

考虑多尺度时空因果特征的排球群体行为识别<sup>\*</sup>

万 宁 张美珠

(中国药科大学体育部 南京 210000)

**摘 要:** 考虑到当前排球群体行为识别方法中交互关系的有向性、延时性以及受时空约束等特点,本文提出一种考虑多尺度时空因果特征的排球群体行为识别方法。首先,基于因果检测模型,对行为交互的因果关系进行分析判断;其次,基于时序因果图推理模型,提取群体行为中的时序因果特征,降低延时性对特征融合过程中产生的误差;然后,将距离和外观约束引入到时序因果特征中,提取出群体行为多尺度时空因果特征。最后,通过公开数据集和自选数据集验证了本文所提算法的有效性,并通过融合对比实验验证了识别框架中模块的重要性。实验表明:本文所提模型能够充分发挥多尺度时空因果图在描述群体交互影响的优势,提取的多尺度时空因果特征图能够有效学习群体行为特征,提升排球群体行为识别性能。

**关键词:** 排球;群体行为识别;因果关系;时空约束;时序推理

**中图分类号:** TN92;TP183 **文献标识码:** A **国家标准学科分类代码:** 510.1040

## Volleyball group behavior recognition method considering causality

Wan Ning Zhang Meizhu

(Department of Physical Education, China Pharmaceutical University, Nanjing 210000, China)

**Abstract:** Considering the characteristics of the interaction in the current volleyball group behavior recognition methods, such as directed, delay and subject to spatio-temporal constraints, this paper proposes a volleyball group behavior recognition method considering multi-scale spatio-temporal causal features. Firstly, the causal relationship of behavior interaction was analyzed and judged based on the causal detection model. Secondly, based on the temporal causality diagram reasoning model, the temporal causal features in crowd behavior were extracted to reduce the error caused by the delay pair feature fusion process. Then, the distance and appearance constraints are introduced into the temporal causal features to extract the multi-scale spatio-temporal causal features of crowd behavior. Finally, the effectiveness of the proposed algorithm is verified by public datasets and self-selected datasets, and the importance of the modules in the recognition framework is verified by fusion and comparison experiments. The experimental results show that the proposed model can give full play to the advantages of multi-scale spatio-temporal causal maps in describing group interaction, and the extracted multi-scale spatio-temporal causal feature maps can effectively learn the characteristics of crowd behavior. It can improve the performance of volleyball group behavior recognition.

**Keywords:** volleyball; crowd behavior recognition; causality; spatio-temporal constraints; temporal reasoning

## 0 引 言

随着信息技术与智慧体育领域的快速发展,面向复杂比赛场景下的群体行为识别作为计算机视觉与体育的交叉领域,能够帮助裁判人员准确掌握比赛场内运行员实时动态,更能够为后续比赛群体行为的分析与预测、运行员异常行为识别预警等功能提供基础技术<sup>[1-5]</sup>。相较于个体行为分析与识别,群体行为具有交互性强、上下文特征明显等特

征,需要重点考虑行为的交互因果关系,针对性设计群体行为识别模型<sup>[6]</sup>。面向排球比赛的群体行为识别主要是利用人工智能等技术<sup>[7-8]</sup>对视频中排球运行员的群体行为进行分析、理解,以此实现复杂比赛场景下对排球运动员群体活动行为进行推断,便于识别意外性行为,维护比赛正常秩序。

当前针对群体行为识别领域国内外已有一定进展,主要是利用深度学习等算法对群体行为特征进行提取,重点

考虑群体间的交互关系,构建特征与群体行为之间的非线性映射关系。Lu 等<sup>[9]</sup>提出一种嵌入图注意力块的图注意力交互模型(graph attention interaction model, GAIM),通过自适应机制,考虑个体和群体层面的不平衡交互关系,学习交互时空变化特征,以此实现对群体行为的预测。Hong 等<sup>[10]</sup>提出了一种基于隐标签挖掘技术的视频群体行为识别算法,该模型通过融合个体行为特征,实现识别准确率比常用算法提高 15.56%。Gao 等<sup>[11]</sup>提出了一种结合卷积神经网络和门控循环单元(convolutional neural networks-gated recurrent unit, CNN-GRU)的混合模型对视频中群体行为进行识别,测试准确率达到 94.8%。

上述方法在群体行为识别领域有一定适用性,但在排球比赛群体行为识别领域,仍有一些特点需要考虑:1)群体行为交互关系具有有向特点,需要区分施加影响行为和受影响行为;2)群体行为交互关系具有延时性,交互双方由于时空位置不同、行为因果不同,引起行为特征时间不同步;3)群体行为受时空约束。不同的群体交互行为的距离往往也不同。

为解决上述问题,本文提出考虑多尺度时空因果特征的排球群体行为识别方法,首先,基于 Grange 因果检测模

型,对群体行为有向交互关系进行分析判断;然后,基于时序因果图推理模型,考虑交互关系延时性特点,保证特征时间对齐,降低特征在融合过程中出现的误差;最后,将距离和外观约束补充到模型中,考虑行为受到的时空约束,提取多尺度时空因果特征,对排球群体行为进行分析识别。

## 1 考虑多尺度时空因果特征的群体行为识别框架

排球群体行为识别主要是对群体交互行为建模和表征好,本文将排球群体行为识别过程分为 3 个任务,分别是因果检测模型、时序因果图推理模型以及多尺度时空因果图推理模型,识别框架如图 1 所示,整个识别过程步骤如下:

步骤 1)检测运动员个体之间的交互影响关系。本文通过时序自回归模型和时序相关回归模型进行回归估计,并基于 F 分布分析判断个体之间是否存在交互影响。

步骤 2)基于时间对齐处理的时序特征融合。通过考虑基于 Granger 的时间延迟情况,保证时间对齐处理,降低个体特征匹配过程中出现的融合误差。

步骤 3)将距离和外观约束引入到时序因果特征图,从而提取出多尺度时空因果特征,实现对群体行为识别。

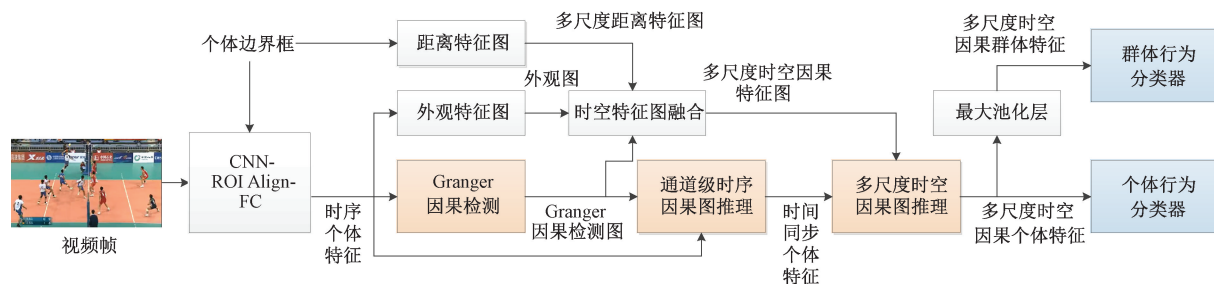


图 1 考虑多尺度时空因果特征的群体行为识别框架

Fig. 1 Group behavior recognition framework considering multi-scale spatio-temporal causal features

### 1.1 因果检测模型

本文采用 Granger 因果条件,对复杂场景中个体行为之间的交互影响进行判断。若个体  $i$  对个体  $j$  产生因果影响,则个体  $i$  与个体  $j$  互为原因和结果。考虑到个体行为的局部动作具有动态时空变化特征,本文基于局部多通道特征机制,表征局部状态,并考虑时序模型的一致性约束,构建基于 Granger 的因果检测模型。该模型包含个体行为受历史行为状态约束以及个体行为受环境影响约束两个子模块<sup>[12]</sup>。

为了对个体行为受历史行为状态约束情况进行描述,本文构建一种时序自回归模型,通过收集个体历史行为状态信息,以此实现对最后行为状态进行预测。为了对个体行为受环境影响约束情况进行描述,本文构建时序相关回归模型,除了自身历史状态,引入环境状态因素,综合预测最后状态信息。若个体行为只受自己历史行为状态影响,则时序自回归模型预测效果明显,时序相关回归模型预测效果一般。反之,若个体行为受环境影响约束,则时序相

关回归模型预测效果较好,时序自回归模型预测效果一般。

为更精准描述个体  $i$  与个体  $j$  的交互影响关系,引入两条假设:1)原假设:个体  $i$  与个体  $j$  存在影响关系;2)对立假设:个体  $i$  与个体  $j$  不存在影响关系。本文设计出基于 Granger 的因果检测模型,以此分析验证个体之间的假设。

如图 1 所示,针对复杂排球比赛中收集的视频帧信息,首先使用 2D-CNN 检测模型对视频帧特征进行学习,然后基于区域兴趣对齐(region of interest align, ROI Align)对个体行为标记框中的特征进行提取,获得个体特征向量,最后基于全连接层(fully connected, FC)对提取出的个体特征向量进行变换,增加模型表征能力。

通过上述操作,能够获得针对视频帧的个体特征信息  $X = \{x_{i,t,c}\}$ ,  $i$ ,  $t$  和  $c$  分别表示个体序号、时间序列和通道序列。如图 2 所示,基于因果检测模型从  $X = \{x_{i,t,c}\}$  中提取 2 个个体特征信息,分别构建时序自回归模型和相关

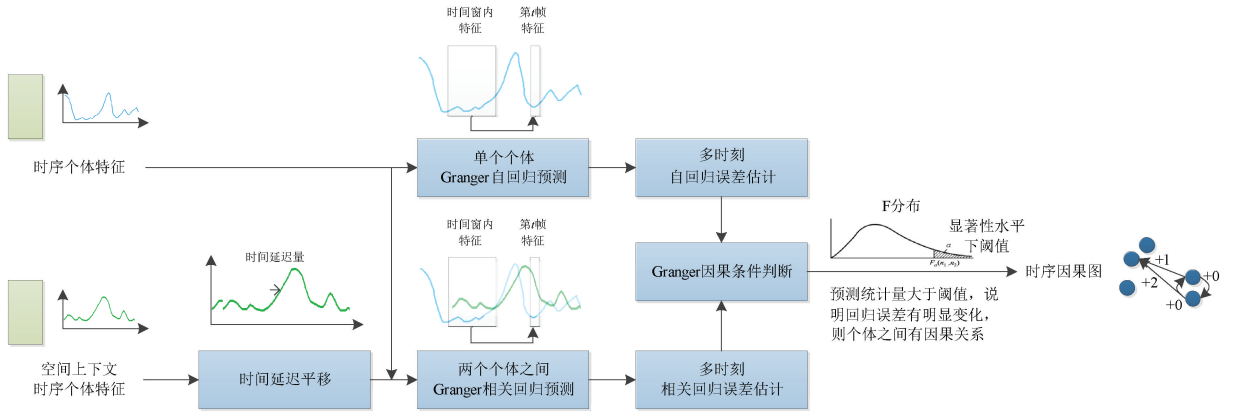


图 2 因果检测模型中的时序自回归模型和相关回归模型

Fig. 2 Time series autoregressive model and correlation regression model in causal detection model

回归模型。

### 1) 时序自回归模型 $\phi_a(\cdot)$

该模型主要是利用自身历史状态信息对个体最新特征进行预测,计算过程为:

$$x_{i,k,c} = \phi_a([x_{i,k-m;k-1,c}], \theta_{m,i}^a) = \sum_{q=1}^m \omega_{m,i,k-q}^a x_{i,k-q,c} + b_{m,i}^a \quad (1)$$

式中:  $i$  为运动员个体序号,  $k$  为当前时刻,  $m$  为时间窗口规模,  $r$  为时间窗口下标, 范围为  $q = [1:m]$ , 个体历史状态信息的时间窗口为  $[k-m:k-1]$ ,  $\theta_{m,i}^a = \{\omega_{m,i,k-q}^a, b_{m,i}^a\}$  为自回归模型中的参数, 表示模型中一次项和常数项。

由于视频中运行员个体行为的目标固定, 需要考虑时序模型的一致性约束, 保证不同通道的参数相同。因此在时序自回归模型  $\phi_a(\cdot)$  中, 将全部通道下不同时刻  $k$  的样本数据都作为训练样本, 模型损失函数采用误差平方和, 优化函数使用最小二乘法, 以此获取最优的参数, 即:

$$\hat{\theta}_{m,i}^a = \underset{\theta_{m,i}^a}{\operatorname{argmin}} \sum_{k,c} \|x_{i,k,c} - \phi_a([x_{i,k-m;k-1,c}], \theta_{m,i}^a)\|^2 \quad (2)$$

### 2) 时序相关回归模型 $\phi_r(\cdot)$

该模型主要是利用自身历史状态以及环境影响因素对个体最新特征进行预测, 模型基于时间延迟效应, 对运动员个体行为之间的影响进行描述, 计算过程如下:

$$x_{i,k,c} = \phi_r([x_{i,k-m;k-1,c}, x_{j,k-delay-m;k-delay-1,c}], \theta_{delay,m,j \rightarrow i}^{r,j \rightarrow i}) = \sum_{q=1}^m \omega_{delay,m,i,k-q}^{r,j \rightarrow i} x_{i,k-q,c} + \sum_{q=1}^m \omega_{delay,m,j,k-delay-q}^{r,j \rightarrow i} x_{j,k-delay-q,c} + b_{m,delay}^{r,j \rightarrow i} \quad (3)$$

式中:  $i$  为运动员个体序号,  $k$  为当前时刻,  $j$  为比赛环境中对个体  $i$  有影响的运动员个体序号,  $m$  为时间窗口规模,  $q$  为时间窗口下标, 被影响的运动员个体时间窗口为  $[k-m:k-1]$ ,  $delay$  为主动影响的运动员个体的时间延迟, 主动影响的运动员个体时间窗口为  $[k-delay-m:k-delay-1]$ ,  $\theta_{delay,m}^{r,j \rightarrow i} = \{\omega_{delay,m,i,k-q}^{r,j \rightarrow i}, \omega_{delay,m,j,k-delay-q}^{r,j \rightarrow i}, b_{m,delay}^{r,j \rightarrow i}\}$  为相关回归模型中的参数, 表述模型中一次项、环境中个体

一次项以及常数项。

在时序相关回归模型训练过程中, 同样需要考虑时序模型的一致性约束, 保证不同通道的参数相同。因此在时序相关回归模型  $\phi_r(\cdot)$  中, 将全部通道下不同时刻  $k$  的样本数据都作为训练样本, 模型损失函数采用误差平方和, 优化函数使用最小二乘法, 以此获取最优的参数, 即:

$$\hat{\theta}_{delay,m}^{r,j \rightarrow i} = \underset{\theta_{delay,m}^{r,j \rightarrow i}}{\operatorname{argmin}} \sum_{k,c} \|x_{i,k,c} - \phi_r([x_{i,k-m;k-1,c}, x_{j,k-delay-m;k-delay-1,c}], \theta_{delay,m}^{r,j \rightarrow i})\|^2 \quad (4)$$

在通过时序自回归模型  $\phi_a(\cdot)$  和时序相关回归模型  $\phi_r(\cdot)$  处理后, 本文采用基于 Granger 的 F 分布对假设进行验证, 以此判断个体  $i$  与个体  $j$  之间是否存在影响交互关系。F 统计量表示为:

$$f_{delay,m}^{r,j \rightarrow i} = \frac{(ssr_{m,i}^a - ssr_{delay,m}^{r,j \rightarrow i})/m}{ssr_{m,i}^{r,j \rightarrow i} / (n_m^r - v_m^r)} \quad (5)$$

$$\begin{cases} ssr_{m,i}^a = \sum_{k,c} \|x_{i,k,c} - \phi_a([x_{i,k-m;k-1,c}], \hat{\theta}_{m,i}^a)\|^2 \\ ssr_{delay,m}^{r,j \rightarrow i} = \sum_{k,c} \|x_{i,k,c} - \phi_r([x_{i,k-m;k-1,c}, x_{j,k-delay-m;k-delay-1,c}], \hat{\theta}_{delay,m}^{r,j \rightarrow i})\|^2 \end{cases}$$

式中:  $ssr_{m,i}^a$  为时间窗口规模为  $m$  下, 基于时序自回归模型的运行员个体  $i$  状态预测误差平方和,  $ssr_{delay,m}^{r,j \rightarrow i}$  为时间窗口规模为  $m$ 、时间延迟  $delay$  下, 基于时序相关回归模型的运行员个体  $i$  状态预测误差平方和。  $n_m^r$  和  $v_m^r$  分别为相关回归模型的样本数量和参数数量, 其中  $n_m^r = kc$ ,  $v_m^r = 2m + 1$ 。  $ssr_{m,i}^a - ssr_{delay,m}^{r,j \rightarrow i}$  服从  $\chi^2$  分布, 其参数数量为时序自回归模型与时序相关回归模型之差, 数值与时间延迟  $m$  相同。

由此获得基于 Granger 的 F 统计量, 该统计数值服从 F 分布  $f_{delay,m}^{r,j \rightarrow i} \sim F(m, kc - 2m - 1)$ , 其自由度为  $(m, kc - 2m - 1)$ 。同时为了对显著度存在差异的运行员群体之间的 Granger 原因进行分析, 构建 Granger 原因概率  $p_{delay,m}^{r,j \rightarrow i}$  的计算公式为:

$$p_{\text{delay},m}^{r,j \rightarrow i} = \int_0^{f_{\text{delay},m}^{r,j \rightarrow i}} \psi_{F \in m, k-2m-1}(z) \quad (6)$$

式中:  $\psi_{F \in m, k-2m-1}$  为概率密度函数,  $z$  为分布中的随机变量。可以看出,若通过时序自回归模型和时序相关回归模型的预测误差之间的差距  $ssr_{m,i}^a - ssr_{\text{delay},m}^{r,j \rightarrow i}$  越大, F 分布统计量数值越大,导致  $p_{\text{delay},m}^{r,j \rightarrow i}$  数值越大,则原假设成立,即个体  $i$  与个体  $j$  存在影响关系。

经过分析,使用  $p_{\text{delay},m}^{r,j \rightarrow i}$  数值具有一定特点:1) 运动员个体之间的交互影响是有向、存在因果关系的。比如在排球比赛中,拦网行为受到扣球行为影响,而扣球行为不受拦网行为影响。2) 运动员个体之间的影响是具有延迟效应的。比如在排球比赛中,拦网行为的运动员需要选择合适的起跳时间,实现与排球的接触,本文需要根据个体之间的影响强度选择合适的延迟量。3)  $p_{\text{delay},m}^{r,j \rightarrow i}$  更便于分析与判断有因果关联的个体之间的影响强度。由于 F 分布中,通过统计量难以直观反映群体交互行为的影响强度,需要对统计量进行转化,便于分析个体之间影响强度。

## 1.2 时序因果图推理模型

运动员群体之间的交互影响程度受时间延迟变量影响,在延迟时间过小或过大时,群体之间的影响程度会减弱,导致个体特征匹配过程中会出现融合误差,因此本文通过考虑基于 Granger 的时间延迟情况,保证时间对齐处理,降低个体特征匹配过程中出现的融合误差。

### 1) 时序因果图

时序因果图如图 3 所示,节点可视为运行员个体,边可视为交互影响关系,箭头方向为施加影响运动员个体指向受影响运动员个体,边上的数值表示运动员群体之间交互影响强度最大时对应的时间延迟量。对于给定的视频帧中的提取的原因概率  $p_{\text{delay},m}^{r,j \rightarrow i}$  构建的矩阵,可分析出运动员个体之间的因果关系。同时为了对运动员个体之间的延迟差异进行衡量,使用 0、1、2、3 来对延迟量进行描述,其中数值 0 表示时间同步,1、2、3 表示不同的延迟量。对于存在延迟量的运动员群体,其最优的时间延迟量可根据最大原因概率数值来确定,即:

$$[\text{delay}_{j \rightarrow i}^*, p_{\text{delay}_{j \rightarrow i}^*, m}^{r,j \rightarrow i}] = \text{argmax}_{\text{delay}_{j \rightarrow i}} p_{\text{delay}_{j \rightarrow i}, m}^{r,j \rightarrow i} \quad (7)$$

式中:  $p_{\text{delay}_{j \rightarrow i}, m}^{r,j \rightarrow i}$  为最优的时间延迟量  $\text{delay}_{j \rightarrow i}^*$  下对应的影响强度,由此获得延迟数值  $\text{delay}_{j \rightarrow i}^*$ 。

为了筛选出显著性强的因果关系,本文设置关系阈值  $\tau$ ,设计因果关系筛选公式为:

$$e_{j \rightarrow i} = \begin{cases} 1, & p_{\text{delay}_{j \rightarrow i}^*, m}^{r,j \rightarrow i} > \tau \\ 0, & \text{其他} \end{cases} \quad (8)$$

由此可获得时序因果图  $E = \{e_{j \rightarrow i}\}$ ,可根据不同延迟量来对排球比赛中复杂时延情况进行描述。

### 2) 时间上下文特征描述

如图 4 所示,为了提取时间同步信息特征,本文设计时间因果图推理模型,对时间关系进行描述。首先,根据

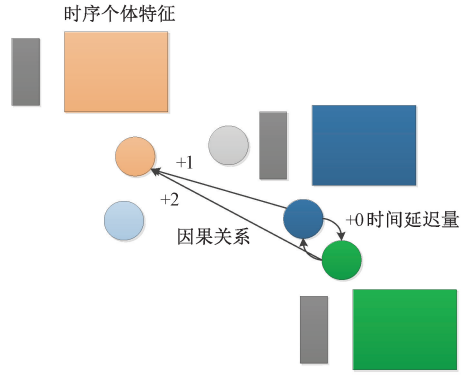


图 3 时序因果图

Fig. 3 Temporal causality diagram

时序因果图反映的个体之间交互关系,选择有交互关系的运动员个体,将施加影响的运动员行为特征按照时间延迟量  $\text{delay}_{j \rightarrow i}^*$  进行平移处理,过程为:

$$x_{i,k,c}^{\text{shift},j \rightarrow i} = \text{shift}(x_{j,k,c}, \text{delay}_{j \rightarrow i}^*) = \begin{cases} x_{j,1,c}, & k - \text{delay}_{j \rightarrow i}^* < 1 \\ x_{j,k - \text{delay}_{j \rightarrow i}^*, c}, & \text{其他} \end{cases} \quad (9)$$

式中:  $k$  表示当前时刻,个体  $j$  为施加影响运动员,  $x_{i,k,c}^{\text{shift},j \rightarrow i}$  为受影响运动员个体  $i$  的因果上下文特征。通过时间平移操作后,能够保证时间对齐,特征融合匹配更为合理。需要说明的是,经过时间平移操作后,初始特征位置会空出,本文选取第 1 帧特征数据对空白地方进行填充。

### 3) 特征融合方式

如图 4 所示,经过时间对齐操作后,获得受影响运动员个体  $i$  的因果上下文特征  $x_{i,k,c}^{\text{shift},j \rightarrow i}$ ,在此基础上,需要设计特征融合方式。本文选择串联方式,兼顾所有个体特征信息,同时为去除通道中的无关特征,设计一种通道级别的特征融合方式,获得时间对齐后的融合特征  $x_{i,k}^{\text{con},j \rightarrow i}$ ,过程为:

$$x_{i,k}^{\text{con},j \rightarrow i} = \text{concat}_k(\omega_i^d x_{i,k}, \omega_{j \rightarrow i}^d x_{i,k}^{\text{shift},j \rightarrow i}) \quad (10)$$

式中:  $\text{concat}_k$  表示维度  $k$  下对个体特征串联处理。同时为了对特征信息进行约束,设计了运动员个体  $i$  和  $j$  分别对应的通道约束参数  $\omega_i^d$  和  $\omega_{j \rightarrow i}^d$ ,运动员个体  $i$  和  $j$  特征处理后,通道规模分别为  $C - C/d$  和  $C/d$ ,其中  $C$  和  $d$  分别表示特征总通道规模和通道系数。可以看出,通道系数  $d$  增大时,会降低施加影响运动员个体  $j$  提供的特征影响,直到施加影响运动员个体影响消失,模型将转换为无交互影响的原始模型。

### 4) 多个体特征平均操作处理

如图 4 所示,通过特征融合方式,能够将个体的影响特征进行融合计算,考虑到每个运动员个体受到的多个体特征影响,需要对多个体特征进行处理。首先,基于因果检测模型,确定对个体  $i$  施加影响的所有交互个体;然后,基于时间对齐操作,获得所有的时间上下文特征描述;最



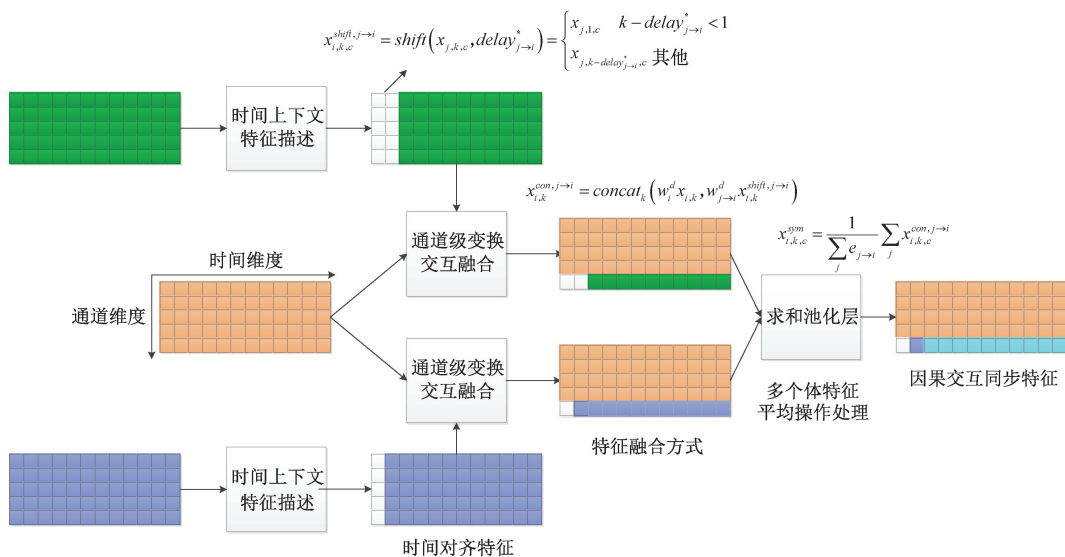


图 4 时序因果图推理过程

Fig. 4 Reasoning process of temporal causality diagram

后,基于平均操作,获得时间对齐特征信息  $x_{i,k,c}^{sym}$ , 过程为:

$$x_{i,k,c}^{sym} = \frac{1}{\sum_j e_{j \rightarrow i}} \sum_j x_{i,k,c}^{con,j \rightarrow i} \quad (11)$$

式中:  $e_{j \rightarrow i}$  为因果关系,在模型中由于只计算有交互影响关系的个体,因此  $e_{j \rightarrow i}$  设置为 1。

经过图 4 中时序因果图推理过程,获得时序因果图推理结果,如图 5 所示。

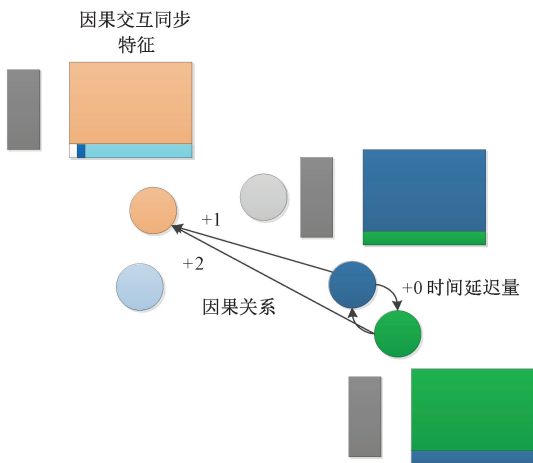


图 5 时序因果图推理结果

Fig. 5 Temporal causality diagram inference results

### 1.3 多尺度时空因果特征图

基于 1.2 节中时序因果图推理模型,能够获得排球运动员群体行为的因果关系,但考虑仍不够全面,需要补充距离约束和外观约束,计算多尺度时空因果特征图,进一步对群体行为类型进行细化描述和区分。距离约束主要用来描述运动员个体之间的距离,外观约束主要用来判断敌我类别。如图 6 所示,距离约束和外观约束可补充到交互特征中,增加模型的代表能力。

互特征中,增加模型的代表能力。

#### 1) 距离约束

针对运动员个体边界框,通过计算多个距离数值,获得多尺度距离特征图,公式为:

$$e_{s,i,j,k}^{dist} = \begin{cases} 1, & dist_{i,j} \leq \lambda_s \cdot width \\ 0, & \text{其他} \end{cases} \quad (12)$$

式中:  $s$  为个体距离序号,  $dist_{i,j}$  为运动员个体  $i$  和  $j$  的边界框之间距离,  $\lambda_s$  为距离尺度参数,  $width$  为图分辨率宽度。

#### 2) 外观约束

针对两个运动员个体,基于点积相似度算法,计算外观特征图中两个运动员个体的外观相似度  $e_{i,j,k}^{app}$ , 公式为:

$$e_{i,j,k}^{app} = \frac{1}{C} (w_1^{app} \cdot x_{i,k,1,C})^T \cdot (w_2^{app} \cdot x_{j,k,1,C}) \quad (13)$$

式中:  $C$  表示特征通道规模,  $x_{i,k,1,C}$  和  $x_{j,k,1,C}$  分别为运动员个体  $i$  和  $j$  在  $k$  时刻下的通道特征数据,  $w_1^{app}$  和  $w_2^{app}$  分别为模块参数,便于特征处理。

#### 3) 多尺度时空因果特征图

通过将因果特征图以及距离特征图、外观特征图进行融合处理,综合考虑多类别约束,获得多尺度时空因果特征图,能够更加清晰的描述特征,该特征图邻接关系计算为:

$$e_{s,i,j,k}^{fuse} = \frac{e_{j \rightarrow i} \cdot e_{s,i,j,k}^{dist} \cdot e_{i,j,k}^{app}}{\sum_j e_{j \rightarrow i} \cdot e_{s,i,j,k}^{dist} \cdot e_{i,j,k}^{app}} \quad (14)$$

式中:分母  $\sum_j e_{j \rightarrow i} \cdot e_{s,i,j,k}^{dist} \cdot e_{i,j,k}^{app}$  为归一化处理。

针对计算出的多尺度时空因果特征图,本文基于图卷积处理,计算所有时刻下的变化特征,公式为:

$$X_{s,k}^{graph} = E_{s,k}^{fuse} E_k^{syn} W_k^{graph} \quad (15)$$

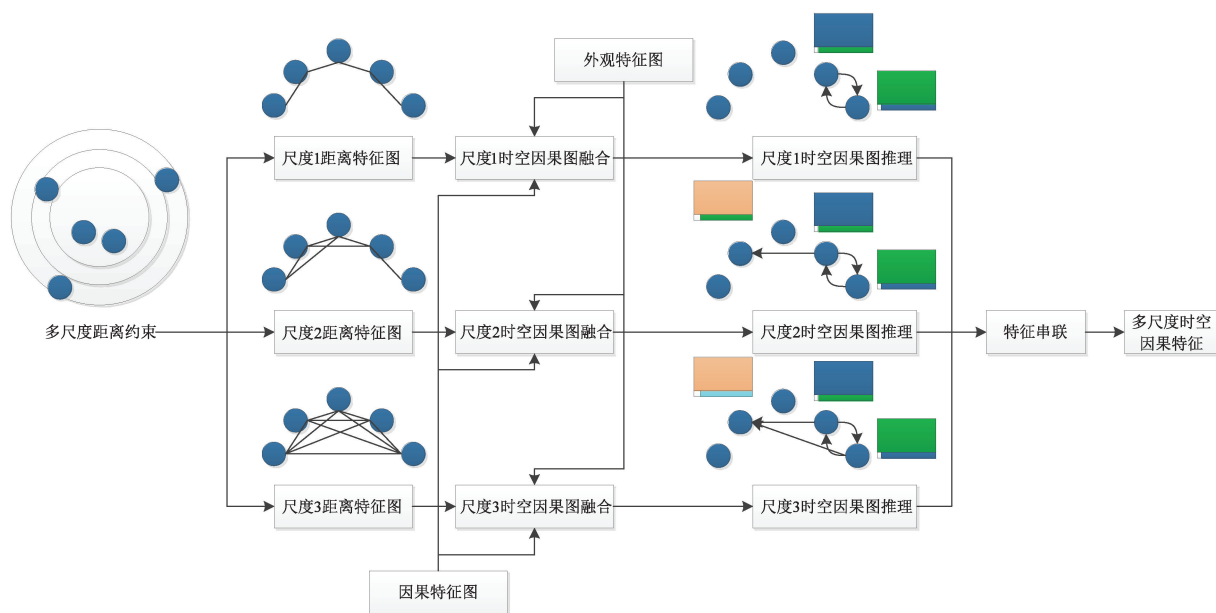


图 6 多尺度时空因果特征计算流程

Fig. 6 Calculation process of multi-scale spatio-temporal causal features

式中:  $X_{s,k}^{graph} = \{x_{s,i,k,c}^{graph}\}$  为图卷积处理后的时空因果特征,  $E_{s,k}^{fuse} = \{e_{s,j \rightarrow i,k}^{fuse}\}$  为时空特征图的邻接关系图,  $X_k^{sym} = \{x_{i,k,c}^{sym}\}$  为时间对齐特征,  $W_k^{graph}$  为图卷积参数。

#### 4) 个体和群体识别机制

针对运动员个体行为识别问题,首先,将图卷积处理后的综合特征在维度  $s$  下串联处理  $concat_s$ ; 然后,基于全连接层  $FCs(\cdot, \omega^{ind})$  处理,获得运动员个体行为类别;最后,通过视频帧的平均结果作为最终的个体行为识别结果,公式为:

$$y_{i,u}^{ind} = \frac{1}{T} \sum_k FCs(concat_s(x_{s,i,k,1,C}^{graph}), \omega^{ind}) \quad (16)$$

式中:  $\omega^{ind}$  为个体行为识别分类器参数。

针对运动员群体行为识别问题,首先,基于  $max pooling$  操作对  $x_{s,i,k,1,C}^{graph}$  进行处理,获得代表性特征  $x_{s,k,1,C}^{graph}$ ; 然后,基于串联处理  $concat_s$  考虑多特征方式;最后基于两层全连接层  $FCs(\cdot, \omega^{group})$  实现群体行为识别,公式为:

$$\begin{cases} y_{i,u}^{group} = \frac{1}{T} \sum_k FCs(concat_s(x_{s,i,k,1,C}^{graph}), \omega^{group}) \\ x_{s,k,1,C}^{graph} = maxpooling_i(x_{s,i,k,1,C}^{graph}) \end{cases} \quad (17)$$

式中:  $\omega^{group}$  为群体行为识别分类器参数。

#### 1.4 损失函数

本文计算出的多尺度时空因果特征能够用于运动员个体和群体行为识别,因此本文在设计损失函数时考虑个体行为和群体行为,公式为:

$$L = \frac{1}{N^{train}} \sum_u \sum_i L_E(y_{i,u}^{ind}, y_{i,u,gt}^{ind}) + \frac{1}{N^{train}} \sum_u L_E(y_{u,gt}^{group}, y_{u,gt}^{group}) \quad (18)$$

式中:  $u$  为训练集视频序号,  $N^{train}$  为训练集视频规模数量,  $y_{i,u,gt}^P$  和  $y_{u,gt}^G$  分别表示分别为个体行为和群体行为的标记,  $L_E(\cdot)$  为交叉熵损失函数,公式为:

$$L_E(y_{i,u}^{ind}, y_{i,u,gt}^{ind}) = y_{i,u}^{ind} \log(y_{i,u,gt}^{ind}) + (1 - y_{i,u,gt}^{ind}) \log(1 - y_{i,u}^{ind}) \quad (19)$$

## 2 实验与结果分析

### 2.1 数据集

为验证本文提出的考虑多尺度时空因果特征的排球群体行为识别方法的有效性与优越性,选择公开排球数据集<sup>[13]</sup>以及从 2024 年巴黎奥运会排球比赛中收集的数据作为研究对象,按照 8:2 的比例划分训练集和测试集,两个数据集信息如表 1 所示。

表 1 排球比赛数据集信息

Table 1 Information of volleyball match dataset

编号	来源	规模	行为种类	性质
数据集 1	文献[13]	55 个视频、4 830 视频帧	个体 9 类、群体 8 类	公开
数据集 2	2024 年巴黎奥运会排球比赛	44 个视频、3 800 视频帧		自选

数据集中行为种类示例如表 2 所示,在图中对个体识别行为种类进行标注,并给出群体行为标签。

2.2 实验参数设置

本文实验中的相关参数如表 3 所示。

表 2 行为种类  
Table 2 Types of behaviors









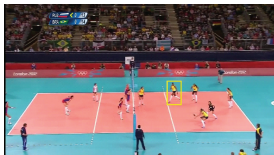








个体行为	图像举例	说明	群体行为	图像举例	说明
Blocking		拦网	Right set		右二传
Digging		垫球	Right spike		右扣球
Falling		跌倒	Right pass		右传球
Jumping		跳跃	Right winpoint		右得分
Moving		移动	Left set		左二传
Setting		二传	Left spike		左扣球
Spiking		扣球	Left pass		左传球
Standing		站立	Left winpoint		左得分
Waiting		等待			

表 3 实验参数设置

Table 3 Experimental parameter settings

参数	设置	参数	设置
操作平台	Pytorch	采样点	2×2
内存	64 G	batch size	0.8
处理器	Intel Core i9-10900X@3.7 GHz	分类器 dropout	0.3
图像特征提取模型	Inception-v3	优化器	Adam
特征通道数	1 024	初始学习率	1×10 <sup>-4</sup>
ROI Align 插值方法	双线性插值	多尺度距离参数	0.1,0.2,0.3,0.4

2.3 模型评价指标

针对运动员个体行为和群体行为识别问题,选择单个类别识别平均正确率  $MPCA$  和多类识别正确率  $MCA$  两个指标作为模型评价指标,公式为:

$$MPCA = \frac{1}{K} \sum_k \frac{Acc_k}{Sum_k} \tag{20}$$

$$MCA = \frac{\sum_k Acc_k}{\sum_k Sum_k} \tag{21}$$

式中:  $Acc_k$  和  $Sum_k$  分别为第  $k$  个行为类别的正确识别数量和样本总数,  $K$  为行为类别总数。

2.4 对比实验与分析

为验证本文提出的考虑多尺度时空因果特征的排球群体行为识别方法的优越性,以 2 个数据集为研究对象,选择最新的一些基线模型作为对比进行实验,并对实验结果进行分析。基线模型如下:

1)多尺度子群体交互关系群体行为识别 (multi-scale sub-group interaction relationships, MSIR)<sup>[14]</sup>:该模型基于 Inception-v3 网络,通过考虑不同尺度下个体交互关系,实现对群体行为进行识别。

2)基于图的残差聚合网络 (graph-based residual aggregation network, GRAIN)<sup>[15]</sup>:来源于 IJCAI2022,该模型基于 GRAIN 网络对群体中个体行为差异进行表征和建模,通过局部残差关系模块捕捉时空特征,从而实现识别。

3)聚类时空网络 (clustered spatial-temporal transformer, CST)<sup>[16]</sup>:来源于 CVPR 2021,该模型通过聚类时空 Transformer,增强表征能力,实现群体行为识别。

4)自监督全局时空交互预训练 (self-supervised global spatio-temporal interaction pre-training, SGSIP)<sup>[17]</sup>:来源于 TCSVT 2023,该模型基于自监督机制,考虑全局时空特征,实现群体行为识别。

5)分层关系推理模型 (hierarchical relation inference, HRI)<sup>[18]</sup>:来源于 TCSVT 2023,该模型基于分层关系推理模型,实现主动空间位置的群体行为识别。

6)赋权关系网络 (empowering relational network, ERN)<sup>[19]</sup>:来源于 ECCV2020,基于自注意力增强条件随机场,构建关系网络群体活动识别模型。

对比实验结果如表 4 所示。

表 4 对比实验结果

Table 4 Compares the experimental results

识别模型	骨干网络	数据集 1			数据集 2		
		群体 $MPCA$	群体 $MCA$	个体 $MCA$	群体 $MPCA$	群体 $MCA$	个体 $MCA$
MSIR	Inception-v3	94.0	94.2	83.7	91.7	92.1	82.1
GRAIN	VGG16	92.3	92.5	82.2	90.1	90.5	80.2
CST	Inception-v3	93.7	94.1	83.5	91.5	91.8	81.3
SGSIP	Inception-v3	93.4	93.7	82.4	90.8	91.0	80.8
HRI	Inception-v3	88.0	92.4	81.3	89.5	89.8	78.6
ERN	I3D	93.6	94.0	83.1	91.1	91.4	81.2
本文模型	Inception-v3	<b>94.3</b>	<b>94.4</b>	<b>84.2</b>	<b>92.4</b>	<b>92.6</b>	<b>82.4</b>

根据表 4 中的对比实验结果,可以看出:

1)GRAIN 和 ERN 识别模型从单独对群体行为建模的角度进行考虑,利用深度学习网络对群体行为特征进行学习,此类方法忽略了群体之间的交互作用,导致潜在的

动作语义难以挖掘,使得效果一般。

2)SGSIP 和 HRI 识别模型从个体行为之间的交互作用进行考虑,结合时空变化对动作的影响,但尺度较为单一,导致特征利用不够明显。MSIR 虽然考虑了群



体行为多尺度的交互关系,但处理方式较为粗略,性能略差。

3)在 2 个数据集上,本文提出的考虑多尺度时空因果特征的排球群体行为识别方法均优于基线模型,在数据集 1 上,3 个评价指标较次优算法分别提升 0.32%、0.21%和 0.60%,在数据集 2 上,3 个评价指标较次优算法分别提升 0.76%、0.54%和 0.37%。表明本文所提模型能够充分发挥多尺度时空因果图在描述群体交互影响的优势,提取的

多尺度时空因果特征图能够有效学习群体行为特征,识别效果较好。

2.5 消融实验与分析

为验证本文提出的考虑多尺度时空因果特征的排球群体行为识别方法中各模块的有效性,以 2 个数据集为研究对象,设计消融实验,并对实验结果进行分析,结果如表 5 所示,需要说明的是,单尺度距离参数  $\lambda_s$  选择 0.4,实验 5 是本文所提出的识别模型。

表 5 消融实验结果  
Table 5 Results of ablation experiments

实验 编号	时序因果 特征	单尺度 距离特征	多尺度 距离特征	外观 特征	数据集 1			数据集 2		
					群体	群体	个体	群体	群体	个体
					MPCA	MCA	MCA	MPCA	MCA	MCA
1	✓	✓		✓	93.3	93.5	83.1	91.3	91.5	81.3
2			✓	✓	88.2	88.5	76.9	86.4	87.2	75.8
3	✓				93.7	93.8	83.6	91.6	91.8	81.6
4	✓		✓		94.0	94.1	83.9	92.1	92.4	81.8
5	✓		✓	✓	<b>94.3</b>	<b>94.4</b>	<b>84.2</b>	<b>92.4</b>	<b>92.6</b>	<b>82.4</b>

根据表 5 中实验结果,可以看出:  
1)实验 1 中识别效果较本文模型差,表明运动员个体与其他多个个体之间的距离数据包含关键特征信息,多尺度距离包含更丰富的行为特征信息,能够更好地对特征进行描述,对识别性能的提升有正向引导作用。  
2)实验 2 中识别效果在实验组中最差,表明时序因果特征是模型中的核心模块,是表征群体交互影响能力的关键手段,该特征能够充分发挥描述群体交互影响的优势,提升综合识别性能。

3)本文识别模型的性能优于实验 4、实验 4 识别模型的性能优于实验 3,表明多尺度距离特征以及外观特征在模型中均发挥积极作用,能够对时序因果特征进行补充,增加模型的表征能力。

2.6 模型稳定性分析

为分析模型稳定性,选择数据集 1 为研究对象,给出个体和群体的损失函数迭代变化图,如图 7 所示,可以看出,针对行为识别问题,本文所提算法的损失值能够在 100 轮次以后趋于稳定,表明模型具有一定稳定性。

2.7 参数影响分析

本节选择时序因果特征提取中的不同延迟量、不同距离尺度等变量作为研究对象,分析变量变化时对识别模型性能的影响。

1)时序因果特征提取中的不同延迟量对识别性能的影响

根据 1.2 节中时间上下文特征描述,时间延迟量  $delay_{j^*-i}$  的选择方式有:(1)0 延迟量,即特征不按照时间平移进行对齐,直接进行拼接(实验 1)。(2)按照固定延迟量(1 帧)进行平移(实验 2)。(3)延迟量自适应选择,即最

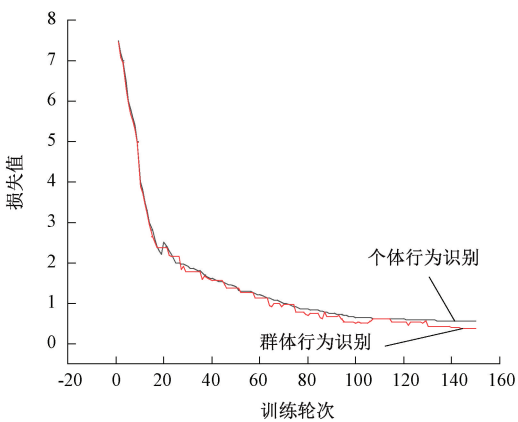


图 7 损失函数迭代变化图  
Fig. 7 Iterative change diagram of loss function

优的时间延迟量可根据最大原因概率数值来确定(本文模型),实验中其他模块保持一致,延迟量变化范围为[0,1,2,3],以 2 个数据集为研究对象,实验结果如表 6 所示。

根据表 6 中实验结果,可以看出:合适的延迟量能够增强模型的识别性能,本文采用的延迟量自适应选择方法,能够确保特征在同一个时间维度上进行处理,降低个体特征匹配过程中出现的融合误差。

2)不同距离尺度参数对识别性能的影响

以 2 个数据集为研究对象,实验中其他模块保持一致,通过调整不同距离尺度参数进行多次实验,分析不同距离尺度参数  $\lambda_s$  对识别性能的影响,实验中其他模块保持一致,本文模型是基于  $\lambda_s$  在[0.1,0.2,0.3,0.4]取值范围内使用 4 个尺度的并联卷积网络,实验结果如表 7 所示。

表 6 不同延迟量下识别性能

Table 6 Recognition performance with different amounts of delay

实验编号	数据集 1			数据集 2		
	群体 MPCA	群体 MCA	个体 MCA	群体 MPCA	群体 MCA	个体 MCA
1	89.8	91.4	79.7	88.4	88.8	77.8
2	92.1	92.5	81.8	89.8	90.1	80.2
本文模型	94.3	94.4	84.2	92.4	92.6	82.4

表 7 不同距离尺度下识别性能

Table 7 Recognition performance under different distance scales

类型	$\lambda_s$	数据集 1			数据集 2		
		群体 MPCA	群体 MCA	个体 MCA	群体 MPCA	群体 MCA	个体 MCA
单尺度	0.1	90.2	92.5	80.6	90.1	90.2	78.8
	0.2	91.0	93.2	82.8	91.1	92.2	79.9
	0.3	93.2	93.4	83.0	91.2	91.4	81.2
	0.4	93.3	93.5	83.1	91.3	91.5	81.3
多尺度	[0.1,0.2,0.3,0.4]	94.3	94.4	84.2	92.4	92.6	82.4

根据表 7 中实验结果,可以看出:1)单尺度情况下,在 [0.1,0.2,0.3,0.4]取值范围内, $\lambda_s$ 取值越大,识别效果越好,表明 $\lambda_s$ 越大,能够帮助模型获取更多的特征信息,因果关系更充分。2)多尺度模型能够包含所有尺度的行为特征信息,保证行为的交互影响关系更为有效,识别效果提升较为明显。

3 结 论

针对排球比赛群体行为特征难以描述、识别性能不高等问题,本文提出了考虑多尺度时空因果特征的排球群体行为识别方法。基于 Grange 因果检测模型,通过时序自回归模型和时序相关回归模型进行回归估计,并基于 F 分布分析判断个体之间是否存在交互影响。融合了时序、距离和外观等特征类型,构建多尺度时空因果特征提取模型,在 2 个数据集上,本文提出的考虑多尺度时空因果特征的排球群体行为识别方法均优于基线模型,提取的多尺度时空因果特征图能够有效学习群体行为特征,识别效果较好。通过融合对比实验,表明时序因果特征是模型中的核心模块,是表征群体交互影响能力的关键手段,该特征能够充分发挥描述群体交互影响的优势。多尺度距离特征以及外观特征在模型中均发挥积极作用,能够对时序因果特征进行补充,增加模型的表征能力。下一步可结合多模态数据,增加行为特征的表征手段,进一步对排球比赛群体行为特征识别进行探索。

参考文献

[1] LI CH, LIU B L. Group non-critical behavior recognition based on joint attention mechanism of sensor data and semantic domain[J]. International

Journal of Advanced Computer Science and Applications, 2024,15(7): 608-617.

[2] LEE Y C, LEE S Y, KIM D Y, et al. GLBRF: Group-based lightweight human behavior recognition framework in video camera[J]. Applied Sciences-Basel,2024,14(6): 2424.

[3] 朱晓林,王冬丽,欧阳万里,等. 基于深度学习的群体行为识别:综述与展望[J/OL]. 控制理论与应用,1-17 [2024-09-25]. <http://kns.sjuku.top/kcms/detail/44.1240.TP.20240812.1126.002.html>.

ZHU X L, WANG D L, OUYANG W L, et al. Group activity recognition based on deep learning: Overview and outlook[J/OL]. Control Theory & Applications, 1-17 [2024-09-25]. <http://kns.sjuku.top/kcms/detail/44.1240.TP.20240812.1126.002.html>.

[4] 陈辰,苏益凡,周伟,等. 实验室人员不安全行为智能识别与预警系统[J/OL]. 电子测量技术,1-10 [2024-12-15]. <http://kns.cnki.net/kcms/detail/11.2175.TN.20241126.1451.002.html>.

CHEN CH, SU Y F, ZHOU W, et al. Intelligent recognition and early warning system for unsafe behaviors of laboratory personnel[J/OL]. Electronic Measurement Technology, 1-10 [2024-12-15]. <http://kns.cnki.net/kcms/detail/11.2175.TN.20241126.1451.002.html>.

[5] 何赞泽,谯灵俊,王洪金,等. 基于改进 DETR 的智慧车间人员典型行为识别算法[J]. 电子测量与仪器学报,2024,38(9):76-84.

- HE Y Z, QIAO L J, WANG H J, et al. Typical behavior recognition algorithm for intelligent workshop personnel based on improved DETR[J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(9): 76-84.
- [6] 周璇, 易剑平. 基于优化 CBAM 改进 ResNet50 的异常行为识别方法[J]. 国外电子测量技术, 2024, 43(5): 36-41.
- ZHOU X, YI J P. Improved abnormal behavior recognition method of ResNet50 based on optimized CBAM[J]. Foreign Electronic Measurement Technology, 2024, 43(5): 36-41.
- [7] 张宏宏, 李文华, 郑家毅, 等. 有人/无人机协同作战: 概念、技术与挑战[J]. 航空学报, 2024, 45(15): 168-194.
- ZHANG H H, LI W H, ZHENG J Y, et al. Manned/unmanned aerial vehicle cooperative combat system: Concepts, technologies, and challenges[J]. Acta Aeronautica et Astronautica Sinica, 2024, 45(15): 168-194.
- [8] 张宏宏, 甘旭升, 李双峰, 等. 复杂低空环境下考虑区域风险评估的无人机航路规划[J]. 仪器仪表学报, 2021, 42(1): 257-266.
- ZHANG H H, GAN X SH, LI SH F, et al. UAV route planning considering regional risk assessment under complex low altitude environment[J]. Chinese Journal of Scientific Instrument, 2021, 42(1): 257-266.
- [9] LU L H, LU Y, YU R ZH, et al. GAIM: Graph attention interaction model for collective activity recognition[J]. IEEE Transactions on Multimedia, 2020, 22(2): 524-539.
- [10] HONG C. Group behavior recognition algorithm for basketball video based on hidden tag mining strategy under gamification teaching and learning [J]. Entertainment Computing, 2024, 50: 100638.
- [11] GAO Y, YAN K, DAI B SH, et al. Recognition of aggressive behavior of group-housed pigs based on CNN-GRU hybrid model with spatio-temporal attention mechanism[J]. Computers and Electronics in Agriculture, 2023, 205: 107606.
- [12] 谢昭, 李骏, 吴克伟, 等. Granger 因果关系时空图推理的群体行为分析[J]. 计算机学报, 2023, 46(4): 856-876.
- XIE ZH, LI J, WU K W, et al. Spatial-temporal graph inference with granger causality relation for group activity analysis [J]. Chinese Journal of Computers, 2023, 46(4): 856-876.
- [13] AMER M R, LEI P, TODOROVIC S. HiRF: Hierarchical random field for collective activity recognition in videos [C]. The 13th European Conference on Computer Vision, 2014, 8694: 572-585.
- [14] 朱丽萍, 吴祀霖, 陈晓禾, 等. 多尺度子群体交互关系下的群体行为识别方法[J]. 电子与信息学报, 2024, 46(5): 2228-2236.
- ZHU L P, WU S L, CHEN X H, et al. Group activity recognition under multi-scale sub-group interaction relationships[J]. Journal of Electronics & Information Technology, 2024, 46(5): 2228-2236.
- [15] LI W, YANG T ZH, WU X, et al. Learning graph based residual aggregation network for group activity recognition[C]. The Thirty-First International Joint Conference on Artificial Intelligence, 2022: 1102-1108.
- [16] LI SH CH, CAO Q G, LIU L B, et al. GroupFormer: Group activity recognition with clustered spatial-temporal transformer [C]. Proceedings of the 2021 IEEE Conference on Computer Vision and Pattern Recognition, Virtual Event, 2021: 13648-13657.
- [17] DU Z X, WANG X, WANG Q. Self-supervised global spatio-temporal interaction pre-training for group activity recognition[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2023, 33(9): 5076-5088.
- [18] WU L F, LANG X L, XIANG Y, et al. Active spatial positions based hierarchical relation inference for group activity recognition[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2023, 33(6): 2839-2851.
- [19] PRAMONO R R A, CHEN Y T, FANG W H. Empowering relational network by self-attention augmented conditional random fields for group activity recognition[C]. The 16th European Conference on Computer Vision, 2020: 71-90.

## 作者简介

万宁(通信作者), 本科, 讲师, 主要研究方向为体育教育学、运动训练学。

E-mail: wanningxh@163.com

张美珠, 本科, 讲师, 主要研究方向为体育教育学、运动训练学。

E-mail: zhangmeizhuxh@163.com