

基于 BlazePose 的学生体育测试智能评测系统的设计

Design of Intelligent Evaluation System Based on BlazePose for Students' physical Examination

张宇超*, 薛筱茜, 张浩然, 韩泽坤, 王文陇
北京邮电大学

*通信作者: yczhang@bupt.edu.cn

【摘要】 人工智能技术的发展,在教育领域产生了很多积极影响。将人工智能技术与特定的场景结合,可以很大程度上辅助和推进传统的教学工作。本文提出了一种基于人体姿态估计方法的学生体育测试智能评测系统,旨在体育教学工作中更准确地评测学生体考分数,并且对测试结果给出针对性建议。系统通过基于轻量级神经网络 BlazePose 得到人体骨关节点信息,根据关节点信息判断动作的合规性并对动作打分。总体上每个检测项目都可以划分为三个功能模块,计数模块、动作检测模块和动作建议模块。本测试平台可以对多达九种体育考试项目进行实地检测,在实际测试中,由于加入量化评判标准使检测结果更具说服力,在确保测试准确率的同时还针对性对学生动作提出改进建议,能够协助学校更好地推行体育健康工作。

【关键词】 深度学习; 人体姿态估计; MediaPipe; 体育健康; 学生体测

Abstract: The development of artificial intelligence technology has had many positive impacts in the field of education. Combining artificial intelligence technology with specific scenarios can greatly assist and advance traditional teaching work. This paper proposes an intelligent evaluation system for students' physical examination based on human body posture estimation method, which aims to evaluate students' physical examination scores more accurately, and give targeted suggestions for test results. The system obtains the information of human bone joint points through the lightweight neural network BlazePose, judges the compliance of the action according to the joint point information, and scores the action. In general, each physical examination item can be divided into three functional modules: counting module, action detection module and action suggestion module. The test platform can conduct on-the-spot testing of up to nine physical examination items. In the actual test, due to the addition of quantitative evaluation standards, the test results are more convincing. While ensuring the accuracy of the test, we also propose improvements to the students' movements. Suggestions can assist schools to better implement sports health work.

Keywords: Deep Learning, Human Body Pose Estimation, MediaPipe, Sports Health, Student Body Measurement

1. 前言

体质是人的生命活动和劳动、工作能力等的物质基础,增强中国人民的体质是我国社会主义现代化建设的重要基础。学生的体能、体质测试,是各院校在学生体育健康教育方面的重要成果。

体育中考是通过统一测试的形式对应届初中毕业生作出体质评价的统一测评模式,即通过测试运动成绩的分数来进行体质评价。采取的测试方式是选取固定的场景以及设备,在测试人员的配合下对体考人员进行动作评测。对于考试内容,各地区的具体要求有所差异,但总体上都包含仰卧起坐、引体向上、坐位体前屈、立定跳远等基础项目。每项体考动作都有具体的评价指标:测试规则、成绩判定和关键动作评判标准。在传统人工评分体系中,往往仅采用工作人员以目测方式获得结果,进而对考生的动作规范进行评价并最终确定成绩,但受制于“人”在评判中存在的诸多不确定性,这种评分方式往往具有评分准确性差、效率低下的特点。对于工作人员自身而言,长时间维持高专注度用眼容易产生视觉和精神上的双重疲劳,进而导致注意力下降,在计数、动作规范判断上出现纰漏的可能性也大大提高。同时,人工检测也更容易受到外部因素的干扰,场地、角度、甚至光线的变化都可能引起工作人员在测试的过程中出现偏差并进而造成误判,例如较暗光线会导致工作人员视力下降,丢失对动作姿态的捕捉,最终工作人员不能进行有效的评估。同时,公平和公正也是测试当中尤为重要的一环,不同的工作人员界定动作规范的标准各不相同,主要依据肉眼的观测,难

以形成统一并且量化的标准，倘若规则界定的尺度参差不齐，那么测试结果也无法服众。此外，对于考生数量较多的测试而言，采用人工评分的方式，检测的时间效率将是一个值得考虑的问题。每个考生的测试过程若全部通过人工方式完成，时间成本将会大大提升，例如对于坐位体前屈、立地跳远等需要读取测量数据的项目，工作人员需要花费数秒来完成成绩评估，而这对于智能评测系统来说仅仅不到一秒，同时还能减少因肉眼观测带来的误差，这将极大提高测评效率，有效节省人力和时间成本。因此针对上述传统人工评分体系呈现的问题，选取一种准确性高、可靠性强的测试平台是非常必要的。

教育数字化转型是数字化转型在教育这一特殊领域的表现形态，教育数字化转型有可能带来极具创造性的新教学策略[5]。随着人工智能、大数据等相关技术的快速发展，各个行业、企业都相继参与了人工智能革命，加速人工智能融入特定的场景中以提供相应服务。这一势头在教育领域中也势不可挡，将人工智能相关技术运用到教育行业中，使其作为传统教育教学方式的高效辅助手段，可以很大程度上提升教学效率。基于人工智能的学生体育测试智能评测系统是这种结合的一个具体实例。

2. 相关研究

近年来智能手机已经成为每个人不可或缺的移动设备，将智能手机作为实验工具和辅助设备，可以为科学教育提供更多的可能性[1]。教育部在2018年4月颁布《教育信息化2.0行动计划》进一步推进了“互联网+教育”的发展，鼓励大力推进人工智能在教学方面的全流程应用[2]。目前，已有研究将人工智能技术结合在日常运动和体育考试中，这类应用大多基于人体姿态估计方法，一种以人体骨关节为研究对象，通过检测关节的位置信息即关节之间的联系进而重构人体肢干的方法。结合人体姿态估计方法和具体应用场景，在体育、健康和领域有重要的影响意义。

融梦科技推出的“云+端 体育AI平台化服务”支持多种体育锻炼运动项目，该产品通过人体姿态估计，识别学生各运动项目的锻炼结果并进行数据化呈现。做到了结合学生个人体育素质提供个性化锻炼建议。FITURE作为家庭科技健身方式的开创者和引领者，推出的首款智能健身产品“FITURE魔镜”采用硬件+内容+服务+AI的商业模式。该产品可以通过内置的AI摄像头实时检测用户动作，采用人体姿态估计关键技术，根据用户选择的健身、运动内容及时给出反馈与建议，帮助用户更好地在家锻炼身体。

当前人体姿态估计方法总体上可以分为单人姿态评估方法和多人姿态评估方法。其中单人姿态估计包括：1) 基于卷积神经网络的方法，例如Toshev等[9]使用AlexNet作为基本网络结构对关节坐标进行回归，通过多次计算获得关节的坐标。2) 基于生成对抗网络的方法，例如Chen等[7]结合人体结构先验知识估计被遮挡的人体部位，用对抗网络来训练姿态生成器。以及3) 基于组合模型的方法，例如Tang等[8]提出一种通过学习人体复杂的结构关系，以精组合各个身体结构部件。多人姿态估计可以分为自上而下的评估方法和自下而上的评估方法，与自上而下方法相比较而言，自下而上的方法运行效率较高，但需要更繁琐的后续处理过程。具体算法的参照和选择需要结合特定的场景，例如适用于单人和多人姿态估计的OpenPose开源框架，在多人场合下具有更高的识别准确率，以及适用于单人姿态估计的解决方案BlazePose，可以在单人场景下展现出更快的速度和更高的性能，帧率是前者几十倍[6]。

BlazePose[6]是谷歌开发的一个轻量级卷积神经网络架构，该模型具有显著加速效果，可以用于在移动设备上实时检测33个人体骨骼关键节点。MediaPipe Pose[6]是基于BlazePose实现的，一种高保真人体姿态跟踪和识别的机器学习解决方案。具有端到端、易部署、实时性强、计算速度快以及开源等优势和特点。大多数流行的目标检测方案适用于自由度很小的刚体，不适用于高清晰、运动自由度大的人体检测。为了克服这个限制，BlazePose专注于检测相对刚性的身体部位（例如人脸或躯干）的边框，达到的效果是对于单人用例，当头部始终清晰可见时，检测效果强大且有效[6]。在体考场景下，保证头部和

躯干清晰可见是必要条件，所以 Mediapipe Pose 框架用于姿态检测和动作判断是很合理的，为实现基于人体姿态识别的体育测试智能评测系统提供了依据和保障。

3. 基于 BlazePose 的学生体育测试智能评测系统的设计

本文设计的学生体育测试智能评测系统测试过程如图 1 所示，首先选择目标检测项目，选定之后通过调取移动端设备的摄像头监测体测同学的动作，将捕捉到的视频流传入基于 BlazePose 实现的关键点识别模块，获得 33 个人体骨关节点，将关节点序列输入关键点处理模块，根据置信度等关键信息筛选出可用于检测的关键点，对关键点序列表示的动作进行检测和评估，及时提示关键动作建议，在检测的过程中控制测试时间，测试完成后计算动作得分，最终将检测结果和建议反馈给测试人员。

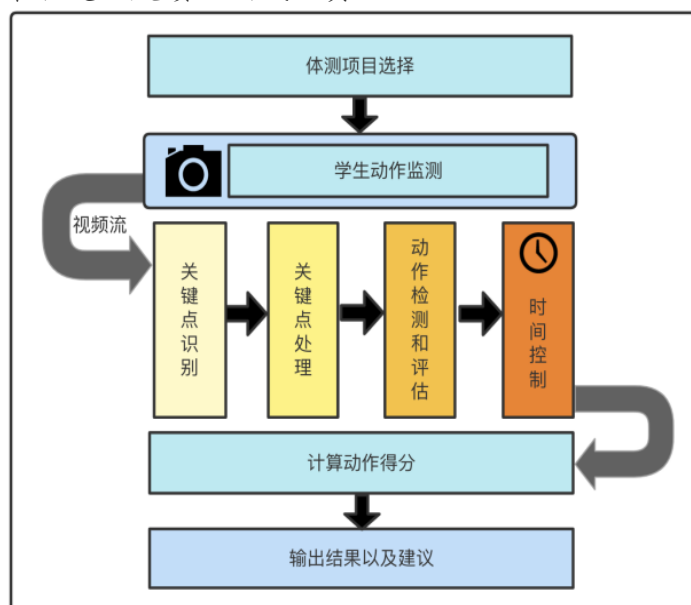


图 1 智能评测系统结构图

3.1. 关键点识别模块

MediaPipe Pose 的主要概念包括数据包 (Packet)、数据流 (Stream)、计算单元 (Calculator)、图 (Graph) 以及子图 (Subgraph)。图 2 是 MediaPipe Pose 的核心框架。

人体姿态估计 pipeline 被实现为一个 MediaPipe graph，如图 2 所示，使用姿态标志模块识别，并使用姿态渲染模块进行渲染。ConstantSidePacket 用来启用分割掩模，而 FlowLimiter 用来数据限流计算单元，它会接收下面计算单元的一个输入信号，如果没有下面单元的输入信号它会丢弃当前的视频帧，这样就可以控制处理过程不会因为输入间隔小于处理时间而出现问题。

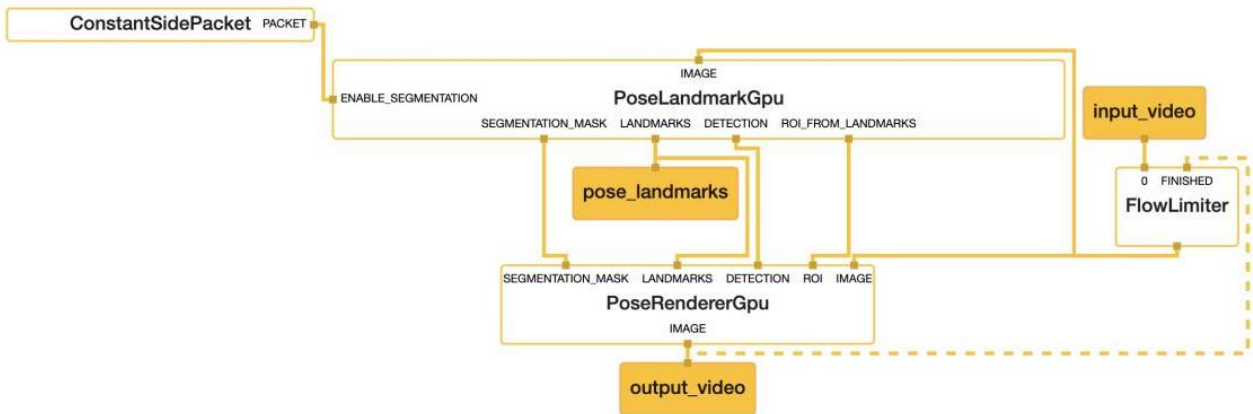


图 2 MediaPipe Pose 核心框架

MediaPipe Pose 的核心神经网络架构 BlazePose 用于实时检测 33 个人体骨骼关键节点，BlazePose 采用热图、偏移和回归组合的方法，仅在训练阶段使用热图和偏移损失，有效地使用热图来监督轻量级 embedding，以便随后被回归编码器网络使用。并且积极利用网络所有阶段之间的跳跃连接来实现高级和低级功能之间的平衡。这不仅可以改善热图预测，还可以显著提高坐标回归的准确性[6]。

3.2. 关键点处理模块

根据体测的项目设计合适的人体检测框。关键点的生成分为自顶向下和自下向上两种构造方式，其中前者针对视频或图片中的单人场景，而后者针对的是多人同时检测的场景。因为本课题项目所针对的是单人体育测试的场景，所以采用的是自顶向下的关键点生成。这要求为体育测试中的学生设立人体检测框。针对于每种体测项目的不同，为了达到识别的精确性，单独训练了较为合适的检测框，旨在让识别效果达到更好。

在体测视频中因为场景的复杂性，并且人体存在遮挡的现象，有时候识别出的骨骼关键点并非是可信的。因此，衡量在关键点识别模块计算出的骨骼点是否可信，有着重要的作用。通过为每一个关键点及它周围的背景与前序帧的差异，判断该关键点的置信度。为每一个场景单独设置了置信度的阈值，如果在体测执行中的关键点，置信区间小于阈值的话，就会将此帧的该点从判别逻辑中移除。从而使得系统更为健全。

3.3. 动作检测以及评价模块

如图 3 所示，各个体测项目功能主要由三个核心模块组成，分别为计数模块、动作检测模块和动作建议模块。当系统成功检测到关节点后，根据不同体测项目的评判标准进行关节点间的特征计算，最后判断计算值是否符合标准，以此进行计数并给定相应的动作建议。下面以斜身引体、引体向上、双杠臂屈伸、仰卧起坐为例，具体介绍系统模块。

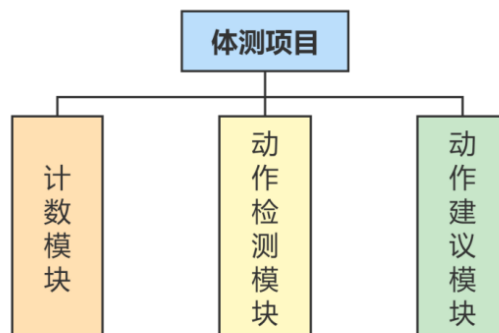


图 3 体测项目结构功能图

3.3.1. 斜身引体项目

斜身引体项目功能模块分别为计数模块、动作检测模块和动作建议模块。在检测过程中，动作检测模块用于判断视频关键帧的动作是否始终满足要求，如果不满足会根据动作建议模块判断所属常见错误动作类型，并及时提醒测试人员。计数模块在检测过程中判断体考测试者是否满足一次完整动作的起始状态和终止状态，并根据计数控制模块决定是否计数。本项目检测关键点为手腕、手肘、肩部、髋部，膝盖以及脚踝。

检测示意图如图 5(a) 所示，准备动作时身体与地面约 45 度角，胳膊与身体约 90 度角，完成动作时下巴需要超过横杠。在测试过程中不得塌腰挺腹或曲腿，并且双腿双脚不得随意移动。完成一次动作后需要回到准备状态。测试流程为，在检测正式开始后首先获取初始状态，以“定制化”获取具有不同身型特征的体考同学的基准关键信息。如果初始准备姿势不正确系统会提醒，直到体考同学满足准备动作的要求才会进入检测。检测开始后获取关键帧并对关键点序列进行判断，在满足“开始”和“完成”的前提下，如果符合动作规范则成功计数，否则输出提醒信息。检测结束后输出计数结果，具体如下：

开始检测时，记录初始状态下手腕、手肘、肩部，髋部，膝盖以及脚踝的关键点，分别记为 $P_{wrist_init}, P_{elbow_init}, P_{shoulder_init}, P_{hip_init}, P_{knee_init}, P_{ankle_init}$ 。后续获取关键点序列 $P_n, n \in \{0, 1, \dots, 33\}$ ，根据横纵坐标 P_x, P_y 进行肢体的角度判断。

准备动作时，躯干胳膊角度，躯干地面角度需要满足条件：

$$85 \leq \theta_{arm_body} \leq 95, 40 \leq \theta_{body_floor} \leq 50$$

检测过程中始终保持躯干挺直，腿部伸直，即满足：

$$\theta_{max_body} \leq 15, \theta_{max_leg} \leq 15$$

完成动作需要满足，身体与地面的夹角与初始状态下手腕、脚踝与地面的夹角的差值的最大值小于阈值，即：

$$\max_{t_k \leq t \leq t_{k+1}} (\theta_{wrist_ankle_floor_init} - \theta_{body_floor_now}) \leq 10, (k \text{ 定义如下})$$

当准备动作和完成动作都符合要求，并且在检测过程中没有犯规行为，则本次动作成功计数，并且继续判断流程。假设 t_k 意味着第 k 次满足准备动作，在时间段 $T_k = t_{k+1} - t_k$ 内，定义动作完成情况如下：

$$a(t_k) = \begin{cases} 0, & \text{未达到完成动作标准或者动作违规} \\ 1, & \text{达到完成动作标准且动作正确} \end{cases}$$

斜身引体检测流程完成后，根据满足准备动作的次数 K ，输出最终计数结果为：

$$\sum_{k=1}^K a(t_k), a(t_k) \in \{0, 1\}$$

3.3.2. 引体向上项目

引体向上项目检测关键点为鼻子、嘴、手腕、手指、手肘、肩部，髋部。

检测功能模块可以划分为：计数模块、动作检测模块和动作建议模块。

计数模块：当体侧者完成一次正确的引体动作后，成功动作次数+1，要求体侧者需下巴过杠，若下巴未过杠则不计数。

错误动作检测模块：开始体侧时，体侧者需双臂伸直，双手与肩同宽或略宽抓杠，若不满足以上条件则会视为未达到开始准备动作条件。

动作建议模块：当以上两个模块发生错误或需要提示信息时，会在界面左上方显示对应的提示信息，保证监测者第一时间了解动作的不足之处并提醒体侧者调整动作。

3.3.3. 双杠臂屈伸项目

双杠臂屈伸项目的检测分为三个大的模块，分别为计数模块、动作检测模块和动作建议模块。在检测过程中，动作检测模块用于判断视频关键帧的动作是否始终满足要求，如果不满足会根据动作建议模块判断所属常见错误动作类型，并及时提醒测试人员。计数模块在检测过程中检测测试人员是否满足一次完整动作的起始状态和终止状态，并根据计数控制模块

决定是否计数。项目检测关键点为手腕、手肘、肩部，髌部，膝盖以及脚踝。

测试流程为，首先在监测开始时，识别每名参加体测同学的初始姿态，一方面这是为了避免在开始识别动作前就检测到了结束的姿态。另一方面，从学生进入画面至达到准备阶段的这段时间，可以让检测系统能更好的针对该名学生的信息进行分析。如果初始姿势不正确系统会提醒，直到体考同学满足准备动作的要求才会进入检测。在检测过程中，如果关键点的置信度较低，则会提示人体拍摄不全。在关键点符合可信度阈值后，根据帧中的关键点信息判断姿势是否正确，是否到达准备状态，是否到达完成状态，是否在满足“开始”和“完成”的前提下符合计数要求。如果是则计数增加，否则继续下一帧的判断或者是输出姿态错误的提示信息。

3.3.4. 仰卧起坐

仰卧起坐项目的检测核心模块为计数模块、动作检测模块和动作建议。此外还用到了人体关节特征计算特征模块作为辅助。在检测过程中，动作检测模块用于判断视频关键帧的动作是否始终满足要求，如果不满足会根据动作建议模块判断所属常见错误动作类型，并及时提醒测试人员。计数模块在检测过程中检测测试人员是否满足一次完整动作的起始状态和终止状态，并根据计数控制模块决定是否计数。

该项目将摄像头捕获的每一帧进行处理，利用人体骨关节检测技术进行目标检测，判断人体是否进入检测画面以及检测到 33 个关节的置信度是否达标，当条件满足后，进入计数状态。在计数过程中，将全程计算臀部与双脚离地距离来判断是否有动作违规，同时在仰卧起坐上升与下降的拐点处分别通过计算膝盖与手肘距离以及头部手肘与地面距离来评估动作是否规范。部分动作规范如图 5(e) 所示，若动作达标则计数加一，否则不进行计数，并将显示不规范动作建议以此提醒测试人员。

4. 实验与讨论

4.1. 视频数据收集

对学生体育测试智能评测系统的设计过程中，需要广泛收集每项动作的标准动作范例和错误动作示范的视频数据集，用以检测过程的设计和测试。我们对九个项目的测试规则、成绩判定规则和关键动作评判标准做了详细的调研和规范，包括引体向上、仰卧起坐、立定跳远、一分钟跳绳、坐位体前屈、原地纵跳摸高、双杠臂屈伸、斜身引体和实心球。并且收集了符合正确计数要求，以及常见错误示例的视频数据集。例如宾夕法尼亚动作数据集和 HMDB51 数据集，以及实地拍摄的动作视频。数据集的来源丰富，包含动作类型广泛，贴合本项目的使用场景。

4.2. 测试环境需求

硬件环境需求：1. 使用装有 Android 操作系统的设备部署和运行体考评测系统（运行内存大于等于 8GB）。2. 拍摄过程中检测设备不能有明显的晃动，需要放置在平稳、高度适宜的设备支架上。

自然环境要求：1. 要求视频中只能出现体测者本人，不能有 2 个及 2 个以上的人同时出现在视频中。2. 体测者身穿衣服颜色应尽量与背景颜色存在较大差别，不应同色或相近颜色。3. 应尽量选择光源充足的场地且不能出现光线过亮导致无法看清人物的情况。4. 检测过程中需要将所需身体关键点全部露出，不得遮挡。

4.3. 检测效果评估

4.3.1. 检测效果评估

对四个体测项目进行准确率评估，针对每个项目的检测算法，都采用准确率 Accuracy 作为性能评估指标[3]，其公示见式为：
$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$$

式中各符号含义见表 1 所示，以斜身引体项目为例，将人工肉眼观测计数的结果作为准确参考，针对每一次动作，当人工计数为 1，并且系统检测计数也为 1 时，TP 累加；当人工计

数为 1, 但是系统检测计数为 0 时, FN 累加; 当人工计数为 0, 但是系统检测计数为 1 时, FP 累加; 当人工计数为 0, 并且系统检测计数也为 0 时, TN 累加; 最后得出每个检测项目的评估结果制作成柱状图, 如图 4 所示。

表 1 TP、FN、FP、TN 的含义解释

检测计数 \ 人工计数	1	0
	1	TP
0	FP	TN

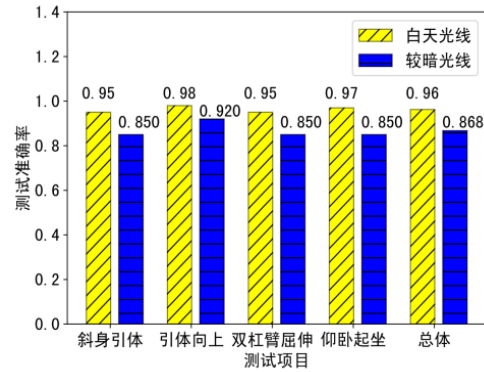


图 4 测试准确率评估

4.3.2. 测试结果展示

下图为实际检测效果, 展现了以斜身引体、引体向上和仰卧起坐为例的智能评测系统的检测结果, 当动作符合测试标准时刷新计数结果, 当动作不规范时给出评测建议, 提醒测试者即时更正。



图 5 测试结果示例 (a)-(e) 为正确动作计数示例, (f)-(j) 为错误动作提示示例

4.4. 实验结果讨论

检测结果显示,在日间光线充足,设备要求满足时,检测准确率皆达到了95%以上,足以辅助体育教学工作者开展体考测试工作。相比于类似产品:Keep和天天跳绳,在相同的项目上达到了接近的检测效果。并且,本智能评测系统考虑到了不同学生的体形特征和考试场景的具体要求,在特定的体考测试环境中会有更好的检测效果。

5. 总结

“体育强则国强”,加强体育锻炼、改善身体素质,是提升我国整体国民体质现状的重要方法。因此,体育考试在中学生综合素质评价中变得越来越重要,也逐步成为了统一的测试模式。传统基于人工的体育成绩评测方法不仅存在着无法避免的测试误差,而且测试效率较低,通常需要多名体育教师同时对一名学生的考试做定性或定量的成绩评价。针对当前智慧课堂在体育中考的实际应用需求,结合当前视频采集技术、人体姿态估计方法等先进技术,本文设计和实现了基于BlazePose的学生体育测试智能评测系统。对体育中考中常见的多个体测项目进行检测,实地测试时,在确保测试准确率的同时对学生动作提出改进建议,能够很好地辅助学校推行体育健康工作,更有助于体育课程的教师团队完善和提升教学内容,适时地转变教学策略[4]。在后续的工作中,我们将进一步考虑视频模糊、遮挡情况下对体测过程造成的影响,使该系统能够精准应用于不同学校差异化的体育环境场景,进而充分利用AI技术赋能体育教学,全面服务体育强国战略。

6. 致谢

这项工作得到了如下项目的支持:国家自然科学基金面上项目(编号:62172054、62072047)、国家重点研发计划国家自然科学基金项目(编号:2019YFB1802603)、北京自然科学基金重点项目(编号:M21030)、北邮-创世云联合实验室(编号:A2022164)。

参考文献

- [1] 何雨泽,余舒雯,张勉,田园心语 & 郑永和(2022).科学教育中以智能手机作为实验工具的研究综述,第二十六届全球华人计算机教育应用大会,64-65.
- [2] 李思颖 & 王露莹.(2022).人工智能在高中教育领域应用的研究热点与展望:第二十六届全球华人计算机教育应用大会,305-306.
- [3] 季利鹏,吴世龙 & 顾小清.(2022).基于MTCNN-AlexNet的学习状态检测系统设计,第二十六届全球华人计算机教育应用大会,381-388.
- [4] 周彦丽,李彤彤,边雨迎,李国涛 & 郭翔宇.(2022).基于情感分析和主题挖掘的课程交互文本分析模型构建与实践,第二十六届全球华人计算机教育应用大会,463-470.
- [5] 舒杭 & 顾小清.(2022).教育数字化转型的现实基础与行动框架:现代教育技术(11),24-33.
- [6] Bazarevsky, V., Grishchenko, I., Raveendran, K., Zhu, T., Zhang, F., & Grundmann, M. (2020). *Blazepose: On-device real-time body pose tracking*. arXiv preprint arXiv:2006.10204.
- [7] Chen, Y., Shen, C., Wei, X. S., Liu, L., & Yang, J. (2017). *Adversarial poseNet: A structure-aware convolutional network for human pose estimation*. In Proceedings of the IEEE International Conference on Computer Vision (pp. 1212-1221).
- [8] Tang, W., Yu, P., & Wu, Y. (2018). *Deeply learned compositional models for human pose estimation*. In Proceedings of the European conference on computer vision (ECCV) (pp. 190-206).
- [9] Toshev, A., & Szegedy, C. (2014). *DeepPose: Human pose estimation via deep neural networks*. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1653-1660).