基于双模态融合引体向上计数方法的设计与实践

郭 辉[1],辛修勇[2],李博瀚[2]

(1.沈阳工业大学体育装备产业技术研究院,辽宁沈阳 110870;2.沈阳工业大学电气工程学院,辽宁沈阳 110870)

摘 要:现有引体向上计数方法多依赖单一模态数据,在复杂场景下存在鲁棒性不足的问题。提出了一种基于双模态融合的引体向上智能计数系统,结合BlazePose姿态估计、FaceMesh人脸识别和单杠检测技术实现高准确率、高鲁棒性的计数功能。通过BlazePose提取人体关键点并结合KNN分类器进行引体向上动作分类,同时利用FaceMesh定位下颌点并通过霍夫变换检测单杠位置,结合双模态数据融合方法提升计数准确率。实验结果表明,该系统在引体向上场景下的计数准确率达到99.2%,明显高于单一模态方法,且基本满足16 fps视频的实时处理需求,具备较高的应用价值。

关键词:引体向上;双模态融合;姿态估计;人脸识别;霍夫变换

DOI:10.19935/j.cnki.1004-2326.2026.10.043

引用格式:郭辉,辛修勇,李博瀚.基于双模态融合引体向上计数方法的设计与实践[J].实验教学与仪器,2026,43(10):152-154.

传统的引体向上计数方法依赖人工记录,存在主观误差大、效率低的问题[1],难以满足大规模学生体质测试的需求。近年来,基于计算机视觉的引体向上自动计数方法逐渐成为研究热点,但其在实际教学中仍面临诸多挑战[2]。例如,单一依赖姿态估计的方法在遮挡或复杂背景下容易失效[3],基于传感器进行计数的方法仅依据传感器是否被遮挡进行计数[4]。针对上述问题,笔者提出一种双模态融合的引体向上智能计数系统,结合姿态估计、人脸识别和霍夫变换[5]技术,实现高准确率、高鲁棒性的引体向上计数功能。具体而言,系统利用BlazePose模型[6]提取人体关键点,通过KNN分类器对动作进行分类;同时,结合FaceMesh模型定位下颌点,并通过霍夫变换检测单杠位置,设计双模态数据融合策略提升计数准确性。通过双模态信息的联合分析,系统能够在复杂场景下保持较高的计数精度,为体育教学中的引体向上测试提供技术支持。

1 系统总体设计

采用双模态融合的设计思路,结合姿态估计、人脸识别和霍夫变换技术,实现引体向上的智能计数系统构建,主要包括5个模块。

(1)数据采集模块:通过摄像头实时捕获视频流,提取RGB图像数据。

(2)姿态估计模块:基于BlazePose模型提取人体关键点,并结合KNN分类器[7-8]进行动作分类。

(3)单杠检测和下颌点检测模块:通过霍夫变换检测单杠位置,并结合FaceMesh模型定位下颌点,为引体向上计数提供空间约束条件。

(4)双模态融合模块:设计数据级和决策级融合策略,结合姿态估计、单杠检测和下颌点检测结果进行计数。

(5)交互界面模块:使用PyQt5设计引体向上测试界面,在交互界面上实时显示计数结果。

2 双模态融合引体向上计数方法的实践

双模态融合计数方法旨在综合利用不同模态的信息源,通过引入冗余信息互补来解决单一模态方法的局限性。构建的系统要结合人体姿态模态(利用BlazePose模型提取人体关键点,跟踪动作轨迹,采用K最近邻方法识别引体向上动作)和空间约束模态(利用FaceMesh模型定位下颌点,以识别受试者面部下颌区域;通过霍夫变换检测单杠位置,结合空间约束进行引体向上计数)两种模态,并通过双模态信息的联合分析,提升双模态融合计数系统的计数准确率和鲁棒性。

2.1 基于BlazePose的人体姿态估计与动作分类计数

BlazePose模型通过轻量级神经网络提取人体的33个关键点,包括肩膀、肘部、手腕等部位。随后基于预测出的人体33个关键点采用K最近邻算法进行引体向上动作分类,以识别出引体向上动作的不同类别(直臂悬垂和下颌过单杠类别)。首先,将通过BlazePose提取的关键点坐标归一化到[0,1]范围内,以消除图像分辨率的影响,根据归一化后的关键点坐标构建特征向量(嵌入向量);其

次，构建引体向上数据集：收集包含直臂悬垂和下颌过单杠两种动作类别的引体向上图片，对所有图片使用BlazePose模型进行人体关键点预测并剔除不合格的图片；最后，将合格的图片中的人体关键点坐标位置进行保存得到引体向上数据集，数据集保存为csv文件格式。

同时，使用KNN分类器对引体向上动作进行分类，通过计算测试样本和训练样本之间的欧氏距离，选择距离最小的前K个值，通过分类决策来确定当前的动作类别。通过交叉验证，认为最合适的K值为15。设置计数阈值为12，若当前动作类别被分类为下颌过单杠类别且数目≥12，则认为当前动作达到下颌过单杠，可以进行计数。

2.2 单杠和下颌点位置检测与空间约束

（1）单杠位置检测。

单杠通常表现为视频画面中的一条水平直线，为了准确检测单杠的水平位置，采用结合Canny边缘检测和霍夫变换技术来检测图像中的水平直线，再结合双手关键点的信息进行筛选来确定单杠位置。先进行边缘检测，将视频帧转为灰度图，并利用Canny边缘检测提取视频帧中的显著边缘信息，其中Canny边缘检测的低阈值和高阈值分别设置为50和150，这可以有效去除噪声并保留候选边缘。再进行霍夫直线变换提取直线段，对直线段进行筛选，选择方向接近水平的直线作为候选单杠线段。再结合双手位置和单杠特点进行单杠选择：首先通过BlazePose检测手腕关键点，得到受试者左右手的像素坐标，并连接成直线，以此作为手部动作轨迹的参考线。针对每条候选单杠线段，判断是否位于参考线的上方一定范围内。去掉该范围外的水平直线段，保留水平位置与手部直线相近的候选线段。如果存在多个候选线段，就通过计算线段的角度差和中心点距离，评估其与手部直线的相似性。最终选择角度差最小且位置最接近的线段作为单杠的位置。

（2）下颌点检测定位及空间约束。

在实际的引体向上测试中，因环境等因素导致图像质量不佳时直接使用FaceMesh模型无法检测到受试者的脸部区域，故在下颌点检测之前通过BlazePose模型来检测受试者全身关键点，并根据这些关键点位置来定位脸部区域，再使用FaceMesh模型检测受试者的下颌点坐标。随后通过从每帧中提取的下颌点位置，结合单杠直线位置，计算下颌点是否超过单杠高度。

空间约束判定：对于每一帧中提取的下颌点位置信息，结合单杠位置进行空间约束判定，即只有当下颌点的坐标大于单杠位置坐标且随后双臂成直臂悬垂状时，才计为一次有效的引体向上动作。

针对下颌点被单杠遮挡导致关键点丢失的问题，通过BlazePose模型检测到的人体框架计算并定位脸部区域，再基于其他脸部关键点进行下颌点位置的推断，最终实现被遮挡下颌点位置的检测。

2.3 双模态融合策略

为提升计数鲁棒性，本系统构建数据级、决策级双层多模态融合策略。数据级融合将BlazePose人体关键点、下颌特征点与单杠位置统一至同一坐标系；决策级融合利用空间约束过滤伪计数，优先采纳双模态一致结果，模态结果冲突时选择置信度更高的输出。两级融合方案使系统能够在复杂场景下实现高鲁棒性动作识别与计数。

2.4 计数逻辑

引体向上计数流程如图1所示。

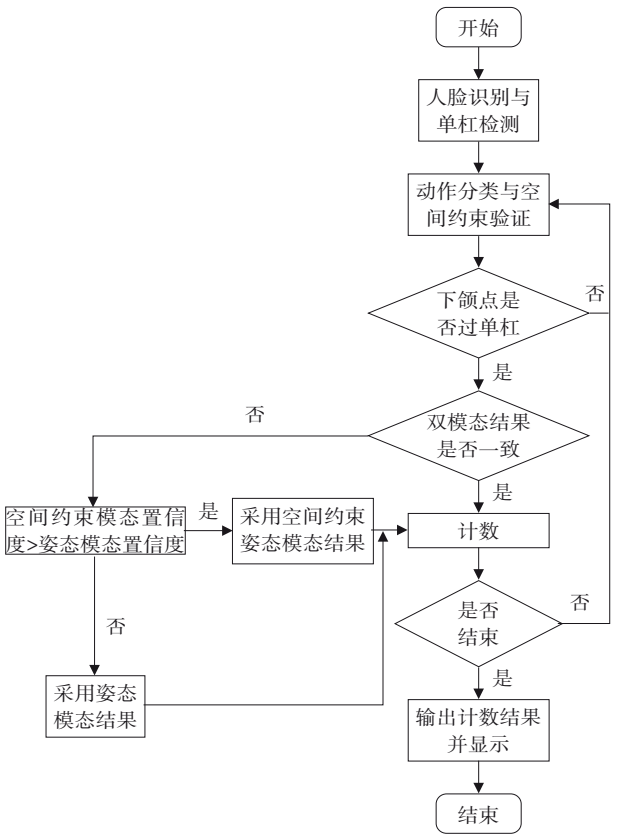


图1 引体向上计数流程

首先，初始化系统，检测单杠位置，检测受试者初始姿态并进行身份识别；随后利用BlazePose提取人体关键点，通过KNN分类器对每帧动作状态分类；再使用FaceMesh分析下颌点位置，通过霍夫变换获取单杠位置，

验证下颌点与单杠的相对位置。结合双模态结果，排除伪计数，累计有效引体向上次数；若双模态结果不一致，则分别计算两种模态各自的当前置信度，选择置信度高的模态结果进行计数判断，最后返回计数结果。

3 实验与分析

3.1 数据集构建与预处理

实验数据来源于沈阳工业大学本科生的引体向上测试视频，覆盖大一至大四年龄段，共 173 名男性受试者，视频拍摄环境为室外。对采集的视频进行处理，提取视频中受试者直臂悬垂和下颌过单杠时的标准动作，分别得到直臂悬垂样本图片 61 张、下颌过单杠样本图片 59 张。将样本中的所有图片使用 BlazePose 模型进行人体关键点预测，将预测出的人体关键点坐标位置进行保存得到引体向上数据集，数据集保存为 csv 文件格式。

3.2 计数结果显示

采用双模态融合智能计数系统开展测试，基于 PyQt5 构建交互界面，系统利用双模态融合策略实现引体向上自动计数，并在图像左上角实时展示计数值。

3.3 评估指标分析

为了验证双模态融合的引体向上智能计数系统的性能，在相同的数据集上与单一模态方法进行对比，采用准确率、召回率、精确率、F1 分数和实时性作为评估指标。结果如表 1 所示。

表 1 不同方法下的模型性能

方法	准确率/%	召回率/%	精确率/%	F1 分数	平均帧率/fps
姿态模态方法	92.6	91.2	94.0	92.6	18
空间约束模态方法	98.7	97.8	99.1	98.4	22
双模态融合方法	99.2	98.7	99.5	99.1	16

由表 1 可知，基于双模态融合的模型方法的准确率达到 99.2%，高于单一模态的准确率，这表明双模态融合计数有效地协同了两种模态的优势，不仅能够精准检测受试者的动作状态，还能通过几何约束条件剔除伪计数，进一步优化了整体计数性能，并且召回率为 98.7%，这表明整体计数漏检极少，在受试者存在短时间遮挡以及下颌点检测失败等复杂场景下，模态融合方法能够通过姿态估计或空间推断补偿单一模态的丢失信息，从而提高整体检测的鲁棒性。F1 分数综合了模型的精确率与召回率，双模态融合方法的 F1 分数达到 99.1，说明它在实际的引体向上场景中能够更好地平衡漏检和误检问题。平均帧率为 16 fps，略高于单一模态方法，但能满足视频实时处理需求，保证系统运行的流畅性。结果表明，双模态融合的引体向上智能计数系统在准确率、鲁棒性上比两种单一模态表现更好，能在实际的引体向上测试中应用。

4 结语

提出了一种基于双模态融合的引体向上智能计数系统，结合 BlazePose 姿态估计、FaceMesh 人脸识别和霍夫变换技术，实现了高精度、高鲁棒性的引体向上计数功能。通过双模态数据融合策略，系统能够有效解决单一模态方法在复杂场景下的局限性，显著提升了计数的准确性和鲁棒性。实验结果表明，双模态融合方法在准确率、召回率、F1 分数等指标上均优于单一模态方法，能够满足实际引体向上测试的需求。系统的实时性和鲁棒性使其适用于各种教学场景，包括室内和室外环境，能够有效提升体育教学的效率和质量。

基金项目：辽宁省教育科研骨干专项重点课题“数智赋能大学生体质健康促进服务体系研究”，编号：JG24QGA34。

作者简介：郭辉（1982—），男，教授，研究方向：体质健康与体育装备工程，E-mail：guohui@sut.edu.cn。

参考文献

[1] 张凯，孙玉国. 基于关节点检测的引体向上评测系统[J]. 智能物联技术，2024，56(01)：58-64.

[2] 洪伟男，王靓涵. 对借力引体向上技术的分析和探究[J]. 体育教学，2021，41(09)：81-82.

[3] 张旭，刘罡，魏志. 基于语义分割和人体姿态估计的引体向上测试平台设计[J]. 传感器与微系统，2024，43(12)：84-88.

[4] 李文博，周磊. 基于体感识别的学生引体向上训练系统[J]. 物联网技术，2018，08(02)：18-20，23.

[5] 孙丰荣，刘积仁. 快速霍夫变换算法[J]. 计算机学报，2001(10)：1102-1110.

[6] Bazarevsky V，Grishchenko I，Raveendran K，et al. Blazepose：On-device real-time body pose tracking[PP/OL]. ArXiv. DOI：10.48550/arXiv.2006.10204.

[7] Zhang S C，Li X L，Zong M，et al. Efficient kNN classification with different numbers of nearest neighbors [J]. IEEE Transactions on Neural Networks and Learning Systems，2018，29(05)：1774-1785.

[8] Arif M，Akram M U，Minhas F A. Pruned fuzzy K-nearest neighbor classifier for beat classification [J]. Journal of Biomedical Science & Engineering，2010，03(04)：380-389.