

基于多通道注意力的人体跳跃屈膝角度识别模型

丁磊¹, 方晨安¹, 胡新荣¹, 李阳¹, 明文凯¹, 林林², 吴渊¹

(1. 武汉纺织大学 计算机与人工智能学院;
2. 武汉纺织大学 体育部, 湖北 武汉 430200)

摘要: 目前,深度学习与可穿戴设备的结合在运动定量评估领域得到了广泛应用,并展现出重要的理论价值和应用潜力。然而,传统神经网络方法在处理具有不同运动质量的细粒度动作时,难以有效提取关键的多维度特征,从而限制了其在精确识别和深入分析个体运动行为时的性能。为了解决这一问题,提出一种多通道注意力模型(Conv+CEA+AG),旨在通过多通道注意力(CEA)机制增强对动作重要特征的捕捉能力,从而提升运动行为分析的精确性。首先,采集不同屈膝角度下的人体跳跃运动数据,并对传感器信号进行预处理,提取有效的跳跃周期信号以构建跳跃运动数据集。其次,将集成多通道注意力(CEA)的神经网络应用于数据集,构建了跳跃运动屈膝角度检测模型。实验结果表明,该模型在分类任务中达到了94.32%的准确率,在回归任务中实现了2.04°的平均绝对误差。与传统的神经网络模型(Conv+GRU)及现有模型(Conv+CIE+AG)相比,准确率分别提升了12.47%和6.08%,平均绝对误差分别下降了3.49°和3.09°,证明了该模型在屈膝角度识别方面表现更优异。

关键词: 人体跳跃;屈膝角度识别;可穿戴传感器;深度学习;多通道注意力

DOI:10.11907/rjtk.251093

扫描二维码阅读全文:



中图分类号:TP391.41

文献标识码:A

文章编号:1672-7800(2026)004-0048-09

Human Jumping Knee-Bending Angle Recognition Model Based on Multi-Channel Attention

DING Lei¹, FANG Chen'an¹, HU Xinrong¹, LI Yang¹, MING Wenkai¹, LIN Lin², WU Yuan¹

(1. School of Computer Science and Artificial Intelligence, Wuhan Textile University;
2. Department of Physical Education, Wuhan Textile University, Wuhan 430200, China)

Abstract: Currently, the combination of deep learning and wearable devices is widely used in the field of quantitative movement assessment and shows important theoretical value and application potential. However, traditional neural network methods are difficult to effectively extract key multidimensional features when dealing with fine-grained actions with different motion qualities, which limits their performance in accurately recognizing and deeply analyzing individual motion behaviors. To address this problem, a multi-channel attention model (Conv+CEA+AG) is proposed, aiming to enhance the ability to capture important features of the action through the multi-channel attention (CEA) mechanism, and thus improve the accuracy of the analysis of motor behavior. Firstly, human jumping motion data under different knee flexion angles are collected, and the sensor signals are preprocessed to extract effective jumping cycle signals in order to construct a jumping motion dataset. Subsequently, a neural network with integrated multi-channel attention is applied to the dataset to construct a jumping motion knee-bending angle detection model. The experimental results show that the proposed model achieves 94.32% accuracy in the classification task and 2.04° mean absolute error in the regression task. Compared with the traditional neural network model (Conv+GRU) and the previous proposed model (Conv+CIE+AG), the accuracy is improved by 12.47% and 6.08%, and the average absolute error decrease by 3.49° and 3.09°, respectively, which prove that the model demonstrate a better performance in knee-bending angle recognition.

Key Words: human jumping; knee-bending angle recognition; wearable sensor; deep learning; multi-channel attention

收稿日期:2025-03-18

基金项目:湖北省自然科学基金一般项目(2022CFB563)

作者简介: 丁磊(1987-),男,博士,武汉纺织大学计算机与人工智能学院副教授,研究方向为智能传感监测、步态分析及模式识别;方晨安(2001-),男,武汉纺织大学计算机与人工智能学院硕士研究生,研究方向为柔性传感、运动检测;吴渊(1991-),男,博士,武汉纺织大学计算机与人工智能学院讲师,研究方向为无线技术、柔性可穿戴技术及智能健康。本文通讯作者:吴渊。

0 引言

目前,深度学习与可穿戴设备相结合的方法在动作定量评估领域发展迅速,这种技术融合为运动分析提供了新思路^[1]。通过深度学习算法对可穿戴设备收集的生理和运动数据进行分析,研究者能够更准确地量化和评估运动表现。该技术的应用不仅提高了运动训练的科学性,而且为运动损伤预防和健康监测提供了有效工具,具有重要的研究意义和实际应用价值。

随着物联网技术的快速发展,目前对跳跃监测系统的研究根据所使用的设备可分为两大类。第一类是基于摄像头的方法。例如,Drazan等^[2]利用Qualisys 300和310运动捕捉系统,通过摄像头收集数据,以获取垂直跳跃过程中下肢在矢状面上的角度信息。然而,这种基于摄像头的监测方法存在一些缺陷,包括可能对用户隐私造成侵犯,以及对环境因素如光照条件和遮挡物较为敏感^[3]。相比之下,基于可穿戴传感器的方案凭借小巧便携、不受环境制约的优势成为新趋势,从而催生了用户对精度高、实时运动跟踪的迫切需求。第二类是基于惯性测量单元(Inertial Measurement Unit, IMU)的方法。IMU是一种体积小、成本低廉的传感器,因其具有穿戴时的非侵入性和运动数据监测的准确性等优点,正逐渐成为研究热点。目前,基于可穿戴IMU的跳跃运动评估受到越来越多人的关注。相关研究涵盖多个方面,包括但不限于跳跃高度估计^[4]、跳跃子阶段识别^[5]、跳跃相关参数估计^[6-7],以及跳跃过程中的关节角度测量^[8]等。

在研究过程中,本文面临的主要挑战在于不同屈膝角度下执行跳跃运动产生的生物力学差异微小,导致跳跃运动的精确识别较为困难。传统的神经网络方法尽管在许多应用中表现出色,但在处理此类相似度高的运动数据时,其特征提取能力往往受限。这些方法难以捕捉到跳跃运动中的关键细微变化,从而影响动作识别的准确性和可靠性。

因此,本文提出一种多通道注意力模型(Convolutional Combined-Enhanced-Attention Attentive-Gate, Conv+CEA+AG)。该模型借鉴深度学习领域的最新进展,特别是注意力机制理论,以增强模型对数据中关键信息的捕捉能力。通过精心设计的多通道架构,该模型能够并行处理来自可穿戴设备的多通道数据流,有效地从原始数据中识别出反映跳跃运动特征的细微信号。基于这种先进的数据处理和特征提取技术,本文能够从复杂的传感器数据中提取出关键运动特征,并对跳跃动作进行更精确的量化和评估。

本文结合3种注意力机制来构造多通道注意力模型(Conv+CEA+AG):①协调注意力机制:将空间坐标与通道特征进行融合,突出位置信息在传感器活动识别中的重要性;②外部注意力机制:将外部记忆单元与传感器数据相

结合,使模型充分利用全局数据特征;③时间注意力机制:以时间上下文为参照,计算传感器数据在各时间点的权重,以此量化并突出不同时间步长在活动识别任务中的贡献度。最后,在自建的跳跃运动数据集上进行实验,本文方法取得了比传统神经网络模型(Convolutional Gated Recurrent Unit, Conv+GRU)和文献[9]提出的模型(Convolutional Cross-Interactive-Encoder Attentive-Gate, Conv+CIE+AG)更好的活动识别效果,验证了该方法的有效性。

本文的主要贡献如下:

(1)提出一种多通道注意力网络,该网络致力于学习丰富的特征表示,捕获细粒度运动和多个传感器通道之间的潜在关联性,以实现更精确的运动识别。

(2)设计一个注意力门控循环单元(Gate Recurrent Unit, GRU)模块,旨在通过进一步挖掘时间上下文信息来增强序列特征的表达能力。

(3)对模型的末端全连接层进行了优化调整,以适应模型在分类和回归任务中的多样化需求,从而提升模型在不同任务场景下的应用效能。

1 相关工作

1.1 基于摄像头的方法

基于摄像头的计算机视觉技术在健身锻炼和体育评估方面展现出应用潜力。Wochatz等^[10]通过微软Kinect V2系统对几种基本运动动作评估的可靠性进行研究,发现Kinect系统在测量关节角度和位置时的可靠性存在问题,并在不同运动任务及运动平面中显示出测量偏差。Freitas-Junior等^[11]使用高速摄像机系统(HSC-Kinovea)对运动员进行间歇性垂直跳跃测试,以评估其下肢的爆发力与耐力。同时,Strutzenberger等^[12]应用一种无标记视频动作捕捉系统(Miqus)测量反动作跳(Countermovement Jump, CMJ)过程中的关键运动学参数,包括跳跃高度和下肢关节的屈曲角度。以上研究表明,尽管基于摄像头的系统在运动评估中具有应用前景,但在精确性和可靠性方面仍需进一步改进。

1.2 基于IMU的方法

基于IMU的可穿戴技术已被证实在跳跃运动分析中具有高效性。Mascia等^[4]开发了一种利用智能手机内置IMU的跳跃高度估计方法,通过训练机器学习模型能够从IMU数据中估计跳跃高度。Lu等^[5]利用多传感器信息融合和机器学习来识别人类的跳跃阶段,这对于辅助跳跃外骨骼的发展至关重要。Jaitner等^[6]设计一种用于监测田径运动员爆发性跳跃事件的可穿戴IMU设备,能够估计跳跃持续时间和反应力指数(Reaction Strength Index, RSI),对运动员表现分析和训练效果评估具有重要价值。此外, Sun等^[7]提出一个模块化的长短期记忆(Long Short-term Memory, LSTM)模型,用于实时监测单腿和双腿跳跃落地

过程中的垂直地面反作用力(Vertical Ground Reaction Force, vGRF)和膝关节伸展力矩(Knee Extension Moment, KEM)。Olawore 等^[8]则使用 IMU 技术测量 CMJ 和侧向滑板跳跃(Lateral Skateboard Jump, LSJ)中的踝关节角度峰值和运动范围(Range of Motion, ROM)。以上研究展示了 IMU 在捕捉跳跃运动动态特征方面的应用潜力。

本文提出的基于多通道注意力网络的跳跃屈膝角度分析方法可以接受平行化输入的 IMU 传感器信息,通过特定的注意力模块精确地揭示各通道在全局中的重要性。同时,该方法提出一种可以充分挖掘 IMU 传感器所捕获角度特征信息的多通道注意力机制,使网络在训练过程中更有效地学习不同屈膝角度的特征,以准确识别跳跃动作中屈膝的细微幅度差异。

2 数据采集及预处理

2.1 设备介绍

本文实验采用 Noitom 公司开发的惯性传感器设备。如图 1 所示,该设备是一个惯性测量单元(IMU),由一个三轴加速度计和一个三轴陀螺仪组成。IMU 固定在弹性护膝上,方便实验人员穿戴。IMU 的穿戴位置和三轴方向如图 1 所示。实验人员穿戴 IMU 执行跳跃时,该设备以 100 Hz 的采样频率(Sampling Rate, SR)收集传感器数据。由于 IMU 具有体积小、实时性强、便于穿戴等特点,能很好地满足实验需求。



Fig. 1 Experimental equipment

图1 实验设备

2.2 数据采集

总共有 10 名志愿者穿戴本文的设备进行实验,包括 6 位男性,4 位女性。本文实验目的是采集人体蹲跳(Squat Jump, SQJ)的屈膝角度标签。屈膝角度被定义为大腿延长线与小腿之间形成的夹角(如图 2 中红色虚线与蓝色实线之间形成的角度)。在实验过程中,要求志愿者以 30°、40°、50°、60°、70°、80° 和 90° 的屈膝角度进行蹲跳。这些角度是通过关节活动测量仪进行衡量的,图 2 中黑色虚线框内即为实物图片。每位志愿者要完成不同角度下的蹲跳实验。在每项实验前,每位志愿者都进行标准化的热身。实验过程中,每位志愿者要做 6 组跳跃,其中每组包括 5 次跳跃动作,每组跳跃间隔 3 ~ 5 min,每次跳跃动作间隔 1 ~ 3 s。最后,经过设备采集得到跳跃动作的加速度计和陀螺

仪原始信号。

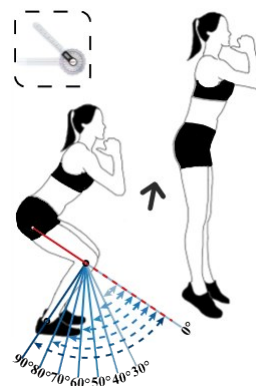


Fig. 2 Schematic diagram of the action

图2 动作示意图

2.3 数据预处理

为了方便提取跳跃周期,需对加速度计的三轴信号进行初步处理。由于 IMU 传感器测量的加速度包含 3 个正交分量,为精确表征运动动态,需对其进行矢量合成以计算加速度幅度 a_m ,即通过三轴分量 a_x 、 a_y 和 a_z 合成得到。具体表达式为:

$$a_m = \sqrt{a_x^2 + a_y^2 + a_z^2} \quad (1)$$

然后,使用一个截止频率为 2.5 Hz 的低通滤波器对 a_m 信号进行平滑处理,以减少噪声并保留信号的低频特性。最后,平滑后的 a_m 信号将用于后续的时间序列分割。这一过程可分解出时间序列数据中的多个有效周期片段,以便于分析和识别其中的事件模式。

为了检测出原始 IMU 信号中存在的跳跃周期,使用窗口峰值算法对 a_m 信号进行分割。该方法通过识别两个加速度峰值来定位每个跳跃周期。如图 3(a)所示,两个峰值代表起飞阶段和着陆阶段的最大峰值,分别记作 $t_{max_{to}}$ 和 $t_{max_{lan}}$ 。该方法的合理性已得到文献[13]和文献[14]研究结果的验证。

因此,首先设定时间窗口大小,依据 a_m 信号的功率谱密度(Power Spectral Density, PSD)确定其主导频率。该频率(f_p)对应 PSD 图中最为显著的峰值,如图 3(b)所示。这一频率峰值指导时间窗口的配置,确保窗口尺寸(Window Size, $WS=(1/f_p)*SR$)与信号的周期性相适应,以优化跳跃周期的识别。之后,先在时间窗口内搜索这两个峰值。在定位到两个峰值点后,从第一峰值 $t_{max_{to}}$ 处开始向前寻找信号幅度不大于重力加速度阈值($g_0 \approx 0.98$)的第一个点,确定为跳跃段的起点,再从第二峰值 $t_{max_{lan}}$ 处开始向后寻找信号幅度不大于 g_0 的第一个点,确定为跳跃段的终点。最后,移动窗口对整个 a_m 时间序列进行搜索。在找到一个跳跃片段后,即从该片段结束后继续搜索。如果在窗口内没有发现跳跃片段,则将窗口整体向后移动一个样本点。搜索过程将持续到提取出所有跳跃片段为止。通过分割算法对 a_m 时间序列的处理,得到所有跳跃片段的起始点和结束点。将这些起始点和结束点应用到加速度计

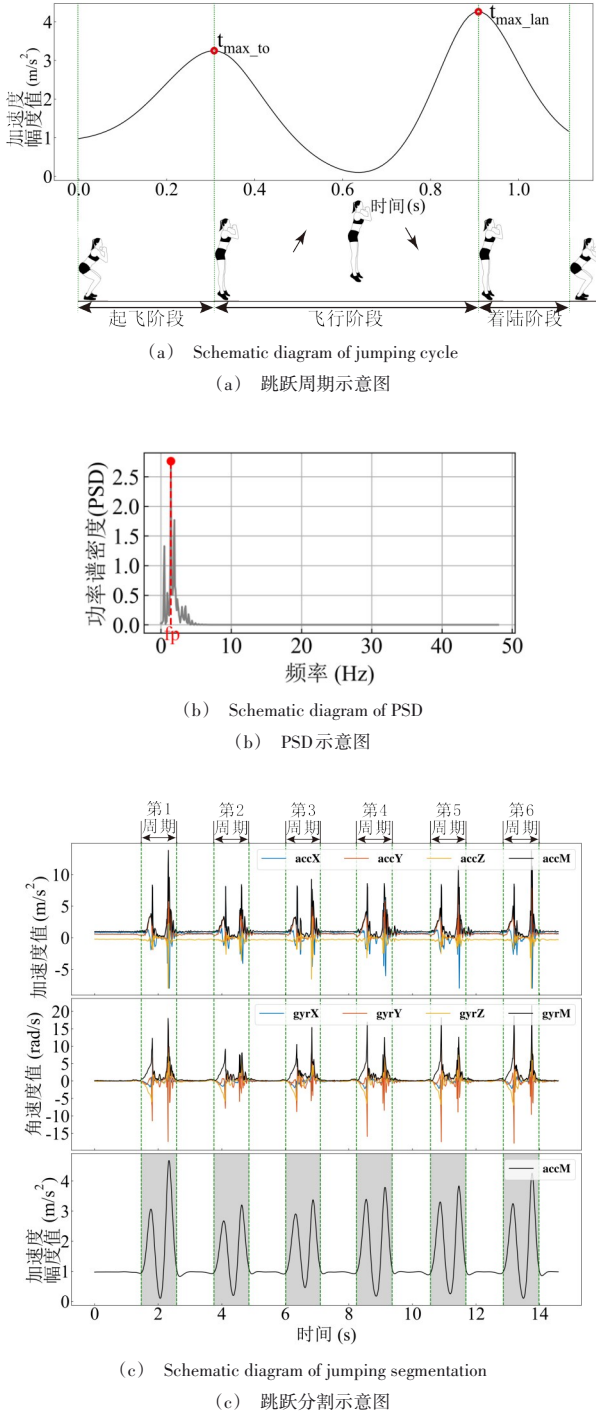


Fig. 3 Schematic diagram of the jump IMU signals

图3 跳跃IMU信号示意图

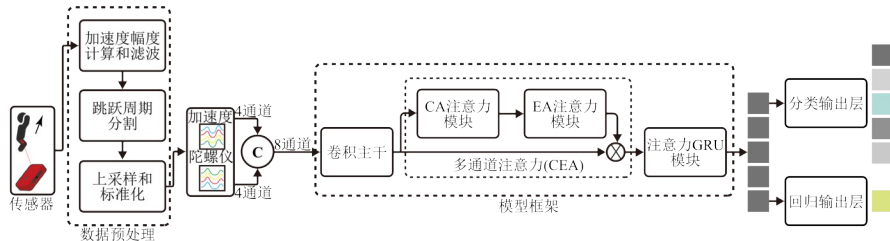


Fig. 4 System framework

图4 系统框架

与陀螺仪原始信号上,即能分割出所有跳跃信号。分割结果如图3(c)所示。

由于屈膝角度的差异,分割得到的片段长度不尽相同。考虑到后续模型输入需要长度一致的片段序列,因此将每个片段的长度上采样到300。最后,采用Z-score标准化对每个片段进行归一化,以减小片段中异常样本点对模型精度的影响。

3 模型设计

本文旨在搭建一个多通道注意力模型。该模型先将加速度计和陀螺仪的原始信号作为输入,再将模型的表征输出作为两个全连接层的输入。最后,这两个全连接层分别输出动作分类和角度回归结果。本文系统实现的整体流程框架如图4所示。首先,模型使用一维卷积神经网络(Convolutional Neural Network, CNN)提取每个传感器通道的初始特征表示。其次,初始特征表示通过多通道注意力(Combined Enhanced Attention, CEA)模块进一步细化成增强特征图,从而使模型具有方向感知和位置敏感的特性,并能隐含地考虑所有数据样本之间的相关性。最后,模型采用注意力GRU处理增强特征图,从而使模型进一步关注到时间域上的注意力权重。以下介绍模型的子模块。

3.1 卷积主干

数据预处理结束后,模型首先得到时间序列切片 $\mathbf{x} \in R^{W \times D}$ 作为输入。其中, W 表示切片时间长度, D 表示传感器通道数量。然后,一维卷积主干沿着输入切片的时间维度自动提取特征。因此,卷积主干网络为每个传感器通道学习到初始特征表示 $\mathbf{x} \in R^{C \times T \times D}$ 。其中, C 代表生成的特征图个数, T 为下采样后的切片时间长度,而 D 仍是传感器通道数。本文采用最先进的(Human Activity Recognition, HAR)模型的卷积层作为主干特征提取器,输入切片被送入由4个卷积层组成的序列中进行处理^[15]。其中,每层卷积都由64个长度为5的一维滤波器组成。

3.2 多通道注意力

将卷积主干生成的一个初始特征图 $\mathbf{F} \in R^{C \times T \times D}$ 作为输入,CEA模块首先生成一个3维协调注意力图 $\mathbf{M}_c \in R^{C \times T \times D}$ 。该协调注意力图是由一对2维方向感知和位置敏感的注意力图 $\mathbf{M}_d \in R^{C \times 1 \times D}$ 与 $\mathbf{M}_t \in R^{C \times T \times 1}$ 相乘得到的。之后,生成一个3维外部注意力图 $\mathbf{M}_e \in R^{C \times T \times D}$ 。

CEA 模块的概述如图 5(c) 所示。整个多通道注意力计算过程可概括为:

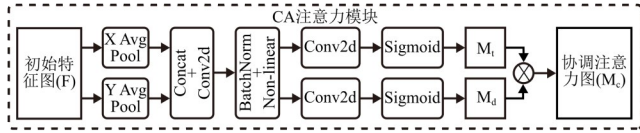
$$F' = M_c(F) = M_d(F) \times M_l(F) \quad (2)$$

$$F'' = M_c(F') \times F \quad (3)$$

式中, $\times$ 表示逐元素乘法, F'' 表示最终的增强输出。

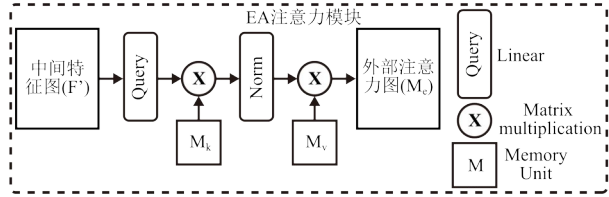
3.3 CA 注意力模块

协调注意力 (Coordinate Attention, CA) 是 Hou 等^[16]提出的一种新的移动网络轻量级注意力机制。ImageNet 分类、目标检测和语义分割领域的相关实验验证了该注意力机制的有效性。将该 CA 模块引入模型是因其能沿着垂直方向对传感器通道之间的关系进行建模, 同时利用水平方向的时间位置信息捕获长程依赖关系。使用该注意力机制来聚合两个空间方向上的特征, 所生成的注意力图可以补充和增强初始特征图的表示。CA 模块的概述如图 5(a) 所示。下文描述 CA 的内部计算过程。



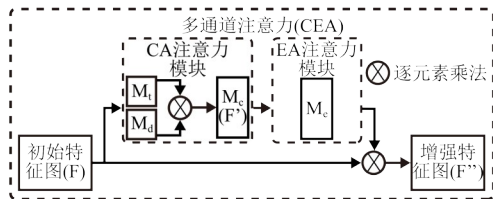
(a) Overview of the CA module

(a) CA 模块概述



(b) Overview of the EA module

(b) EA 模块概述



(c) Overview of the CEA module

(c) CEA 模块概述

Fig. 5 Overview of the different sub-modules

图 5 不同子模块概述

给定输入 X , CA 模块首先使用两个空间大小为 $(T, 1)$ 和 $(1, D)$ 的池化核, 分别沿着水平和垂直方向对每个通道进行编码。因此, 在时间 t 处第 c 个通道的输出可以表示为:

$$z'_c(t) = \frac{1}{D} \sum_{0 \leq i < D} x_c(t, i) \quad (4)$$

在传感器通道 d 处第 c 个通道的输出可以表示为:

$$z'_c(d) = \frac{1}{T} \sum_{0 \leq i < T} x_c(j, d) \quad (5)$$

经过上述两个公式的转换, 生成一对带有方向感知和

位置敏感的特征图。之后, CA 模块将这对特征图拼接起来, 并将其输入共享的 1×1 卷积变换函数 F_1 。进而得到:

$$f = \delta(F_1([z', z''])) \quad (6)$$

式中, $[\cdot, \cdot]$ 表示沿空间维度的拼接操作, δ 是非线性激活函数, $f \in R^{C_{lr} \times (T+D)}$ 是在两个空间方向上编码生成的中间特征图。之后, CA 模块沿着空间维度将 f 分解为两个独立的 $f^t \in R^{C_{lr} \times T}$ 和 $f^d \in R^{C_{lr} \times D}$ 张量。参数 r 作为缩减率, 用于控制块尺寸的缩放比例。此外, CA 模块使用两个 1×1 卷积变换 F_t 和 F_d , 将 f^t 和 f^d 分别转换为与输入 X 具有相同通道数的张量。从而生成:

$$g^t = \sigma(F_t(f^t)) \quad (7)$$

$$g^d = \sigma(F_d(f^d)) \quad (8)$$

式中, σ 是 sigmoid 函数。之后, 生成的 g^t 和 g^d 被展开, 并分别用作注意力权重。最后, CA 模块的输出 Y 可以表示为:

$$y_c(i, j) = g^t(i) \times g^d(j) \quad (9)$$

3.4 EA 注意力模块

外部注意力 (External Attention, EA) 是由 Guo 等^[17]提出的一种新型轻量级但有效的注意力机制, 可用于各种视觉任务。相较于自注意力 (Self-attention), EA 模块具备线性计算复杂度的显著优势, 能够有效克服自注意力在长序列处理中的计算瓶颈, 同时可以隐含地关注所有数据样本之间的相关性。因此, 本文在模型中添加该 EA 模块。EA 模块的概述如图 5(b) 所示。EA 模块通过式 (10)、式 (11) 计算输入样本点与外部存储单元 $M \in R^{S \times d}()$ 之间的注意力。其中, d, S 分别表示特征维度和存储单元中的元素数量。

$$A = \text{Norm}(FM_k^T) \quad (10)$$

$$F_{out} = AM_v \quad (11)$$

式中, F 是输入特征图, M_k, M_v 是分别用作键与值的不同存储单元, A 是从该学习的数据集级先验知识推断出的注意力图, F_{out} 则是输出的外部注意力图。Norm 是在文献 [18] 中提出的双归一化, 其定义了从原始相似度 $\tilde{\alpha}$ 到最终注意力权重 α 的双归一化流程: 先经过指数归一化, 再经过二次比例缩放。具体公式如下:

$$(\tilde{\alpha})_{i,j} = FM_k^T \quad (12)$$

$$\alpha_{i,j} = \exp(\tilde{\alpha}_{i,j}) / \sum_k \exp(\tilde{\alpha}_{k,j}) \quad (13)$$

$$\alpha_{i,j} = \alpha_{i,j} / \sum_k \alpha_{i,k} \quad (14)$$

3.5 注意力 GRU 模块

鉴于利用注意力机制建模活动的时间上下文已被证实具有显著优势, 本文模型最后采用 2 层注意力 GRU (AG) 模块对生成的增强表征序列进行时序特征提取与整合^[19]。通常, GRU 编码器在查看整个时间序列后只是输出最后时刻的隐藏状态, 而该 AG 模块能够学习并生成所有时刻隐

藏状态 $(h_t)_{t=1}^T$ 的软注意力权重。通过该方式,模型能够自适应地量化各时间步长的相对重要性,并最终总结出一个包含关键时序信息的全局摘要表示:

$$z = \sum_t \beta_t h_t \quad (15)$$

式中, β_t 是时间步长 t 的计算注意力权重。

4 实验与结果分析

4.1 实验环境

实验在Windows10专业版64位操作系统下进行,采用Pycharm编译环境,使用Python3.8.18、Pytorch1.13.1、CUDA11.6搭建网络模型,并在NVIDIA GeForce GTX 1050显卡上进行模型训练。

4.2 实验参数设置

实验数据为本文第2章收集的数据集,共4880个样本数据。其中,3525个样本作为训练集、623个样本作为验证集,732个样本作为测试集。在训练过程中,为了防止过拟合,在AG模块和最后输出层的前面加入dropout层(保留概率取值为0.5)。此外,将AG模块内部GRU神经元的衰减率设置为0.25。在卷积主干中,四层卷积选择整流线性单元(Rectified Linear Unit, ReLU)作为激活函数。在分类输出层中,线性层使用softmax函数输出跳跃动作的分类概率。在回归输出层中,线性层不经过激活函数直接输出屈膝角度预测值。本文选择Pytorch库中的Adam优化器来训练模型,因为Adam优化器能自适应调整学习率,从而提高模型的收敛速度和泛化能力。另外,本文参考了相关领域的研究进行网络模型的训练参数设置。本文所有网络模型均采用小批量梯度下降法,其中批次大小设定为64。该参数配置参照了Wang等^[20]的HAR实验参数设计,旨在平衡GPU内存限制与梯度更新的稳定性。图6为本文模型训练的损失曲线。在整个训练过程中,模型在约

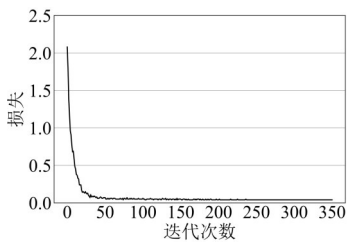


Fig. 6 Model training loss curve

图6 模型训练损失曲线

200个迭代次数后进入稳定收敛阶段,损失值波动低于1%(0.05±0.01)。选择350个迭代次数是为了确保所有网络结构充分拟合复杂特征,避免早停导致的潜在欠拟合风险。本文的初始学习率设置为0.001,并且每10个迭代次数按0.9倍指数衰减。该策略基于HAR领域经典研究中对深度HAR模型的训练方案,以确保模型收敛的稳定性。

在分类实验的训练中,分类损失由交叉熵损失和中心损失联合确定^[21]。联合分类损失公式如下:

$$\Theta^* = \arg \min_{\Theta} \mathcal{L} + \gamma \mathcal{L}_c \quad (16)$$

式中, $\mathcal{L}$ 是交叉熵损失, $\mathcal{L}_c$ 是中心损失, γ 是两个损失函数之间的平衡系数(本实验中 γ 设为0.3), Θ 是所有可训练参数的集合。虽然交叉熵损失能使模型学习到可分离的类间活动表征,但其并没有明确让模型学习最小化的类内变化表示。因此,为了弥补这一不足,中心损失被纳入分类训练中。在回归实验的训练中,均方误差(Mean Square Error, MSE)是更适用的回归损失选择,其表示模型预测值与样本真实值之间差值平方和的平均值。因此,角度估计的损失函数采用MSE。

4.3 模型评估指标

本文采用准确率(Accuracy)、类平均F分数(F_m)作为评估分类结果的指标,以及平均绝对误差(Mean Absolute Error, MAE)、均方根误差(Root Mean Square Error, RMSE)作为评估回归结果的指标。 F_m 的计算过程是先计算每个类别的 F_1 分数,然后对所有类别的 F_1 分数取平均值。其中,Accuracy是指模型正确预测的样本数占总样本数的比例; F_1 分数是指对模型精确度(Precision Rate, PR)和召回率(Recall Rate, RR)的结合度量,计算公式为 $2(PR \cdot RR)/(PR + RR)$;RMSE是指均方误差(MSE)的开平方;MAE是指预测值与真实值之间距离的平均误差幅度。

4.4 实验结果

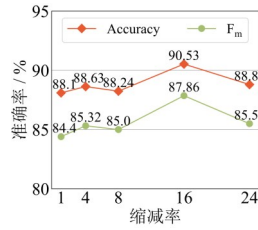
本节展示本文模型在分类和回归上的实验结果。首先,通过实验确定CA模块中超参数 r 的最佳值。其次,探究EA模块中超参数 S 的最佳值。最后,在前两个实验的基础上,考虑这两个注意力模块的排列方式,通过实验寻找最佳的排列策略。此外,将本文模型与文献[9]提出的模型进行对比,并对本文模型进行对比和消融实验。另外,还讨论了男女性别、穿戴部位对总体识别率的影响。

4.4.1 不同缩减率

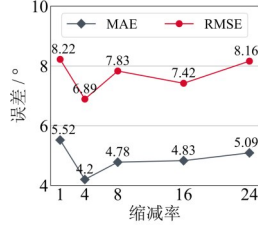
为了研究CA模块中缩减率 r 对模型性能的影响,本文首先定义 r 的可能范围,然后观察模型性能的变化。如图7所示,从实验结果可以看出,通过降低缩减率可使模型获得更好的性能。然而,当模型达到最佳性能后继续降低缩减率则会导致模型性能下降。具体来说,当 r 为16时,模型在分类任务上的效果最好。此时,Accuracy为90.53%, F_m 为87.86%。当 r 为4时,模型在回归任务上的效果最好。此时,MAE为4.2°,RMSE为6.89°。

4.4.2 不同存储单元数

接下来研究EA模块中存储单元的元素数量 S 对模型性能的影响。本文同样先定义 S 的可能范围,然后观察模型性能的变化。如图8所示,从实验结果可以看出,通过增加存储单元中的元素数量能够使模型获得更好的性能。然而,当模型达到最佳性能后继续增加存储单元中的元素数量则会导致模型性能下降。具体来说,当 S 为192时,模型在分类和回归任务上的效果最好。此时,Accuracy为93.02%, F_m 为90.83%,MAE为2.53°,RMSE为4.49°。



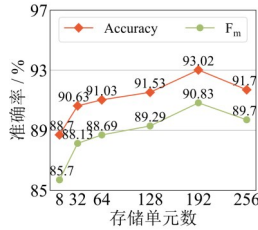
(a) Classification results
(a) 分类结果



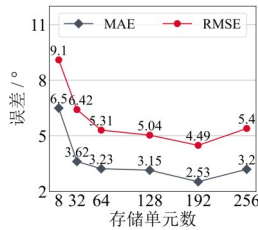
(b) Regression results
(b) 回归结果

Fig. 7 Different reduction rates

图7 不同缩减率



(a) Classification results
(a) 分类结果



(b) Regression results
(b) 回归结果

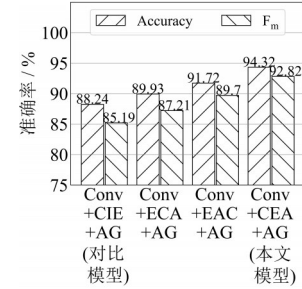
Fig. 8 Number of different memory cells

图8 不同存储单元数

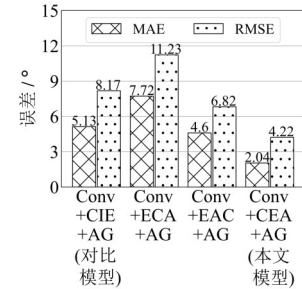
4.4.3 不同排列方式

通过前面的实验可以确定超参数 r 和 S 的最佳值。在分类任务中,CA模块中的 r 取16,EA模块中的 S 取192;在回归任务中,CA模块中的 r 取4,EA模块中的 S 取192。由于CA和EA模块的作用不同,其不同的排列方式可能会影响模型性能。因此,本文最后比较CA和EA模块的3种不同排列方式:先CA后EA(CEA)、先EA后CA(ECA)、EA和CA并行(EAC)。如图9所示,从实验结果可以看出,CEA串行排列方法优于另外两种排列方法,EAC并行排列方法优于ECA串行排列方法。实验结果表明,对CA和EA模块进行合理安排可以进一步提高模型性能。所以,本文最终

采用融合二者优势的CEA模块作为核心框架。另外,将本文提出的CEA模块与Abedin等^[9]设计的跨通道交互编码器(Cross-channel Interactive Encoder, CIE)进行比较。CIE采用自注意机制以有效地处理初始特征表示,并揭示潜在的通道交互。实验结果表明,本文提出的CEA模块优于CIE方法。



(a) Classification results
(a) 分类结果



(b) Regression results
(b) 回归结果

Fig. 9 Different arrangements

图9 不同排列方式

4.4.4 对比实验

为了评估多通道注意力机制(CEA)与其他注意力机制的性能差异,本文在跳跃运动数据集上针对Accuracy、F_m、MAE和RMSE 4个指标进行了对比实验。图10展示了多通道注意力(CEA)与其他注意力机制(SA-Net^[22]、triplet-attention^[23]和SK-Net^[24])在分类与回归任务上的对比实验结果。

在动作分类任务中,多通道注意力(CEA)的Accuracy分别比SA-Net、triplet-attention和SK-Net高4.79%、4.49%与3.29%;CEA的F_m值分别比SA-Net、triplet-attention和SK-Net高6.28%、5.93%与4.23%。在角度回归任务中,多通道注意力(CEA)的MAE分别比SA-Net、triplet-attention和SK-Net低2.06°、2.66°与1.76°;CEA的RMSE分别比SA-Net、triplet-attention和SK-Net低2.28°、2.78°与1.38°。实验结果表明,SA-Net、triplet-attention和SK-Net由于通道相关性不足,导致提取的数据特征不够完整,整体性能弱于多通道注意力(CEA)。

综上所述,本文提出的多通道注意力机制(CEA)能够有效增加特征提取的多样性,充分考虑数据特征的通道相关性,从而提高特征利用率,在分类和回归任务中均表现

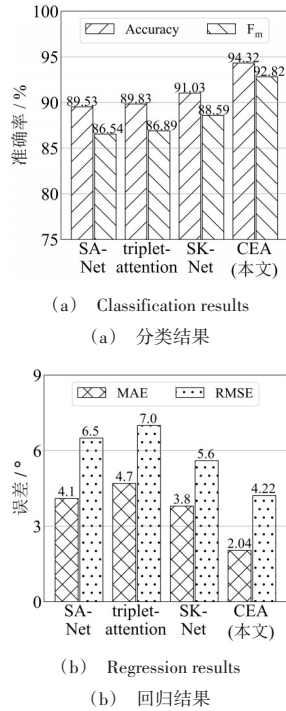


Fig. 10 Comparative experiment results

图 10 对比实验结果

出优异的性能。

4.4.5 消融实验

以下将通过消融实验来了解本文方法中各种组件的重要性。在消融实验中,通过删除框架中的特定模块来检验其对模型性能的贡献。消融实验结果如图 11 所示。首先,模型中不添加注意力机制,仅使用 Conv+GRU 进行实验。该 Conv+GRU 模型作为基线参考,实现了 81.85% 的 Accuracy 和 5.53° 的 MAE。其次,模型中加入 AG 模块,在 GRU 的输出序列上构建了时间注意力机制。该模块通过学习从时间步特征到注意力权重的非线性映射,能够自适应聚焦于序列中的关键时刻。集成 AG 模块后的 Conv+AG 模型性能得到了提高,此时 Accuracy 达到 90.03%, MAE 下降到 5.14°。再次,模型分别与 CA 和 EA 模块进行结合,以进一步改善模型性能,而且 EA 模块带来的性能提升比 CA 模块更明显。此时,Conv+CA+AG 模型的 Accuracy 达到 90.53%, MAE 下降到 4.2°; Conv+EA+AG 模型的 Accuracy 达到 93.02%, MAE 下降到 2.53°。最后,模型同时结合 CA 和 EA 模块,即 CEA 模块。本文模型 Conv+CEA+AG 实现了最优性能,此时 Accuracy 达到 94.32%, MAE 下降到 2.04°。

4.4.6 男女性别的影响

由于男性和女性在骨骼肌与生理结构上的差异,本文通过实验探讨这些差异对跳跃运动中膝关节屈曲角度检测的影响。如图 12 所示,实验结果表明,男性的 Accuracy 达到 95.02%,比女性高 0.25%;男性的 MAE 为 2.70°,比女性低 1.29°,表明男性在膝关节屈曲角度检测准确性方面可能略有优势。

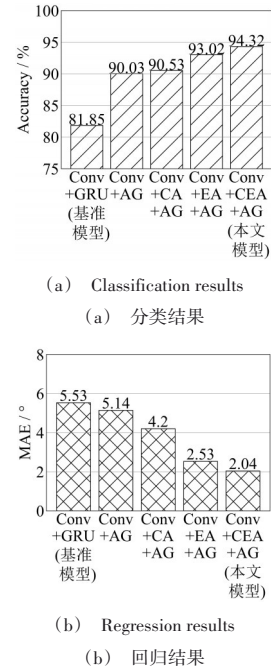


Fig. 11 Ablation experiment

图 11 消融实验

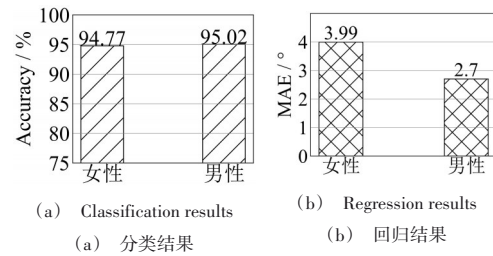


Fig. 12 Impact of gender

图 12 男女性别的影响

4.4.7 穿戴部位的影响

由于 IMU 在不同身体部位的佩戴位置可能会影响传感器捕捉到的运动数据质量,因此本文进一步分析了 IMU 在不同大腿部位佩戴对跳跃动作中膝关节屈曲角度检测准确性的影响。如图 13 所示,结果显示 IMU 佩戴于右大腿外侧时,Accuracy 达到 95.96%,相比佩戴于左大腿外侧高 1.94%, MAE 为 3.75°,相比佩戴于左大腿外侧高 0.72°,表明无论 IMU 佩戴于左大腿还是右大腿外侧,均能获得较好的膝关节屈曲角度检测性能。

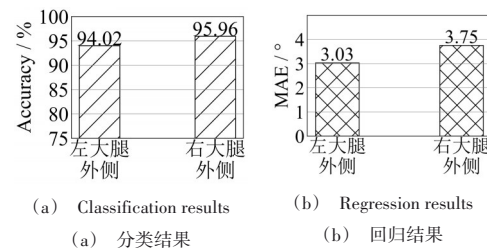


Fig. 13 Impact of wearing parts

图 13 穿戴部位的影响

5 结语

基于可穿戴设备的人体活动记录与追踪技术已成为实现运动辅助监测等许多现实应用的关键。本文旨在利用可穿戴 IMU 和有效的深度学习模型对跳跃练习的执行质量进行评估。为此,本文提出一种基于多通道注意力网络的模型,通过深入挖掘传感器通道间的潜在相关性,以进一步丰富活动的特征表示,使得对更复杂且细粒度的跳跃运动的识别成为可能。这种多通道注意力机制的引入显著提高了模型对跳跃运动中细微差别的捕捉能力,从而提升了识别的准确性。

实验结果表明,本文模型(Conv+CEA+AG)在分类任务中达到 94.32% 的准确率(Accuracy),在回归任务中实现了 2.04° 的平均绝对误差(MAE)。与传统神经网络模型(Conv+GRU)相比,Accuracy 提高了 12.47%,MAE 降低了 3.49°;与现有模型(Conv+CIE+AG)相比,Accuracy 提高了 6.08%,MAE 降低了 3.09°。该结果不仅证明了多通道注意力网络在细粒度运动识别中的有效性,而且为运动训练监测、运动表现分析及质量评估等领域的应用提供了新的技术途径和研究思路。未来的研究将进一步探索模型的优化和泛化能力,以期在更广泛的运动识别任务中获得更好的性能。

参考文献:

- [1] ZHANG Y W, LI R X, DAI J W, et al. Activity recognition algorithm based on multi-feature fusion [J]. *Software Guide*, 2022, 21(3): 72-77. 张耀威, 李瑞祥, 戴健威, 等. 基于多特征融合的活动识别算法[J]. *软件导刊*, 2022, 21(3): 72-77.
- [2] DRAZAN J F, PHILLIPS W T, SEETHAPATHI N, et al. Moving outside the lab: markerless motion capture accurately quantifies sagittal plane kinematics during the vertical jump [J]. *Journal of Biomechanics*, 2021, 125: 110547.
- [3] ARULAMPALAM J, PIERREPONT J, KARK L. Markerless motion capture: validity of Microsoft Kinect cameras and iPosoft [J]. *Gait & Posture*, 2015, 42: S76.
- [4] MASCIA G, DE LAZZARI B, CAMOMILLA V. Machine learning aided jump height estimate democratization through smartphone measures [J]. *Frontiers in Sports and Active Living*, 2023, 5: 1112739.
- [5] LU Y Z, WANG H, HU F, et al. Effective recognition of human lower limb jump locomotion phases based on multi-sensor information fusion and machine learning [J]. *Medical & Biological Engineering & Computing*, 2021, 59: 883-899.
- [6] JAITNER T, SCHMIDT M, NOLTE K, et al. Vertical jump diagnosis for multiple athletes using a wearable inertial sensor unit [J]. *Sports Technology*, 2015, 8(1): 51-57.
- [7] SUN T, LI D X, FAN B F, et al. Real-time ground reaction force and knee extension moment estimation during drop landings via modular LSTM modeling and wearable IMUs [J]. *IEEE Journal of Biomedical and Health Informatics*, 2023, 27(7): 3222-3233.
- [8] OLAWORE I. Measurement of ankle joint kinematics using IMUs during countermovement jumps and lateral skater jumps [J]. *Open Access Library Journal*, 2024, 11(7): 1-13.
- [9] ABEDIN A, EHSANPOUR M, SHI Q, et al. Attend and discriminate: beyond the state-of-the-art for human activity recognition using wearable sensors [J]. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2021, 5(1): 1-22.
- [10] WOCHATZ M, TILGNER N, MUELLER S, et al. Reliability and validity of the Kinect V2 for the assessment of lower extremity rehabilitation exercises [J]. *Gait & Posture*, 2019, 70: 330-335.
- [11] FREITAS-JUNIOR C G D, PAES P P, FORTES L D S, et al. Reliability of the high-speed camera-based system (HSC-Kinovea) for lower-limb explosive strength endurance assessment in athletes [J]. *Journal of Physical Education*, 2021, 32(1): 1-8.
- [12] STRUTZENBERGER G, KANKO R, SELBIE S, et al. Assessment of kinematic CMJ data using a deep learning algorithm-based markerless motion capture system [J]. *ISBS Proceedings Archive*, 2021, 39(1): 236.
- [13] QUAGLIARELLA L, SASANELLI N, BELGIOVINE G, et al. Evaluation of standing vertical jump by ankles acceleration measurement [J]. *The Journal of Strength & Conditioning Research*, 2010, 24(5): 1229-1236.
- [14] GRAINGER M, WEISBERG A, STERGIOU P, et al. Comparison of two methods in the estimation of vertical jump height [J]. *Journal of Human Sport and Exercise*, 2020, 15(3): 623-632.
- [15] ORDÓÑEZ F J, ROGGEN D. Deep convolutional and LSTM recurrent neural networks for multimodal wearable activity recognition [J]. *Sensors*, 2016, 16(1): 115.
- [16] HOU Q B, ZHOU D Q, FENG J S. Coordinate attention for efficient mobile network design [C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021: 13713-13722.
- [17] GUO M H, LIU Z N, MU T J, et al. Beyond self-attention: external attention using two linear layers for visual tasks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 45(5): 5436-5447.
- [18] GUO M H, CAI J X, LIU Z N, et al. PCT: point cloud transformer [J]. *Computational Visual Media*, 2021, 7: 187-199.
- [19] MURAHARI V S, PLÖTZ T. On attention models for human activity recognition [C]//*Proceedings of the 2018 ACM International Symposium on Wearable Computers*, 2018: 100-103.
- [20] WANG J, ZHENG V W, CHEN Y, et al. Deep transfer learning for cross-domain activity recognition [DB/OL]. <https://arxiv.org/pdf/1807.07963>.
- [21] WEN Y D, ZHANG K P, LI Z F, et al. A discriminative feature learning approach for deep face recognition [C]//*European Conference on Computer Vision*, 2016: 499-515.
- [22] ZHANG Q L, YANG Y B. SA-Net: shuffle attention for deep convolutional neural networks [C]//*2021 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2021: 2235-2239.
- [23] MISRA D, NALAMADA T, ARASANIPALAI A U, et al. Rotate to attend: convolutional triplet attention module [C]//*Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2021: 3139-3148.
- [24] LI X, WANG W H, HU X L, et al. Selective kernel networks [C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019: 510-519.

(责任编辑:黄健)