

中文引用格式:赵荣泳,韦炳宇,朱文杰,等.公共场所行人异常行为识别方法综述[J].中国安全科学学报,2024,34(2):83-93.

英文引用格式:ZHAO Rongyong, WEI Bingyu, ZHU Wenjie, et al. Overview of recognition methods of pedestrian abnormal behaviors in public places[J]. China Safety Science Journal, 2024, 34(2): 83-93.

公共场所行人异常行为识别方法综述*

赵荣泳 副教授, 韦炳宇**, 朱文杰, 郑程元, 李浩男

(同济大学 电子与信息工程学院 CIMS 研究中心, 上海 201804)

中图分类号: X915.2; TP391

文献标志码: A

DOI: 10.16265/j.cnki.issn1003-3033.2024.02.1125

基金项目: 国家自然科学基金资助(72374154)。

【摘要】 为明确公共场所行人异常行为识别理论与技术研究进展,首先,借助中国知网(CNKI)和Web 版引文数据库(WOS),给出公共场所行人异常行为广义定义与泛在特征,将常见异常行为划分为危害行为、不合群行为和违规行为3类;其次,从数据和技术基础视域,将现有异常行为识别方法划分为人工设计法、人体骨架法、红绿蓝(RGB)图像法和可穿戴传感器法4类;然后,梳理国内外主流人群异常行为数据集,分析相关算法在数据集上的性能表现;最后,从可用数据集和数据融合检测等方面总结现有研究方法局限性,给出未来研究方向与优化建议。研究结果表明:4类异常行为识别方法各有其优缺点;异常行为识别领域缺乏行为种类丰富、定义清晰、高质量的人群异常行为数据集;未来研究应聚焦稳健性强、准确率高的异常行为识别方法、模型及算法;探索多维数据融合互补检测方法,提升异常行为识别理论成果的应用场景的自治性和自适应性,提高公共场所人群安全治理水平。

【关键词】 公共场所; 行人异常行为; 识别方法; 可穿戴传感器; 异常行为数据集

Overview of recognition methods of pedestrian abnormal behaviors in public places

ZHAO Rongyong, WEI Bingyu, ZHU Wenjie, ZHENG Chengyuan, LI Haonan

(School of Electronic and Information Engineering, Tongji University, Shanghai 201804, China)

Abstract: The purpose of this research is to clarify the research progress of the theory and technology of pedestrian abnormal behavior recognition in public places. Firstly, with the help of China National Knowledge Infrastructure (CNKI) and the Web of Science (WOS), a broad definition and universal characteristics of abnormal pedestrian behavior in public places were given. The existing research results related to abnormal behaviors were divided into three categories: harmful behaviors, dissociable behaviors and violations. Then, from the perspective of data and technological foundations, the existing abnormal behavior recognition methods were divided into four categories: artificial design, human skeleton, Red Geen Blue (RGB) images and wearable sensors. Secondly, this study sorted out the abnormal behavior datasets of mainstream populations both domestically and internationally, and analyzed the performance of relevant algorithms on the datasets. Finally, the limitations of existing research methods in available datasets and data fusion detection were summarized, and future research directions and optimization

* 文章编号:1003-3033(2024)02-0083-11; 收稿日期:2023-08-12; 修稿日期:2023-11-18

** 通信作者:韦炳宇(2001—),男,广西贵港人,硕士研究生,研究方向为公共安全、异常行为识别、人群动力学等。E-mail:2331988@tongji.edu.cn.

suggestions were provided. The results indicate that these four types of abnormal behavior recognition methods have their own advantages and disadvantages. It is necessary to construct a diversified, well-defined and high-quality international benchmark dataset of abnormal behaviors among the crowd. Future research should focus on robust and accurate methods, models, and algorithms for identifying abnormal behaviors, explore multi-dimensional data fusion complementary detection methods, improve the application scenario consistency and adaptability of the theoretical results of abnormal behavior recognition, and eventually enhance the level of public place crowd safety governance.

Keywords: public places; pedestrian abnormal behaviors; recognition methods; wearable sensor; abnormal behavior dataset

0 引言

在机场、高铁站等人流密集场所,由行人异常行为引发的踩踏、恐怖袭击、群体恐慌等突发事件是公共安全领域社会关注的焦点,也是公共安全理论研究热点和难点。特别是近年来,伴随着城镇化发展,公共场所数量日益增多,人群聚集性大型活动规模也具增长态势。研究表明:人群活动中的行人异常行为是造成突发安全事件的重要原因^[1-2]。全球范围内由行人异常行为引发的突发安全事件所造成的人群伤亡和财产损失巨大。为识别行人异常行为,及时采取人群疏导和控制措施,研究人员从人工设计特征、人体骨架姿态特征和红绿蓝(Red Green Blue, RGB)图像特征、可穿戴传感器等角度展开理论和技术研究,探索行人异常行为的识别理论、方法、模型、算法和装置,并取得了明显成果。随着智能监控网络摄像机在人流密集场所的普遍应用,以及可穿戴传感器(如手环等)的普及,从海量数据中分析行人的行为并分辨出正常行为与异常行为,成为解决人流密集场所安全疏导和控制问题的关键环节^[3-4]。

早期基于视觉的行人异常行为检测方法从视频中制作人工设计的特征,表示行人的外观和动作特征,通过传统机器学习算法检测异常行为,其中,人工设计的特征包括梯度直方图、运动流、光流特征、轨迹特征等低级视觉特征^[5]。借助深度学习技术的快速发展,研究人员使用深度学习技术从视频中提取行人的外观和动作等高级特征,从而提高检测准确率。由于监控摄像机容易受到环境光和遮挡的影响,而可穿戴传感器具有不受环境、天气等外部条件限制以及数据真实完整的优点^[6],可用于步行、奔跑、跌倒等行人常见行为的检测^[7-9],进而识别异常行为。

近年来,学者们针对公共场所下行人异常行为

识别开展了研究,SOEMANN 等^[10]从监测目标、异常行为的定义与假设、使用的传感器类型和特征提取过程、学习方法和建模算法等5方面,介绍了智能监控中的异常行为检测方法;POPOOLA 等^[11]讨论了异常行为的上下文特征、场景建模以及行为抽象和表示;MABROUK 等^[12]重点研究了行为表示和行为建模,对行为表示的特征提取和描述技术进行了综述,给出行为建模的分类方法和框架;DHIMAN 等^[13]总结了现有人类异常活动识别的手工特征方法和基于神经网络的方法,以及用于异常行为识别的输入数据类型(包括RGB图像,深度图像、骨骼数据);徐涛等^[5]重点探讨了卷积神经网络(Convolutional Neural Networks, CNN)、自编码网络和生成对抗网络在人群异常行为检测任务中的应用研究现状;张晓平等^[3]讨论了特征提取方法,并从异常行为识别和异常行为检测2个角度分析和讨论了异常行为。综上,现有研究尚缺乏对公共场所下异常行为的清晰定义;仅分析所使用的技术方法优缺点,而未结合数据类型分析;同时,存在对已有算法的分析和梳理不全面等问题。

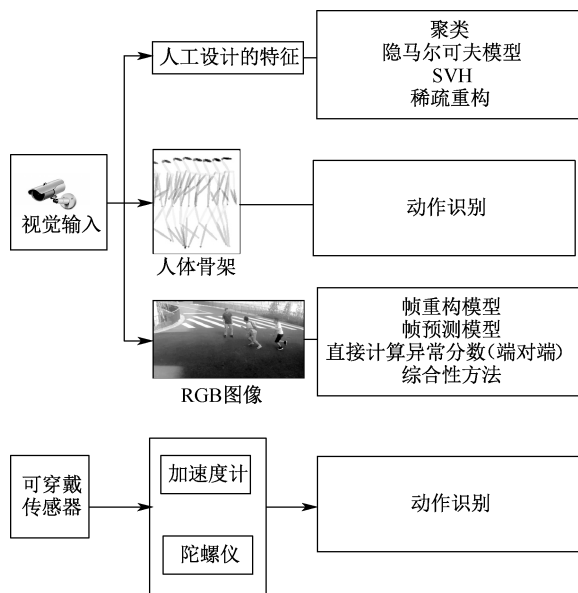
鉴于此,笔者拟从识别方法所涉及的数据类型和技术基础角度,划分当前主要的异常行为识别方法;统计、分析相关算法在异常行为数据集上的性能;并分析与总结目前异常行为识别相关研究存在的问题,以期深化异常行为识别技术研究,为保障公共场所人群安全提供参考依据。

1 公共场所中行人异常行为

1.1 行人异常行为特征

广义上的异常行为通常指很少发生的、明显区别于其他的普遍行为。研究人员将偏离正常行为称为异常行为^[14]。进一步,公共场所下的异常行为是指与所处情景不符、可能对自身或周围行人造成伤害,甚至造成一定规模安全事件的行为。公共场所

下的异常行为泛在特征包括:①随机性。行人的异常行为往往无法预测。②危害性。行人的异常行为对自身或周围环境、行人产生危害。③情景依赖性。异常行为的界定需要依赖于特定的场所。④辨识性。异常行为需要具有一定的辨识性,如轻微的肢体动作辨识度低,不列入考虑范围。如在校园人行道、商城、火车站等公共场所中的异常行为,包括抢劫、追逐、斗殴、跌倒、骑行等。异常行为识别方法分类如图 1 所示。



注:支持向量机(Support Vector Machine, SVM)。

图 1 异常行为识别方法分类

Fig. 1 Classification of abnormal behavior recognition method

各类场景中常见的异常行为见表 1。与正常行为相比,异常行为具有不可预知、无周期、持续时间短、突发性强等特点^[15]。同时,由于人群中的异常行为是高度多样化的,需要为不同场景定义异常行为。例如:在运动场地奔跑不会被视作异常行为,而在人行道中奔跑、跳跃等行为会被视作可能引发安全事件的异常行为。综上,把行人异常行为划分为 3 类:①危害行为。是指可能对自身或周围行人产生危害的行为。②不合群行为。是指与周围多数人不相同的行为。③违规行为。是指违反公共场所行为规范的行为。公共场所异常行为类型如图 2 所示。

表 1 各类场景中常见的异常行为

Tab. 1 Common abnormal behaviors in different scenarios

场景	异常行为	行为类型	相关文献
校园人行道	奔跑、跌倒、追逐等	危害行为	[14, 16-17]

续表 1

场景	异常行为	行为类型	相关研究文献
(高速)公路	行走、骑行、奔跑等	不合群行为	[18-19]
火车站、地铁站	跌倒、奔跑等	危害行为	[20-22]
超市、商店	偷窃等	违规行为	[16, 23]

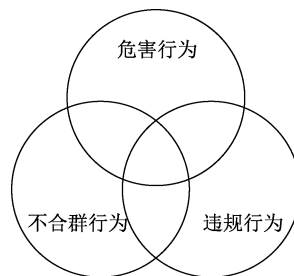


图 2 公共场所异常行为类型

Fig. 2 Category of abnormal behaviors in public places

1.2 异常行为研究方法现状

在文献搜集过程中,采用文献计量分析方法和内容分析法 2 种文献分析方法。在中国知网(China National Knowledge Infrastructure, CNKI)和 Web 版引文数据库(Web Of Science, WOS)中,采用关键字搜索 2000—2023 年相关文献并分析。首先,在 CNKI 以中文关键词:异常行为识别、检测,行人、人群进行搜索;在 WOS 以英文关键词组:Anomaly, or Abnormal, or Anomalous, Pedestrian behavior, Crowd activity 展开搜索;然后,人工阅读标题和摘要,删除重复文献和排除内容严重偏离的文献;最后,选取 195 篇文献用于后续分析。在检索和文献筛选过程中发现,由于公共场所下的异常行为类型复杂多样,有些研究只针对单类异常行为(跌倒^[24-25]、奔跑^[26-27]、逆行^[28-29]),有些研究针对多类异常行为或更广泛意义的异常行为。

近年来,行人异常行为逐渐成为研究热点。2007—2023 年行人异常行为相关研究的论文发表数量与趋势如图 3 所示。从图 3 可以看出,自 2017 年以来,异常行为研究论文数量明显增加,正成为研究热点。以跌倒行为为例,2007—2023 年基于不同数据类型实现跌倒行为检测相关研究的趋势如图 4 所示。由图 4 可知:基于传感器的异常行为识别研究分别在 2014、2019 年出现 2 次明显增长;基于视频的异常行为识别研究则从 2016 年开始呈现逐年增长的趋势,日益成为研究热点。

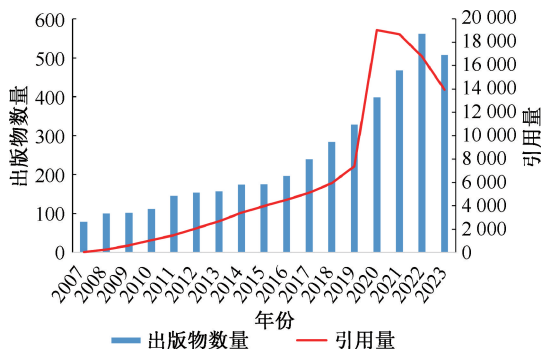


图3 2007—2023年关于异常行为研究的论文出版数量

Fig. 3 Number of papers published on abnormal behavior research from 2007 to 2023

2 行人异常行为识别方法

目前,按照数据类型角度,行人异常行为识别的输入数据包括视觉数据和可穿戴传感器数据。从数据和技术基础视域,将现有异常行为识别方法划分为人工设计法、人体骨架法、RGB图像法和可穿戴传感器法4类。其中,基于可穿戴传感器的方法多由获取的传感器数据识别行人的行为动作。

2.1 人工设计法

人工设计法是指基于人工设计特征的异常行为检测方法。早期行人异常行为识别方法多注重检测

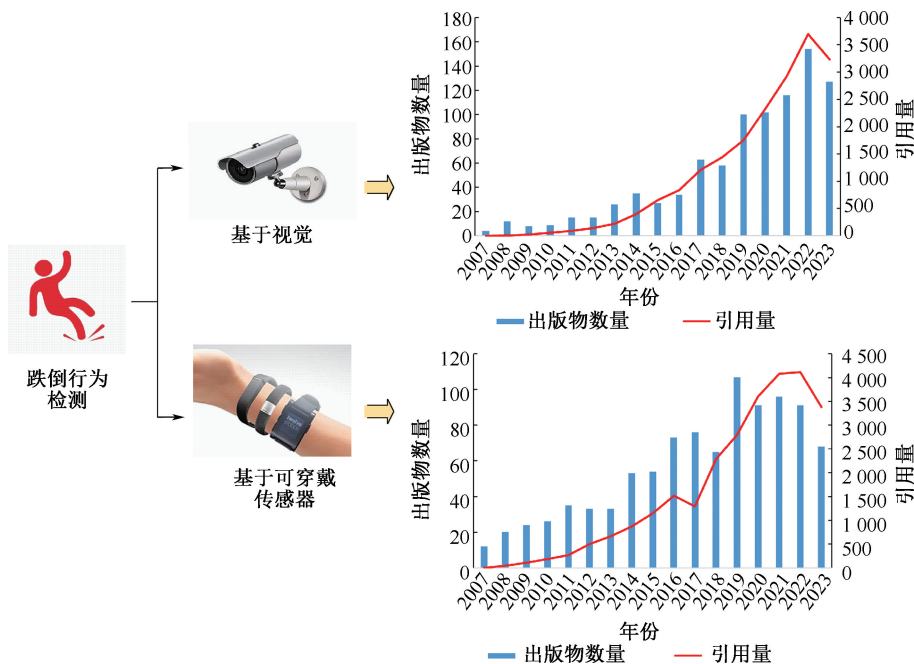


图4 跌倒行为研究趋势

Fig. 4 Research trends of fall behaviors

方面,并且多基于轨迹特征、光流特征、时空特征、梯度直方图等人工设计的特征。其基本思想在于通过这些人工设计的特征描述行为的外观和动作特征,以及正常行为与异常行为的差异。在生成这些特征后,研究人员通过聚类^[30-31]、隐马尔可夫模型^[21,32]、SVM^[18]和稀疏重构^[16,22-23]等模型或方法检测并定位出异常行为在图像中的位置。

轨迹特征用来表示视频序列中物体的位置信息和速度信息,其由运动向量组成。LI Ce等^[23]通过物体跟踪算法从视频中收集一组正常行为的轨迹,通过观察这些轨迹手动划分为不同子集;再对所有收集到的轨迹提取最小二乘三次样条曲线近似(Least-squares Cubic Spline Curves Approximation, LCSCA)特征,并建立字典集。在测试阶段,每个测

试轨迹也将用LCSCA特征表示,引入稀疏重构分析并设立一个经验性的阈值来区分正常行为和异常行为。

与基于轨迹的行为理解方式不同,ROSHTKHARI等^[33]提出一种基于逐像素时空特征分析的方法,使用时空视频卷(Spatio-Temporal Video Volumes, STVs)描述符来表示视频帧,并通过聚类的方法分别构建时域码本和空间域码本,从而学习每个像素的行为模式。CHATHURAMALI等^[18]提出2种时空描述符,其中,一种描述符结合光流值和轮廓特征的直方图,另一种为基于轨迹特征的描述符,通过这2种描述符使用SVM机对正常行为和异常行为分类。其中,轮廓特征是通过每一帧使用背景减法提取人物和物体轮廓得到。

视频序列中的动态纹理 (Dynamic Texture, DT)^[34]是指一种具有空间不变性和时间一致性的统计量的平稳随机过程的实现。LI Weixin 等^[14]使用混合动态纹理 (Mixed Dynamic Texture, MDT) 模型解释复杂人群场景的外观和运动。总的来说,这些基于人工设计特征的方法容易受环境条件影响,且精度较差。如基于 MDT 的方法存在异常行为定位结果不准确等问题^[35]。

2.2 人体骨架法

人体骨架法是指基于人体骨架的异常行为识别方法。行人的异常行为伴随着相应的肢体动作,通过识别行人的行为动作根据预定义的异常行为库即可判定是否为异常行为。人体骨架的空间位置信息能有效表征人的姿态信息。而连续的、时变的姿态信息构成了人的动作。并且,随着姿态估计算法(如 Openpose^[36], AlphaPose^[37])的改进,2D 人体骨架的获取变得更为简便,不少基于人体骨架的异常行为识别方法被提出。这类方法的基本流程为:使用姿态估计开源库从视频数据中提取行人的人体骨骼关键点形成对应行为骨架序列,接着使用卷积 CNN^[38-39]、循环神经网络 (Recurrent Neural Networks, RNN)^[40]、图 CNN (Graph CNN, GCNN)^[41-43]等方法对骨架序列分类。

ZHENG Yin 等^[44]使用 Openpose 从视频中提取人体骨骼关键点,接着使用时空图 GCNN (Spatial Temporal GCNN, ST-GCNN)^[41]对输入的骨架序列分类,从而判断视频中的行人是否跌倒。刘耀等^[45]采用类似的做法,用于识别施工现场工人的不安全动作,其在识别爬梯危险动作上达到较高的准确率。NAR 等^[46]探索了一种利用 Kinect 3D 相机提供的骨骼信息识别姿势技术,并将其用于有效的实时自动取款机智能监控。

在面对多人场景时,由于需要使用行人重识别算法匹配连续帧中的同一行人才能获得行人对应行为骨架序列,匹配误差将对动作识别效果产生影响。另外,这类方法往往在单人的异常行为识别中较为有效,而对于追逐、斗殴等多人异常行为的识别较为困难。因此,这类方法目前尚未在主流的异常行为数据集中大范围使用。

2.3 RGB 图像法

RGB 图像法是基于 RGB 图像的异常行为识别方法。随着深度学习技术的发展,从图像中自动学习特征受到越来越多的关注。基于 RGB 图像的异

常行为检测方法直接将图像或视频片段作为输入数据,并使用神经网络提取特征对行为建模。目前,主流的基于 RGB 图像的异常行为识别方法主要包括帧重构、帧预测、端对端的异常分数计算和一些综合性方法。

2.3.1 帧重构方法

研究人员认为,在训练阶段先对正常行为建模,在测试阶段则将偏离正常行为模式的行为判定为异常行为,这类在模型训练时只使用正常行为数据的方法称为弱监督的异常行为检测方法^[5]。帧重构的基本假设是:正常行为相比于异常行为更容易重构^[47]。当训练数据只包含正常行为时,正常行为的重构误差较小,而异常行为的重构误差将非常大^[48-50]。重构的对象是视频帧序列中的运动特征^[48],也可以是单个或若干个图像帧^[51-53]。

目前,基于深度学习的帧重构多基于自编码器 (Auto-Encoders, AEs)^[54]。AEs^[54]是一种无监督全连接单隐藏层网络,用于从无标记数据中学习,可用于解决单分类问题。CHONG YongShean 等^[51]将视频帧序列作为输入,使用 AEs 重构输入的帧序列和欧式距离衡量重构误差,通过归一化重构误差进而计算输入帧序列的异常分数。其使用的自动编码器包含用于空间维度编解码的 2D 卷积层,以及用于时间维度编解码的卷积-长短期记忆层 (Convolutional Long Short-term Memory, ConvLSTM),较为有效提取行为的时空特征。在时空特征提取方法的选择上,ZHAO Yiru 等^[53]采用被广泛应用于视频分析的 3D 卷积网络并集成到自动编码器中,其效果要明显优于只使用 2D 卷积的方法。尽管上述方法的效果优于先前基于人工设计特征的方法,但是,GONG Dong 等^[49]指出,AEs 的泛化能力很好,以至于即使是异常行为也可能很好地被重构。如果某些异常行为与用于训练的正常行为有着相似的样式(局部轮廓等),则 AEs 也能有效重构这些异常行为。为解决这一问题,GONG Dong 等^[49]为 AE 添加了一个 Memory 模块,称之为 MemAE 模型。这一模型能够在一定程度上保证正常行为的重构误差较小,而异常行为的重构误差较大。

2.3.2 帧预测方法

LIU Wen 等^[17]指出,由于深度神经网络具有强大的拟合和泛化能力,基于帧重构方法的假设并不总是成立的,即异常行为的重构误差并不总是更大。因此,提出使用视频帧预测的方法检测异常行为。具体而言,对于输入一个视频片段,通过一个预测器

得到预测帧。如果当前输入帧与预测帧相一致,则表示该输入帧不包含异常行为,否则,预测帧与输入帧差异较大的区域可被认定为发生异常行为的区域。LIU Wen 等^[55]在后续的研究中将边缘学习嵌入到帧预测网络中,用于学习更紧凑的正常行为数据分布,并扩大正常行为和异常行为之间的边界。此外,由于异常行为或异常事件是无界的,其提出的方法能更有效地检测从未观察到的异常行为。WANG Xuanzhao 等^[56]提出一种基于多路径卷积的帧预测网络,其能更好地处理不同尺度的语义信息和区域,并捕获正常行为的时空依赖关系。

2.3.3 端对端的异常分数计算方法

帧重构和帧预测方法的共同点在于都会先学习新的数据表示之后再定义异常分数。然而,PANG Guansong 等^[57]认为,当表示学习和异常检测方法分开时将会产生次优甚至是与异常检测方法无关的表示。

SULTANI 等^[19]使用深度多实例排序框架学习异常行为,将只包含正常行为和异常行为的视频视为包,将视频片段视为多实例学习中的实例,并自动学习深度异常排序模型。该模型将计算异常视频片段的异常分数。将异常行为检测作为一个回归问题,希望异常视频片段比正常视频片段具有更高的异常分数。

PANG Guansong 等^[58]提出一种基于自训练有序回归的端到端无监督视频异常检测方法,其在没有手动标记的正常、异常数据下,端对端学习一组视频帧的异常分数。该方法主要包括3个模块:①使用现有的无监督异常检测方法检测没有异常行为的正常帧样本和存在异常行为的异常帧样本;②这些样本被输入到端对端评分模块中用于优化异常评分;③生成相应的新异常分数集,用于更新正常帧样本和异常帧样本。

肖进胜等^[59]以视频输入形成视频包的形式,使用三维卷积核提取视频包的单帧空间特征和多帧时间特征后再作加权融合;使用全连接层压缩特征数量并使用注意力机制对各特征加权处理;利用包级池化操作间视频包级别的特征映射为该视频包的异常分数。

2.3.4 综合性方法

除上述提到的帧重构、帧预测和端对端的异常分数计算方法,一些综合性的方法也被提出。GEORGESCU 等^[60]通过对象级别的自监督和多任务学习实现视频中的异常行为检测,利用一个预先

训练好的检测器用于检测物体,通过联合学习多个代理任务(3个自监督任务和一个基于知识蒸馏的代理任务)来训练3D CNN产生用于判定异常行为的信息。这3个自监督任务分别是:①判断物体是向前/向后移动;②检测连续帧中物体的运动不规则性;③重构物体的外观信息。在推理过程中,通过平均每个任务的预测分数来计算异常分数。

TANG Yao 等^[61]把帧重构和帧预测方法结合在一起,将2个U-Net模块串联起来作为生成器。其网络模型的输入是4个堆叠帧 $I_t, I_{t+1}, I_{t+2}, I_{t+3}$,第1个U-Net网络的输出为中间结果 I_m ;中间帧 I_m 包含未来帧的信息,并将其传递给下一个U-Net网络,用于生成重构帧 \hat{I}_{t+4} 。优化的目标值是最小化 I_m 和 \hat{I}_{t+4} 之间的差异。

2.4 可穿戴传感器法

可穿戴传感器法是指基于可穿戴传感器的异常行为识别方法。加速度计、磁力计、陀螺仪、血氧检测器等传感器可以嵌入到手环或手机等人们可穿戴或便携电子设备。研究人员通过收集这些传感器采集得到的信号数据,对行人的动作或身体健康状态建模,进而识别行人的异常行为^[62-64]。

周晓芳等^[65]使用加速度计、磁力计、陀螺仪等多传感器组成动作感知模块,对收集到的数据使用主成分分析(Principal Component Analysis, PCA)降维,使用SVM分类算法对动作类别分类,实现管控区域内个人异常行为(闯入、脱管、滞留和暴力动作)的识别。KIPRIJANOVSKA 等^[66]提出一种基于腕带装置和深度神经网络的步态异常检测方法,使用集成CNN和双向长短期记忆神经网络的网络学习惯性传感器(Inertial Measurement Unit, IMU)传感器采集得到信号的时空特征,实现对异常步态和跌倒行为的检测。CHAKRABORTY 等^[67]采用2个惯性传感器采集健康儿童和脑瘫儿童的步行步态数据,使用离散小波变换分解收集到的信号,接着使用分解后的信号数据训练多通道一维CNN模型,实现异常步态的检测。

3 人群异常行为数据集及试验性能

3.1 人群异常行为主流数据集

目前,在人群异常行为识别领域,常用的国际公开数据集包括香港中文大学校园人行道异常行为数据集(Chinese University of Hong Kong, CUHK Avenue)^[16],加州大学圣迭戈分校行人异常行为数

据集 (University of California, San Diego, UCSD Peds1, Peds2)^[14], 上海科技大学校园异常行为数据集 (ShanghaiTech Campus)^[17], 明尼苏达大学异常

行为数据集 (University of Minnesota System, UMN) 等。4 个异常行为数据集中的部分异常行为样例如图 5 所示。

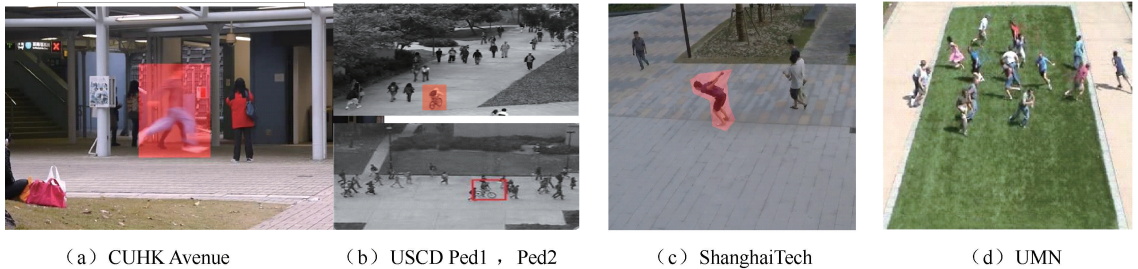


图 5 异常行为数据集中的异常行为样例

Fig. 5 Examples of abnormal behaviors in abnormal behavior datasets

3.2 异常行为识别评价指标

人群异常行为识别评估有帧级标准和像素级标准。帧级标准是指如果一帧图像有一个及以上的异常像素, 则该帧为异常帧; 像素级标准是指检测到的异常行为区域与真实异常行为区域的交集超过预定阈值时才判定为成功检测出异常行为。行人异常行为识别任务常用的性能指标包括曲线下面积 (Area Under Roc Curve, AUC)、等错误率 (Equal Error Rate, EER)、准确率 (Accuracy)、 F_1 分数。其中, AUC 是指接受者操作特性 (Receiver Operating Characteristic, ROC) 下的 AUC, AUC 数值越高代表

检测效果越好; EER 指错误接受率 (False Acceptance Rate, FAR) 和错误拒绝率 (False Rejection Rate, FRR) 相等时的错误率。

3.3 异常行为识别试验性能对比

基于人工设计的特征、人体骨架和 RGB 图像的异常行为识别方法在常见公共数据集上的试验性能及对比见表 2。其中, CUHK 为 CUHK Avenue 数据集; SH. Tech 为 ShanghaiTech Campus 数据集; (f) 表示帧级标准; (p) 表示像素级标准; 未注明 (f) 和 (p) 的 AUC 默认指的是帧级标准; “—” 表示文献中未提及。

表 2 异常行为识别试验性能对比

Tab. 2 Comparison of abnormal behavior recognition performances

类别	方法	年份	数据集	性能/%
人工设计的特征	STVs 描述符+聚类 ^[33]	2013	UCSD Ped1	ERR-15.00% (f), 29.00% (p)
	轨迹特征+稀疏重构 ^[23]	2013	CAVIAR ^[68]	Accuracy-90.42%
	光流特征+轮廓特征+SVM ^[18]	2014	UMN	Accuracy-95.83%
人体骨架	Openpose+ST-GCN ^[44]	2019	自建数据集	Accuracy-100.00%
	Alphapose+ST-GCN ^[45]	2022	自建数据集	Accuracy-98.48%
	Kinect 3D+逻辑回归 ^[46]	2016	自建数据集	—
RGB 图像	帧重构+ST-AEs ^[51]	2017	CUHK, UCSD Ped1, Ped2	AUC-80.30%, 89.90%, 87.40%
	帧重构+MemAE ^[49]	2019	UCSD Ped1, CUHK, SH. Tech	AUC-94.10%, 83.30%, 71.20%
	帧预测+U-Net ^[17]	2018	CUHK, UCSD Ped1, Ped2	AUC-84.90%, 83.1%, 95.40%
	帧预测+边际学习 ^[55]	2019	CUHK, SH. Tech	AUC-92.80%, 76.80%
	端对端+深度多实例排序 ^[19]	2018	自建数据集	AUC-75.41%
	端对端+自训练学习 ^[58]	2020	UCSD Ped1, Ped2, UMN	AUC-71.70%, 83.20%, 97.25%
	自监督+多任务学习 ^[60]	2021	CUHK, UCSD Ped2, SH. Tech	AUC-92.80%, 99.8%, 92.80%
可穿戴传感器	帧重构+帧预测 ^[61]	2020	UCSD Ped1, Ped2, CUHK, SH. Tech	AUC-82.60%, 96.20%, 83.70%, 71.50%
	惯性传感器+SVM ^[65]	2022	自建数据集	F_1 -96.50%
	惯性传感器+CNN ^[67]	2020	自建数据集	Accuracy-96.40%

4 研究不足及展望

公共场所行人异常行为识别方法虽然已在公共数据集上取得了不错的效果,但还存在以下不足:

① 现有识别方法或聚焦于广义的异常行为而不关注异常行为的具体类别,或聚焦于单个类别的异常行为,未能实现多类型、精确的异常行为识别;② 现有算法计算量大,不适用于低成本的实际应用;③ 基于计算机视觉的方法容易受遮挡、光线变化等干扰。为此,提出以下展望:

1) 亟待一个种类丰富、定义清晰和高质量的人群异常行为数据集。目前可供使用的人群异常行为数据集仍面临异常行为定义模糊、异常行为种类稀缺和类别不明确等问题。

2) 多数据融合的异常行为识别方法。在公共场所下,建议采集一定区域内的视频、人群音频、运动学传感器数据,分别识别人类表情、动作、轨迹和音频特征与内容,实现行人表情、人体动作姿态和音频与对话等多维数据融合的识别。

3) 设计遮挡判断模型和反遮挡推算模型。为减少密集人群遮挡带来的干扰,可结合人体平衡条件与关节运动约束条件,设计遮挡判断模型和反遮挡推算模型。

5 结论

1) 在划分的4种行人异常行为识别方法中,人工设计法检测效率高、计算相对简单,但存在难以建模复杂人体行为特征、稳健性弱等问题。

2) 人体骨架法能识别具体的行为类别,但需要精确的姿态估计算法予以配合。

3) RGB图像法充分发挥深度神经网络自动提取图像中高级特征的优势,能够检测更多类型的异常行为,但不能识别具体的行为类别;

4) 可穿戴传感器采集法的优势在于可精确、连续地采集个体姿态信息,但应用成本较高。未来可根据人群遮挡率、行为识别容错率、行为风险程度等要素综合选择对应的识别方法。

参考文献

- [1] KOK V J, LIM M K, CHAN C S. Crowd behavior analysis: a review where physics meets biology [J]. *Neurocomputing*, 2016, 177: 342-362.
- [2] 张萌, 韩豫, 刘泽峰. 深度学习下建筑工人高空安全防护装备检测方法 [J]. *中国安全科学学报*, 2022, 32(5): 140-146.
ZHANG Meng, HAN Yu, LIU Zefeng. Detection method of high-altitude safety protective equipment for construction workers based on deep learning [J]. *China Safety Science Journal*, 2022, 32(5): 140-146.
- [3] 张晓平, 纪佳慧, 王力, 等. 基于视频的人体异常行为识别与检测方法综述 [J]. *控制与决策*, 2022, 37(1): 14-27.
ZHANG Xiaoping, JI Jiahui, WANG Li, et al. Overview of video based human abnormal behavior recognition and detection methods [J]. *Control and Decision*, 2022, 37(1): 14-27.
- [4] 李建更, 谢海征. 基于姿态估计的人体异常行为识别算法 [J]. *北京工业大学学报*, 2022, 48(7): 710-720.
LI Jiangeng, XIE Haizheng. Human abnormal behavior recognition algorithm based on pose estimation [J]. *Journal of Beijing University of Technology*, 2022, 48(7): 710-720.
- [5] 徐涛, 田崇阳, 刘才华. 基于深度学习的人群异常行为检测综述 [J]. *计算机科学*, 2021, 48(9): 125-134.
XU Tao, TIAN Chongyang, LIU Caihua. Deep learning for abnormal crowd behavior detection: a review [J]. *Computer Science*, 2021, 48(9): 125-134.
- [6] 郭毅博, 孟文化, 范一鸣, 等. 基于可穿戴传感器数据的人体行为识别数据特征提取方法 [J]. *计算机辅助设计与图形学学报*, 2021, 33(8): 1 246-1 253.
GUO Yibo, MENG Wenhua, FAN Yiming, et al. Wearable sensor data based human behavior recognition: a method of data feature extraction [J]. *Journal of Computer-Aided Design & Computer Graphics*, 2021, 33(8): 1 246-1 253.
- [7] 张烈平, 匡贞伍, 李昆键, 等. 基于加速度传感器和神经网络的人体活动行为识别 [J]. *现代电子技术*, 2019, 42(16): 71-74.
ZHANG Lieping, KUANG Zhenwu, LI Kunjian, et al. Human activity behavior recognition based on acceleration sensor and neural network [J]. *Modern Electronics Technique*, 2019, 42(16): 71-74.
- [8] 陈波, 余秋婷, 陈铁明. 基于传感器人体行为识别深度学习模型的研究 [J]. *浙江工业大学学报*, 2018, 46(4): 375-381.

- CHEN Bo, YU Qiuting, CHEN Tieming. On deep-learning-model-based sensor activity recognition [J]. Journal of Zhejiang University of Technology, 2018, 46(4): 375-381.
- [9] 艾达, 王倩, 樊炜鑫, 等. 智能手机传感器的人体行为识别技术 [J]. 西安邮电大学学报, 2020, 25(1): 42-48.
AI Da, WANG Qian, FAN Weixin, et al. A review of human behavior recognition based on multi-sensor of smart phone [J]. Journal of Xi'an University of Posts and Telecommunications, 2020, 25(1): 42-48.
- [10] SODEMANN A A, ROSS M P, BORGHETTI B J. A review of anomaly detection in automated surveillance [J]. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, 2012, 42(6): 1 257-1 272.
- [11] POPOOLA O P, KEJUN W. Video-based abnormal human behavior recognition: a review [J]. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, 2012, 42(6): 865-878.
- [12] MABROUK A B, ZAGROUBA E. Abnormal behavior recognition for intelligent video surveillance systems: a review [J]. Expert Systems with Applications, 2018, 91: 480-491.
- [13] DHIMAN C, VISHWAKARMA D K. A review of state-of-the-art techniques for abnormal human activity recognition [J]. Engineering Applications of Artificial Intelligence, 2019, 77: 21-45.
- [14] LI Weixin, MAHADEVAN V, VASCONCELOS N. Anomaly detection and localization in crowded scenes [J]. IEEE Trans Pattern Anal Mach Intell, 2014, 36(1): 18-32.
- [15] 杜鉴豪, 许力. 基于区域光流特征的异常行为检测 [J]. 浙江大学学报:工学版, 2011, 45(7): 1 161-1 166.
DU Jianhao, XU Li. Abnormal behavior detection based on regional optical flow [J]. Journal of Zhejiang University: Engineering Science, 2011, 45(7): 1 161-1 166.
- [16] LU Cewu, SHI Jianping, JIA Jiaya. Abnormal event detection at 150 fps in Matlab [C]. 2013 IEEE International Conference on Computer Vision, 2013: 2 720-2 727.
- [17] LIU Wen, LUO Weixin, LIAN Dongze, et al. Future frame prediction for anomaly detection: a new baseline [C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018: 6 536-6 545.
- [18] CHATHURAMALI K G M, RAMASINGHE S, RODRIGO R. Abnormal activity recognition using spatio-temporal features [C]. The 7th International Conference on Information and Automation for Sustainability, 2014: 1-5.
- [19] SULTANI W, CHEN C, SHAH M. Real-world anomaly detection in surveillance videos [C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018: 18-23.
- [20] LIU Shuo, XUE Hao, XU Chunjie, et al. Abnormal behavior detection based on the motion-changed rules [C]. 2020 15th IEEE International Conference on Signal Processing (ICSP), 2020: 146-149.
- [21] KRATZ L, NISHINO K. Anomaly detection in extremely crowded scenes using spatio-temporal motion pattern models [C]. 2009 IEEE Conference on Computer Vision and Pattern Recognition, 2009: 20-25.
- [22] CONG Yang, YUAN Junsong, LIU Ji. Sparse reconstruction cost for abnormal event detection [C]. 2011 IEEE Conference on Computer Vision and Pattern Recognition, 2011: 20-25.
- [23] LI Ce, HAN Zhenjun, YE Qixiang, et al. Visual abnormal behavior detection based on trajectory sparse reconstruction analysis [J]. Neurocomputing, 2013, 119: 94-100.
- [24] 马露, 裴伟, 朱永英, 等. 基于深度学习的跌倒行为识别 [J]. 计算机科学, 2019, 46(9): 106-112.
MA Lu, PEI Wei, ZHU Yongying, et al. Fall action recognition based on deep learning [J]. Computer Science, 2019, 46(9): 106-112.
- [25] 杨晨晨, 马春梅, 朱金奇. 基于智能手机的跌倒行为识别算法研究 [J]. 计算机工程, 2019, 45(2): 178-183.
YANG Chenchen, MA Chunmei, ZHU Jinqi. Study of fall behavior identification algorithm based on smartphone [J]. Computer Engineering, 2019, 45(2): 178-183.
- [26] 桑海峰, 陈禹, 何大阔. 基于整体特征的人群聚集和奔跑行为检测 [J]. 光电子·激光, 2016, 27(1): 52-60.
SANG Haifeng, CHEN Yu, HE Dakuo, et al. Crowd gathering and running behavior detection based on overall features [J]. Journal of Optoelectronics · Laser, 2016, 27(1): 52-60.
- [27] 何鹏, 陈跃跃, 扈啸. 基于智能手表加速度传感器的人体行为识别 [J]. 电脑与信息技术, 2015, 23(5): 6-8, 33.
HE Peng, CHEN Yueyue, HU Xiao. Human behavior recognition based on the accelerator sensor of intelligent watch [J]. Computer and Information Technology, 2015, 23(5): 6-8, 33.
- [28] 王冰, 康增建, 吕晓军. 铁路客运站视频监控系统中的行人逆行异常事件检测算法研究 [J]. 铁路计算机应用, 2012, 21(4): 19-22.
WANG Bing, KANG Zengjian, LYU Xiaojun. Research on algorithm of abnormal event detection to pedesorian opposing flow in video monitoring system of railway passenger station [J]. Railway Computer Application, 2012, 21(4): 19-22.

- [29] 柏万胜, 孙鹏, 郎宇博, 等. 复杂动态背景下视频中逆行检测技术 [J]. 警察技术, 2023(3): 44-47.
- [30] MO Xuan, MONGA V, BALA R, et al. Adaptive sparse representations for video anomaly detection [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2014, 24(4): 631-645.
- [31] GUO Shuqiang, BAI Qianlong, GAO Song, et al. An analysis method of crowd abnormal behavior for video service robot [J]. IEEE Access, 2019, 7: 169 577-169 585.
- [32] CHANDRAKALA S, DEEPAK K, VIGNESH L K P. Bag-of-Event-Models based embeddings for detecting anomalies in surveillance videos [J]. Expert Systems with Applications, 2022, 190: 1-10.
- [33] ROSHTKHARI M J, LEVINE M D. Online dominant and anomalous behavior detection in videos [C]. 2013 IEEE Conference on Computer Vision and Pattern Recognition, 2013: 2 611-2 618.
- [34] DORETTO G, CHIUSO A, WU Y N, et al. Dynamic textures [J]. International Journal of Computer Vision, 2003, 51: 91-109.
- [35] SINGH K, RAJORA S, VISHWAKARMA D K, et al. Crowd anomaly detection using aggregation of ensembles of fine-tuned ConvNets [J]. Neurocomputing, 2020, 371: 188-98.
- [36] CAO Zhe, SIMON T, WEI Shihen, et al. Realtime multi-person 2d pose estimation using part affinity fields [C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 1 302-1 310.
- [37] FANG Haoshu, LI Jiefeng, TANG Hongyang, et al. AlphaPose: whole-body regional multi-person pose estimation and tracking in real-time [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(6): 7 157-7 173.
- [38] WANG Pichao, LI Wanqing, LI Chuankun, et al. Action recognition based on joint trajectory maps with convolutional neural networks [J]. Knowledge-Based Systems, 2018, 158: 43-53.
- [39] BO Li, DAI Yuchao, CHENG Huahui, et al. Skeleton based action recognition using translation-scale invariant image mapping and multi-scale deep CNN [C]. 2017 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), 2017: 601-604.
- [40] WANG Hongsong, WANG Liang. Learning robust representations using recurrent neural networks for skeleton based action classification and detection [C]. 2017 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), 2017: 591-596.
- [41] YAN Sijie, XIONG Yuanjun, LIN Dahua. Spatial temporal graph convolutional networks for skeleton-based action recognition [C]. AAAI Conference on Artificial Intelligence, 2018: 7 444-7 452.
- [42] YANG C L, SETYOKO A, TAMPUBOLON H, et al. Pairwise adjacency matrix on spatial temporal graph convolution network for skeleton-based two-person interaction recognition [C]. 2020 IEEE International Conference on Image Processing (ICIP), 2020: 2 166-2 170.
- [43] FENG Dong, WU Zhongcheng, ZHANG Jun, et al. Multi-scale spatial temporal graph neural network for skeleton-based action recognition [J]. IEEE Access, 2021, 9: 58 256-58 265.
- [44] ZHENG Yin, ZHANG Dongping, YANG Li, et al. Fall detection and recognition based on GCN and 2D Pose [C]. 2019 6th International Conference on Systems and Informatics (ICSAI), 2019: 558-562.
- [45] 刘耀, 焦双健. St-Gcn 在建筑工人不安全动作识别中的应用 [J]. 中国安全科学学报, 2022, 32(4): 30-35.
LIU Yao, JIAO Shuangjian. Application of ST-GCN in unsafe action identification of construction workers [J]. China Safety Science Journal, 2022, 32(4): 30-35.
- [46] NAR R, SINGAL A, KUMAR P. Abnormal activity detection for bank ATM surveillance [C]. 2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI), 2016: 2 042-2 046.
- [47] PANG Guansong, SHEN Chunhua, CAO Longbing, et al. Deep learning for anomaly detection [J]. ACM Computing Surveys, 2021, 54(2): 1-38.
- [48] HASAN M, CHOI J, NEUMANN J, et al. Learning temporal regularity in video sequences [C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 733-742.
- [49] GONG Dong, LIU Lingqiao, LE Vuong, et al. Memorizing normality to detect anomaly: memory-augmented deep autoencoder for unsupervised anomaly detection [C]. 2019 IEEE/CVF International Conference on Computer Vision (ICCV), 2019: 1 705-1 714.
- [50] LI Jing, HUANG Qingwang, DU Yingjun, et al. Variational abnormal behavior detection with motion consistency [J]. IEEE Trans Image Process, 2022, 31: 275-286.
- [51] CHONG Y S, TAY Y H. Abnormal event detection in videos using spatiotemporal autoencoder [M]. Berlin: Advances in Neural Networks, 2017: 189-196.

- [52] RAVANBAKSH M, NABI M, SANGINETO E, et al. Abnormal event detection in videos using generative adversarial nets [C]. 2017 IEEE International Conference on Image Processing (ICIP), 2017: 1 577–1 581.
- [53] ZHAO Yiru, DENG Bing, SHEN Chen, et al. Spatio-temporal autoencoder for video anomaly detection [C]. 25th ACM International Conference on Multimedia, 2017: 1 933–1 941.
- [54] RUMELHART D E, HINTON G E, WILLIAMS R J. Learning representations by back-propagating errors [J]. Nature, 1986, 323: 533–536.
- [55] LIU Wen, LUO Weixin, LI Zhengxin, et al. Margin learning embedded prediction for video anomaly detection with a few anomalies [C]. International Joint Conference on Artificial Intelligence, 2019: 3 023–3 030.
- [56] WANG Xuanzhao, CHE Zhengping, JIANG Bo, et al. Robust unsupervised video anomaly detection by multipath frame prediction [J]. IEEE Trans Neural Netw Learn Syst, 2022, 33(6): 2 301–2 312.
- [57] PANG Guansong, SHEN Chunhua, VAN DEN HENGEL A. Deep anomaly detection with deviation networks [C]. 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2019: 353–362.
- [58] PANG Guansong, YAN Cheng, SHEN Chunhua, et al. Self-trained deep ordinal regression for end-to-end video anomaly detection [C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020: 12 170–12 179.
- [59] 肖进胜, 申梦瑶, 江明俊, 等. 融合包注意力机制的监控视频异常行为检测 [J]. 自动化学报, 2022, 48(12): 2 951–2 959.
- XIAO Jinsheng, SHEN Mengyao, JIANG Mingjun, et al. Abnormal behavior detection algorithm with video-bag attention mechanism in surveillance video [J]. Acta Automatica Sinica, 2022, 48(12): 2 951–2 959.
- [60] GEORGESCU M I, BĂRBĂLĂU A, IONESCU R T, et al. Anomaly detection in video via self-supervised and multi-task learning [C]. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021: 12 737–12 747.
- [61] TANG Yao, ZHAO Lin, ZHANG Shanshan, et al. Integrating prediction and reconstruction for anomaly detection [J]. Pattern Recognition Letters, 2020, 129: 123–130.
- [62] SABOOR A, KASK T, KUUSIK A, et al. Latest research trends in gait analysis using wearable sensors and machine learning: a systematic review [J]. IEEE Access, 2020, 8: 167 830–167 864.
- [63] WU Donghui, WANG Zhelong, CHEN Ye, et al. Mixed-kernel based weighted extreme learning machine for inertial sensor based human activity recognition with imbalanced dataset [J]. Neurocomputing, 2016, 190: 35–49.
- [64] MOHAMMADIAN R N, VAN LAARHOVEN T, FURLANELLO C, et al. Novelty detection using deep normative modeling for imu-based abnormal movement monitoring in parkinson's disease and autism spectrum disorders [J]. Sensors (Basel), 2018, 18(10): 3 533–3 550.
- [65] 周晓芳, 吴松洋, 韩玮, 等. 管控区域人员异常行为检测系统与方法研究 [J]. 计算机应用与软件, 2022, 39(5): 92–97.
- ZHOU Xiaofang, WU Songyang, HAN Wei, et al. Detection system and method for abnormal behavior of personnel in control areas [J]. Computer Applications and Software, 2022, 39(5): 92–97.
- [66] KIPRIJANOVSKA I, GJORESKE H, GAMS M. Detection of gait abnormalities for fall risk assessment using wrist-worn inertial sensors and deep learning [J]. Sensors (Basel), 2020, 20(18): 5 373–5 394.
- [67] CHAKRABORTY J, NANDY A. Discrete wavelet transform based data representation in deep neural network for gait abnormality detection [J]. Biomedical Signal Processing and Control, 2020, 62: 102 076–102 088.
- [68] CAVIAR. Context aware vision using image-based active recognition [EB/OL]. (2005-09-30). <https://homepages.inf.ed.ac.uk/rbf/CAVIAR/>.

作者简介: 赵荣泳 (1976—), 男, 山东济南人, 博士, 副教授, 主要从事公共安全系统工程和复杂系统优化等方面的研究。E-mail: zhaorongyong@tongji.edu.cn。

