

融入时序激励机制的人体行为识别

梁 鸿, 张兆雷⁺, 李传秀, 钟 敏

(中国石油大学(华东) 计算机科学与技术学院, 山东 青岛 266580)

摘 要: 为解决现有人体行为识别方法不能有效融合视频局部特征的问题, 提高行为识别的准确率, 对视频的一般特点进行分析, 提出一种时序激励机制。以 BN-Inception 作为基础模型分别提取 RGB 图像序列与光流图像序列的特征, 在模型中嵌入时序激励模块实现对视频局部特征序列的动态加权, 突出局部特征序列中对行为识别有益的特征的作用, 使融合得到的全局特征更具鉴别力。在数据集 HMDB51 和自建的油田生产现场行为识别数据集 OilField-7 上进行实验, 准确率分别可达 71.6% 和 92.8%, 验证了所提方法的有效性。

关键词: 行为识别; 局部特征; 时序激励机制; 加权; 融合

中图分类号: TP391.4 **文献标识码:** A **文章编号:** 1000-7024 (2020) 10-2907-06

doi: 10.16208/j.issn1000-7024.2020.10.035

Human action recognition with temporal excitation mechanism

LIANG Hong, ZHANG Zhao-lei⁺, LI Chuan-xiu, ZHONG Min

(College of Computer Science and Technology, China University of Petroleum (East China), Qingdao 266580, China)

Abstract: To deal with the problem that the existing human action recognition method cannot effectively integrate local features of the video and to improve the accuracy of human behavior recognition, the general characteristics of the video were analyzed, and a temporal excitation mechanism was proposed. BN-Inception was used as the basic model to extract the features of RGB image sequences and optical flow image sequences. The temporal excitation model was embedded to realize the dynamic weighting of the video local feature sequences, highlighting the features of local feature sequences that were useful for action recognition. The role of the fusion of the global features was more discriminating. Experiments were carried out on the dataset HMDB51 and the self-built oilfield production action recognition dataset OilField-7. Results show that the recognition accuracy is up to 71.6% and 92.8%, and the validity of the proposed method is then verified.

Key words: action recognition; local feature; temporal excitation mechanism; weighting; fusion

0 引 言

基于视频的人体行为识别是计算机视觉领域的热点和富有挑战性的问题。由于深度卷积神经网络强大的特征学习能力, VGGNet^[1]、ResNet^[2] 等优秀的网络模型被应用于人体行为识别任务中。

将深度卷积神经网络应用到人体行为识别中的关键问题是如何融合视频中的局部特征。Karpathy 等^[3] 在人体行为识别任务中引入深度卷积神经网络并研究了连续视频帧作为输入的特征融合方法。Simonyan 等^[4] 把光流作为表示运动的特征提出了双流卷积神经网络, 在行为识别任务上

具有重要的意义。双流卷积神经网络虽然有效地利用了时间维度信息, 但是对长时序的运动缺乏特征表示, 导致最后识别准确率并不高。Wang 等^[5] 提出基于长时间范围结构建模的网络改进了传统双流 CNN 的缺点, 但是各视频段特征的简单平均融合使不同的特征发挥相同的作用。Ng 等^[6] 针对视频的时序问题设计了双流长短时记忆网络, 但是效果不理想。

针对以上分析, 本文总结了视频的一般特性: ①相邻的帧之间具有高度相似性; ②某些行为顺序的改变不会对识别结果产生明显影响; ③某些行为只用少数几帧就可识别。因此本文提出了时序激励卷积神经网络 (temporal ex-

收稿日期: 2019-07-11; 修订日期: 2019-10-21

基金项目: 中央高校基本科研业务费基金项目 (18CX02138A)

作者简介: 梁鸿 (1966-), 男, 山东东营人, 博士, 教授, 研究方向为计算机图像处理、计算机视觉; 十通信作者: 张兆雷 (1995-), 男, 山东德州人, 硕士研究生, 研究方向为计算机视觉、人体行为分析; 李传秀 (1979-), 女, 山东临沂人, 博士, 讲师, 研究方向为计算机图形图像处理; 钟敏 (1978-), 女, 四川广汉人, 博士, 讲师, 研究方向为计算机视觉。E-mail: 549373834@qq.com

citation based convolutional neural network, TEN), 通过时序激励模块 (temporal excitation model, TE Model) 建模局部特征之间的关系, 提高人体行为识别的准确率。时序激励就是对局部特征的加权, 使对行为识别有益的特征发挥更大的作用同时抑制冗余和噪声特征。另外, 与近期提出的基于注意力机制的方法^[7]不同的是, 本文方法不仅利用单个局部特征, 还在一定程度上利用了时序上不同位置局部特征的关系信息。

1 时序激励卷积神经网络

本文在提出时序激励机制时将人体行为识别任务分为 3 个主要阶段: 局部特征提取阶段、局部特征融合阶段和全局特征分类阶段, 每个阶段都可以通过合适的神经网络来解决。其中局部特征提取使用现有的深度卷积神经网络,

例如 ResNet、BN-Inception 等, 全局特征分类可以用全连接层。本文主要贡献在于局部特征融合阶段, 即探索如何在给定一组局部特征的情况下生成具有辨别性的视频全局特征表示。为此本文设计了时序激励模块, 并将其融入到卷积神经网络中, 此模块能根据输入的局部特征序列自适应地为每一个局部特征分配权重。对行为识别有益的局部特征会得到比较大的权重, 而冗余和噪声特征得到较小的权重, 以此达到增加有益特征的作用并且抑制无用特征的目的。时序激励卷积神经网络模型结构如图 1 所示, 模型的优势在于依然能够进行端到端的训练, 这为整个模型的优化提供了方便。本文采用空间流与光流结合的方法进行人体行为识别, 其中的每一个流都能输入到提出的方法中优化学习, 最后的预测结果是由两流上预测的行为类别分数加权融合得到。

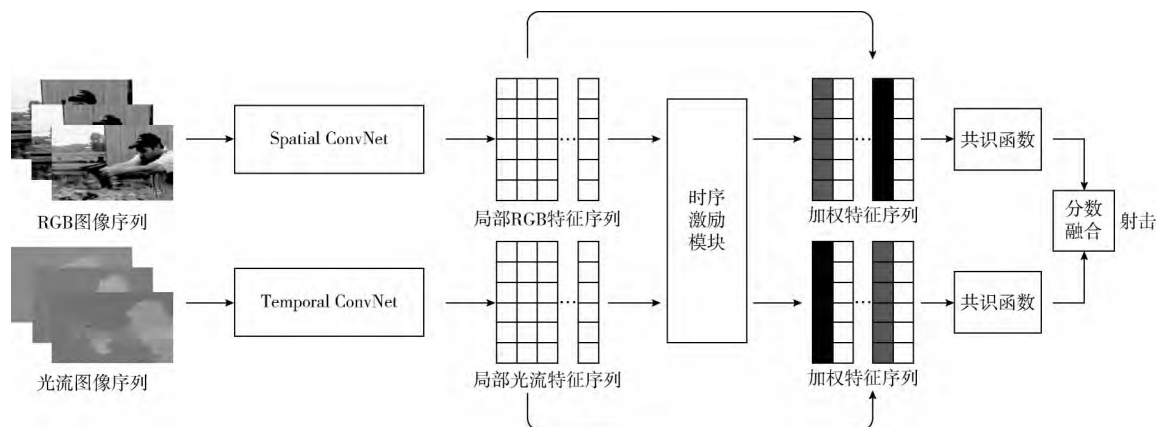


图 1 时序激励卷积神经网络模型结构

将时序激励深度卷积神经网络 (TEN) 应用于人体行为识别可以分为以下 6 个步骤:

步骤 1 将视频平均分成若干片段, 在每个片段中随机挑选一帧 RGB 图像和光流图像的堆叠作为模型的输入;

步骤 2 利用深度卷积神经网络提取每一帧 RGB 图像或光流图像的堆叠深度特征并且组成局部特征序列;

步骤 3 将局部特征序列输入到设计的时序激励模块中, 生成与局部特征相应的权重向量;

步骤 4 将生成的权重向量赋值给第 2 步中的局部特征序列, 完成局部特征的加权;

步骤 5 分别将两个流的加权特征序列输入到共识函数中, 生成视频的全局特征表示;

步骤 6 将空间流与光流的行为类别得分加权融合得到最终的视频行为类别预测。

2 局部特征序列的提取

本文提出的时序激励深度卷积神经网络 (TEN) 结构如图 1 所示。为了得到视频中的局部特征序列, 将每一段

带有标签的输入视频 V 平均分为 K 段, 即 $V = \{S_1, S_2, \dots, S_K, y\}$, S_i 代表视频中的第 i 个片段, y 是此视频包含的行为类别标签。提取视频的局部特征采用的深度卷积神经网络是在 ImageNet 图像分类数据集上预训练过的 BN-Inception, 因为 Wang 等^[5]已经验证了 BN-Inception 能在速度与准确率之间取得比较好的平衡。假设在视频中每个片段随机提取的帧 T_i 输入到卷积神经网络后映射成的深度特征可表示为

$$x_i = \mathbf{F}(T_i, W), i = 1, 2, \dots, K \quad (1)$$

其中, W 为卷积神经网络 \mathbf{F} 的参数, x_i 为 M 维的向量。对于 RGB 流, $T_i = T_i^{RGB}$ 并且 $x_i = x_i^{RGB}$; 对于光流, $T_i = T_i^{FLOW}$ 并且 $x_i = x_i^{FLOW}$ 。

局部特征序列定义为同一视频中不同部分的帧经卷积神经网络所映射的一组深度特征的序列。为了方便, 本文使用一个 $M \times K$ 矩阵 X 来表示包含 K 个局部特征序列

$$X = (x_1, x_2, \dots, x_K) \quad (2)$$

值得注意的是, 本文中提取的局部特征是有序的, 即局部特征的顺序须和对应的视频帧顺序一致。

3 时序激励模块

本文设计的时序激励模块的结构如图2所示。本文提出的时序激励模块中包含对局部特征序列的压缩与激励操作, 并且嵌入在网络的最后一个全连接层之后。

在局部特征序列输入到时序激励模块后, 首先进行特征的压缩操作。本文采用与 SENet^[8]中一样的深度特征压缩方法, 即平均池化方法, 如式(3)所示

$$a_i = \frac{1}{M} \sum_{j=1}^M x_{ji}, i = 1, 2, \dots, K \quad (3)$$

其中, a_i 表示视频中第 i 个局部特征经过压缩后的输出值, x_{ji} 表示局部特征序列中第 i 个局部特征向量 x_i 的第 j 个元素。因此, 能够表示时序上局部特征依赖性的特征向量可以写为

$$A = (a_1, a_2, \dots, a_K) \quad (4)$$

在局部特征融合的过程中, 时序激励模块的输入是代表局部特征序列的向量。它的任务是分析每个局部特征的重要性, 然后赋予全局特征尽可能多的有用信息, 同时抑制冗余的信息和噪声。这种时序激励输出可以看作是局部特征序列上的加权, 加权之后的视频局部特征序列 V 可表示为

$$V = cX \quad (5)$$

其中, c 是 K 维的权向量, 由权重函数决定。权重函数的选择是时序激励模块的关键所在。它的输入是能够代表局部

特征序列的向量 A , 输出是权向量 c , 权向量每一维的数值都对应其所在位置的局部特征。

使用权重函数计算局部特征权重的方法很多, 全局平均可以看作是时序激励机制的一种退化形式, 相应的权重函数可以表示为

$$c = \frac{1}{K} L \quad (6)$$

其中, L 是维度为 K 并且所有元素都等于 1 的向量。Lippmann 在 1987 年的论文^[9]中发现具有两个隐藏层的多层感知器足以表示任何所需形状的分类区域。受此启发, 对于更具延展性的时序激励权重函数, 本文首先使用包含一个隐藏层的全连接层 (FC2)

$$c = \text{sigmoid}(\omega_2 \text{relu}(\omega_1 A^T + b_1) + b_2) \quad (7)$$

其中, ω_1, ω_2 分别为维度 $5 * K$ 的参数向量。类似地, 也可以使用包含两个隐藏层的全连接层 (FC3)

$$c = \text{sigmoid}(\omega_3 \text{relu}(\omega_2 \text{relu}(\omega_1 A^T + b_1) + b_2) + b_3) \quad (8)$$

其中, ω_1, ω_2 是维度 $2 * K$ 的参数向量, ω_3 是维度 K 的参数向量。

视频中的局部特征经过时序激励模块之后输出经加权的局部特征序列进入到全局特征融合的阶段, 此处本文选择平均融合的方法将 K 个不同位置的局部特征序列 V 融合为代表整个视频的全局特征, 最后将全局特征送入到时序激励模块之后的分类器中进行优化学习。

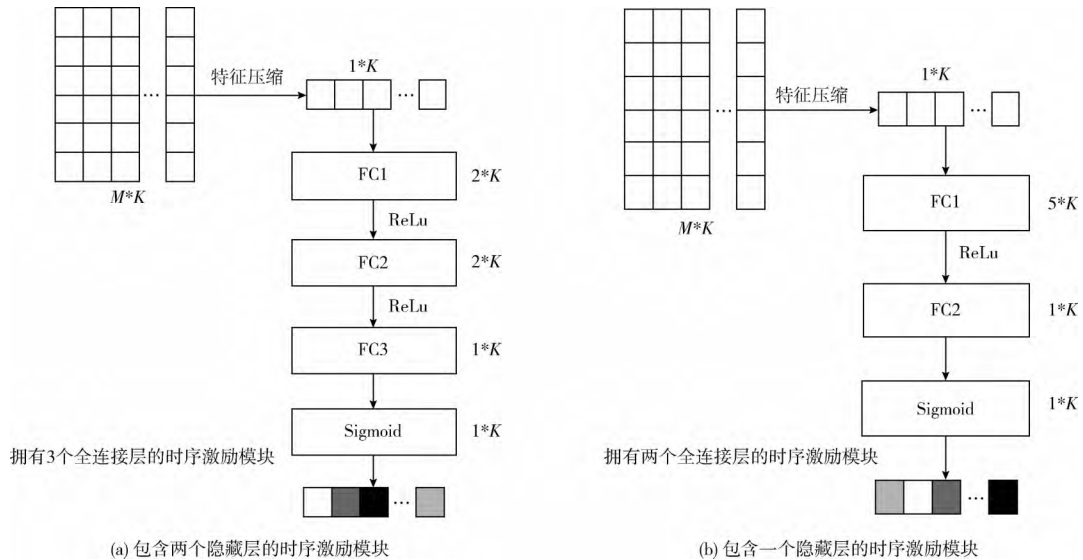


图2 时序激励模块结构

4 实验及结果分析

首先介绍了本文使用的两个人体行为识别数据集:

HMDB51^[10]和自建的油田生产现场行为识别数据集 Oil-field-7, 并且在这两个数据集上评估了本文提出的方法。通

过与一些优秀方法的比较, 验证了本文方法的有效性, 最后对时序激励模块输出的局部特征权重进行可视化。

4.1 数据集

HMDB-51: 包含来自 51 个动作类的 6766 个视频剪辑。数据集中的视频多数来源于电影片段, 还有一小部分来自

公共数据库以及 YouTube 等网络视频库。这些动作类包括一般的面部动作、带有对象操作的面部动作、一般身体动作以及带有人体互动的身体动作。这个数据集具有挑战性，因为它包含了许多摄像机运动造成的低质量视频。

Oilfield-7: 本文自建的油田生产现场行为识别数据集包含 7 类动作: Tank Clean, Rig Walk, Room Operate, Site Walk, Room Stand, Tank Construction, Tank Walk, 囊括了日常工作中的常见行为: 擦拭, 走, 作业等。数据集中所有的视频均来自油田生产现场的监控视频, 裁剪时确保每个视频片段中只包含一个行为类别。最终形成的数据集共包含剪辑的视频 200 个, 平均每个视频 300 帧。本文借鉴公共数据集的标准评估准则^[11], 将数据集随机划分为 3 种不同的训练集和测试集。

4.2 实验细节

训练时, 本文使用在 ImageNet 图像分类数据集上预先训练的 BN-Inception 模型初始化网络的权重, 将在视频中采样得到的 RGB 图像和光流图像裁剪到 224×224 的大小作为网络的输入。训练过程遵循 TSN^[7]中所使用的训练策略(冻结除第一个之外的所有批量标准化层和在全局池化后加入 dropout)。TEN 中处理每个流的卷积神经网络单独训练, 基于标准交叉熵损失函数, 采用随机梯度下降算法(SGD)优化损失函数, 批处理大小为 32。在 HMDB51 数据集上, 对于空间流, 总共训练 45 个 Epochs。使用 0.001 的初始学习率, 每 20 个 Epochs 学习率下降到原来的 0.1。对于时间流, 本文也使用 0.001 的初始学习率, 总共训练 100 个 epochs, 每 40 个 epochs 学习率下降到原来的 0.1。所有 SGD 动量都固定在 0.9, 为减小过拟合的风险将 dropout 设置为 0.8。对于自建的 Oilfield-7 数据集, 模型的训练过程与 HMDB51 的过程相同, 只是根据数据集大小调整训练的次数。测试时, 对每个测试视频, 均匀采样固定数量的片段(本文的实验中空间流为 25 段, 时间流为 20 段)。使用中心裁剪, 水平移动, 角裁剪和刻度抖动等技术对采集的图像做增强处理。使用加权平均融合空间流与时间流的预测得分生成视频中行为类别的预测, 分别在两个数据集的 3 个测试集上评估本文的方法并报告了实验结果。

在整个训练和测试的过程使用到了 Pytorch 深度学习框架和 CUDA8.0, 并且使用 Opencv 提取视频中的光流, 本文的实验在 NVIDIA GTX1080Ti 上进行。

4.3 结果分析

首先, 比较了本文提出的方法与另外两种时序特征融合方法: ①最大片段融合; ②平均片段融合。与本文的方法一样, 最大片段融合与平均片段融合都在网络的最后一个全连接层之后分类层之前。实验结果见表 1, 本文在 HMDB51 数据集的第一个 split 上进行实验, 训练时视频的片段数 K 设置为 5。通过引入时序激励模块, 本文将只是用 RGB 图像作为输入的平均准确度提高到了 54.9%。比平

均片段融合方法高 1.2%, 比最大片段融合方法高 1.5%。由此可见通过引入时序激励模块, 在融合视频中的局部特征时, 使得对行为识别有益的特征能够获得更多的关注, 并且在一定程度上抑制冗余和噪声特征, 使得到的全局特征更具鉴别力。

表 1 HMDB51 上不同的融合方法比较 (RGB 帧, split1)

融合方法	准确率/%
Max	53.4
Average	53.7
Average with TE Model(本文)	54.9

另外, 本文对提出的时序激励模块中不同的权重函数对全局特征的影响进行分析, 考虑了第 3 节中描述的含有一个隐藏层的 FC2 和含有两个隐藏层的 FC3。为保证权重函数能有效地为局部特征分配权重, 在隐藏单元的设置上, FC2 中的隐藏单元数设置为输入层单元的 5 倍, FC3 中两层隐藏单元的数量均设置为输入层单元数量的两倍。虽然 FC3 较 FC2 多出一层隐藏单元, 但是实际的参数量 FC3 要比 FC2 减少约 20%, 同时使用 FC3 权重函数输出的权重要优于使用 FC2 输出的权重。使用不同权重函数在 HMDB51 上的实验结果见表 2, 本文观察到, 当输入的局部特征数量比较小时, 使用 FC2 的表现要好于 FC3 的表现, 随着输入的局部特征数量的增加, FC2 的性能要比 FC3 差。经过分析本文认为随着局部特征数量的增加, FC2 的表达能力逐渐趋向饱和, 而拥有两个隐藏层的 FC3 的表达能力逐渐显现出来。

表 2 HMDB51 上两种权重函数随着局部特征数量的增加产生的准确率变化 (RGB 帧, split1)

K	3	4	5	6	7	8
FC2	53.0%	54.2%	54.2%	53.7%	52.8%	53.2%
FC3	52.8%	53.4%	54.9%	54.2%	54.3%	54.6%

为了进一步评估本文所提方法的有效性, 本文在 HMDB51 数据集中 3 种不同的测试集上进行验证, 最后取这 3 个数据集上准确率的平均值作为评价指标, 实验结果见表 3, 本文提出的方法展示了较好的性能, 在只用 RGB 图像作为输入的情况下比 TSN 提高了 2.9%, 比 Attention Pooling^[14]方法提高了 1.7%。此外, 通过与时间流分数的加权融合, 本文方法比 Two-Stream 高出 12.2%, 比 TSN 高出 1.3%, 见表 4, 本文提出的方法达到了目前最先进的性能, 表中还可以看出本文的方法相比于 Attention 系列^[7,18]只能依据单个局部特征分配权重的方法具有一定的优越性, 这种优越性体现在能在一定程度上关注到局部特征之间的时序关系使生成的权重更具区分性, 增加有益特征的作用从而使局部特征融合成的全局特征更具代表性。

表 3 不同方法在 HMDB51 上的比较(只用 RGB 帧)

方法	Split1	Split2	Split3	Avg
TSN ^[5]	54.4%	49.5%	49.2%	51.0%
ActionVLAD ^[12]	51.2%	—	—	49.8%
ResNet50(RGB) ^[13]	—	—	—	48.9%
Attention Pooling ^[14]	54.4%	51.1%	50.9%	52.2%
TEN(本文,FC2)	54.2%	52.2%	53.5%	53.3%
TEN(本文,FC3)	54.9%	53.5%	53.4%	53.9%

接下来本文在自建的油田生产现场行为识别数据集 OilField-7 上评估提出方法的有效性,本文选择基线模型 Two-stream 以及 TSN 作为比较的对象,这两个模型是双流法中的比较有代表性的模型。在 Oilfield-7 上,实验设置同 HMDB51 一样,对于空间流选择 25 个片段,时间流选择 20 个片段进行测试,在最后的双流分数融合阶段本文把权重设置为 1 和 1.5。结果如表 5 所示,本文的方法在准确率上较 Two-Stream 与 TSN 分别提升了 5.3%与 2.2%。每类行为的识别准确率见表 6,在表中可以看出‘Tank Clean’和‘Tank Construction’两类的准确率较低,因为这两个行为不明显并且距离监控摄像头较远,而‘Tank Walk’的准确率较高说明模型加入时序激励模块之后能关注到行为发生的帧,增加其在全局特征中的影响。

最后,本文对提出的时序激励模块对时序上不同位置

表 6 在 OilField-7 数据集每个行为类别上 3 种方法的比较

方法	Tank Clean	Rig Walk	Room Operate	Site Walk	Room Stand	Tank Construction	Tank Walk
Two-Stream	81.2%	89.3%	93.7%	90.4%	92.1%	80.8%	85.0%
TSN	81.0%	92.1%	94.3%	95.6%	94.0%	88.6%	91.0%
TEN(本文)	83.5%	94.6%	96.8%	98.1%	95.6%	84.7%	96.0%

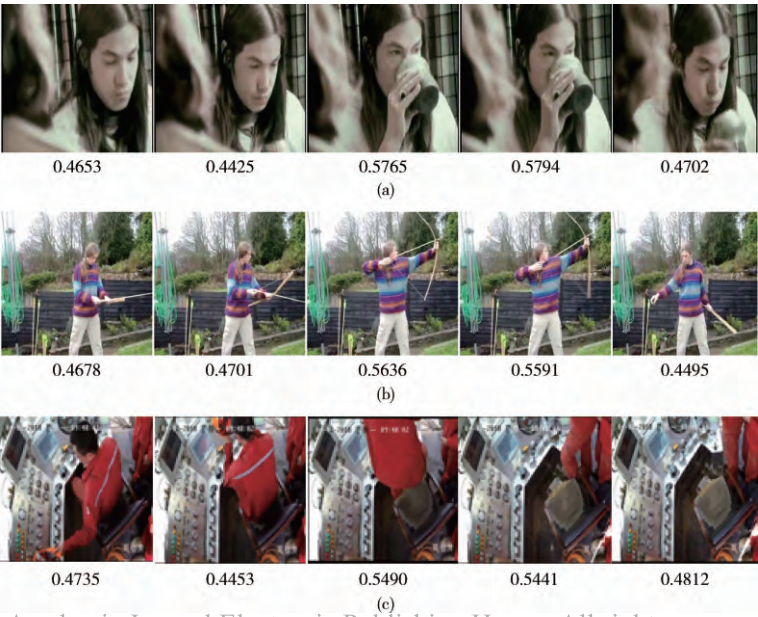
的局部特征的响应进行可视化,以便更好展示网络在学习过程中对时序局部特关注程度。如图 3 所示,本文在

表 4 不同方法在 HMDB51 数据集上的比较(RGB 帧+光流)

方法	HMDB51(RGB+Flow)
iDT+FV ^[15]	57.2%
iDT+HSV ^[16]	61.1%
Two-stream ^[4]	59.4%
TSN(2 modalities,3 segments) ^[5]	68.5%
TSN(3 modalities,3 segments) ^[5]	69.4%
TS-LSTM ^[17]	69.0%
ActionVLAD ^[12]	69.8%
Attention Cluster ^[7]	69.2%
ATW ^[18]	70.5%
TSN(2 modalities,5 segments)(本文)	70.3%
TEN(2 modalities,5 segments)(本文)	71.6%

表 5 在 oilfield-7 数据集上 3 种方法的比较

方法	空间流	时间流	双流
Two-Stream	87.3%	75.0%	87.5%
TSN	91.3%	83.3%	90.6%
TEN(本文)	92.9%	87.5%	92.8%



随机挑选的几段视频上测试提出的方法并将时序激励模块产生的权重可视化。图 3 (c) 是本文自建数据集中工作人员的站起行为, 底部的数字是时序激励模块为采样得到的帧所分配的权重, 可以看出在人还没有站起时分配的权重比较低, 当有明显的站起特征时, 时序激励模块会给予更多的关注即分配更大的权重。

5 结束语

本文为了改进现有的基于双流模型的人体行为识别方法不能有效融合局部特征的不足, 提出了融入时序激励机制的卷积神经网络, 设计的时序激励模块可以为局部特征序列生成具有区别力的权重, 并且嵌入时序激励模块的卷积神经网络依然能够端到端的学习。通过与经典方法在 HMDB51 和自建的 OilField-7 数据集上的结果对比, 验证了本文提出的方法能有效提高人体行为识别的准确率。此外, 本文的方法还可以用于未裁剪视频的人体行为识别, 将在未来的工作中在其它行为识别数据集上评估和改进本文的方法。

参考文献:

- [1] Simonyan K, Zisserman A. Very deep convolutional networks for large scale image recognition [C] //International Conference on Learning Representations, 2015: 1-14.
- [2] He Kaiming, Zhang Xiangyu, Ren Shaoqing, et al. Deep residual learning for image recognition [C] //Computer Vision and Pattern Recognition. IEEE, 2016: 770-778.
- [3] Karpathy A, Toderici G, Shetty S, et al. Large-scale video classification with convolutional neural networks [C] //Computer Vision and Pattern Recognition. IEEE, 2014: 1725-1732.
- [4] Simonyan K, Zisserman A. Two stream convolutional networks for action recognition in videos [J]. Computational Linguistics, 2014, 1 (4): 568-576.
- [5] Wang Limin, Xiong Yuanjun, Wang Zhe, et al. Temporal segment networks: Towards good practices for deep action recognition [C] //European Conference on Computer Vision, 2016: 20-36.
- [6] Ng Y H, Hausknecht M, Vijayanarasimhan S, et al. Beyond short snippets: Deep networks for video classification [C] //Computer Vision and Pattern Recognition. IEEE, 2015: 4694-4702.
- [7] Long Xiang, Gan Chuang, Gerard de Melo, et al. Attention clusters: Purely attention based local feature integration for video classification [C] //Computer Vision and Pattern Recognition. IEEE, 2018: 7834-7843.
- [8] Hu Jie, Shen Li. Squeeze-and-excitation networks [C] //Computer Vision and Pattern Recognition. IEEE, 2018: 7132-7141.
- [9] Richard P, Lippmann. An introduction to computing with neural nets [J]. IEEE ASSP, 1987, 4 (2): 4-22.
- [10] Kuehne H, Jhuang H, Garrote E, et al. HMDB: A large video database for human motion recognition [C] //International Conference on Computer Vision. IEEE, 2012.
- [11] Haroon Idrees, Amir R Zamir, Jiang Yugang, et al. The THUMOS challenge on action recognition for videos “in the wild” [J]. Computer Vision and Image Understanding, 2017, 155: 1-23.
- [12] Girdhar R, Ramanan D, Gupta A, et al. ActionVLAD: Learning spatio-temporal aggregation for action classification [C] //Computer Vision and Pattern Recognition. IEEE, 2017: 971-980.
- [13] Feichtenhofer C, Pinz A, Wildes R P. Spatiotemporal residual networks for video action recognition [C] //Neural Information Processing Systems, 2016: 6433-6441.
- [14] Girdhar R, Ramanan D. Attentional pooling for action recognition [C] //Neural Information Processing Systems, 2017: 6609-6620.
- [15] Wang Heng, Schmid C. Action recognition with improved trajectories [C] //International Conference on Computer Vision. IEEE, 2014: 3551-3558.
- [16] Peng Xiaojiang, Wang Limin, Wang Xingxing, et al. Bag of visual words and fusion methods for action recognition: Comprehensive study and good practice [J]. Computer Vision and Image Understanding, 2016, 150 (C): 109-125.
- [17] Ma C Y, Chen M H, Kira Z, et al. TS-LSTM and temporal-inception: Exploiting spatiotemporal dynamics for activity recognition [EB/OL]. [2019-09-27]. <https://arxiv.org/abs/1703.10667>.
- [18] Zang Jinliang, Wang Le, Liu Ziyi, et al. Attention-based temporal weighted convolutional neural network for action recognition [C] //Artificial Intelligence Applications and Innovations, 2018: 97-108.