance in AUC (%) using different unsupervised algorithms employed in combination with iForest. Section 3 depicts the supervision-based AUC performance of SOTA methods. In Section 5, we have provided ROC plots using three datasets.

1. Selection of Detection Threshold (τ)

Table 1. Comparison of the proposed method with Zhong *et al*. [48] using three different τ values. Performance is shown in AUC (%) and FAR is shown in brackets.

	CCTV-Fights		**UBI-Fights**		**UCF-Crime**	
Threshold (τ)	Zhong *et al*.	**Ours**	Zhong *et al*.	**Ours**	Zhong *et al*.	**Ours**
0.40	78.22 (6.08)	83.89 (2.89)	81.66 (8.91)	87.20 (2.67)	81.69 (4.78)	85.23 (0.89)
0.45	76.23 (4.32)	82.12 (1.60)	80.34 (7.67)	86.75 (1.95)	79.23 (3.14)	84.55 (0.60)
0.50	74.78 (5.09)	**81.01 (1.70)**	82.43 (5.78)	**86.31 (1.41)**	81.08 (2.80)	**84.50 (0.52)**

2. Generating Pseudo-Anomaly Score

We have experimented with the following unsupervised anomaly detection algorithms: (i) MCD (Minimum Covariance Distance), (ii) PCA (Principle Component Analysis), (iii) LOF (Local Outlier Factors), (iv) One-class SVM, and (v) Isolation Forest. Tab. 2 shows the AUC (%) using the algorithms combined with iForest.

Table 2. Performance of the proposed method in terms of AUC (%) with different unsupervised algorithms combined with iForest to generate pseudo-anomaly score.

Algorithm	**CCTV-Fights** [31]	**UBI-Fights** [6]	**UCF-Crime** [38]
MCD	77.24	84.07	81.11
PCA	**79.94**	**85.13**	**83.58**
LOF	77.60	84.86	82.02
OCSVM	**81.12**	**86.75**	**84.55**

3. Performance Comparisons of Varying Levels of Supervision

Tab. 3 shows the performance of all methods. We have compared the performance of the proposed method with semi-supervised, weakly-supervised, and unsupervised approaches. Table 1 has been revised in the original paper in Section 4.4.

4. Additional Results

We have added four additional results on testing videos taken from the CCTV-Fight [31], UBI-Fights [6], and UCF-Crime [38] datasets. Fig.1 depicts a fighting scene from CCTV-Fight dataset. In the fifth segment, Ω regressor has detected an anomalous activity; however, ψ has generated a very low dynamicity score for the same segment avoiding false alarms.

Table 3. Frame-level AUC scores (in %) of the state-of-the-art methods on three video anomaly datasets, D1: CCTV-Fights [31], D2: UBI-Fights [6], and D3: UCF-Crime [38]. The top two results are shown in red and blue.

Year	**Method**	**D1**	**D2**	**D3**	**Superv.**
2016	Hasan *et al*. [11]	52.43	64.87	50.6	Semi.
2017	Hinami *et al*. [12]	56.70	67.12	57.10	Semi.
2018	Ravanbaksh *et al*. [35]	60.37	69.45	61.61	Unsuper.
2018	Sultani *et al*. [38]	72.55	78.70	75.41	Weak.
2019	Ionescu *et al*. [13]	73.86	78.49	76.20	Unsuper.
2019	Nguyen *et al*. [27]	76.43	77.18	75.65	Semi.
2019	Zhu *et al*. [49]	75.20	81.02	79.0	Weak.
2020	Degardin *et al*. [6]	77.14	84.60	76.90	Weak.
2020	Ramachandra *et al*. [34]	73.81	82.45	75.46	Semi.
2020	Pang *et al*. [28]	76.78	84.65	78.50	Unsuper.
2021	Feng *et al*. [9]	**81.43**	**85.19**	**82.30**	Weak.
2021	Kopuklu *et al*. [15]	74.90	79.63	75.12	Weak.
2022	Doshi *et al*. [8]	75.86	80.71	79.46	Semi.
2022	Park *et al*. [29]	73.28	77.23	75.40	Unsuper.
2022	Leroux *et al*. [17]	76.20	78.06	76.78	Unsuper.
	Ours (Farneback Flow)	79.31	84.12	81.40	Unsuper.
	Ours (SelFlow [20])	**81.01**	**86.31**	**84.50**	Unsuper.

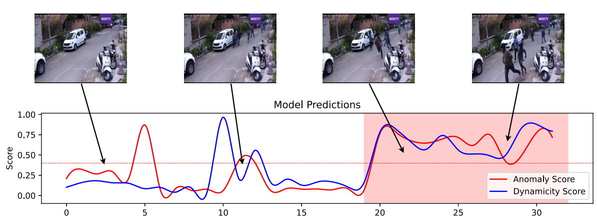

Figure 1. **Results Visualization**: Qualitative results on test videos taken from the CCTV-Fight [31] dataset. Each image represents a frame in a temporal segment. The shaded portions are ground truths and the horizontal line represents the threshold.

Fig. 2 depicts a fighting scene from UBI-Fights [6] dataset. When the fight started, both regressors have produced high scores. Moreover, they have generated a very low score for the normal (post-fighting scene) scene.

The UCF-Crime [38] dataset contains a large number of normal videos. We visualize the prediction of the proposed model to study FAR. Fig. 3 depicts two categories of UCF-Crime: normal activity and road accidents. Note that, no shaded portion (ground truth) is produced for normal activity. It can be observed that the Ω regressor is not generating any anomaly score. However, ψ is detecting the activities

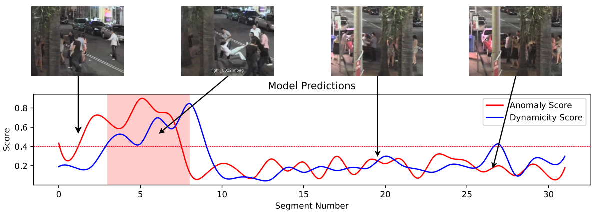

Figure 2. **Results Visualization**: Qualitative results on test videos taken from the UBI-Fight [6] dataset.

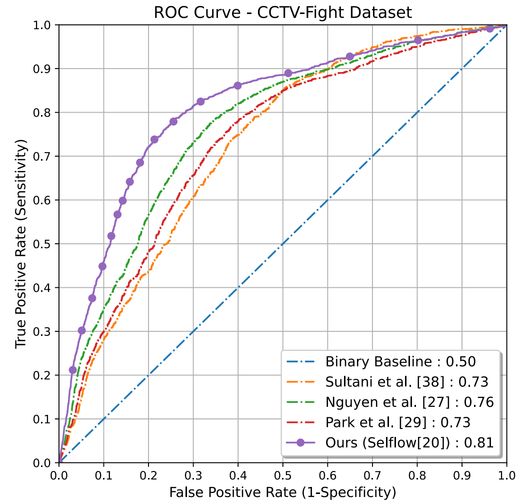

Figure 4. ROC results on CCTV-Fight [31] dataset.

such as cars and pedestrians moving, resulting higher dynamicity scores.

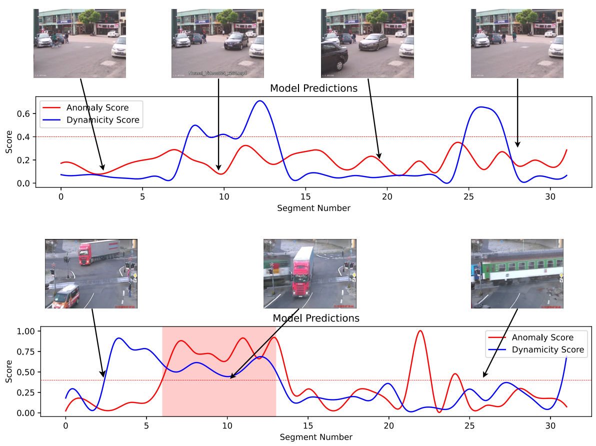

Figure 3. **Results Visualization (Normal Activity and Road Accident)**: Qualitative results on test videos taken from the UCF-Crime dataset.

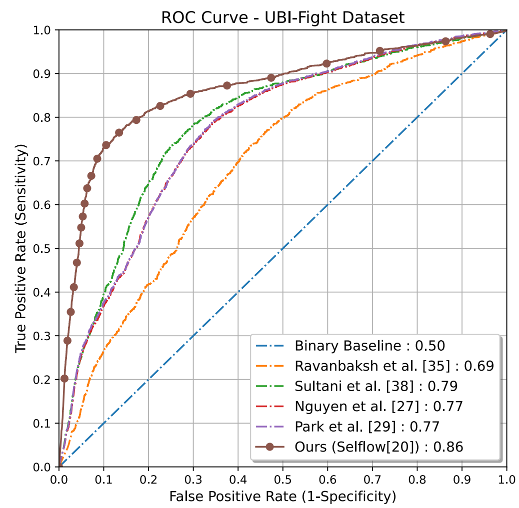

Figure 5. ROC results on UBI-Fight [6] dataset.

Both regressors agree on the ground truths for the road accident category. We have observed similar prediction performance on anomaly types such as shooting, burglary, vandalism, abuse, explosion, assault, etc.

5. ROC Curves on Three Datasets

All the references mentioned in this document are taken from the original paper.

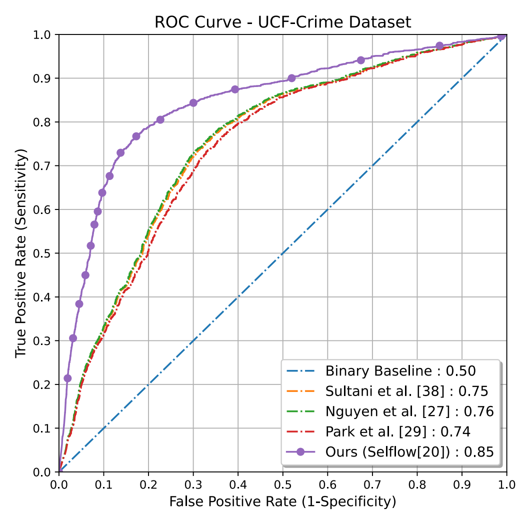

Figure 6. ROC results on UCF-crime [38] dataset.