

通过类别特定帧聚类增强动作显著性的 弱监督时序动作检测

夏惠芬^{1,3}, 詹永照^{1,2}, 刘洪麟¹, 任晓鹏¹

¹ 江苏大学计算机科学与通信工程学院, 中国镇江市, 212013

² 大数据泛在感知与智慧农业应用工程研究中心, 中国镇江市, 212013

³ 常州机电职业技术学院, 中国常州市, 213164

摘要: 时序动作检测任务是指在未裁剪的视频中检测出动作的开始时间和结束时间, 并对动作实例进行分类。随着视频中动作类别的增多, 现有仅提供视频级别标签的弱监督时序动作检测方法已无法提供足够的监督。单帧标注方法引起了人们兴趣。但现有单帧标注方法仅从视频片段序列的角度对标注的单帧建模, 而忽略了标注单帧的动作显著性, 并且没有充分考虑它们在同一动作类别中的相关性。考虑到在同一动作类别中, 带标注的单帧能表现出独特的外观特征和清晰的动作模式, 本文提出一种新颖的通过类别特定帧聚类来增强动作显著性的弱监督时序动作检测方法。该方法采用 K -均值聚类算法对同一动作类别的帧聚合, 将其作为该动作类别的特征表示。通过计算每帧与各个动作类别之间的相似度, 得到类激活分数。特定于类别的单帧表征建模可以为主线中的视频片段序列建模提供补充性的指导。因此, 针对标注的帧和其对应的视频片段序列, 提出凸组合融合机制, 用于增强动作显著性的一致性特性, 从而生成更加鲁棒的类激活序列, 进行精确的动作分类和动作定位。由于动作显著性增强的补充指导, 该方法优于现有的基于单帧标注的动作检测方法。在 THUMOS14、GTEA 和 BEOID 3 个数据集上进行的实验表明, 所提方法具有更高的检测性能。

关键词: 弱监督; 时序动作检测; 单帧标注; 类别特定; 动作显著性

<https://doi.org/10.1631/FITEE.2300024>