

SVD-GAN for Real-Time Unsupervised Video Anomaly Detection

1. INTRODUCTION

The supplementary material includes the analysis and results obtained from the proposed SVD-GAN architecture for anomaly detection. The material elaborates the efficiency of trained model parameters, stable and robust reconstructed images, Receiver Operating Characteristic (ROC) Curve, Area under roc Curve (AUC) and Equal Error Rate (EER). The document includes the additional sample of images for Figure 4 in the main paper, detailed information about the number of anomalous and non-anomalous frames used for training and testing, and a graph plotting the regularity score for datasets, supporting the arguments in section 4.2 of the main paper.

2. REDUCTION IN PARAMETERS

The generator of the SVD-GAN incorporates a depthwise separable convolutions for feature extraction which improves the computational efficiency of the model. The conceptual idea of depth-wise separable convolution is shown in Figure S1.

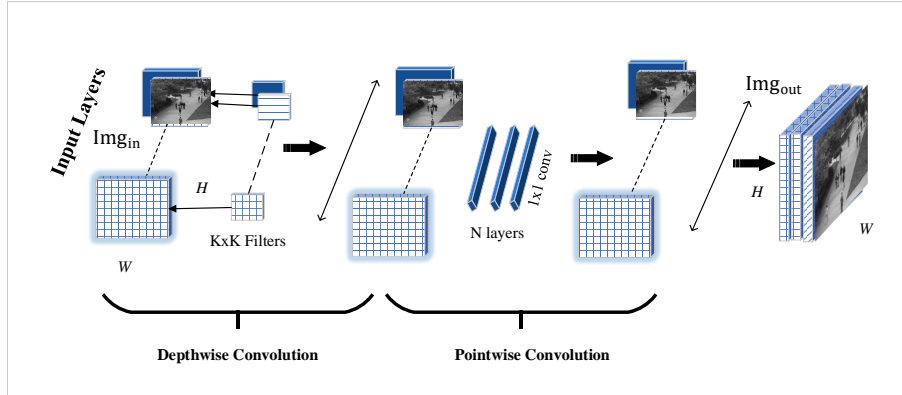


Fig. S1. Structure of depth-wise separable convolution layers illustrating the convolutional and pointwise operations

Here, a spatial convolution is performed independently over every channel of an input, followed by a pointwise convolution. Thus, depthwise separable convolution reduces the number of parameters and computations in convolutional operations while increasing representational efficiency. The usage of depthwise separable convolutions also improves convergence of GAN by considerably reducing size of the model. As discussed in the main paper, the parameters in the encoder is reduced by 4.3%, decoder by 3%, compared to 2D convolution as tabulated in Table S1.

3. ROBUST RECONSTRUCTION

It is crucial for the GAN architecture to learn spatio-temporal feature representations from the input sample space and reconstruct/generate images similar to real ones. To achieve this, a SVD loss is proposed in our architecture for stable reconstruction. Figure S2 shows the robustness of our proposed SVD-GAN in reconstructing images from different source of cameras, angles, scenes and lightning environment.

4. EVALUATION

The proposed system is evaluated with the benchmark datasets namely UCSD dataset, CHUK Avenue dataset and ShanghaiTech Campus dataset. The end-to-end system is evaluated frame-

Layers	Kernel Size	Output Size	Parameters
Encoder			
Depthwise Layer1	5x5	10,64,64,128	281
Depthwise Layer2	3x3	10,32,32,64	9408
ConvLSTM 1	3x3	10,16,16,64	295168
ConvLSTM 2	3x3	10,16,16,32	110720
ConvLSTM 3	3x3	10,16,16,64	221440
ConvLSTM 4	3x3	10,16,16,128	885248
Decoder			
Conv2DTranspose	3x3	10,32,32,64	73792
Conv2DTranspose	5x5	10,64,64,128	204928
Conv2DTranspose	5x5	10,128,128,128	409728
Depthwise Layer	5x5	10,128,128,1	3329

Table S1. Reduced model parameters of our proposed architecture



Fig. S2. Sample of reconstructed images for complex anomaly detection

level for any occurrence of anomalous events. These datasets consists for both normal frames and complex anomalous frames for testing and training the model. The detailed information about the anomalous events, source of data and number of frames are tabulated in Table S2.

The supplementary material also includes the anomaly scores and ROC curves obtained for the benchmark datasets. The Receiver Operator Characteristic (ROC) curve is an evaluation metric for finding the thresholds for anomaly detection. It is a probability curve that plots the TPR against FPR at various threshold values and essentially separates the anomalous and non-anomalous data. Finally, Area Under the Curve (AUC) measure is used to distinguish frames with a summary of the ROC curve.

The frame level evaluation of testing data consisting of anomalous and non-anomalous events is visualised in graphs plotted against the anomaly score. Figure S4 illustrates the anomaly detection graph with false positives and true positives for ShanghaiTech dataset. In addition, few more graphs visualised for datasets UCSD, Avenue and ShanghaiTech datasets is shown in Figure S3.

Benchmark Datasets	Anomalous Events	Sources	Normal Frames	Anomalous Frames	Training Frames	Testing Frames
UCSD Ped1	40	1	9995	4005	6800	7200
UCSD Ped2	12	1	2924	1636	2550	2010
CHUK Avenue	47	1	26832	3820	15328	15324
ShanghaiTech	130	13	300308	17090	274515	42883

Table S2. Complexity of the datasets

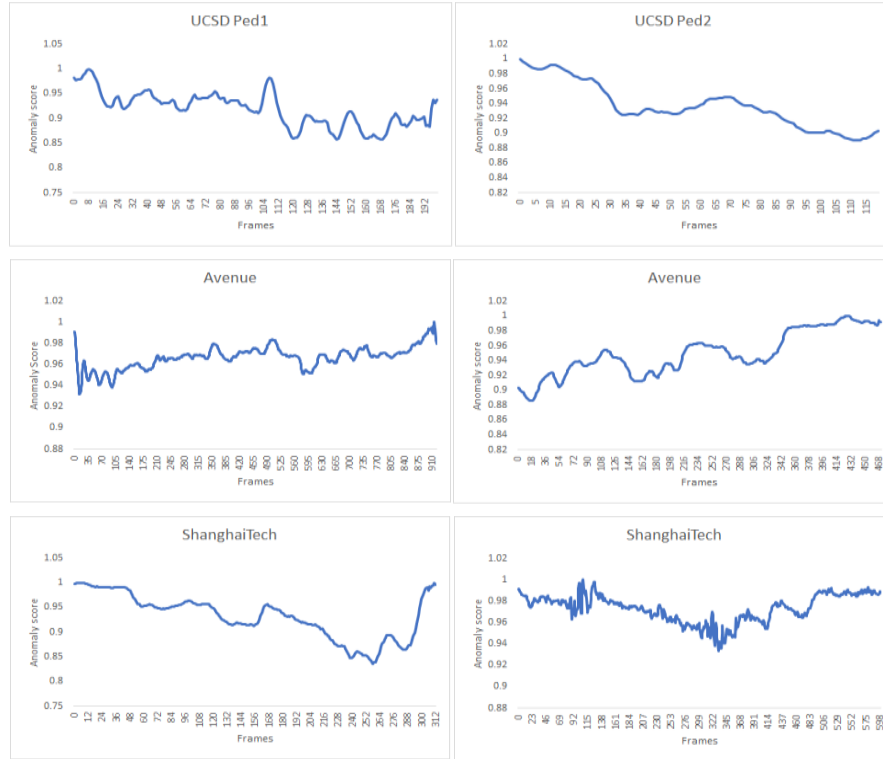


Fig. S3. Visualisation of anomaly scores for benchmark datasets

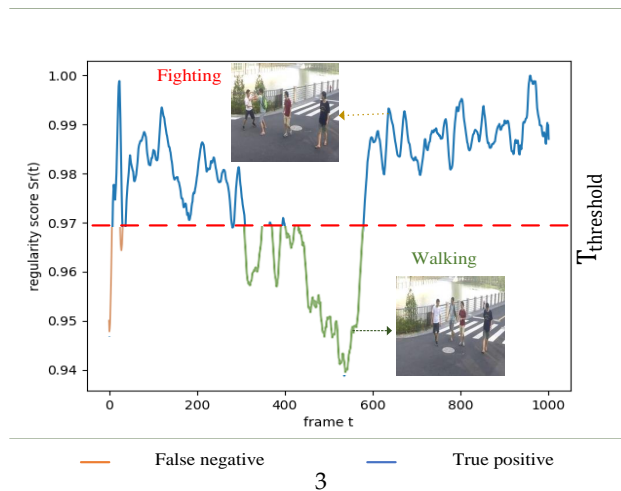


Fig. S4. A sample graph showing variations in anomaly score for a test data having normal and anomalous frames