

Spatio-temporal action localization consists of three levels of tasks: spatial localization, action classification, and temporal segmentation. In this work, we propos