
心理所提出长视频中多尺度检测微表情片段的卷积神经网络

作者：writer 来源：中国科学院

本文原地址：<https://www.iikx.com/news/progress/13450.html>

本文仅供学习交流之用，版权归原作者所有，请勿用于商业用途！

微表情是一种短暂微小的面部表情，通常出现在个体试图隐藏真实感受的时刻。微表情的分析具有潜在的应用价值，如在医疗关怀等领域。与常见的普通表情相比，微表情有三个显著特征，即持续时间短（<500ms）、强度低和局部运动。人类用肉眼较难发现和识别短暂而微弱的表情，因而需要借助计算机视觉来开展相关分析。

微表情检测的目标是定位微表情在视频中发生的时刻，这一研究具有实用价值。若能在一段视频中准确地检测和定位到某个时间点有微表情出现，那么就说明这个人在这个时刻或有异常。然而，在真实场景中，从大量头部动作和人脸表情中检测分离出微小短暂的微表情是极具挑战性的任务。因此，目前微表情检测的研究相对较少。同时，由于微表情数据库的样本量较少，限制了深度学习技术与微表情检测方法的结合。近年来，更适合用于微表情检测研究的长视频数据集CAS (ME)2和SAMM已发布。研究长视频中微表情的检测方法将有助于实现微表情分析在真实场景的应用。

受平面人脸检测技术启发，中国科学院心理研究所行为科学重点实验室黄昌兵研究组尝试将相应技术扩展至三维时-空上的人脸微表情检测，提出一个多尺度的卷积神经网络（CNN）：MESNet。MESNet包含一个两阶段的预测框架和几种针对微表情检测任务设计的特殊的架构和策略。研究设计出一个基于微表情识别任务的轻量级的2+1D-时空卷积网络，通过分别提取空间和时间上微表情特征来区分固定长度的视频片段是否为微表情或非微表情样本（图1）。在MESNet中，2+1D-时空卷积网络被用于提供主干网络和预先训练的参数。此外，为了更准确地检测微表情片段，研究增加了一个片段候选网络（Clip Proposal Network, CPN）和一个分类回归网络（CRNet）。CPN是一个全卷积网络，它被添加到2+1D-时空卷积网络预训练的卷积层中，并将在长视频中所有可能的微表情区间列入候选（图2）。CRNet模块对候选视频片段进行进一步判断，并回归它们的时间边界，从而得出最终的微表情片段检测结果（图3）。同时，为了提升网络的检测性能，视频片段的光流特征作为高级特征被输入到MESNet中。

此外，研究还改进了之前的微表情检测结果的评估方法，重新定义了更适应基于微表情区间检测的指标。在最近发布的长视频数据集CAS (ME)2和SAMM中进行实验，并使用留一交叉验证法来评估定位性能。对比结果表明，该MESNet算法的微表情片段检测性能表现优异，特别是在SAMM数据集上，其检测结果明显优于其它已发表的方法。

相关研究结果在线发表在IEEE Transactions on Image Processing (DOI: 10.1109/TIP.2021.3064258) 上。研究工作得到国家自然科学基金项目、中国博士后科学基金及国家重点研发计划重点专项的资

助。

图1.2+1D-时空卷积网络

图3.分类回归网络 (CRNet)

研究团队单位：心理研究所

更多 科学进展 请访问 <https://www.iikx.com/news/progress/>

本文版权归原作者所有，请勿用于商业用途，[爱科学iikx.com](http://www.iikx.com)转发