Online Spatio-Temporal Action Detection in Long-Distance Imaging Affected by the Atmosphere

by

Eli Chen (Prof. Yitzhak Yitzhaky)

Current state-of-the-art approaches for spatio-temporal action detection deal with stable videos and quite sterilized environments, as seen in the UCF-101 benchmark. In addition, the objects of interest are typically relatively close to the camera, and therefore fairly clear and easily distinguished. This study presents an approach method for online human action detection in long-distance imaging affected by atmospheric distortions. We created a unique dataset of typical actions in long-range imaging. Various CNN frameworks were examined for the initial moving object detection phase, including 2D, 3D, one stream, and two-stream (RGB frames and optical flow). The basic object detection methods examined within these frameworks include the YOLOv3 and an extension of the inflated 3D ConvNet with a Feature-Fused Single Shot Multibox Detector (FFSSD) to improve small object detection. To cope with the harmful effect of the spatio-temporal random movements induced by atmospheric effects on motion estimation, we first fit the optical flow stream characteristics to a temporally noisy turbulent environment. A significant improvement of the action detection quality under such noisy conditions was obtained by constructing an online tracking algorithm that incrementally constructs and labels the objects' tracks from the network's frame-level detections. Experimental results show that our approach outperforms the state-of-the-art on our dataset in terms of the mAP measure.