Supplementary Material:

The visualizations for test cases where the model performed well and poorly in detecting anomalous frames, on the ShanghaiTech dataset. (4 random cases in total).

Good Performance:

Example #1

Fig.1(a): (Video_ID#01_0053) Ground truth labelled frames. Red bounding box implies anomaly.

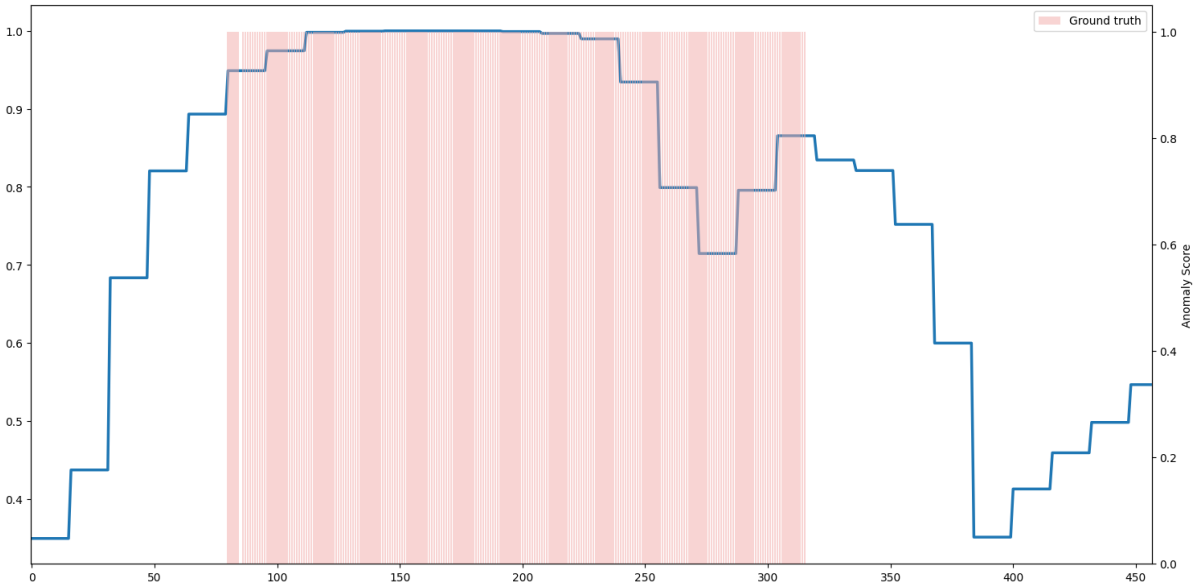

Fig.1(b): Blue plot represents the model predicted anomaly score for each frame.

In this first example, anomaly occurs as a rickshaw passes through 'only walking zone'. We can see that the ground truth anomaly scores (horizontal top edge of red box in Fig.1.(b)) and high predicted abnormal scores are overlapping mostly. Suggesting good performance of the model.

Example #2

Fig.2(a): (Video_ID#06_0153) Ground truth labelled frames. Red bounding box implies anomaly.

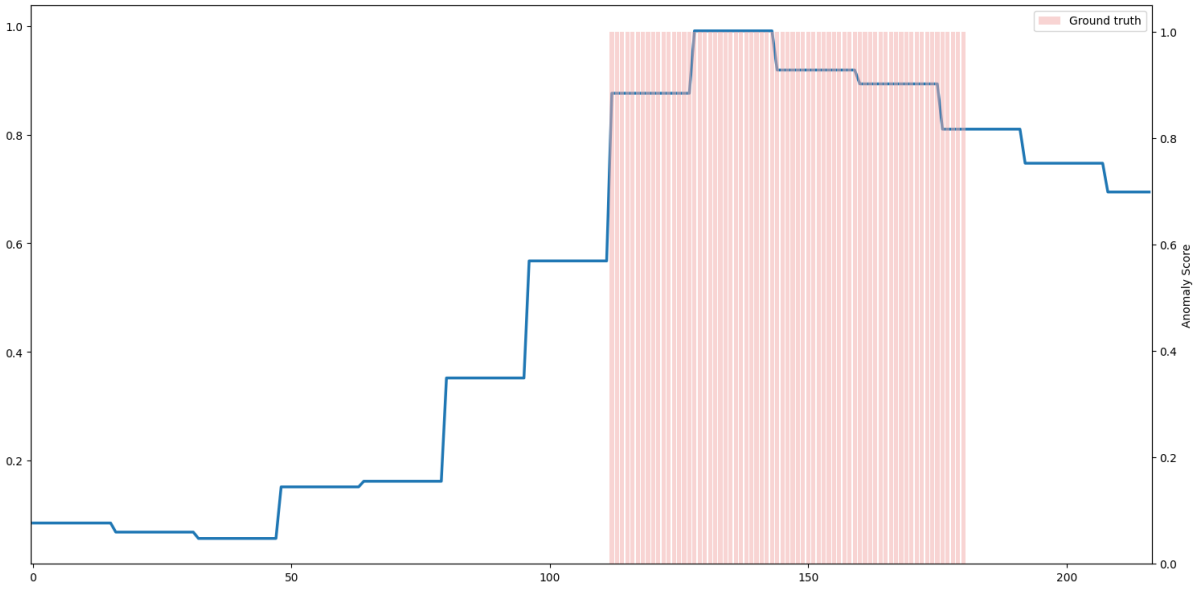

Fig.2(b): Blue plot represents the model predicted anomaly score for each frame.

In this example, anomaly occurs as a bicycle passes through a pedestrian zone. And the model gives high anomaly scores for all of the ground truth.

Bad Performance:

Below are the cases where the model performed poorly for some test cases. The model gave poor results when the scooter/cycle was far away. (i.e. very tiny in the image, even normal humans would find it difficult to detect it). However, the ground truth labeled a frame as anomalous, even if only part of the scooter is visible and far away from the camera viewpoint.

We can infer that model isn't able to detect faraway and occluded objects that causes Anomaly.

Example #3

Fig.3(a): (Video_ID#10_0042) Ground truth labelled frames. Red bounding box implies anomaly.

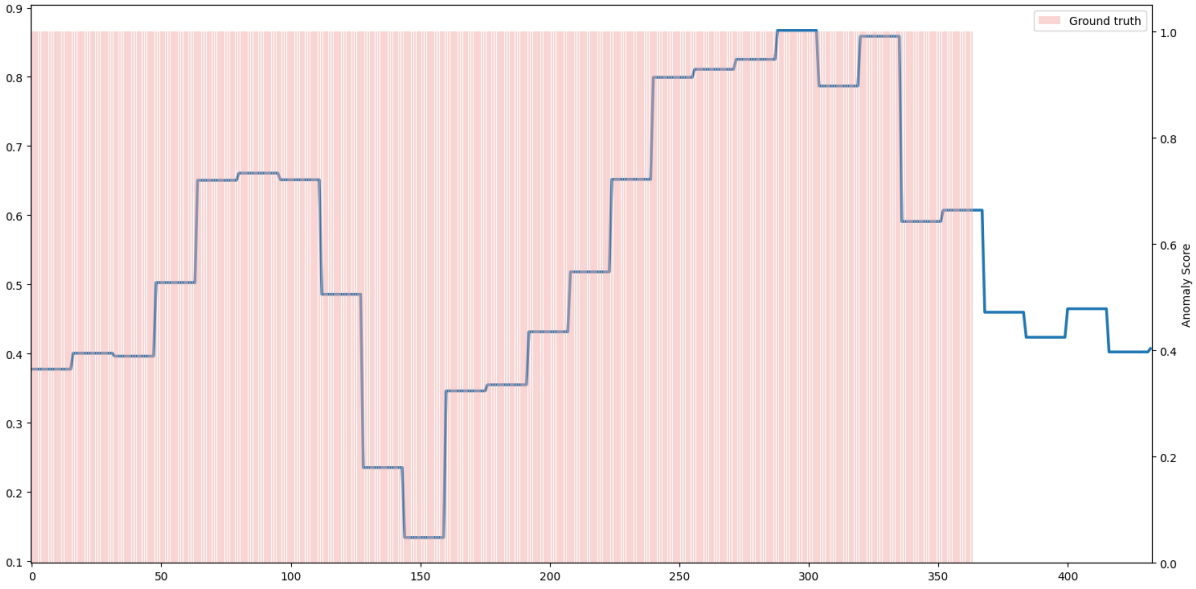

Fig.3(b): Blue plot represents the model predicted anomaly score for each frame.

In Fig.3.b) Its clearly visible that for frames below 250 (especially frame-150), the model predicts low anomaly scores whereas the ground truth score for all these frames are 1.

Example #4

Fig.4(a): (Video_ID#12_0174) Ground truth labelled frames. Red bounding box implies anomaly.

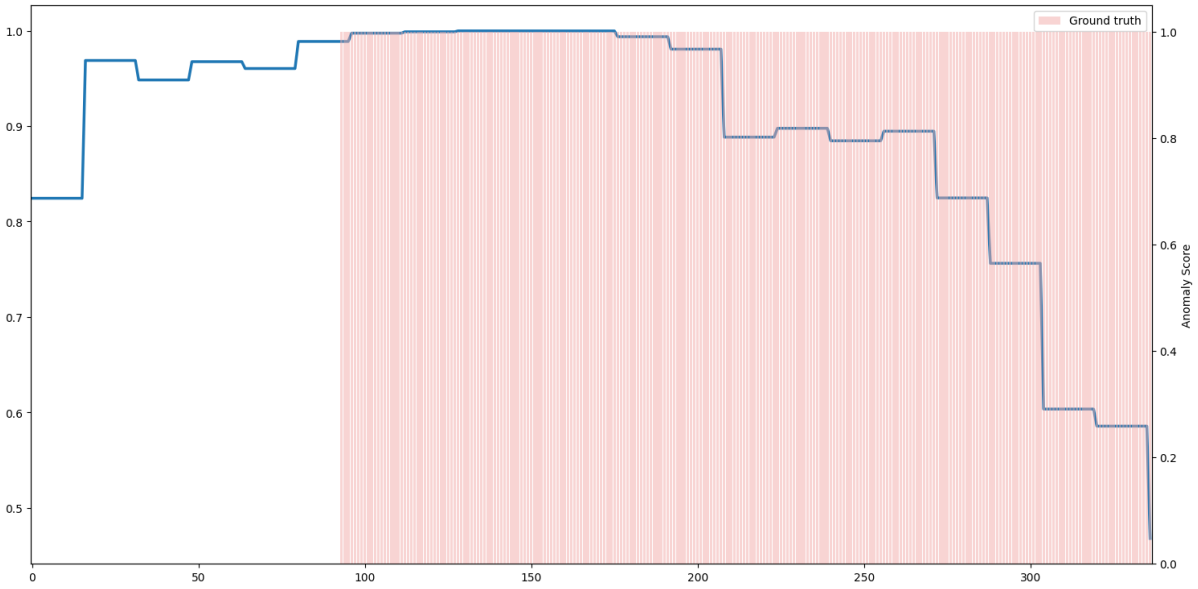

Fig.4(b): Blue plot represents the model predicted anomaly score for each frame.

In Fig.4(b) for higher frames(frames>200), the anomaly scores are low. From Fig.4.(a) it is clear that the bike leaves further away from camera as the video progresses.