

基于 BlazePose 和 KNN 的健身计数系统设计与实现

孔亚琪, 刘宇

(南京邮电大学教育科学与技术学院, 江苏 南京 210023)

✉ 1021163408@njupt.edu.cn; yliu@njupt.edu.cn



摘要:目前,健身动作的识别与计数受大模型的训练以及动作种类繁多等影响,少有实时性、准确性、稳定性等各方面均表现优异的健身动作识别与计数系统。该系统利用 BlazePose 进行动作识别,以满足健身动作识别的实时性和稳定性的要求,使用 KNN 算法实现动作的分类与计数功能,利用 Tkinter 实现交互式界面,使系统具备可交互性。该系统仅需较少的自采集数据集即可实现动作识别与计数功能,测试结果表明该系统达到 95.5% 的计数准确率和 30 FPS 的实时检测速度,可广泛应用于健身场所和线上健身平台。

关键词:健身计数系统;动作识别;动作计数;自采集数据集

中图分类号:TP311.5 **文献标识码:**A

Design and Implementation of a Fitness Counting System Based on BlazePose and KNN

KONG Yaqi, LIU Yu

(College of Education Science and Technology, Nanjing University of Posts and Telecommunications, Nanjing 210023, China)

✉ 1021163408@njupt.edu.cn; yliu@njupt.edu.cn

Abstract: Currently, recognition and counting of fitness pose are limited by factors such as the training of large models and the diversity of movement types, which often results in poor real-time performance, accuracy, and stability in fitness pose recognition and counting systems. The proposed system aims to address these shortcomings by using BlazePose for action recognition, thereby meeting the requirements of real-time and stable identification of fitness poses. The K-Nearest Neighbors (KNN) algorithm is employed to facilitate movement classification and counting. Additionally, Tkinter is used to realize an interactive interface, which enhances the system's interactivity. A noteworthy aspect of this system is its ability to perform pose recognition and counting functions with a minimal self-collected dataset. Test results demonstrate that the system achieves a count accuracy rate of 95.5% and real-time detection speed of 30 Frames per Second (FPS), making it suitable for broad applications in fitness venues and online fitness platforms.

Keywords: fitness counting system; pose recognition; pose counting; self-collected dataset

0 引言 (Introduction)

目前,健身运动已经成为人们的一种生活方式,AI 健身技术可以帮助用户更好地完成健身动作,从而达到更好的健身效果。健身动作具有多样性和复杂性的特点,现有的健身动作识别与计数方法尚存在一些问题,例如识别精度低、计数不准等。

因此,如何提高健身动作识别与计数的速度和准确率,成为当前研究的热点和难点。

近年来,国内外学者对于健身动作识别与计数进行了广泛的研究。其中,基于深度学习的方法得到了广泛应用。例如, YANG 等^[1]使用卷积神经网络(CNN)识别深蹲和俯卧撑等动

作,但是此方法需要大量的数据集与计算量才能达到较高的准确度。基于传感器的方法是一种常见的健身计数器的设计方案,主要使用加速度计、陀螺仪等传感器测量用户的运动状态,并统计用户完成的动作次数。例如, MURO-DE-LA-HERRAN 等^[2]使用陀螺仪和加速度计计数深蹲、引体向上等动作。尽管此方法具有较高的计数准确度,但传感器的佩戴会对用户运动过程的舒适度和体验感造成一定的影响。为解决以上问题,本文使用 BlazePose 模型对人体的 33 个关键点进行检测,运用 KNN 算法作为动作分类器,对悬停动作与当前动作关键点之间的距离进行相似度度量,用于判断用户是否处于健身的动作中,加入计数器统计动作次数,并对其健身动作进行实时记录。

1 系统简介(System introduction)

本文所提健身动作检测与计数系统基于 BlazePose 模型和 KNN 算法,可以自动检测和计数多种常见的健身动作,如深蹲、俯卧撑、引体向上、仰卧起坐等。该系统利用 BlazePose 模型实时跟踪人体骨骼关节,提取人体运动动作的特征,并使用 KNN 算法判断当前的健身动作类型并加入计数器进行计数;并且,该系统具有良好的计数准确率和实时性,可以实现对健身动作的高效计数和检测,使用户能够更加方便且有效地进行健身训练。此外,该系统可以记录用户的运动数据,并将运动数据保存到本地文件资源管理器,方便用户进行健身记录查看和分析。此健身计数系统的功能流程图如图 1 所示。

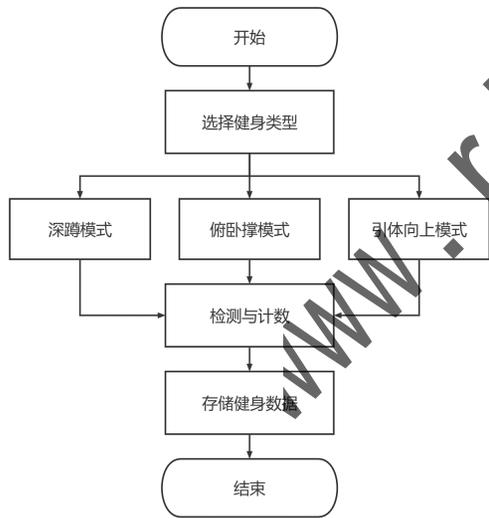


图 1 系统功能流程图

Fig. 1 System functional flowchart

2 系统关键模块设计 (Design of key system modules)

2.1 健身动作检测与计数模块设计

健身动作检测与计数模块为该系统的核心模块。首先,该模块使用 BlazePose 检测并跟踪健身者的身体姿势;其次,根据检测的姿势生成人体相应的 33 个关键点坐标,针对每个动作类别,系统会预先收集许多相应类别的健身动作样本并进行模型训练,每个样本对应一个正确执行该动作的健身者姿势序列,对于新检测到的动作,该模块通过摄像头传入的人体姿势

关键点并计算其 33 个关键点与每个训练样本关键点之间的相似度,使用 KNN 算法确定最合适类别的标签;最后,加入计数器用于对每个动作的执行次数进行计数。健身动作检测与计数的算法流程图如图 2 所示。

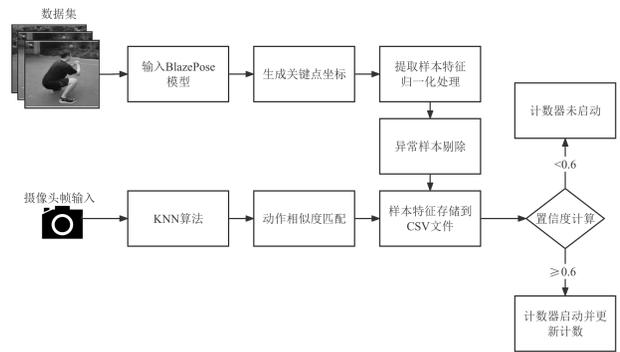


图 2 健身动作检测与计数算法流程图

Fig. 2 Flowchart of the algorithm for detecting and counting fitness pose

2.2 BlazePose 和 KNN 简介

BlazePose 是由 Google 研发的一种人体姿势检测技术,能够实时检测和跟踪人体的 33 个关键点位置,它通过使用卷积神经网络(CNN)进行人体姿势检测,其模型结构相对轻量化,并且检测速度非常快,因此适用于实时应用^[3]。该模型具有较高的准确性和鲁棒性,并且能够在不同环境下进行运动捕捉,包括普通摄像头和移动设备摄像头等。该模块中,使用 BlazePose 模型实时检测人体姿势并生成 33 个关键点的三维坐标数据,从而确定健身动作是否正确执行。如图 3 所示为 BlazePose 模型的 33 个人体姿势关键点位置与序号。

K 最近邻算法(KNN)是一种基于实例的学习算法,可用于解决分类和回归问题^[4]。在该模块中,KNN 算法通过计算样本之间的距离对数据进行分类,从而识别健身者正在执行的动作。

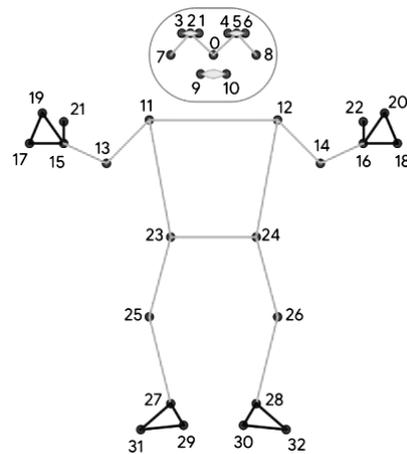


图 3 BlazePose 模型关键点位置

Fig. 3 Key point positions of the BlazePose model

2.3 训练样本

2.3.1 提取样本关键点坐标

首先,将数据集样本图片分为两类放入两个不同的文件夹中,以深蹲为例,将数据集图片分为两类,分别为完全蹲下状态

和完全起立状态,放入 squat_down 和 squat_up 两个文件夹中,然后调用一个名为 bootstrap 的函数,该函数可以从一组图像中提取人体姿势关键点坐标,并保存这些坐标。

2.3.2 提取样本特征

特征提取的过程是通过将姿势的关键点转换为特征向量实现的。这些特征向量可以被看作是数字表示,它们捕捉到了关键点的相对位置和方向等重要信息,并且可以用于训练姿势相关的深度学习模型^[5]。将已保存的关键点数据复制到一个名为 samples 的列表中。使用 PCA 算法(一种常见的数据降维方法,通过线性变换将高维数据转化为低维数据,并且保留数据的主要特征)对 samples 列表中的姿态关键点数据进行降维处理。降维后的姿态关键点数据称为特征向量,将这些特征向量写入 CSV 文件中,即可得到每个动作对应的特征向量。

2.3.3 归一化处理

对于已经提取的样本特征,进行归一化处理可以使不同特征之间的比较更加准确和公平,从而提高机器学习算法的性能和模型的泛化能力^[6]。归一化处理后,可以将不同身高、体型的人的姿势映射到相同的尺度下,从而方便进行特征提取、姿势分类和计数等操作。如果没有做归一化处理,由于不同的姿势关键点取值范围可能不同,导致不同姿势的特征向量之间难以比较。并且,归一化处理可以有效地降低噪声对关键点特征提取的影响,这是因为一些关键点的取值范围可能与其他关键点不同,导致它们对特征向量的贡献有所偏差。通过归一化处理,可以将这类偏差降至最低,从而减少噪声对关键点特征的干扰^[7]。因此,对人体姿势关键点进行归一化处理,可以提高模型的鲁棒性和泛化能力。

本文采用最大最小归一化(Min-Max Scaling)方法进行归一化处理,第 i 个关键点坐标的原始取值为 (x_i, y_i) ,那么它的归一化处理后的值为 (\hat{x}_i, \hat{y}_i) ,具体归一化处理公式如下:

$$\hat{x}_i = \frac{x_i - \bar{x}}{s} \quad (1)$$

$$\hat{y}_i = \frac{y_i - \bar{y}}{s} \quad (2)$$

其中, \bar{x} 和 \bar{y} 分别表示所有关键点在 x 方向和 y 方向上的平均值, s 表示所有关键点在 x 和 y 方向上的标准差。对于一个关键点序列,需要对所有关键点都进行上述公式的归一化处理,得到一个新的序列作为归一化后的人体姿势关键点特征并写入 CSV 文件中。

2.3.4 异常样本识别与剔除

识别异常样本的方法是通过计算每个样本的嵌入向量与输入姿势嵌入向量的欧氏距离,将距离大于某个阈值的样本识别为离群点^[8]。距离越大,表示两个特征向量之间的差异越大,越有可能是离群点。通常,阈值的取值范围是在 0~1,该范围内的值可以表示距离的比例关系。当阈值为 1.0 时,意味着输入姿势嵌入向量与识别为离群点的样本的嵌入向量之间的欧氏距离为 1 以上,这意味着异常样本与输入样本特征向量之间的差异较大,可能为错误数据,应该被排除或者进一步分析处理。将阈值设置为 1.0 后,检测到 3 张异常样本图片,定义一个 remove_outliers 函数即可剔除这个图片文件及其相应的关键点数据。剔除异常样本后会排除一些异常噪声或错误

数据,可以进一步提高模型的鲁棒性和泛化能力。

2.4 构建动作分类器

KNN 动作分类器是机器学习模型的其中之一,可用于分类人体动作。该模型通过将人体特征转化为向量形式,并计算不同向量之间的相似度,利用 KNN 算法根据最近邻标准对相似度最高的几个向量进行分类^[9]。KNN 动作分类器具体实现算法如下所示:

```
class KNNPoseClassifier:
    def __init__(self, n_neighbors=3):
        self.classifier = KNeighborsClassifier(n_neighbors=n_neighbors)
    def fit(self, X, y):
        self.classifier.fit(X, y)
    def predict(self, X):
        return self.classifier.predict(X)
```

将已训练的 CSV 样本文件输入 KNN 模型,KNN 模型使用基于欧氏距离的相似性度量方法,计算数据集中所有样本和当前摄像头输入样本的特征向量之间的距离,这个距离就是摄像头输入的当前帧与已训练的样本之间的相似度。在这个过程中, k 个训练样本与测试样本之间的相似度可以被解释为置信度,因为该值反映了分类器输出结果的可信程度^[10]。例如,当 k 个训练样本中有 n 个样本属于 down(蹲下)类, m 个样本属于 up(起立)类时,该测试样本被分类为 down(蹲下)类的置信度可以表示如下:

$$\text{confidence}(\text{down}) = \frac{n}{k} \quad (3)$$

其中, k 表示 KNN 分类器选择的最近邻数。如果 n 越大,则置信度越高,这是因为更多的训练样本被投票选为 down(蹲下)类。同样,被分类为 up(起立)类的置信度可以表示如下:

$$\text{confidence}(\text{up}) = \frac{m}{k} \quad (4)$$

将检测 down(蹲下)的阈值设为 0.6,即如果检测到当前动作置信度大于 0.6,即可以认定健身者处于深蹲健身动作中,对于健身者进入深蹲健身状态,标记为 True,则计数器可以运行;否则不会进行检测与计数。图 4 为一个连续做 5 个深蹲的健身者的处于 down(蹲下)状态的置信度检测结果图(图中斜线阴影部分为经过平滑处理后的置信度检测结果),横轴为帧数,竖轴为置信度。

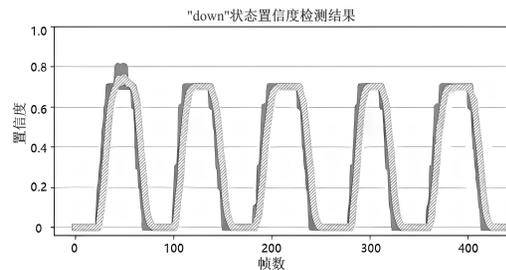


图 4 连续深蹲的健身者处于“down”状态置信度检测图

Fig. 4 Confidence interval detection graph of a fitness enthusiast in a continuous “down” state squat

2.5 计数器实现

计数器主要用于计算给定目标姿势类的重复次数。通过

输入动作分类器的结果,获取给定动作的置信度。如果置信度超过进入姿势的阈值,则开始计算该姿势的重复次数。当置信度低于退出姿势的阈值时,即完成一次姿势,并记录该动作的开始时间、结束时间、动作持续时间和与上一次动作的间隔时间等。检测过程中,会将结果写入 Excel 表格并按照一定格式保存。计数器实现算法具体如下所示:

```
# 获取姿势的置信度.
pose_confidence = 0.0
if self._class_name in pose_classification:
    pose_confidence = pose_classification[self._class_name]
# 如果处于 down(蹲下)状态并且正在退出它,则计数器更新状态
if pose_confidence < self._exit_threshold:
    self._n_repeats += 1
    self._pose_entered = False
return self._n_repeats
```

3 实现与验证 (Implementation and verification)

3.1 数据采集

本文选取了三名身高体重相差较大的体育生作为数据采集对象,通过拍摄其在多种不同场景下的健身动作,尽可能地保证数据收集的全面性。三名体育生的身高和体重分别为 170 cm、64 kg,179 cm、75 kg,185 cm、86 kg。三名体育生分别进行持续的深蹲、俯卧撑、引体向上动作,使用 25 FPS 的录像设备在不同光线、不同场地、不同拍摄距离的情况下进行水平状态下的 360°循环拍摄。拍摄的健身数据将用于后续对运动姿态识别的研究与分析,为更准确和可靠的运动姿态识别算法的开发提供基础数据支持。

3.2 数据预处理

为了更方便地采集运动者动作的图像数据信息,本设计利用 Python 编写图片帧提取的程序,具体要求为将 MP4 格式的视频以每秒 5 帧的速度进行图片帧提取,将数据集录像输入程序即可快速提取录像的帧。对于提取的视频帧进行手动筛选后,共得到 300 张图像,即每种运动类型的图像 100 张作为数据集,提取的部分深蹲数据集图片如图 5 所示。



图 5 深蹲数据集
Fig. 5 Squat dataset

3.3 运动检测与计数准确率测试

本文召集另外两名身高和体重分别为 175 cm、70 kg 和 183 cm、83 kg 的体育生作为测试者,用于测试实际计数的准确率。两名测试者分别进行两组持续的深蹲、俯卧撑、引体向上动作,深蹲和俯卧撑每组 40 个动作,引体向上每组 20 个动作,每名体育生做 100 个动作,深蹲、俯卧撑、引体向上动作总数共计 200 个。在此过程中,计数器漏计共 5 个动作,实际计数动作详见表 1,实际计数准确率为 95.5%,已具备一定的应用价值。

表 1 计数器准确率统计表

Tab. 1 Statistics table of counter efficiency

测试者	运动类型	实做动作数/个	检测动作数/个	实际计数准确率/%
测试者 1	深蹲	40	40	100.0
	俯卧撑	40	38	95.0
	引体向上	20	17	85.0
测试者 2	深蹲	40	39	97.5
	俯卧撑	40	39	97.5
	引体向上	20	18	90.0
汇总		200	191	95.5

3.4 系统实现

本系统基于 BlazePose 动作检测模型和 KNN 算法,旨在实现一个稳定、实时、高计数准确率的健身动作计数系统,系统主界面使用 Tkinter 进行开发,Tkinter 是 Python 的标准图形用户界面(GUI)库,它提供了一种直观、简洁的方式使得开发人员能够快速创建包括窗口、按钮、菜单、文本框等基本元素在内的可视化应用程序界面。该系统的主界面如图 6 所示。



图 6 系统主界面

Fig. 6 System main interface

进入主界面后,有三种健身类型可供用户选择,选择其中一种即可自动调用本地摄像头进行实时的健身动作检测,检测窗口左上角显示置信度检测图,右上角数字显示当前动作数,例如选择俯卧撑进行检测与计数后,检测与计数效果如图 7 所示。

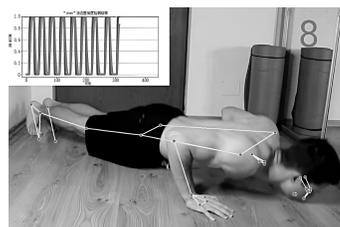


图 7 俯卧撑检测与计数效果

Fig. 7 Push-up detection and counting effect

本次实现设备基于 Windows 11 操作系统环境,使用 AMD Ryzen 7 5800H 处理器,NVIDIA GeForce RTX 3060 显卡,16 GB 内存和分辨率为 1 280×720 的 HD 检测摄像头,实现了基于 BlazePose 模型和 KNN 算法的健身动作计数器,其检测速度可达 30 FPS,能够有效地检测用户的健身动作,提供准确的计数统计并实时记录用户的健身数据。

4 结论(Conclusion)

本文基于 BlazePose 动作识别模型和 KNN 算法为核心,实现了一个计数准确率较高的健身动作计数器系统的开发,该计数器结合计算机视觉技术和机器学习算法,可以准确地检测用户的运动姿态并提供准确率较高的计数统计。通过详细的数据采集、数据处理、算法设计和系统验证等步骤,验证了该健身计数系统在不同环境下,针对不同健身者的实际动作检测准确率可达 95.5%,并具有实时的检测功能和较高的计算速度,检测速度可达 30 FPS,有望在健身场所和线上健身平台等实际应用场景中得到推广与使用。

参考文献(References)

- [1] YANG L, LI Y, ZENG D, et al. Human exercise posture analysis based on pose estimation[C]//IEEE. 2021 IEEE 5th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC). Piscataway: Institute of Electrical and Electronics Engineers, 2021: 1715-1719.
- [2] MURO-DE-LA-HERRAN A, GARCIA-ZAPIRAIN B, MENDEZ-ZORRILLA A. Gait analysis methods: an overview of wearable and non-wearable systems, highlighting clinical applications[J]. Sensors (Basel), 2014, 14(2): 3362-3394.
- [3] MOHAMMED S W, GARRAPALLY V, MANCHALA S,

et al. Recognition of yoga asana from real-time videos using Blaze-pose[J]. International Journal of Computing and Digital Systems, 2022, 12(1): 1304-1295.

- [4] GUO G, WANG H, BELL D, et al. KNN model-based approach in classification[C]//Springer. On The Move to Meaningful Internet Systems 2003: OTM 2003 Workshops. Berlin Heidelberg: Springer, 2003: 986-996.
- [5] 陈星宏. 基于 3D 骨架的肢体动作识别研究[D]. 成都: 电子科技大学, 2019.
- [6] 李华. 基于深度学习轻量高效的人体姿态估计模型研究[D]. 成都: 电子科技大学, 2022.
- [7] 刘勇, 李杰, 张建林, 等. 基于深度学习的二维人体姿态估计研究进展[J]. 计算机工程, 2021, 47(3): 1-16.
- [8] 徐玲玲, 迟冬祥. 面向不平衡数据集的机器学习分类策略[J]. 计算机工程与应用, 2020, 56(24): 12-27.
- [9] NUNEZ-GODOYS, ALVEAR-PUERTAS V, REALPE-GODOY S, et al. Human-sitting-pose detection using data classification and dimensionality reduction [C]// IEEE. 2016 IEEE Ecuador Technical Chapters Meeting. Piscataway: Institute of Electrical and Electronics Engineers, 2016: 1-5.
- [10] 吴建超, 王利民, 武港山. 视频群体行为识别综述[J]. 软件学报, 2023, 34(2): 964-984.

作者简介:

孔亚琪(1998-),男,硕士生. 研究领域:深度学习,教育人工智能.

刘宇(1971-),女,博士,副教授. 研究领域:软件工程,数字化学习,学习分析,课程与教学理论.

(上接第 43 页)

参考文献(References)

- [1] 中共中央国务院印发《中国教育现代化 2035》[N]. 人民日报, 2019-02-24(1).
- [2] 苏超, 王国中. 基于改进 OpenPose 的学生行为识别研究[J]. 计算机应用研究, 2021, 38(10): 3183-3188.
- [3] 魏艳涛, 秦道影, 胡佳敏, 等. 基于深度学习的学生课堂行为识别[J]. 现代教育技术, 2019, 29(7): 87-91.
- [4] 曾劭伦. 基于视频的学生课堂行为检测算法研究[D]. 哈尔滨: 哈尔滨工业大学, 2020.
- [5] 黄勇康, 梁美玉, 王笑笑, 等. 基于深度时空残差卷积神经网络的课堂教学视频中多人课堂行为识别[J]. 计算机应用, 2022, 42(3): 736-742.
- [6] 郭俊奇, 吕嘉昊, 王汝涵, 等. 深度学习模型驱动的师生课堂行为识别[J]. 北京师范大学学报(自然科学版), 2021, 57(6): 905-912.
- [7] UDDIN M Z, HASSAN M M, AISANAD A, et al. A body sensor data fusion and deep recurrent neural network-based

behavior recognition approach for robust healthcare[J]. Information Fusion, 2020, 55: 105-115.

- [8] 皮连生, 蔡维静. 超越布卢姆: 试论“知识分类与目标导向”教学中的学习结果测量与评价[J]. 华东师范大学学报, 2000(2): 40-49.
- [9] 王立辉, 杨贤昭, 刘惠康, 等. 基于 GhostNet 与注意力机制的行人检测跟踪算法[J]. 数据采集与处理, 2022, 37(1): 108-121.
- [10] 孙泽龙, 杨国兴, 温静远, 等. 基于多域 VQGAN 的文本生成国画方法研究[J]. 软件学报, 2023, 34(5): 2116-2133.
- [11] 李坤, 侯庆. 基于注意力机制的轻量型人体姿态估计[J]. 计算机应用, 2022, 42(8): 2407-2414.

作者简介:

王禹钧(1997-),女,硕士生. 研究领域:智慧教育.

马致明(1964-),男,硕士,教授. 研究领域:计算机辅助教育,软件开发与应用.