

基于新颖性检测的跌倒风险预测^①



刘雅秦¹, 叶宁^{1,2}, 徐康^{1,2}, 王汝传^{1,2}, 唐震³

¹(南京邮电大学 计算机学院、软件学院、网络空间安全学院, 南京 210023)

²(江苏省无线传感网高技术研究重点实验室, 南京 210023)

³(南京邮电大学 材料科学与工程学院, 南京 210023)

通信作者: 刘雅秦, E-mail: lyq645626379@163.com

摘要: 跌倒是 65 岁及以上人群因伤害致死的第一位原因。结合受试者个体信息的个性化特征, 提出一种基于 Kinect 三维骨架数据的步态特征提取方法, 对老年人的跌倒风险进行评估和预测。将跌倒风险分为高跌倒风险和低跌倒风险两类, 考虑数据采集的成本问题, 采用新颖性检测模型在不平衡数据集下对特征数据进行训练和评估。实验结果表明, OC-SVM (one-class SVM) 检测准确率达 86.96%, *F1-score* 为 88.55%, 能够有效地区分低跌倒风险受试者和高跌倒风险受试者。同时, 证明了基于 Kinect 三维骨架数据预测老年人跌倒风险的潜力。

关键词: 跌倒风险; 新颖性检测; Kinect; 步态分析; 单类支持向量机

引用格式: 刘雅秦, 叶宁, 徐康, 王汝传, 唐震. 基于新颖性检测的跌倒风险预测. 计算机系统应用, 2022, 31(4): 163-170. <http://www.c-s-a.org.cn/1003-3254/8400.html>

Falls Risk Prediction Based on Novelty Detection

LIU Ya-Qin¹, YE Ning^{1,2}, XU Kang^{1,2}, WANG Ru-Chuan^{1,2}, TANG Zhen³

¹(School of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing 210023, China)

²(Jiangsu High Technology Research Key Laboratory for Wireless Sensor Networks, Nanjing 210023, China)

³(Institute of Advanced Materials, Nanjing University of Posts and Telecommunications, Nanjing 210023, China)

Abstract: Falls are the first cause of injury-related deaths in people over the age of 65. A gait feature extraction method based on Kinect 3D skeleton data is proposed. This method can assess and predict the falls risks of the elderly according to the personalized features of the individual information of the subjects. The falls risks are divided into two classes: high falls risks and low falls risks. Considering the cost of data collection, the novelty detection model is used to train and access the feature data on an unbalanced data set. The experimental results show that the accuracy of one-class support vector machine (OC-SVM) detection is 86.96% and the *F1-score* is 88.55%, which means the proposed method can effectively distinguish subjects with low falls risks from those with high falls risks. These results also demonstrate the potential of predicting the falls risks of the elderly with Kinect 3D skeleton data.

Key words: falls risk; novelty detection; Kinect; gait analysis; one-class support vector machine (OC-SVM)

跌倒是老年人健康风险占比最大的事故。根据世界卫生组织的报告, 全球每年约有 28%–35% 的 65 岁以上的老年人跌倒, 而 70 岁以上的老年人跌倒比例增加至 32%–42%^[1]。《中国伤害预防报告》显示: 老年人

跌倒发生率为 20.7%^[2], 跌倒是 65 岁及以上人群因伤害致死的第一位原因^[3]。对个人而言, 跌倒使老年人机能衰退, 活动能力下降, 出现致伤甚至致残; 对社会来说, 跌倒会增加医疗支出, 耗费社会资源^[4]。因此, 开展

① 基金项目: 江苏省科技厅重点研发计划 (社会发展)(BE2020713)

收稿时间: 2021-06-04; 修改时间: 2021-07-07; 采用时间: 2021-07-13; csa 在线出版时间: 2022-03-22

老年人群跌倒风险的评估和预测工作是必要的。

人的步态参数包含个人身体活动能力的重要信息,老年人由于生理机能下降及各种神经系统、肌肉骨骼疾病、药物、心理等多种因素的综合作用,步态经常出现各种复杂的问题^[5]。并且步态和平衡问题可以量化和远程评估,因此跌倒风险可通过测量老年人的步态变化来检测。

目前,国内外已经提出了许多基于不同传感器用于步态评估的技术方法。常见的方法之一是使用可穿戴传感器,如加速度计、陀螺仪和惯性传感器^[6-8]。这些设备大多能够提供准确的步态参数,具备小型化、高灵敏度和低成本等特征,可实时地检测老年的活动状态^[9]。但是,这些传感器需要技术人员来帮助受试者佩戴以及经常性的维护,它们日常的放置需保证正确和安全。用于运动跟踪的光学系统(如 Vicon 系统),其基于固定在对象身体某些关键点上的标记。另外,存在无需穿戴的环境传感器,例如,测力平台、活动地毯(如 GAITRite),当人们在上面行走时,它们可以检测双脚与地面接触的准确时刻,并计算速度和步长^[10]。

以上这些设备及系统必须安装在特定的场所中(如医疗机构和实验室),并且只能由专业人员操作,不便于对老年人的步态状况进行长期和连续的监测与评估。因此,显然需要一种成本低并且易于使用的系统,该系统允许在居家和社区环境下对步态模式进行连续的定量分析。

RGB-D 相机具有低成本和对可见光不变性的优点,由于基于 RGB-D 相机的系统始终处于开启状态,因此无需受试者佩戴和充电维护。使用 Kinect 传感器进行身体活动能力评估首次出现在文献^[11]中,在这项工作中,评估了使用微软 Kinect 获得时间和空间步态参数的测量结果,并与现有的基于网络摄像头的系统以及 Vicon 运动捕捉系统进行了比较。实验结果表明, Kinect 传感器测量步态参数具有足够的精度。

新颖性检测使用单一类别的数据构建预测模型,它属于异常检测领域,是指在新数据中发现不符合预期行为模式数据的问题。

大多数新颖性检测方法都是针对特定用例的,需要了解该方法适用的场景以及相关领域的专家知识^[12],具体可分为 3 大类别。一是经典的统计或密度估计方法:对“正态”类别进行建模,以确定测试样本是否来自

相同的分布;二是将训练集中样本点与特征空间中所有其他可能的未知数据点分开,如对训练数据集中的各类构建最小球体,测试集数据不落入任何球体即视为新颖点;三是基于 encode-decode 思想,通过压缩数据表示和测量输入的重构误差来检测新颖性。

近年来,深度学习模型在解决不同领域具有挑战性检测问题中的表现引人注目,其通过神经网络学习特征表示或异常分数进行新颖性检测^[13]。

异常步态数据属于异常数据中的模式异常或称集体异常。在先前的研究工作中,已经将新颖性检测应用于步态病理学的分析^[14-16],特别地,在文献^[16]中使用智能手机的惯性步态数据,用于训练和评估单类支持向量机(OC-SVM)和支持向量数据描述(SVDD)这两种新颖性检测分类器,结果证明了使用新颖性检测从智能手机数据中预测跌倒风险的潜力,但该项工作并未对单类支持向量机和其他不同类别新颖性检测模型的预测性能进行比较。

本文提出了一种步态特征提取方法,对基于 Kinect 获得的三维骨架数据进行时空和运动学分析,提取出相应的步态参数特征,并结合受试者个体信息生成的个性化特征,作为特征集合共同输入到分类器中训练以预测跌倒风险。同时,考虑到数据收集工作中可能遇到的样本不平衡问题,即低跌倒风险人群的数量可能远超过高跌倒风险人群,使用异常检测领域的新颖性检测模型单类支持向量机(OC-SVM)对提取的特征集合进行训练和评估。

1 基于 Kinect 跌倒风险预测方法流程

1.1 实验设置

本实验使用的深度相机为美国微软公司研发的 Kinect v2,其拍摄出的 RGB-D 行为数据主要包含 3 种模式:深度数据、RGB 数据和三维骨架数据。由于 Kinect 获取的数据包含比较多的噪声,相比基于深度信息数据的方法,基于三维骨架的方法信息根据准确,还可省去轮廓线提取等过程带来的错误和误差^[17],因此本文使用三维骨架数据。同时,相比基于 RGB 数据的方法,使用三维骨架数据可保护用户的隐私,未来更适合应用于居家和社区场景。

如图 1 所示,在实验场景中, Kinect 被放置在离地面 0.8 m 的高度,考虑到 Kinect 探测距离的限制,受试

者开始距设备 3.5 m, 要求沿直线行走 3 m, 行走结束时距离设备 0.5 m.

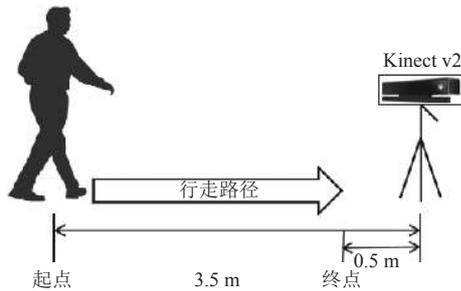


图1 实验数据采集设置

1.2 跌倒风险预测方法流程

如图2所示, 本文跌倒风险预测方法流程为: 将采集到的原始三维骨架数据经数据去噪处理后, 通过步态特征提取方法提取出受试者的步态参数, 结合受试者自身个性化特征共同输入到跌倒风险预测模型中, 最后输出受试者的风险等级(高/低).

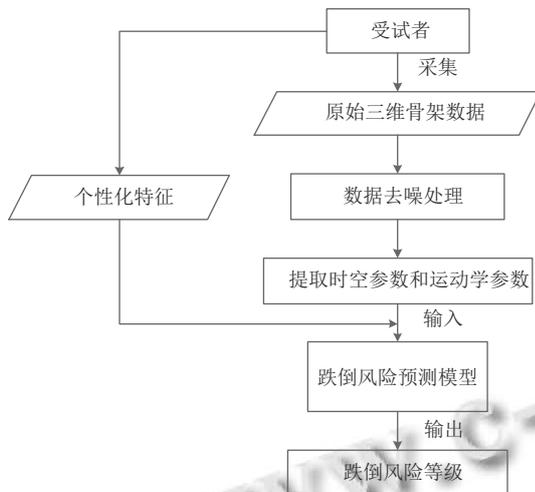


图2 跌倒风险预测方法流程图

其中, 提出的基于三维骨架数据的步态特征提取方法可提取出相应的时空参数和运动学参数如步幅、步速、膝关节角度等; 并且引起跌倒的因素众多, 因此本文也结合了受试者的个体信息如病史、跌倒史等作为共同输入. 具体内容在第2节介绍.

考虑到高跌倒风险样本不易获得的情况, 由于新颖性检测器不需要平衡数据集, 提出的跌倒风险预测模型使用新颖性检测领域中的单类支持向量机(OC-SVM). 具体内容在第3节介绍.

2 基于三维骨架数据的步态特征提取方法

2.1 三维骨架数据去噪

由于 Kinect 测量的关节点存在噪声, 所以需要获取的三维骨架数据进行去噪处理, 提高测量的精度.

首先, 需要去除原始样本数据时间序列中的高频扰动, 应用一次移动平均滤波 (single moving average, SMA). 一次移动平均的本质是一种低通滤波, 对时间序列直接求等权重均值. 本文将每段三维骨架步态样本中关节点的某一方向 ($x/y/z$) 信号看作一组一维时间序列, 其中, n 为参与 SMA 的帧数的个数, 则第 $t+1$ 帧信号 SMA 后的值 S_{t+1} 计算公式为式 (1):

$$S_{t+1} = \frac{S_t + S_{t-1} + \dots + S_{t-n+1}}{n} \quad (1)$$

同时, 为了更准确地获取步态时空参数, 在上述一次移动平均滤波的基础上, 对足跟关节 15 和 19 的 z 轴方向数据应用中值滤波器. 中值滤波器在图像处理领域应用广泛, 其基本思想是, 如果一段信号是平缓变化的, 那么某一点的输出值可以用该点某个窗口大小内所有值的中值来代替. Kinect 三维骨架数据布局如图3所示.

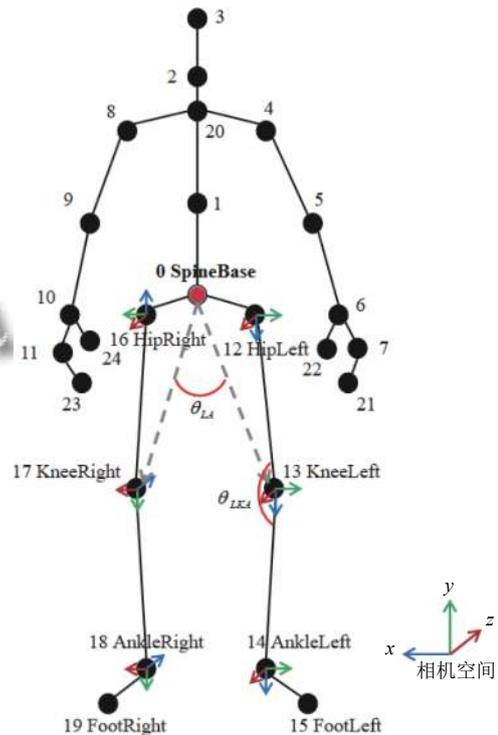


图3 Kinect 三维骨架数据的布局

去噪处理后效果如图4和图5所示(图4为左膝关节角度原始数据和滤波后的效果; 图5为 $z_{15} - z_{19}$ 原始数据和滤波后的效果).

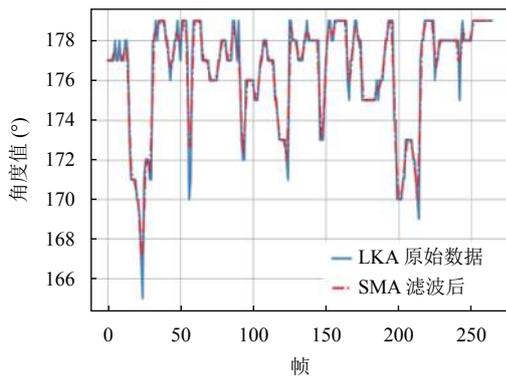


图4 左膝关节角度原始数据和滤波后效果对比

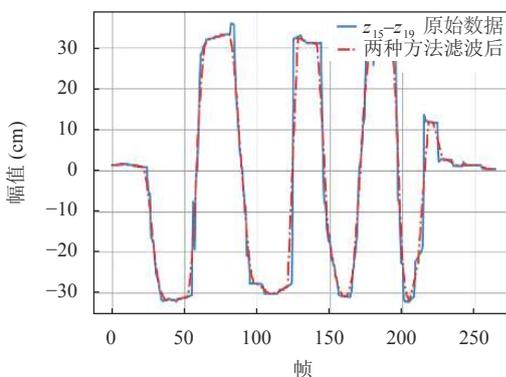


图5 z₁₅-z₁₉原始数据和滤波后效果对比

2.2 步态参数提取

若通过步态分析预测跌倒风险,适当地选择所需的步态参数是必要的.本文将所需的步态参数分为两类:时空参数和运动学参数.

时空参数包括步幅、步速、步频.步幅指一足着地至同一足再次着地的距离.对于每个骨架数据帧 t ,Kinect提供一个时间戳.为了计算步幅,本文使用两个足跟关节15和19的 z 坐标差值^[18],如图6所示,同时这两个坐标差值的极值给出了每一步的开始帧 t_s 和结束帧 t_f .步幅 $stride$ 的计算公式为式(2):

$$stride = |z_{15} - z_{19}| \quad (2)$$

步速为一个步幅周期($t_s \sim t_f$)的平均速度,是预测老年人跌倒最常用的步态参数.若Kinect v2以 u fps的速度更新数据,则每个数据帧 t 间隔时间为 $\frac{1}{u}$ s.步速 $speed$ 的计算公式为式(3):

$$speed = \frac{|z_{15} - z_{19}|}{t_f - t_s} \quad (3)$$

步频同样也对老年人跌倒有预测作用,步频由步幅决定,在本文中计算一个行走样本数据内的平均步

频.若一个样本中的步幅数为 m ,骨架数据帧数为 n ,则步频 $cadence$ 的计算公式为式(4):

$$cadence = \frac{m}{n \times \frac{1}{u}} \quad (4)$$

运动学参数是反映行走过程中下肢运动和平衡能力的指标,包括左右膝关节角度和两腿间角度.在Kinect获取到关节的三维坐标点后,将骨骼看作两个关节点间的直线,通过余弦定理可计算关节间角度.

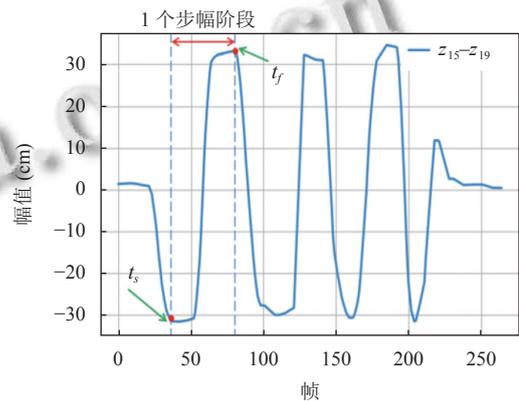


图6 行走阶段的步态样本步幅分割示例

膝关节角度定义为每条腿的大腿和小腿间的夹角.左膝关节角度LKA(left knee angle)为连接左腿关节12至13及13至14的两个向量之间的角度,右膝关节角度RKA(right knee angle)为连接右腿关节16至17及17至18的两个向量之间的角度.

两腿间的角度LA(leg angle)定义为连接关节0至13及0至17的两个向量之间的角度.

以左膝关节角度为例,如图3所示,计算公式为式(5):

$$\theta_{LKA} = \arccos \frac{L_{12,13}^2 + L_{13,14}^2 - L_{12,14}^2}{2L_{12,13} \cdot L_{13,14}} \quad (5)$$

其中, $L_{i,j}$ 为关节 i 和 j 间的长度, $L_{i,j} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$.

为了能表示受试者的平衡能力变化的稳定性,将每一步幅范围中的左膝关节角度值取方差,则得到特征值 $var(LKA)$.

2.3 个性化特征

本文重点关注的是基于老年人的步态特征预测跌倒风险,但受试者的自身因素也因适当考虑.以性别因素为例,根据中国健康与养老追踪调查(CHARLS)2015年数据,研究发现无论在城镇还是农村地区,无论在哪个年龄段,中国老年女性的跌倒发生率始终显著高于老年男性^[19].同时,跌倒史被认为与跌倒风险增强

有很强的联系^[20]。

参考文献 [21] 提出使用三维骨骼关节时域特征和元数据特征结合预测跌倒风险等级的方法, 本文选择性别 (“1”和“0”分别代表男性和女性)、是否被诊断具有某些与平衡相关的损伤 (“1”和“0”分别代表是和否) 以及过去 6 个月是否存在跌倒史 (“1”和“0”分别代表是和否) 这 3 个二值特征, 和步态特征共同作为输入。表 1 列出了本实验用于预测跌倒风险的特征。

表 1 预测跌倒风险特征列表

特征类型	特征名称
步态特征	步幅 <i>stride</i>
	步速 <i>speed</i>
	步频 <i>cadence</i>
	左膝关节角度方差值 <i>var(LKA)</i>
	右膝关节角度方差值 <i>var(RKA)</i>
	两腿间角度方差值 <i>var(LA)</i>
个性化特征	性别 <i>gender</i>
	是否被诊断具有某些与平衡相关的损伤 <i>diseases</i>
	过去6个月是否存在跌倒史 <i>fall history</i>

3 基于新颖性检测的跌倒风险预测模型

3.1 新颖性检测

新颖性检测是典型的一分类问题, 指对大量正常数据中包含异常值或未预见值的自动识别。新颖性检测在很多领域有着广泛的实际应用, 如医学诊断、工业系统故障检测、电子安全系统入侵检测、文本挖掘等^[22]。

新颖性检测的实现方法包括频率论方法、贝叶斯方法、信息论、极值统计、支持向量法、其他核方法以及神经网络。一般使用训练数据构建模型, 训练数据可包含非常少的新颖点 (异常点) 或不包含新颖点。根据算法得出数据 x 的新颖性分数 $z(x)$, 并根据通常称为新颖性阈值 $z(x) = k$ 的决策边界来检测其偏离正态性的情况。本文选择单类支持向量机训练和评估基于第 2 节特征提取方法生成的特征数据。

考虑到未来面向社会开展老年人跌倒风险预测实际工作的初期, 进行步态样本数据采集时, 由于高跌倒风险样本不易获得, 为降低数据收集的成本, 低跌倒风险的样本数量可能远超过高跌倒风险数量。可将低跌倒风险样本看作正常数据, 高跌倒风险样本看作异常数据。由于新颖性检测算法中训练数据允许包含极少

量异常数据, 对上述不平衡数据情况下进行模型的构建与训练是适用且有效的。

3.2 单类支持向量机

本节主要介绍单类支持向量机 (one-class SVM, OC-SVM) 的基本概念。

根据文献 [23] 中的描述, OC-SVM 的原理是将特征空间中的原点与数据点分离开, 以最大化间隔作为优化目标。该算法返回一个二元决策函数 $f(x)$, 能够获取特征空间中数据的概率密度区域。 $f(x)$ 在捕获大部分数据点的区域中取值+1, 在其他地方取值-1。

给定训练集 $X = \{x_1, x_2, \dots, x_N\} \subseteq R^n$, N 为训练集样本数量。为了将训练数据集与原点分离, 并且与原点的距离最大, 需求解二次规划问题 (6):

$$\begin{cases} \min_{w, \rho, \zeta} & \frac{1}{2} w^T w + \frac{1}{vN} \sum_{i=1}^N \zeta_i - \rho \\ \text{s.t.} & w^T \cdot \Phi(x_i) \geq \rho - \zeta_i, \zeta_i \geq 0, i = 1, \dots, N \end{cases} \quad (6)$$

其中, ζ_i 为松弛变量, $\Phi(x_i)$ 为原始训练样本 x_i 在特征空间 Φ 中的投影, w 和 ρ 表征超平面。

引入拉格朗日乘数 $[\alpha_1, \dots, \alpha_N]^T$ 后, 决策函数 $f(x) = \text{sgn}(w^T \cdot \Phi(x) - \rho)$ 变为 $f(x) = \text{sgn}(\sum_{i=1}^N \alpha_i k(x_i, x) - \rho)$ 。这里 $k(x_i, x) = (\Phi(x_i) \cdot \Phi(x))$ 是一个核函数, 在实际应用中, 经常使用的是 RBF 核函数 $k(x_i, x) = \exp(-\gamma \|x_i - x\|^2)$ 。

此时, 需求解对偶问题 (7):

$$\begin{cases} \min_{\alpha} & \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j k(x_i, x_j) \\ \text{s.t.} & 0 \leq \alpha_i \leq \frac{1}{vN}, \sum_{i=1}^N \alpha_i = 1 \end{cases} \quad (7)$$

需要注意参数 $v \in (0, 1)$, 当其趋近于 0 时为训练数据集中作为支持向量的样本数量的下限, 当其趋近于 1 时为训练数据中新颖性分数的上限。也就是说, v 的值直接决定了单类支持向量机新颖性检测算法的灵敏度。

4 实验与分析

4.1 数据采集

数据采集和开发基于 Visual Studio 2019 和 JetBrains PyCharm Community Edition 2017。在实验室条件下, Kinect v2 以 30 fps 的速度进行数据更新。

共 12 名志愿者参与实验步态数据的采集工作, 分别模拟不同跌倒风险的老年人步态, 其中男性 9 名, 女

性3名。在每次实验中,受试者距设备3.5 m,沿直线行走3 m,每名受试者重复6~8次。在12名受试者中,4名被标记为有高跌倒风险(女性1名,男性3名),8名被标记为有低跌倒风险(女性2名,男性6名),具体信息见表2。

表2 参与实验的12名受试者个性化特征信息及跌倒风险等级

受试者ID	性别 <i>gender</i>	是否被诊断具有某些 与平衡相关的损伤 <i>diseases</i>	过去6个月 是否存在 跌倒史 <i>fall history</i>	跌倒风险 等级
1	男	无	有	高
2	女	有	有	高
3	男	有	有	高
4	男	有	无	高
5	女	无	无	低
6	男	无	无	低
7	男	有	无	低
8	男	无	无	低
9	男	有	无	低
10	男	无	无	低
11	男	无	无	低
12	女	有	无	低

总体而言,实验数据集包含60个标记为低跌倒风险的原始步态片段和29个标记为高跌倒风险的原始步态片段。

如图7所示,采用第2节的步态特征提取方法对三维骨架数据集进行预处理及特征提取后,基于滑动窗口算法对每个受试者单次行走样本进行再划分,时间连续的步幅被分组以形成包含1个步幅的数据窗口,按时间顺序排列,滑动步长为1。这样,可以从单个对象的数据中提取多个数据窗口,以生成多个样本。最终生成338条低跌倒风险样本,236条高跌倒风险样本。

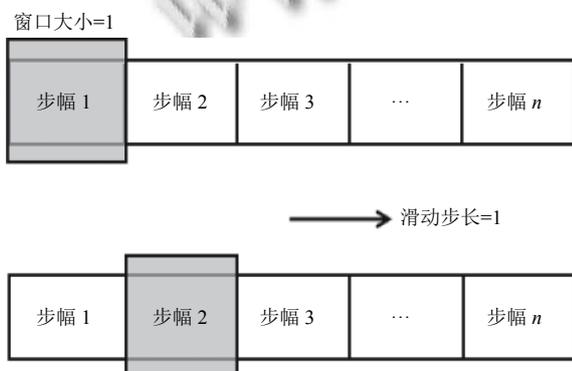


图7 基于滑动窗口为1的步态样本数据扩充

这种数据扩充方法的优点是^[24]:它在数据中引入了个体对象的变化,这有助于提高机器学习分类器的泛化性能;该方法允许频繁和实时的推断,这对于跌倒风险预测中的及时干预十分重要。

4.2 不平衡数据集下的对比实验

在文献[16]的数据集划分方法基础上,本文的实验数据集划分如下:随机选取80%的低跌倒风险样本放入训练数据中,同时随机选取与训练集中低跌倒风险样本比例为1:90的高跌倒风险样本;剩余20%的低跌倒风险样本用于构建测试集的一部分,而剩下部分由随机选取的百分比与完整数据集百分比相同的高跌倒风险样本构成。

K近邻(K-nearest neighbors, KNN)算法和神经网络在异常检测领域都有相应的研究和应用,神经网络方法可处理非线性和非高斯数据,KNN通过计算测试样本与其邻近样本之间的相似度来检测测试样本是否异常。

孤立森林(isolation forest, iForest)是一种基于决策树的分类算法,已被证实在大部分工业异常检测场景中效果较好。

因此,本文选择K近邻(KNN)算法和多层感知机(multi-layer perceptron, MLP)、孤立森林(iForest)和OC-SVM进行对照实验。

实验使用了超参数空间上的网格搜索来执行超参数优化。为了提高模型的稳健性和避免过拟合,使用了K折交叉验证法(K=5),找到最优超参数后,在训练集上重新训练模型,并使用独立测试集对模型性能做出最终的评价。

本节使用的模型及方法均通过Scikit-learn库^[25]实现。

4.3 实验结果分析

新颖性检测技术的有效性可以根据有多少新数据点被正确识别以及有多少正常数据被错误地分类为新数据来评估^[22]。并且,需考虑到在使用不平衡数据集的实验下,正样本数量远远超过负样本,模型分类准确度并不能完全衡量和评价模型的好坏,更合适的指标是F1-score,其为综合考虑精确率(Precision)和召回率(Recall)的调和值。

实验使用正确率(ACC)、虚警率(FPR)和F1-score评估模型的检测性能,定义如式(8):

$$\begin{cases} ACC = \frac{t_p + t_n}{t_p + t_n + f_p + f_n} \\ FPR = \frac{f_p}{t_n + f_p} \\ F1-score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \end{cases} \quad (8)$$

其中,

$$Precision = \frac{t_p}{t_p + f_p}, \quad Recall = \frac{t_p}{t_p + f_n} \quad (9)$$

其中, t_p , t_n , f_p , f_n 分别代表真阳性、真阴性、假阳性、假阴性的样本数量。

神经网络算法具有随机性, 使用同样的数据训练同一个网络会得到不同的结果. 孤立森林算法和随机森林算法类似, 每次选择划分属性和划分点(值)时都是随机的。

因此对 MLP 及 iForest 的训练和测试过程完整重复 10 次, 其评价标准的结果均取平均值. 模型性能评估结果如表 3 所示。

表 3 不平衡数据集下分类器性能评估结果

分类器	ACC	FPR	F1-score
MLP	0.7731	0.5468	0.8394
KNN	0.8	0.4894	0.8553
OC-SVM	0.8696	0.1064	0.8855
iForest	0.8435	0.1253	0.8293

新颖性检测技术应该以高检测率为目标, 并且需要保持低虚警率. 接收者操作特征 (ROC) 曲线通常用于表示检测率和虚警率间的权衡. 所有 4 种分类器的 ROC 曲线如图 8 所示. 在上述模型中, OC-SVM 检测的准确率达 86.96%, 虚警率为 10.64%, 其 $F1-score$ 为 88.55%, AUC 为 0.96, 表现最佳. 结果表明, OC-SVM 可用于不平衡数据集下步态数据的跌倒风险预测, 并在中小型数据集中有较好的表现力. 同时, 证明了基于 Kinect 三维骨架数据预测老年人跌倒风险的潜力。

4.4 个性化特征的消融实验

为了验证引入个性化特征对跌倒风险预测的作用, 本节进行消融实验。

去除个性化特征, 仅选择数据集的跌倒风险特征(表 1) 中 5 个步态特征对 4 种模型进行训练和测试, 模型性能评估结果如表 4 所示。

由表 4 可知, OC-SVM 在消融实验中预测跌倒风险的能力仍是最好的. 去除个性化特征后, 大部分模型的准确率及 $F1-score$ 下降, 虚警率上升, 影响较大的

为 MLP 和 iForest.

综上, 证明引入个性化特征对于提升跌倒风险预测方法的精度和鲁棒性有积极作用, 为未来增加受试者个体信息特征及进一步构建跌倒风险预测模型提供参考。

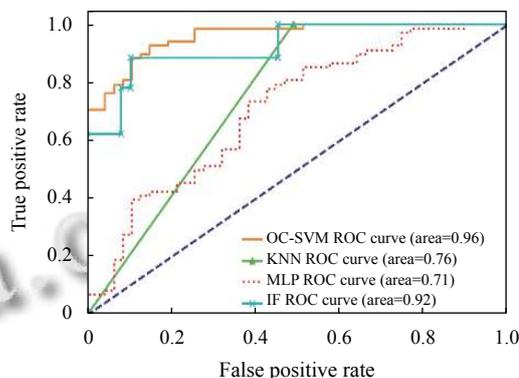


图 8 4 种分类器 ROC 曲线比较

表 4 消融实验下分类器性能评估结果

分类器	ACC	FPR	F1-score
MLP	0.7304	0.6382	0.8121
KNN	0.8	0.4894	0.8553
OC-SVM	0.8503	0.1157	0.8735
iForest	0.8174	0.1335	0.8154

5 结束语

本文提出了一种基于 Kinect 三维骨架步态数据的特征提取方法, 并结合老年人自身个性化特征以预测跌倒风险. 同时, 考虑到数据采集工作的现实情况, 采用新颖性检测模型 OC-SVM 在不平衡数据集下进行数据的训练与评估, 获得了良好的表现, 有效地区分出低跌倒风险受试者和高跌倒风险受试者。

在今后的工作中: 考虑增加更多的个性化特征用于跌倒风险评估; 在更多数据的支持下, 提出平衡数据集下的跌倒风险预测模型, 以此提高模型在不同数据集下对老年人跌倒风险预测的准确度和稳健性。

参考文献

- United Nations. World population prospects: The 2015 revision: Key findings & advance tables. New York: United Nations, 2015.
- 卫生部疾病预防控制局, 卫生部统计信息中心, 中国疾病预防控制中心. 中国伤害预防报告. 北京: 人民卫生出版社, 2007.

- 3 中国疾病预防控制中心慢性非传染性疾病预防控制中心, 国家卫生健康委员会统计信息中心. 中国死因监测数据集 2018. 北京: 中国科学技术出版社, 2019.
- 4 丁志宏, 杜书然, 王明鑫. 我国城市老年人跌倒状况及其影响因素研究. 人口与发展, 2018, 24(4): 120–128.
- 5 王莉, 于卫华. 步态分析在老年人跌倒中的应用进展. 中华护理杂志, 2016, 51(3): 347–351. [doi: [10.3761/j.issn.0254-1769.2016.03.019](https://doi.org/10.3761/j.issn.0254-1769.2016.03.019)]
- 6 Culhane KM, O'Connor M, Lyons D, *et al.* Accelerometers in rehabilitation medicine for older adults. *Age and Ageing*, 2005, 34(6): 556–560. [doi: [10.1093/ageing/afi192](https://doi.org/10.1093/ageing/afi192)]
- 7 Greene BR, O'Donovan A, Romero-Ortuno R, *et al.* Quantitative falls risk assessment using the timed up and go test. *IEEE Transactions on Biomedical Engineering*, 2010, 57(12): 2918–2926. [doi: [10.1109/TBME.2010.2083659](https://doi.org/10.1109/TBME.2010.2083659)]
- 8 Gietzelt M, Feldwieser F, Gövercin M, *et al.* A prospective field study for sensor-based identification of fall risk in older people with dementia. *Informatics for Health and Social Care*, 2014, 39(3–4): 249–261. [doi: [10.3109/17538157.2014.931851](https://doi.org/10.3109/17538157.2014.931851)]
- 9 朱连杰, 陈正宇, 田晨林. 基于可穿戴设备的跌倒检测方法综述. 计算机工程与应用, 2019, 55(18): 8–14. [doi: [10.3778/j.issn.1002-8331.1905-0269](https://doi.org/10.3778/j.issn.1002-8331.1905-0269)]
- 10 Dubois A, Charpillat F. A gait analysis method based on a depth camera for fall prevention. 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Chicago: IEEE, 2014. 4515–4518. [doi: [10.1109/EMBC.2014.6944627](https://doi.org/10.1109/EMBC.2014.6944627)]
- 11 Stone EE, Skubic M. Evaluation of an inexpensive depth camera for passive in-home fall risk assessment. 2011 5th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops. Dublin: IEEE, 2011. 71–77.
- 12 Cook AA, Mısırlı G, Fan Z. Anomaly detection for IoT time-series data: A survey. *IEEE Internet of Things Journal*, 2020, 7(7): 6481–6494. [doi: [10.1109/JIOT.2019.2958185](https://doi.org/10.1109/JIOT.2019.2958185)]
- 13 Pang GS, Shen CH, Cao LB, *et al.* Deep learning for anomaly detection: A review. *ACM Computing Surveys*, 2021, 54(2): 38. [doi: [10.1145/3439950](https://doi.org/10.1145/3439950)]
- 14 Schutte LM, Narayanan U, Stout JL, *et al.* An index for quantifying deviations from normal gait. *Gait & Posture*, 2000, 11(1): 25–31. [doi: [10.1016/S0966-6362\(99\)00047-8](https://doi.org/10.1016/S0966-6362(99)00047-8)]
- 15 Macellari V, Giacomozzi C, Saggini R. Spatial-temporal parameters of gait: Reference data and a statistical method for normality assessment. *Gait & Posture*, 1999, 10(2): 171–181. [doi: [10.1016/S0966-6362\(99\)00021-1](https://doi.org/10.1016/S0966-6362(99)00021-1)]
- 16 Martinez M, De Leon PL, Keeley D. Novelty detection for predicting falls risk using smartphone gait data. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. New Orleans: IEEE, 2017. 2202–2206. [doi: [10.1109/ICASSP.2017.7952547](https://doi.org/10.1109/ICASSP.2017.7952547)]
- 17 李茜楠. 基于 Kinect 的异常步态检测 [硕士学位论文]. 济南: 山东大学, 2016. [doi: [10.7666/d.Y3035625](https://doi.org/10.7666/d.Y3035625)]
- 18 Kargar BAH, Mollahosseini A, Struempf T, *et al.* Automatic measurement of physical mobility in Get-Up-and-Go Test using Kinect sensor. 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Chicago: IEEE, 2014. 3492–3495. [doi: [10.1109/EMBC.2014.6944375](https://doi.org/10.1109/EMBC.2014.6944375)]
- 19 袁亚运, 李红芳. 中国城乡老年人跌倒的性别差异研究——基于 CHARLS 2015 数据. 人口与社会, 2020, 36(4): 46–57. [doi: [10.14132/j.2095-7963.2020.04.004](https://doi.org/10.14132/j.2095-7963.2020.04.004)]
- 20 Deandrea S, Lucenteforte E, Bravi F, *et al.* Risk factors for falls in community-dwelling older people: A systematic review and meta-analysis. *Epidemiology*, 2010, 21(5): 658–668. [doi: [10.1097/EDE.0b013e3181e89905](https://doi.org/10.1097/EDE.0b013e3181e89905)]
- 21 Tripathy SR, Chakravarty K, Sinha A. Eigen posture based fall risk assessment system using Kinect. 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Honolulu: IEEE, 2018. 1–4.
- 22 Pimentel MAF, Clifton DA, Clifton L, *et al.* A review of novelty detection. *Signal Processing*, 2014, 99: 215–249. [doi: [10.1016/j.sigpro.2013.12.026](https://doi.org/10.1016/j.sigpro.2013.12.026)]
- 23 Schölkopf B, Williamson R, Smola A, *et al.* Support vector method for novelty detection. *Proceedings of the 12th International Conference on Neural Information Processing Systems*. Denver: MIT Press, 1999. 582–588.
- 24 Tunca C, Salur G, Ersoy C. Deep learning for fall risk assessment with inertial sensors: Utilizing domain knowledge in spatio-temporal gait parameters. *IEEE Journal of Biomedical and Health Informatics*, 2020, 24(7): 1994–2005. [doi: [10.1109/JBHI.2019.2958879](https://doi.org/10.1109/JBHI.2019.2958879)]
- 25 Buitinck L, Louppe G, Blondel M, *et al.* API design for machine learning software: Experiences from the scikit-learn project. arXiv: 1309.0238, 2013.