

# 基于双流卷积与双中心 loss 的行为识别研究

毛志强<sup>1</sup>，马翠红<sup>1</sup>，崔金龙<sup>2</sup>，王毅<sup>1</sup>

(1 华北理工大学 电气工程学院, 河北 唐山 063210; 2 北京交通大学 海滨学院, 河北 沧州 061100)

**摘要:** 针对行为视频中相似动作类内差异大、类间差异小, 识别准确率不高的问题, 提出了一种基于双流卷积网络与双中心 loss 的行为识别方法. 该方法首先构建双流卷积网络结构, 以 C3Dnet 模型作为双流结构的基础模型, 分别提取多尺度 RGB 视频帧中的表观短时运动信息和堆叠光流图中的长时运动信息; 然后将双流结构提取的深度信息经长短时记忆(LSTM)网络解析后进行特征融合; 最后, 利用基于双中心 loss 的 2C-softmax 目标函数, 来最大化类间距离和最小化类内距离, 从而实现相似动作的分类与识别. 在数据集 KTH 上的实验结果表明, 该方法能够准确识别相似动作, 识别准确率可达 98.2%, 具有很好的识别效果.

**关键词:** 双流卷积网络; 中心 loss; 长短时记忆(LSTM); 光流图

## Research on action recognition based on two-stream convolution and double center loss

MAO Zhi-qiang<sup>1</sup>，MA Cui-hong<sup>1</sup>，CUI Jin-long<sup>2</sup>，WANG Yi<sup>1</sup>

(1 College of Electrical Engineering, North China University of Science and Technology, Tangshan 063210, China;

2 Beijing Jiaotong University Haibin College, Cangzhou 061100, China)

**Abstract:** Aiming at the problem of large difference in similar action classes, small difference between classes in action video and low recognition accuracy, a action recognition method based on two-stream convolution network and double-center loss is proposed. The method first constructs a two-stream convolutional network structure, and uses the C3Dnet model as the basic model of the two-stream structure to extract the apparent short-term motion information in the multi-scale RGB video frame and the long-term motion information in the stacked optical flow map respectively; Then, the depth information extracted by the two-stream structure is parsed by a long and short time memory (LSTM) network to perform feature fusion; Finally, the 2C-softmax objective function based on dual-center loss is used to maximize the distance between classes and minimize the distance within the class, so as to classify and identify similar actions. The experimental results on the data set KTH show that the method can accurately identify similar actions, and the recognition accuracy can reach 98.2%, which has a good recognition effect.

**Key words:** two-stream convolution network; center loss; long short-term memory (LSTM); optical flow map

**作者简介:**

毛志强 男, (1991-), 硕士研究生. 研究方向为计算机视觉、目标检测与人体行为分析.

马翠红(通讯作者) 女, (1960-), 教授. 研究方向为复杂工业系统的建模与控制、图像识别与视频分析. E-mail: 864404484@qq.com.

崔金龙 男, (1989-), 硕士, 助教. 研究方向为钢成分测量.

王毅 男, (1994-), 硕士研究生. 研究方向为计算机视觉、目标检测与视频分析.