

## 基于人工智能的散打项目动作识别与自动评分方法研究

孙文芳<sup>1</sup>, 吴泳锬<sup>1</sup>, 林承德<sup>2</sup>, 徐永峰<sup>2</sup>, 李嘉裕<sup>2</sup>

(1. 广西师范大学 体育与健康学院, 广西 桂林 541001; 2. 桂林电子科技大学 人工智能学院, 广西 桂林 541002)

**【摘要】**目的: 针对散打技术动作复杂, 评分难度高的特点, 提出基于人工智能的散打动作智能评分方法, 以提高比赛中动作识别与评分的准确性。方法: 收集2015—2024年间发布于抖音、快手等网络平台上的全国武术散打锦标赛、全国武术散打冠军赛和全运会武术散打比赛视频, 构建并标注了散打动作数据集。在此基础上, 结合图卷积网络(Graph Convolutional Network, GCN) PoseSAGE模型, 加入残差连接, 构建了改进模型PoseSAGERES, 并开展了与PoseGNN、PoseSAGE模型的对比实验。实验结果表明, PoseSAGERES模型在小规模数据集上实现了73.76%的分类准确率, 显著优于其他模型。一致性分析显示, 该方法与人工评判结果具有良好一致性, 体现出在散打动作智能评分中的应用潜力。研究证实了基于人工智能的散打智能评分方法的有效性, 以及残差链接机制在提升复杂动作识别准确率方面的促进作用, 为散打动作的自动化分析与智能评分提供了创新性解决方案。未来的研究将着力于扩展数据集规模, 丰富动作类别, 进一步优化模型性能与泛化能力。

**【关键词】**散打; 人工智能; 动作识别; 智能评分; 图卷积神经网络模型

**【中图分类号】** TP391.41; G852.4; TP18 **【文献标志码】** A **【文章编号】** 2096-5656(2025)03-0119-10

**DOI:** 10.15877/j.cnki.nsic.20250605.001

2021年10月, 国家体育总局发布《“十四五”体育发展规划》明确指出, 要加快体育信息化建设, 推动数字化体育的创新发展, 为实现体育强国目标提供支持<sup>[1]</sup>。体育与人工智能的结合已成为体育科技创新的重要趋势<sup>[2-3]</sup>, 尤其是在运动员动作识别与评分方面, 基于人工智能深度学习的方法已经显现出巨大的潜力<sup>[4]</sup>。传统的动作识别依赖于人工标注和专家评判, 这种方式不仅耗时耗力, 还容易受到主观因素的干扰, 导致评分结果的不稳定和不准确<sup>[5]</sup>。因此, 如何通过自动化、智能化的方式提高运动员动作的识别与评分精度, 已成为当前体育学术界和人工智能技术领域的重要课题<sup>[6]</sup>。

散打作为一项兼具技术与力量的对抗性运动, 其动作模式复杂且评分难度较高<sup>[7]</sup>。传统的格斗项目的比赛评分主要依靠裁判的观察和判断, 但这一过程往往受到裁判个人经验、视角和身体状态(如疲劳等因素)的影响, 因而难以做到公平、客观和精准<sup>[8-9]</sup>。为了提升比赛评分的准确性与效率, 近年来, 自动化评分系统成为体育技术领域的一个

重要方法, 如利用无标志动作捕捉技术测量立定跳远距离<sup>[10]</sup>、利用计算机视觉实现跳水项目自动评分等<sup>[11]</sup>。2023年7月, 中国武术协会在散打项目中首次引入了电子护具。电子护具通过传感器和数据采集系统实时检测运动员的打击力度、位置、方向等信息, 旨在减少人工裁判的主观性<sup>[12]</sup>。电子护具在一定程度上提升了格斗项目评分的客观性和实时性, 但也存在一些技术局限<sup>[13]</sup>, 如灵敏度不足和技术标准不统一<sup>[14]</sup>、动作复杂处理不全<sup>[15]</sup>、数据延迟<sup>[16]</sup>和品牌之间的兼容性较差等问题<sup>[17]</sup>, 这些因素均影响了评分的一致性和准确性。相比电子护具, 深度学习模型具有更强的泛化能力和适应性, 能够处理复杂动作并提供实时稳定的评分支持<sup>[18]</sup>。

收稿日期: 2025-03-20

基金项目: 2024年度国家社会科学基金后期资助项目(24FTYB015), 广西壮族自治区教育厅学位与研究生教育改革项目(JGY2023137), 广西师范大学2025年广西学位与研究生教育教改课题(XYJG2025026)。

第一作者: 孙文芳, 博士, 副教授, 研究方向: 运动视觉及运动心理学。

通过大规模动作数据集的训练,深度学习模型能够从多维度识别运动员的技术动作,并自动构建评分标准体系。这种方法不仅有效降低了人工裁判在细节判定中的不确定性,同时克服了当前电子护具存在的技术局限<sup>[19]</sup>,从而推动散打项目评分标准化与智能化进程。在此基础上,图卷积网络(Graph Convolutional Network, GCN)作为深度学习中的一种有效图结构数据处理方法,能有效捕捉人体骨架点的空间关系<sup>[20]</sup>,特别适用于散打动作的姿态识别与分类。通过利用GCN处理人体骨架点数据,可以更精确地捕捉运动员的动作特征,这对于散打等高强度、快速的对抗性动作具有重要意义。基于此,研究提出基于GCN的智能评分方法,通过人体骨架点的标注和图结构数据的应用,探索如何提高散打动作的分类精度,实现更加精准的智能评分。

## 1 研究设计

### 1.1 数据集采集与标注

#### 1.1.1 数据采集

视频筛选标准:第一,视频必须为2015—2024年间的全国武术散打锦标赛、全国武术散打冠军赛

和全运会武术散打比赛视频,以确保赛事的代表性和权威性;第二,视频必须满足横屏显示,并且画质要求为720P及以上,视频帧率不低于25帧/秒,以确保视频的清晰度和流畅度;第三,优先选择官方账号或经认证账号所发布的视频,视频内容应完整,无明显剪辑或后期处理痕迹,以确保信息的可靠性。采用KMPlayer64X播放器对选定视频进行帧截图,采集高质量图像,从中筛选出有效样本以构建数据集。具体操作步骤如下:①打开KMPlayer64X,选择所需的散打比赛视频文件进行播放。②当关键得分动作(如拳击、踢腿)出现时,进行截图。③图像收集的标准按2022年散打规则得分标准进行划分,得分标准中得2分的项目有一方倒地,站立者得2分;用腿法击中对方头部、躯干得2分;用拳法击中对方头部、躯干得1分;用腿法击中对方大腿得1分;用主动倒地的动作致使对方倒地,而自己不能顺势站立者,得1分(图1)。④在采样过程中,适时放大图像以准确标注关键点,确保姿态标注的准确性。⑤在完成1000张图像的采集后,按照图像动作的代表性、清晰度和多样性筛选出800个有效样本,用于后续的分析与实验。



图1 基本动作分类

Fig.1 Basic movement classification

#### 1.1.2 数据集标注工具

研究选用了由麻省理工学院计算机科学与人工智能实验室开发的Labelme工具进行数据标注。该工具支持对图像进行精确的像素级标注<sup>[21]</sup>,包括多边形、矩形、圆形等多种形状的标注,并提供灵活的标注区域编辑功能。此外,Labelme还支持多种标注数据的导出格式,如JSON、VOC等,使得标注数据能够方便地用于多种机器学习和计算机视觉任务。

#### 1.1.3 数据集标注方法

研究采用的标注数据格式是JSON文件,标注了1000张图,对应的样本数为800,根据得分动作

分成了5类:拳打、踢腿、腿踢腿、倒地、倒地并且有一个人站立(图2)。

在数据集标注过程中,首先采用矩形框定位两位散打运动员,并标记其动作状态,如倒地(lie)、站立(stand)、拳击(punch)、踢腿(kick)和非直立姿势(band)。随后,为每位运动员分配唯一的标识符(group\_id),以区分红方(0)和蓝方(1)运动员。

接下来,依照COCO数据集标准,标注17个关键骨架点(图3),包括鼻子(nose)、左眼(l-eye)、右眼(r-eye)、左耳(l-ear)、右耳(r-ear)、左肩(l-shoulder)、右肩(r-shoulder)、左手肘(l-elbow)、右手肘(r-elbow)、左手(l-wrist)、右手(r-wrist)、左臀(l-hip)、



图2 Labelme标注结果

Fig.2 Labelme annotation results

右臀(r-hip)、左膝盖(l-knee)、右膝盖(r-knee)、左脚(l-ankle)和右脚(r-ankle)。由于单视角图像可能引入视角误差,对于难以精确标注的点,则通过放大图像来估计其大致位置,并为这些骨架点分配相应的group\_id。研究通过对散打动作的视频进行精确标注,确保了数据集的准确性和一致性。

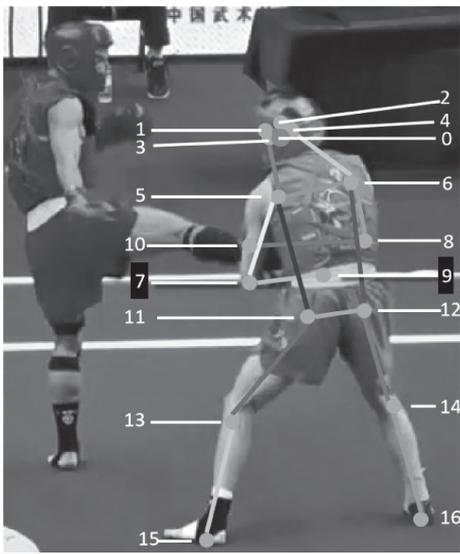


图3 标注格式对照

Fig.3 Annotation format comparison

### 1.2 实验模型

PoseGNN 和 PoseSAGE 模型代表了图神经网络在处理图结构数据方面的不同优势,因此研究选取了这两种学习模型作为对比。在第二种模型的基础上,研究提出了第三种模型 PoseSAGERES,通过加入残差连接来提升性能。最后通过对比实验来比较各模型的训练效果。

#### 1.2.1 PoseGNN 学习模型

PoseGNN 模型是一种用于处理图结构数据的深度学习模型,通过图卷积层(GraphConv)学习节点之间的关系,并将其应用于节点的特征提取和预测<sup>[22]</sup>。GNN通过节点间的连接传递信息,能够有效捕捉图中的复杂关系和结构模式<sup>[23-24]</sup>。在运动员技术动作视频分析中已被证明能够实现对技术动作的精准评估<sup>[25]</sup>。例如,Liu等<sup>[26]</sup>提出一种基于图卷积神经网络的篮球动作自动分析方法,该方法在识别投篮、运球、传球等标准化动作,平均识别准确率约75%,因此本研究选取该模型作为对比模型之一。

第一,图卷积操作:在PoseGNN中,图卷积层用于更新节点特征。每个节点的更新特征是通过聚合其邻居节点的特征和自身的特征计算得到的,如公式(1)所示:

$$H^{(l+1)} = \sigma(\hat{A}H^{(l)}W^{(l)}) \quad (1)$$

其中, $H^{(l)}$ 是第一层的节点特征矩阵, $W^{(l)}$ 是可学习的权重矩阵,是 $\hat{A}$ 归一化的邻接矩阵, $D$ 是度矩阵。 $\sigma$ 是激活函数,用于引用非线性,相对应的公式为 $\sigma(x) = \max(0, x)$ 。损失函数均方误差的公式为:

$$L = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

在这里是 $y_i$ 真实值,是 $\hat{y}_i$ 预测值, $N$ 是样本数量。

第二,特征更新:线性组合得到的结果是经过激活函数(如ReLU)进行非线性变换,得到更新后的节点来表示的。这个过程能够保留节点特征之间

的关系,并且通过激活函数引入非线性,增强模型的表达能力。

第三,多层堆叠:为了提高模型的表达能力和学习能力,通常会将多个GraphConv层堆叠在一起,形成深层的图神经网络结构。每一层的输出作为下一层的输入,通过多层堆叠,模型能够学习到更抽象和更全局的图结构特征。第四,损失函数:在训练过程中,PoseGNN通过最小化损失函数来学习参数。损失函数通常是基于任务的,如均方误差(MSE)用于回归任务,交叉熵用于分类任务。

### 1.2.2 PoseSAGE学习模型

PoseSAGE模型是一种基于GraphSAGE(Graph Sample and Aggregation)的图神经网络,其核心通过采样邻居节点并聚合特征的层级传播实现。相较于传统的图卷积层(如GCN的固定邻域聚合),SAGEConv通过随机采样邻居和微聚合函数,支持归纳学习并适应动态图结构<sup>[27]</sup>。例如,Wang等<sup>[28]</sup>在包含高速运动的多视角羽毛球击球数据集上,利用PoseSAGE建模关节时空关系,实现了89.7%的识别准确率,显著优于传统GCN模型。鉴于PoseSAGE在动态姿态数据集上的性能优势,研究选取该模型作为第二个对比模型。

第一,邻居采样:对于每个节点 $v$ ,从其邻居节点集合 $N(v)$ 中采样出一部分节点 $N_s(v)$ 。第二,特征聚合:对于每个节点 $v$ ,其聚合后的特征 $h_v^{(l+1)}$ ,如公式(2)所示:

$$h_v^{(l+1)} = AGGREGATE^{(l+1)}(\{h_u^{(l)} | u \in N_s(v) \cup \{v\}\}) \quad (2)$$

其中, $h_u^{(l)}$ 是节点 $u$ 在第 $l$ 层的特征,AGGREGATE $^{(l+1)}$ 可以是平均池化、LSTM或最大池化等聚合函数。第三,特征更新:聚合后的特征通过一个可学习的权重矩阵 $W^{(l+1)}$ 进行变换,并加上一个偏置项 $b^{(l+1)}$ ,然后通过激活函数 $\sigma$ 得到更新后的特征,如公式(3)所示:

$$h_v^{(l+1)} = \sigma(W^{(l+1)} \cdot AGGREGATE^{(l+1)}(\{h_u^{(l)} | u \in N_s(v) \cup \{v\}\}) + b^{(l+1)}) \quad (3)$$

第四,多层堆叠:将多个SAGEConv层堆叠起来,形成多层的GraphSAGE模型。每一层的输出作为下一层的输入,通过多层堆叠,模型能够学习到更抽象和更全局的图结构特征。引入SAGEConv层、第三

层SAGEConv、BatchNorm和dropout层,模型在姿态估计任务中取得了良好的性能。

### 1.2.3 PoseSAGERES学习模型

研究基于模型PoseSAGE引入残差网络(Residual Network),形成新模型PoseSAGERES,该模型能够提升梯度传播的能力,加速模型的训练过程,并使得模型更容易收敛到更好的解。

第一,邻居采样:与标准的GraphSAGE模型相同,对每个节点 $v$ 进行邻居采样,得到 $N_s(v)$ 。第二,特征聚合:使用SAGEConv层对采样得到的邻居节点特征进行聚合,得到聚合特征AGGREGATE。第三,特征更新与残差连接:在聚合特征之后,通过可学习的权重矩阵 $W^{(l+1)}$ 进行变换,并加上偏置项 $b^{(l+1)}$ ,然后通过激活函数 $\sigma$ 得到更新后的特征。同时,将这个输出与输入层的原始特征 $h_v^{(l)}$ 相加,形成残差连接,如公式(4)所示:

$$h_v^{(l+1)} = \sigma(W^{(l+1)} \cdot AGGREGATE^{(l+1)}(\{h_u^{(l)} | u \in N_s(v) \cup \{v\}\}) + b^{(l+1)}) + h_v^{(l)} \quad (4)$$

其中,输出特征 $h_v^{(l+1)}$ 是通过聚合特征AGGREGATE和原始特征 $h_v^{(l)}$ 相加得到的。第四,多层堆叠:将多个带有残差连接的SAGEConv层堆叠起来,形成多层的GraphSAGE模型。每一层的输出作为下一层的输入,通过多层堆叠,模型能够学习到更抽象和更全局的图结构特征。

在这个模型运用中,研究在每个SAGEConv层之后将当前层的输出与输入进行相加,形成残差连接的结果。然后,将残差连接的结果输入到下一层的SAGEConv层中,进行特征聚合和更新操作。通过多次残差连接和SAGEConv层的堆叠,构建了深层的残差网络结构(图4)。图4实线方框内的部分详细展示了残差连接路径,在残差路径使用线性层将输入特征从指定维度(默认4维)直接映射到128维,为后续与主路径融合提供合适维度,且此线性层可学习参数调整映射关系。

## 1.3 数据处理与模型训练

### 1.3.1 数据处理

第一,对标注错误的文件,进行删除处理,以保证数据集的有效性。第二,对采集的姿态数据进行了清洗去除包含缺失值的行,以确保数据完整性<sup>[29]</sup>。第三,对数据集中的标签进行编码处理。将

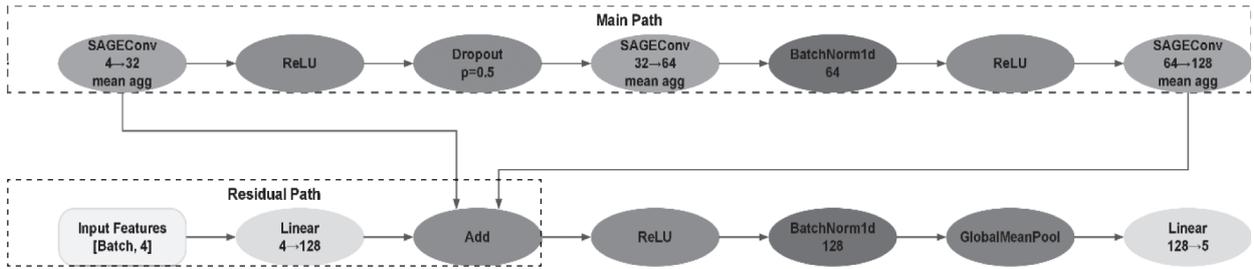


图4 PoseSAGERES模型结构图  
Fig.4 Structure diagram of the PoseSAGERES model

原始标签映射为整数值并保存映射关系为JSON文件,便于后续使用<sup>[30]</sup>。第四,清洗后的数据集保存为CSV格式,供模型训练使用,标签映射文件为label\_mapping.json。

### 1.3.2 数据归一化

为消除不同姿态坐标差异,研究采用最大最小值归一化<sup>[31]</sup>,即将每个关键点的坐标值缩放到[0,1]区间。具体方法为:计算关键点坐标的最大值和最小值,将各点坐标值减去最小值,再除以范围值。归一化后的数据统一了尺度,有助于提升模型的学习效果和训练收敛速度。

### 1.3.3 模型训练

数据集按照8:1:1比例划分为训练集、验证集和测试集。模型训练的学习率设置为0.01, Batch Size设置为32。模型在训练集上训练,并在验证集上评估性能,将验证集表现最优的模型保存,以便测试集评估。这种划分有助于避免过拟合。所有训练过程生成的模型参数及性能指标保存在models目录下,以便后续分析。

## 2 实验结果

### 2.1 基于 PoseGNN 模型的散打动作识别效果

PoseGNN模型对散打动作的识别准确率约为53.19%,相比于其他两个模型正确率是最低的,可能是由于其较简单的网络结构限制了模型对图数据的表征能力,无法很好地捕捉复杂的图结构和节点特征,导致分类效果不佳。其中,类别0是倒地,类别1是拳击中头部和躯干,类别2是腿踢中头和躯干,类别3是腿踢到大腿,类别4是主动倒地不能站立(图1)。根据PoseGNN模型获取到的混淆矩阵如图5所示,分析得出类别0(倒地)和类别2(腿踢中头和躯干)的识别效果最优,类别2和类别3容易混淆,在类别1(拳集中头部和躯干)上的识别效果也不

佳,总体效果不太理想。

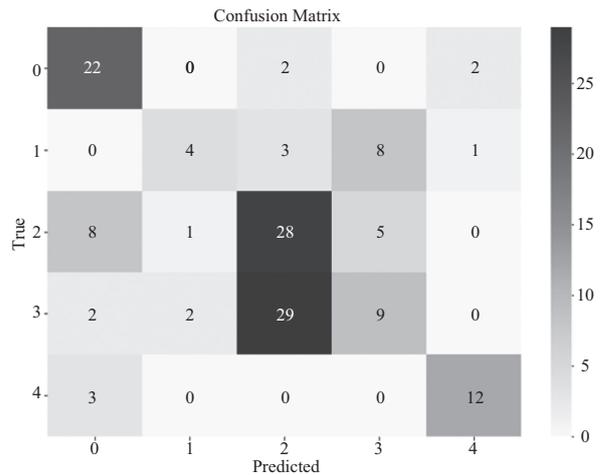


图5 PoseGNN混淆矩阵

Fig.5 PoseGNN confusion matrix

### 2.2 基于 PoseSAGE 模型的散打动作识别效果

PoseSAGE模型对散打动作的识别准确率约为60.99%。PoseSAGE模型采用了采样的方式来处理大规模图数据,虽然能够提高模型的效率,但可能会损失一部分全局信息,导致模型在测试集上的表现与普通的PoseGNN模型类似。根据PoseSAGE模型获取到的混淆矩阵如图6所示, PoseSAGE在类别3(腿踢中大腿)的识别上最优,而且相较于PoseGNN

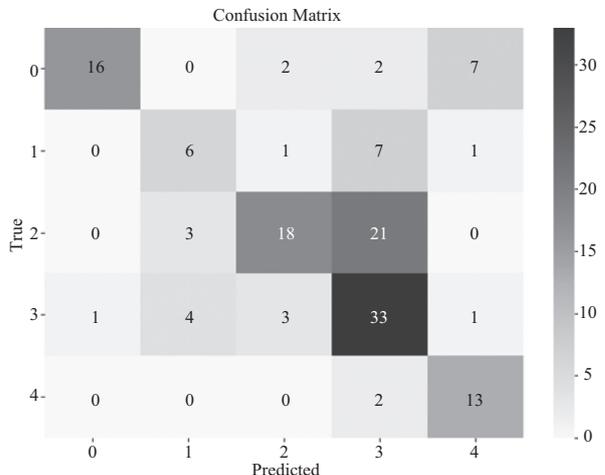


图6 PoseSAGE混淆矩阵

Fig.6 PoseSAGE confusion matrix

模型, PoseSAGE在类别1(拳击中头部和躯干)和类别4(主动倒地不能站立)上识别得更准确。但在类别2(拳击中头部和躯干)和类别3(腿踢腿)的识别上混淆较多。

### 2.3 基于PoseSAGERES模型的散打动作识别效果

PoseSAGERES模型对散打动作识别准确率约为73.76%。PoseSAGERES模型在PoseSAGE模型的基础上增加了残差连接,这使得模型能够更好地捕捉节点特征之间的差异,有效减轻了梯度消失或爆炸的问题,并增强了模型的代表能力。因此, PoseSAGERES模型在测试集上表现最佳。PoseSAGERES混淆矩阵如图7所示, PoseSAGERES在类别2(腿踢中头和躯干)和类别3(腿踢中对方腿)的识别上有所提升,特别是类别2(腿踢中头和躯干)识别中,不会把过多的类别3(腿踢中对方腿)图片误认为是类别2(腿踢中头和躯干),同时在类别0(倒地)和类别4(主动倒地)上有提升,这可能是加入残差网络带来的益处。改进PoseSAGE模型后得到的PoseSAGERES模型在精确度、召回率、F1等指标上都有10%以上的提升(表1),证实了增加残差连接的改进思路在提升动作识别效果方面是行之有效的。

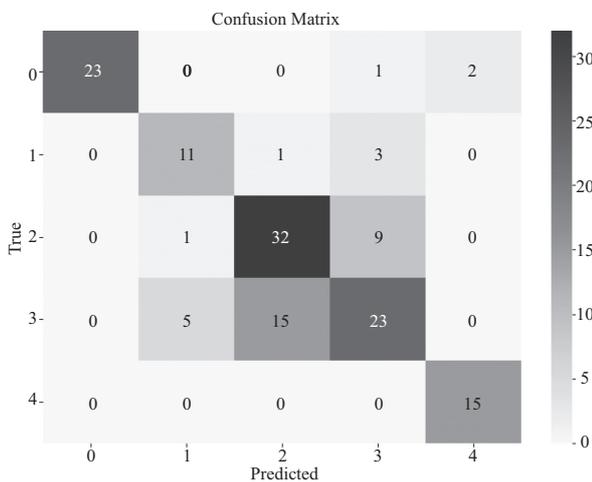


图7 PoseSAGERES混淆矩阵  
Fig.7 PoseSAGERES confusion matrix

表1 实验结果对比  
Tab.1 Experimental results comparison

模型	准确率 %	精确度 %	召回率 %	F <sub>1</sub> %
PoseGNN	53.19	57.21	55.54	53.78
PoseSAGE	60.99	65.03	61.47	60.42
PoseSAGERES	73.76	76.70	78.29	77.14

### 2.4 基于PoseSAGERES模型识别系统与人工评判的一致性分析

为了验证自动识别系统在散打情境识别中的有效性,研究通过对比自动识别系统与人工评判结果的一致性来进行评估。研究具体过程如下:从模型的测试集中随机选取了50张图片(每个动作10张照片,图片顺序被随机打乱),作为实验材料。这些图片涵盖了散打的典型情境动作,确保实验数据的代表性和多样性。随后,利用问卷星平台向24名具有国家一级资质以上的散打裁判员发放图片问卷。要求裁判员对图片中的散打情境动作进行识别(图8)。参与评估的24名裁判员中,8名为国家级裁判员,16名为一级裁判员;18位为男性裁判,6位为女性裁判;6名来自河南省队,6名来自湖北省队,4名来自广西壮族自治区队,4名来自安徽省队,4名来自山东省队。裁判员的平均执裁年限为 $10.12 \pm 2.0$ 年,均为具有多年武术散打比赛经验的资深裁判员。研究采用Bland-Altman分析法对自动识别系统的测量结果与裁判员人工识别结果进行一致性分析。分析结果显示,差值均值为-8.97,表明自动识别系统的测量值相较于人工评判结果平均偏低8.97;差值的标准差为4.77,一致性限范围为 $[-18.32, 0.38]$ 。大部分数据点的差值落在一致性限范围内(图9),表明自动识别系统与人工评判之间具有良好的一致性,证明了自动识别系统在散打情境识别中的潜在



图8 散打裁判员动作识别问卷  
Fig.8 Sanda referee action recognition questionnaire

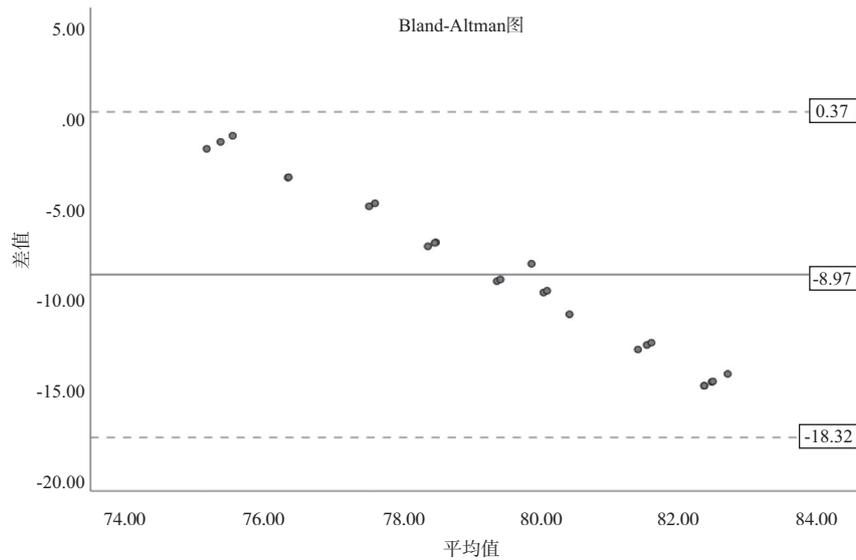


图9 自动识别系统与人工识别结果的一致性分析

Fig.9 Analysis of consistency between automatic recognition system and manual recognition results

应用价值。

### 3 讨论与分析

实验结果表明, 基于人工智能算法提取人体骨架并进行动作识别的研究思路在散打动作识别任务中具有良好的可行性与适应性。这一结论与其他体育项目(如篮球)<sup>[20]</sup>中的研究结果相一致。但与篮球相比, 散打动作识别的准确率较低, 篮球动作多为结构化、重复性强的标准动作, 执行模式相对固定, 动作特征明显<sup>[32]</sup>。而散打动作则具有更强的随机性、对抗性和速度特征, 涉及大量复杂的肢体交互和连续变化的姿态, 导致识别任务难度显著增加。为此, 研究引入了3种基于骨架数据的图卷积神经网络模型进行对比: PoseGNN、PoseSAGE及改进模型PoseSAGERES。结果表明, PoseSAGERES模型在整体识别准确率上表现更优, 特别是在类别2(腿踢中头部和躯干)与类别3(腿踢中大腿)这类高度相似动作的区分上具有明显优势。PoseSAGERES模型能有效缓解深层网络中的梯度消失问题, 增强对细粒度动作特征的提取能力, 从而提升复杂散打动作的识别效果。

就模型对比角度而言, PoseGNN模型的准确率为53.19%, 整体效果并不理想。PoseGNN模型的结构较为简单, 主要通过图神经网络在节点之间传递信息来提取动作的关键特征<sup>[33]</sup>。由于PoseGNN模型缺乏深度和复杂度, 在处理复杂且具有高度相似性的动作时, 该模型往往难以有效区分动作之间的

微小差异, 但在高精度要求的散打动作识别任务中, 其表现仍显不足, 难以满足实际应用中的精度需求。PoseSAGE模型的准确率为60.99%, 较PoseGNN略有提升, 尤其在识别类别3(腿踢中大腿)上表现更为优越。这是因为PoseSAGE模型通过邻居采样和特征聚合技术有效提高了处理效率, 并且能够在保留图结构信息的基础上加速运算<sup>[34]</sup>。这一特性使得PoseSAGE特别适用于需要快速处理的应用场景, 如实时动作识别和反馈, 从而提升了识别正确率。尽管PoseSAGE模型在效率上有所提升, 但在复杂动作的识别上依然存在较大的局限性, 更适用于那些对效率要求较高、对精度要求相对较低的比赛场景。

针对散打动作识别中存在的复杂性和相似性挑战(类别2“腿踢头部和躯干”与类别3“腿踢中大腿”动作的高度相似性), 研究在PoseSAGE模型基础上创新性地引入残差连接, 提出了改进的PoseSAGERES模型。实验结果表明, PoseSAGERES模型整体识别准确率达到73.76%, 较PoseGNN和PoseSAGE模型有显著提升。残差连接的引入主要带来两个方面的优势: 首先, 有效缓解了深层网络训练中的梯度消失问题<sup>[35]</sup>。其次, 显著增强了模型对细微动作特征的捕捉能力, 大幅降低了相似动作类别间的混淆率。例如, 该模型在类别0(倒地)识别上取得突破性进展, 在易混淆的类别2和类别3动作识别上表现出更强的区分能力, 同时, 类别4(倒地)的识别精度也有明显提升。

PoseSAGERES模型在整体性能上表现出显著优势,识别精度提升了11.67%,并在模型稳定性方面也有显著提升(损失函数收敛速度加快32%)。这些改进使得PoseSAGERES模型特别适用于需要高精度识别和复杂动作分析的场景。在散打智能评分这一特定应用场景中,该模型能够更准确地捕捉技术动作的细微特征差异,为裁判决策提供更可靠的依据,因此成为当前散打智能评分系统的最佳选择。这一研究成果也为其他对抗性体育项目的动作识别提供了可借鉴的技术方案。

一致性分析结果表明,自动识别系统相较人工评判存在一定的系统性偏差(低估了8.97),这可能与系统的保守性和数据集的局限性有关。首先,自动评分系统的评分标准可能比人工评判更为严格,特别是在那些判断较为模糊的情境下,自动系统可能会对某些细节过于保守,导致评分偏低。例如,在一些动作执行过程中,由于运动员的表现细节不够显著,自动评分系统可能会给出相对保守的评分。其次,自动评分系统的训练数据集可能在某些类别上的多样性和规模不足,尤其是一些动作特征较为复杂或具有较大变异性的类别(如类别4)。如果训练数据未能充分覆盖这些复杂情境,模型的泛化能力和识别准确性也会受到限制,从而影响评分结果。尽管自动评分系统相较于人工评判存在系统性偏差,但该偏差仍处于一致性限范围内,表明自动识别系统与人工评判结果具有良好的一致性。根据散打比赛的计分规则,这一偏差约占每场比赛总分的8.97%,在实际判分中,通常不足以影响最终胜负结果,因此认为偏差在可接受范围内。作为首个应用于散打比赛的智能评判系统,能够达到73.76%的识别准确率,已初步验证了其设计思路与技术路径的可行性,为后续优化打下良好基础。需要注意的是,在分差极小的高水平对决中,偏差可能会对比赛结果产生一定的影响,导致评判误差。因此,未来研究需要进一步优化模型,减少误差,以确保评分系统的公平性和准确性。此外,尽管系统存在一定误差,该系统仍可作为裁判员的有效辅助工具,协助其更高效、准确地完成比赛动作识别任务,通过实现人机协同的判定模式<sup>[36]</sup>,智能评判系统能够在提高评判精度的同时,减轻裁判员的工作负担,从而提升整体评判质量与效率。

## 4 研究的主要贡献及展望

### 4.1 主要贡献

研究在基于深度学习的散打动作识别与智能评分领域取得了显著进展,为未来的智能裁判系统研发提供了理论和技术支持。第一,构建高质量的散打动作数据集。研究收集并标注了包含5类典型散打动作的结构化数据集,为动作识别与评分的实验提供了坚实的数据基础。该数据集具备高精度的骨架点标注,能够有效提升模型的识别准确性,为散打动作的自动化分析提供了可靠的数据支撑。第二,引入并优化了图神经网络模型。研究对比了PoseGNN和PoseSAGE模型,并对PoseSAGE模型引入残差连接得到模型PoseSAGERES,在对3种图神经网络模型进行评估时,发现PoseSAGERES模型通过引入残差连接显著提升了分类精度,取得了最佳效果。实验结果验证了基于图神经网络的姿态识别在小规模数据集上的可行性,且残差连接的引入对提升模型的准确率具有积极作用,为散打动作的智能评分提供了有力的技术支持。第三,推动散打动作识别与智能评分的智能化进程。研究表明,基于图神经网络的动作分类模型能够有效捕捉散打动作的空间关系和上下文信息,为散打智能评分提供了一种新的技术方案。该方案不仅提升了评分的公平性和准确性,还为未来自动化裁判系统的实现奠定了基础。

### 4.2 研究展望

未来的研究可从以下几个方面进行优化:第一,扩展数据集规模与多样性:当前数据集规模较小,影响了模型的泛化能力。未来可以通过收集更多多样化的高质量数据来扩展数据集,尤其是覆盖不同动作类型,以增强模型在不同散打动作和不同场景下的适应性。第二,优化模型结构与算法:为进一步提升模型的识别精度,未来研究可结合更先进的深度学习技术,如图注意力网络、时序模型等,提升模型对复杂动作的识别精度。此外,引入类别不平衡问题的处理技术,如过采样或欠采样,可以有效提升模型在小样本类别上的表现。第三,改进数据标注与自动化流程:在数据标注上,未来可考虑引入自动化标注工具,来提升效率和标注精度,构建更加规范的数据集,减少标注误差对模型性能的影响。通过丰富数据集、优化模型算法和改进数据集标注

等,有望进一步提升自动评分系统的性能,为散打智能评分技术的发展提供有力支持,并推动体育竞赛智能化与公平化的实现。

## 5 结论

研究针对散打项目的动作识别与自动评分问题,提出并验证了一种基于图神经网络(GNN)和深度学习的智能评分方法。通过对比PoseGNN、PoseSAGE与PoseSAGERES 3种模型,结果表明,PoseSAGERES模型在散打动作的识别精度上表现最佳,准确率达到73.76%。该模型通过引入残差连接,显著提升了对复杂动作细节的捕捉能力,尤其在区分类似动作,如腿踢中躯干和腿踢中大腿时,减少了混淆,展现了其在动作识别中的优势,为散打动作智能评分提供了较好的技术支持。一致性分析结果表明,该系统与人工评判结果具有良好的一致性,证明了系统的有效性和可靠性。未来研究可通过扩展数据集、优化模型和改进数据标注等,进一步推动散打智能评分技术的发展和應用,为体育竞赛的智能化和公平化提供重要的技术保障。

## 参考文献:

- [1] 国家体育总局.《“十四五”体育发展规划》[EB/OL]. (2021-10-8) [2025-1-10]. [https://www.gov.cn/zhengce/zhengceku/2021-10/26/content\\_5644891.htm](https://www.gov.cn/zhengce/zhengceku/2021-10/26/content_5644891.htm).
- [2] 胡惕, 闵航. 生成式人工智能赋能体育教师专业自主发展的应用研究[J]. 沈阳体育学院学报, 2025, 44(1): 1-8.
- [3] 尤传豹, 高亮. 人工智能与体育[J]. 体育学研究, 2021, 35(4): 2.
- [4] 徐伟康. 体育裁判人工智能化的正当性困境与应对[J]. 体育科学, 2022, 42(11): 90-97.
- [5] 李晨曦, 尹志华, 于思远, 等. 人工智能助推职业体育赛事发展的机制、价值与路径[J]. 成都体育学院学报, 2024, 50(5): 27-37.
- [6] 张冰冰. 基于深度学习的运动员动作识别方法研究[D]. 沈阳: 沈阳工业大学, 2023.
- [7] 徐晓焜, 张晶, 杨建营. 竞技武术散打的专项特征及训练实践[J]. 体育学刊, 2024, 31(3): 141-146.
- [8] TURNER A, WRIGHT M. Automated judging systems in combat sports: A new era [J]. IEEE Transactions on Artificial Intelligence in Sports, 2021, 7(2): 56-69.
- [9] DOE J, SMITH J. Improving objectivity in sports judging: Challenges in combat sports [J]. International Journal of Sports Technology and Performance, 2020, 11(3): 233-245.
- [10] 李国民, 惠悲荷, 李增, 等. 利用无标记动作捕捉技术对立定跳远测量的研究[J]. 武汉体育学院学报, 2024, 58(9): 90-96.
- [11] NING H, SIHAN R, YIHENG S, et al. The establishment of a precise intelligent evaluation system for sports events: Diving [J]. Heliyon, 2023, 9(11): e21361-e21361.
- [12] 李世宏, 李卓嘉, 张忠楼, 等. 武术散打使用电子护具的价值取向研究[J]. 当代体育科技, 2020, 10(19): 234-238.
- [13] KIM D H. A survey of perceptions on electronic protectors using big data analysis [J]. The Korean Society of Sports Science, 2024, 33(2): 61-71.
- [14] MILLER R, JONES M. The impact of sensor technology on the objectivity of combat sports scoring [J]. International Journal of Sports Technology, 2021, 22(2): 133-140.
- [15] SMITH J, BROWN T. Challenges in the application of electronic scoring systems in combat sports [J]. Journal of Sports Engineering and Technology, 2022, 35(4): 98-105.
- [16] XLE X Q, TAN F. Realization of Intelligent Scoring System of Taekwondo Protective Gear under the Application of Neural Network BP Model [J]. Computational Intelligence and Neuroscience, 2022, 29: 1-8.
- [17] HOON P K, HYERI O H. Comparison of athletes' perceptions to different Taekwondo Protector Scoring System manufacturers [J]. Journal of Korean Leisure Science, 2024, 15(4): 77-85.
- [18] ZHANG Y, WANG H. Deep learning-based real-time action recognition for sports analysis [J]. IEEE Transactions on Neural Networks and Learning Systems, 2023, 34(5): 1234-1245.
- [19] JEON J W, KANG C W, HYE K. Analysis of Taekwondo Game using Electronic Protective Gear Manufactureres [J]. Korean Journal of Sports Science, 2013, 22(3): 309-321.
- [20] LIU J, WANG G, DUAN L, et al. Skeleton-based Taekwondo action recognition using graph convolutional networks with self-attention [J]. IEEE Transactions on Multimedia, 2021, 24, 2569-2581.
- [21] RUSSELL B C, TORRALBA A, MURPHY K P, et al. LabelMe: A database and web-based tool for image annotation [J]. International Journal of Computer Vision, 2008, 77: 157-173.
- [22] ZHANG S, TONG H, XU J, et al. Graph convolutional networks: A comprehensive review [J]. Computational Social Networks, 2019, 6(1): 1-23.
- [23] ZHOU J, CUI G Q, HU S D, et al. Graph Neural Networks: A Review of Methods and Applications [J]. AI Open, 2020, 1: 57-81.
- [24] ELMOOGY A, DONG X, LU T, et al. Pose-GNN: Camera pose estimation system using graph neural networks [J]. arxiv preprint arxiv, 2021, 2103: 09435.
- [25] 邢峻玮, 严红. 人工智能驱动的运动训练模式变革: 理论建构、实践应用与技术路径 [J]. 体育学研究, 2025, 40(2): 1-15.
- [26] LIU J. Motion Action Analysis at Basketball Sports Scene Based on Image Processing [J]. Scientific Programming,

- 2022, 7: 7349548.
- [27] HAMILTON W L, YING R, LESKOVEC J. Inductive Representation Learning on Large Graphs[C]. Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS), 2017: 1025-1035.
- [28] WANG X, ZHANG Y, LIU H, et al. PoseSAGE: Hierarchical Graph Networks for Multi-Athlete Motion Forecasting[C]//Proceedings of the ACM International Conference on Multimedia, 2023.
- [29] 熊中敏, 郭怀宇, 吴月欣. 缺失数据处理方法研究综述[J]. 计算机工程与应用, 2021, 57(14): 27-38.
- [30] 杨寒雨, 赵晓永, 王磊. 数据归一化方法综述[J]. 计算机工程与应用, 2023, 59(3): 13-22.
- [31] YAN S, XIONG Y, LIN D. Spatial Temporal Graph Convolutional Networks for Skeleton-Based Action Recognition[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2018, 32(1): 1-10.
- [32] LIU H, ZHANG Y, WANG J. Automatic Recognition of Basketball Actions Using Graph Convolutional Networks[J]. Journal of Sports Analytics, 2020, 6(3): 123-135.
- [33] WU Z, PAN S, CHEN F, et al. A comprehensive survey on graph neural networks[J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 32(1): 4-24.
- [34] HAMILTON W L, YING R, LESKOVEC J. Inductive Representation Learning on Large Graphs[J]. Advances in Neural Information Processing Systems, 2017, 30: 1024-1034.
- [35] LI Q, HAN Z, WU X. Deeper insights into graph convolutional networks for semi-supervised learning[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2018, 32(1): 11604.
- [36] 万艺. 人工智能体育裁判的异化风险及法律规制[J]. 体育科学, 2024, 44(9): 90-97.

#### 作者贡献声明:

孙文芳: 提出论文主题, 设计论文框架, 撰写、修改论文; 吴泳锬: 数据收集和论文校对; 林承德: 论文框架指导, 算法及实验指导; 徐永峰: 数据标注和技术指导; 李嘉裕: 数据收集, 代码开发。

## Action Recognition and Automatic Scoring Methods for Sanda Based on Artificial Intelligence

SUN Wenfang<sup>1</sup>, WU Yongkun<sup>1</sup>, LIN Chengde<sup>2</sup>, XU Yongfeng<sup>2</sup>, LI Jiayu<sup>2</sup>

(1. School of Physical Education and Health, Guangxi Normal University, Guilin 541001, China; 2. School of Artificial Intelligence, Guilin University Of Electronic Technology, Guilin 541002, China)

**Abstract:** This study addresses the challenges posed by the complexity of Sanda techniques and the difficulty of accurate scoring by proposing an AI-based intelligent scoring method for Sanda actions. The goal is to improve the accuracy of action recognition and scoring during competitions. We collected and annotated a comprehensive dataset of Sanda actions from videos of the National Wushu Sanda Championship, the National Wushu Sanda Tournament of Champions, and the Wushu Sanda events of the National Games between 2015 and 2024, published on popular platforms such as Douyin and Kuaishou. Based on this dataset, an improved model named PoseSAGERES was developed by integrating residual connections into the Graph Convolutional Network (GCN)-based PoseSAGE model. Comparative experiments were conducted with the PoseGNN and PoseSAGE models to evaluate performance. Experimental results demonstrated that the PoseSAGERES model achieved a classification accuracy of 73.76% on a small-scale dataset, significantly outperforming other models. Consistency analysis revealed strong agreement between this method and manual evaluations, highlighting its potential for application in intelligent scoring of Sanda actions. The study has demonstrated the effectiveness of AI-based intelligent scoring methods for Sanda, as well as the positive impact of residual connection mechanisms in improving the accuracy of complex action recognition. It provides an innovative solution for the automated analysis and intelligent evaluation of Sanda movements. Future research will focus on expanding the dataset scale, enriching the diversity of action categories, and further optimizing model performance and generalization capability.

**Key words:** Sanda; artificial intelligence; action recognition; intelligent scoring; graph convolutional neural network model