Anomaly Detection in Video Sequence with Appearance-Motion Correspondence

Abstract

- Learning a correspondence between common object appearances (e.g. pedestrian, background, tree, etc.) and their associated motions.

- A combination of a reconstruction network and an image translation model that share the same encoder.
    - a reconstruction network : the most significant structures that appear in video frames
    - an image translation model : associate motion templates to such structures.
- The training stage is performed using only videos of normal events and the model is then capable to estimate frame-level scores for an unknown input.

Introduction