

## Supplementary Material for Weakly-Supervised Multi-Person Action Recognition in 360° Videos

Visualizations of person detection results using Mask-RCNN on different panoramic frames. There exist some inaccurate or missing detections, which can mostly be corrected by temporal max-pooling of the detection masks from adjacent frames.

