第 11 卷 第 4 期		集	成	技	术
2022 年 7 月	JOURNAL	OF IN	TEGR	ATION	TECHNOLOGY

#### 引文格式:

刘语诗,赵秀栩,冯鹤云,等.基于拓扑非线性动态建模的神经退行性疾病异常步态识别 [J].集成技术,2022,11(4):92-105.

Liu YS, Zhao XX, Feng HY, et al. Abnormality detection in neurodegenerative disease analysis based on topological nonlinear and dynamic modelling of gait [J]. Journal of Integration Technology, 2022, 11(4): 92-105.

# 基于拓扑非线性动态建模的神经退行性疾病异常 步态识别

刘语诗<sup>1,2,3</sup> 赵秀栩<sup>1,2\*</sup> 冯鹤云<sup>3</sup> 颜 延<sup>3\*</sup>

<sup>1</sup>(武汉理工大学机电工程学院 武汉 430070) <sup>2</sup>(数字制造湖北省重点实验室 武汉 430070) <sup>3</sup>(中国科学院深圳先进技术研究院 深圳 518055)

**摘 要** 通过对步态节律变化的观察,可以实现对人体复杂系统的观测,在行走过程中,步幅间隔时间序列的动态特性,能有效反映人体系统的状态变化,可用于步态异常检测及相关疾病辨识。人体步态传感信息的时间序列相空间重建,是一种有效表征系统非线性动力学特性的建模方法。相空间的几何建模和统计分析,是异常步态识别典型的分析方法,被广泛应用于神经退行性疾病检测等临床研究中。该文从空间拓扑特性分析的角度出发,提出一种基于拓扑非线性动态建模的分析方法,用于神经退行性疾病的异常步态识别。该文首先采用延时嵌入的相空间重构方法,将步态的波动时间序列转化为抽象相空间的状态点云;其次,采用基于计算拓扑中的持续同调工具,提取状态点云所在空间的拓扑描述信息;再次,采用基于拓扑描述的持续态势图,构建时间序列的拓扑非线性动态特征;最后,融合步态周期中左右足的步幅间隔、站立间隔、摆动间隔时间序列的拓扑非线性动态特征,作为分类器输入,构建出异常步态的机器学习识别模型。对患有肌硬化症、享延顿病和帕金森病的神经退行性疾病患者,进行 5 min 异常步态的连续行走数据(50 步滑动窗数据)测试,步态识别的准确率分别为0.875 0 (0.914 6)、0.940 6 (0.962 3)和 0.958 3 (0.961 4)。因此,拓扑非线性动态建模分析是一种有效的神经退行性疾病异常步态检测方法,为基于步态分析的神经退行性疾病检测和可穿戴数据分析提供了一种新的思路。

关键词 异常步态; 拓扑数据分析; 时间序列分析; 持续同调; 机器学习; 非线性动力学 中图分类号 TP 394.1 文献标志码 A doi: 10.12146/j.issn.2095-3135.20211215001

收稿日期: 2021-12-15 修回日期: 2022-01-27

基金项目: 国家重点研发计划项目(2020YFC2007203)

作者简介:刘语诗,硕士,研究方向为精益生产与智能制造、信号处理;赵秀栩(通讯作者),博士生导师,研究方向为生产运作与企业信息化、机 械工况监测与智能诊断等,E-mail: zhaoxiuxu@whut.edu.cn;冯鹤云,硕士,研究方向为嵌入式软件开发及应用、硬件电路等;颜延(通讯作者),博 士,助理研究员,主要研究方向为人工智能与拓扑数据分析,E-mail: yan.yan@siat.ac.cn。

# Abnormality Detection in Neurodegenerative Disease Analysis Based on Topological Nonlinear and Dynamic Modelling of Gait

LIU Yushi<sup>1,2,3</sup> ZHAO Xiuxu<sup>1,2\*</sup> FENG Heyun<sup>3</sup> YAN Yan<sup>3\*</sup>

<sup>1</sup>(School of Mechanical and Electrical Engineering, Wuhan University of Technology, Wuhan 430070, China) <sup>2</sup>(Hubei Key Laboratory of Digital Manufacturing, Wuhan 430070, China) <sup>3</sup>(Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China)

\*Corresponding Authors: zhaoxiuxu@whut.edu.cn; yan.yan@siat.ac.cn

**Abstract** The complex system of human body can be observed by observing the change of gait rhythm. The dynamic characteristics of the time series of stride interval during walking can effectively reflect the state change of human system, which can be used for abnormal gait detection and related disease identification. Time series phase space reconstruction of human gait sensor information is an effective modeling method to characterize nonlinear dynamics of the system. Geometric modeling and statistical analysis of phase space are typical analysis methods for abnormal gait recognition, which are widely used in clinical research such as neurodegenerative disease detection. In this paper, a topological nonlinear dynamic modeling method is proposed to identify abnormal gait in neurodegenerative diseases from the perspective of spatial topological analysis. Firstly, a phase-space reconstruction method based on time-delay embedding is used to transform the wave time series of gait into an abstract phase-space state point cloud. Secondly, the continuous homology tool based on computational topology is used to extract the topology description information of the space where the state point cloud resides. Thirdly, the topological nonlinear dynamic characteristics of time series are constructed by using the continuous situation graph based on topological description. Finally, a machine learning recognition model of abnormal gait was constructed by integrating the topological nonlinear dynamic characteristics of the time series of stride interval, standing interval and swing interval of left and right feet in the gait cycle as the input of classifier. The results showed that the area under receive-operator-curve was 0.875 0 (0.914 6), 0.940 6 (0.962 3) and 0.958 3 (0.961 4) in the 5-minute continuous walking data (50-step sliding window data) of abnormal gait in patients with neurodegenerative diseases of muscular sclerosis, Huntington's disease and Parkinson's disease, respectively. Therefore, topological nonlinear dynamic modeling analysis is an effective gait detection method for neurodegenerative diseases, and provides a new idea for neurodegenerative disease detection and wearable data analysis based on gait analysis.

**Keywords** abnormal gait; topological data analysis; time series analysis; persistent homology; machine learning; nonlinear dynamics

Funding This work is supported by Key Research and Development Program of China (2020YFC2007203)

# 1 引 言

神经退行性疾病(Neural Degenerative Diseases, NDDs)患者步态异常,严重影响日常

行动,临床症状表现为行走过程中震颤、动作 迟缓等步态异常。其中,最常见的疾病患者有3 种:肌硬化症(Amyotrophic lateral sclerosis, ALS) 患者(步态如图1所示)、亨廷顿病(Huntington's Disease, HD) 患者 (步态如图 2 所示) 以及帕金 森病 (Parkinson's Disease, PD) 患者 (步态如图 3 所示)。由图 1~3 可知,在行走模式上,不同 的 NDDs 患者会表现出不同的异常特征,因此, 从病人步态信息中提取 NDDs 评估特征和生物 标记,研究正常受试者 (步态如图 4 所示)和患有 NDDs 受试者的步态差异,进行异常步态的检测 和识别,对 NDDs 的早期监测、健康风险预警和 病程评估都具有重要意义。

在行走过程中,人体从同一只脚跟离地跨出 到再次脚跟着地的行进过程被称为一个步态周 期,重复该动作的时间间隔被称为步幅。一个步态周期可进一步地被划分为支撑相(下肢接触地面及承重力的时间)和摆动相(脚离开地面向前迈步到再次落地之间的时间)<sup>[1-2]</sup>。步态动力学能够揭示不同 NDDs 患者的运动障碍与行走模式。在行走过程中,ALS 患者主要表现为肌无力、肌肉萎缩、束缚等引起的肌肉运动障碍<sup>[3]</sup>;HD 患者的特征则是表现为不自主地抽搐<sup>[4]</sup>;PD 患者的步态障碍尤为明显,主要表现为静态震颤、步态冻结等<sup>[5]</sup>。由于步态异常模式的多样化,相关研究者们将许多生物标志物和分析方法用于异常步

图 1 肌硬化症步态示意图 Fig. 1 Gait diagram of muscular sclerosis

图 2 亨廷顿病步态示意图

Fig. 2 Gait diagram of Huntington's disease

图 3 帕金森病步态示意图

Fig. 3 Gait diagram of Parkinson's disease

Mun manus man when a second when the second when the second secon

图 4 正常受试者步态示意图

Fig. 4 Schematic diagram of normal subjects' gait

态的检测与评估。

利用人体传感信息进行异常步态识别,是 健康检测领域快速发展的重要课题之一。近年 来,学者们主要从4个方面开展异常步态的识 别: 基于计算机视觉的步态识别、基于肌电信号 的步态识别、基于惯性传感器的步态识别以及 基于足底压力的步态识别。研究发现,可靠、精 确的异常步态监测手段,能辅助医生诊断疾病, 也可以用于日常生活中保健与智能疾病的监测。 1994 年,麻省理工学院的 Niyogi 等<sup>[6]</sup>从步态图 像中提取时域、频域特征用来识别步态,发现不 同个体的步态能够得到有效鉴别,从此基于计 算机视觉的步态识别开启了新篇章。Tang 等<sup>[7]</sup> 将步态周期按不同的阶段输入卷积神经网络模 型,并结合视频提取有效的连续轮廓差异图, 实现了步态检测,在公共数据集上取得了较高 的精度。徐俞弢等<sup>[8]</sup>利用 Vicon 系统获取左右脚 踝的运动轨迹,进而分析个人的左右步态对称 性。2019年,赵雨佳<sup>[9]</sup>提出了一种基于稀疏学习 的方法,该方法可筛选出有效的疾病特征,实现 对 NDDs 的早期检测,对不同疾病的分类可达到 90% 左右的准确率。21 世纪初, Anderson 等<sup>[10]</sup> 首次将肌电信号加入步态动作的研究中,从此, 基于肌电信号的步态识别成为一个新的研究方 向。2019年, Becker 等<sup>[11]</sup>采集了实验者在跑步 时的6块下肢肌电信号,并通过评估发现,6块 肌肉中的比目鱼肌的肌电信号与步速具有良好的 相关性,可通过肌电信号的强弱来判断步速的大 小。Zawawi 等<sup>[12]</sup>提取了步态肌电信号的电压均 方根、频率平均值、方差以及标准差4个特征, 并对肌肉骨骼疾病健康检查任务中的肌电进行分 类,实验结果表明,基于支持向量机的分类准确 率超过80%,能够帮助康复中心进行疾病诊断。 目前,基于惯性传感器的步态识别是研究较为广 泛的步态识别。Mileti 等<sup>[13]</sup>提出一种可对 PD 患 者腿部的加速度数据进行分类的马尔可夫模型,

实验结果表明,该方法在异常步态的检测上效果 显著。Camps 等<sup>[14]</sup>提出了基于深度学习的模型, 通过从腰部采集传感器数据,来进行异常步态的 识别,实验结果证明,腰部的传感器数据可以很 好地区分不同类型的步态。Lee 等<sup>[15]</sup>利用从惯性 传感器获取三轴加速度信息和角速度信息来区分 不同步态,并在正常人与偏瘫患者的两种步态中 得到了验证。通过可穿戴式鞋子中的传感器提取 行走过程中的力信号,基于足底压力传感器可以 准确获得步幅、支撑相以及摆动相的足底压力信 息。Hausdorff<sup>[16]</sup>提出一种基于分形的分析方法 来显示步幅波动时间序列的变化,并利用步幅间 隔波动对步态节奏进行统计分析,以利于 PD 检 测,该方法主要利用了信号轮数和步态节律的标 准差(σ)等特征<sup>[17]</sup>。此外,还有学者利用统计参 数对 NDDs 进行分析<sup>[18-19]</sup>, Wahid 等<sup>[20]</sup>对 PD 患 者和健康受试者之间的时空模型进行了研究,提 出了一种混合时空模型<sup>[21]</sup>,该模型可用于步态 数据中的 PD 检测,还可对其严重程度进行等级 划分。2015年, Baratin 等<sup>[22]</sup>运用小波特征证明 步态的不对称性,并利用步态的规律性识别步 态,验证了步态规律在步态识别中的有效性。 2018 年,Ye 等<sup>[23]</sup>提出一种神经网络自适应能力 与模糊逻辑定性方法相结合的自适应神经模糊推 理模型,用于神经退行性步态节奏波动的分析。 Zeng 等<sup>[24]</sup>提出一种具有较强判别性的确定性学习 方法,用于 NDDs 的分类,取得了良好的分类性 能。基于双足运动数据融合的时域、频域分析方 法,不仅反映了运动的生理相关性,而且还能够 获得各种 NDDs 患者运动不对称的差异特征<sup>[25]</sup>。 2020年,江俊敏<sup>[26]</sup>从足底压力信号中提取了多 种步态信息,并从非线性动力学的角度判断不同 信号的混沌特性以实现对 NDDs 的正常分类。 Ren 等<sup>[27]</sup>提取步态时间序列中的相位同步指数以 及条件熵两种特征进行 NDDs 分类。经验模态分 解<sup>[28]</sup>、基于规则的方法<sup>[29]</sup>、基于距离度量的稀疏 编码<sup>[30]</sup>、多重分形去趋势互相关分析<sup>[31]</sup>、基于 模糊递归图的纹理分析<sup>[32]</sup>等也是 NDDs 识别中步 态节奏波动分析的重要工具。通过对不同特征的 表征能力进行大量研究,进一步证明了 NDDs 是 引起步态异常的主要因素之一。

本文主要从以下3个方面进行研究:

(1)基于足底压力的 NDDs 步态数据库,采 用一种非线性动力学特征分析的步态建模方法, 从步态周期的不同阶段进行特征提取,该方法能 对步态周期间隔时间序列的相空间重构进行拓扑 建模,从而提取具有非线性动力学的特征;

(2)在 NDDs 数据集中对本实验方法进行验证,采用留一交叉验证,在 ALS、HD 以及 PD 患者与正常组的分类识别实验中,模型的工作特征曲线面积(Area Under Receive-Operator-Curve, AUC)值分别为 0.875 0、0.940 6、0.958 3,具有较高的识别水平;

(3)此外,本实验还采用固定周期滑动窗对数据集进行划分,再次验证了本实验方法的有效性。实验结果表明,在基于 50 个步态间隔的时间序列的异常步态识别中,模型的 AUC 值分别为 0.914 6、0.962 3、0.961 4,明显优于相同数据集的其他非线性特征识别的准确率。

# 2 材料与方法

#### 2.1 数据集及预处理

本实验采用的 NDDs 步态数据库由美国哈 佛医院提供,是目前被广泛使用的步态数据集 (具体内容见 http://www.physionet.org/physiobank/ database/gaitndd)<sup>[33]</sup>。该数据库主要包括 64 位受 试者:健康受试者 16 名,年龄(39.3±18.5)岁, 男性 2 名、女性 14 名;ALS 患者 13 名,年龄 (54.9±13.4)岁,男性 10 名、女性 3 名;HD 患 者 20 名,年龄(47.7±12.2)岁,男性 6 名、女 性 14 名;PD 患者 15 名,年龄(66.8±10.9)岁, 男性 10 名、女性 5 名。实验中对不同类型人群 在行走过程中的步态数据进行采集,以实现对 NDDs 病理生理学机制的研究。该数据集中的步 态信号由足底压力传感器采集,传感器内置的模 式转换器有 12 位,采样频率为 300 Hz。实验过 程中,每个受试者在长 77 m 的走廊中以正常速 度行走 5 min。数据集还提供了步态周期的时间 信息、左右足的跨步间隔、左右摆动间隔、左右 站姿间隔以及其对应的比例信息。本文主要基于 步态跨步、摆动以及站姿间隔时间序列 3 种步态 阶段研究其动态变化信息。

本实验时长为 5 min, 但走廊长度为 77 m, 为保证时长, 被试人员在行走过程中需要多次折 返, 而折返过程会导致数据异常, 给实验带来干 扰。因此, 本实验对异常值进行了处理, 方法如 下: 对于每一个时间序列, 分别计算出数据中的 中值 ( $\mu$ )、标准差 (s), 对于时间序列中的每个值  $t_i$ , 当  $t_i < \mu + 2s$  或  $t_i > \mu + 2s$  时,  $t_i$  会被替换成中 值, 其他数值不变。此外,由于步态起步过程中 所采集的时间序列不够稳定,所以为了最大限度 地减少步态数据的启动效应, 在构建数据集时, 本实验将步态时间序列前 20 s 的数据删除。经 上述步骤后,再对数据进行归一化处理,可以大 大提高数据在运算过程中的收敛速度, 从而提高 效率。

#### 2.2 相空间重构

步态节律波动是非线性、非平稳的时间序 列<sup>[34]</sup>。当进行拓扑非线性动力学分析时,可将步 态的时间序列转化为相应的动力学系统在相空间 的轨迹,该过程被称为相空间重构<sup>[35]</sup>,重构的相 空间被称为拓扑空间。在相空间重构的过程中, 步态时间序列通过延时嵌入的方式可转换为抽象 相空间的点云,由于状态相空间重构是通过延迟 嵌入实现的,所以可将延迟嵌入后动态系统的时 间序列在相空间中形成的点云看作是基于向量的 点。在数学上,对于时间序列  $t=\{t_1,t_2,t_3,\dots,t_k\}$ , 可通过公式(1)表示延迟嵌入的过程:

$$X_{K} = \left(t_{K}, t_{K+\tau}, \cdots, t_{K+(d-1)\tau}\right)$$
(1)

其中,  $X_{\kappa}$  为延迟嵌入后的时间序列;  $\tau$  为相空间 延迟参数; d 为相空间的嵌入维数,选择适当的 嵌入参数  $\tau$  和 d 能够提高模型的鲁棒性。在以往 的研究过程中,也有许多相空间重建的方法被提 出,如平均互信息法<sup>[36]</sup>、选择最佳延迟参数的相 关策略<sup>[37-38]</sup>以及基于几何信息选择延迟参数的方 法<sup>[39]</sup>和伪近邻的算法<sup>[35]</sup>等。基于相关研究<sup>[40]</sup>发 现,固定  $\tau \times d$  的尺寸不变,不会影响实验结论 的有效性。因此,基于平均互信息法原则及固定 的  $\tau \times d$  尺寸,本实验对时间序列进行最优延迟 参数的选择。先设置不同的延时参数,然后利用 分类器分别进行测试,最后选择  $\tau=5$ 、d=3 作为 实验的最佳参数。如图 5 所示,将时间序列映射 到空间中形成点云,构建出最佳的相空间系统。





#### 2.3 持续同调与拓扑描述

持续同调是拓扑分析的主要工具之一,对于 点云所处的抽象拓扑空间,可通过该工具对其进行 分析。分析原理为利用单纯复形序列来描述数据 点之间的关系,提取有效拓扑特征<sup>[41]</sup>。其中,单 纯复形是一种利用点(0-单纯形)、线(1-单纯形)、 三角形面(2-单纯形)、四面体(3-单纯形)等几何 对象来拼接或粘合成整个拓扑空间的数据结构(如 图 6 所示)。同理,在高维空间中,由两个及两个 以上的单纯形连接而成的形状被称为单纯复形。





持续同调的建模首先是将拓扑空间中的单个 点视为具有半径的 *ε*-小球,不同半径参数 *ε* 的拓 扑性质可以采用不同的单纯复形表示。在半径增 大的过程中,形成了不同的基于 *ε* 的单纯复形序 列(如图 7 所示),在这个过程中能够获取不同拓 扑对象的生存信息,从而实现对点云所处抽象拓 扑空间的特征描述。具体构建过程如下:

(1)将点云空间中点的半径都设置为 ε,当半径较小时,空间中的点相互独立,此时,数据点之间通过 0-单纯形进行构建;

(2)随着 ϵ 的不断增大,空间中的 ϵ-小球慢慢产生交集,产生交集的个体的数量开始增多,数据点之间开始通过 1-单纯形进行构建。随着半径继续增大,更多的 ϵ-小球慢慢产生交集,之前产生交集的数据点会不断地消失,并生成新的融合体,构建出新的单纯复形序列。随着产生交集的数据点越来越多,融合体越来越大,单纯复形的构建也随之复杂;

(3)当 *ϵ* 增加到无穷大时,所有连接个体融合,之前的所有单纯形也合并成一个更加复杂的单纯复形结构。

在单纯复形构建的过程中会形成不同的拓扑特征:0-阶同调孔洞、1-阶同调孔洞以及 n-阶同调孔洞,其中,0-阶同调孔洞如图 8 中的 B<sub>0</sub> 所示,1-阶同调孔洞如图 8 中的 B<sub>1</sub> 所示。在点云的变化过程中,通过持续条码记录不同拓扑特征的持续状态<sup>[42]</sup>,如图 8 蓝色直线所示,直线的



图 7 持续同调的二维点云建模过程





图 8 拓扑空间分析以及拓扑印记的提取

Fig. 8 Topological space analysis and topological imprint extraction

长度代表拓扑特征的生存时间,当半径参数不同 时,会有不同的拓扑特征出现和消失,且每个特 征的生存时间不同。对于每个条形码,都可以用 单个序对(b,,d)进行表示。对于点云所处的拓扑 空间,可以用如下集合来表示其特征:

. . .

11

$$\left\{ \left\{ b_0^1, d_0^1 \right\}, \left\{ b_0^2, d_0^2 \right\}, \cdots, \left\{ b_0^M, d_0^M \right\} \right\}$$
(2)

$$\left\{ \left\{ b_0^1, d_0^1 \right\}, \left\{ b_0^2, d_0^2 \right\}, \cdots, \left\{ b_0^N, d_0^N \right\} \right\}$$
(3)

>>

其中, M为 0-阶同调孔洞的个数; N为 1-阶同调 孔洞的个数:每个序对为不同状态时 0-阶同调、 1-阶同调出生和死亡的时间。

#### 2.4 拓扑特征提取

为了提取特征方便,将每个拓扑特征条形码 的出生时间(b,)和死亡时间(d,)分别作为坐标轴的 x 轴和 y 轴, 转换为如图 9 所示的持续散点图。 正常步态和 PD 异常步态的持续散点图信息如 图 9 所示,通过散点图可以观察到不同步态提取 的拓扑特征具有显著差异,基于该散点图也可以 获取不同的信息来进行步态识别。

近年来,由于拓扑描述构造的持续景观图 (Persistence Landscape, PL)可以实现对步态间 隔序列对应的动力系统的动态描述,且能通过数 学的方式进行表达,使较抽象的孔洞信息更易理 解,计算更为方便,在运算过程中,能够提高模 型的收敛速度,因此,PL 特征常常作为数据的 特征向量,应用于模式分析任务中。在几何上,

PL 通过将持续散点图进行 45° 顺时针旋转,再将每个分布点的信息进行累加;在数学上,PL 可以通过一个连续的分段线性函数集合进行表示。假设每个点  $Z_a = (b_a, d_a)$ ,考虑一个分段线性函数:

$$\Lambda_{p}(t) = \begin{cases}
t - b_{a} & t \in \left(b_{a}, \frac{b_{a} + d_{a}}{2}\right) \\
-t + d_{a} & t \in \left(b_{a}, \frac{b_{a} + d_{a}}{2}\right) \\
0 & t \notin \left(b_{a}, d_{a}\right)
\end{cases} \tag{4}$$

将  $Z_a$  对应的点设置为(x,y), 令(x,y)=

$$\left(\frac{b_{a}+d_{a}}{2}, \frac{-b_{a}+d_{a}}{2}\right)$$
, 上式可以表达为:  
 $\Lambda_{p}(t) = \begin{cases} t-x+y & t \in [x-y,x] \\ x+y-t & t \in (x,x+y] \\ 0 & t \notin (x-y,x+y) \end{cases}$  (5)

将不同点的图像重叠起来即为分段线性曲线,则 可转化为如式(6)的集合:

$$\lambda(k,t) = K \max_{p \in D} \Lambda_p(t) \qquad t \in [0,T], k \in \mathbb{N}$$
(6)

本研究选取 k=1 作为最大值,提取不同步态的 PL 特征,并将其与机器学习相结合进行分类。提取 PL 特征如图 10 所示,以帕金森步态为例,其中蓝色代表 0-阶同调的 PL 特征,红色代表 1-阶同调的 PL 特征。通过观察图 10 可知,正常步态和异常步态的 0-阶同调的 PL 特征 并无明显区别,而 1-阶同调的 PL 特征图存在显著区别,所以,本文利用基于拓扑分析的 1-阶 同调的 PL 特征来区分健康组与异常组的步态时间序列。

#### 2.5 算法评估

本研究基于非线性拓扑特征的机器学习算法



Fig. 9 Scatter plot of birth and death in asynchronous state



Fig. 10 PL features of order 0 homology and order 1 homology of normal gait and Parkinson's abnormal gait

对正常步态与异常步态进行评估,主要包括正常 组与 ALS 组、正常组与 HD 组、正常组与 PD 组 的 3 个二分类问题,由于数据集中包含的对象数 量有限,所以本实验采用了留一交叉验证的方 法。文献<sup>[3,16,24]</sup>中存在相同的识别问题,即采用 数据集中的一个单独样本作为该任务的测试样 本,其余则作为训练样本,以更好地提取有效特 征。为进一步说明该方法的有效性,本实验将每 个时间序列分割成更短的切片,以生成更大的数 据集,利用滑动窗的方法将一个完整步态周期分 为左右脚步幅间隔、左右脚摆动间隔以及左右脚 站立间隔6个时间序列。实验过程中发现,当窗 口长度少于 50 个点时,相空间重构就会失败, 所以为保证相空间成功构建,提取有效特征, 且 创建的数据集更大,本实验选择每段的时间序列 为 50 个点。通过上述方法对每个类别都作相应 拓展后,采用训练集:测试集为1:1的方式进 行实验。

由于每个时间序列都是相互独立的,不同时 间序列所构建的拓扑特征是不同的。如图 11 所 示,在不同步态周期中,左右脚与地面接触的面 积是不同的,但人体体重不会随着位置的移动而 发生改变,因此随着面积的改变,足底压力所获 取的信息是不同的。此外,步态周期虽然具有一定的规律,但是并不稳定,将不同脚跨步间隔的 足底压力信息列入考虑范围,将能够更全面地获 取受试者在行走过程中不同状态的步态特征。本 实验采用多特征融合的方法,即波动融合(将左 脚和右脚的步幅间隔、摆动间隔以及站立间隔都 进行融合),然后使用朴素贝叶斯(Naive Bayes, NB)、支持向量机(Support Vector Machine, SVM)和随机森林(Random Forest, RF)3种具有 代表性的分类器对模型进行综合评估,并利用受 试者 AUC 值来衡量每个模型的识别能力,AUC 值越大,模型的鲁棒性越好。

2022 年



#### 3 结 论

本研究基于步态时间序列特征的融合,以验

证所提取方法的有效性。首先,基于原始4种时间为5min的步态时间序列进行评估,在提取1-阶同调的PL特征后使用留一验证法,然后利用NB、SVM、RF3个分类器分别对正常与ALS、正常与HD、正常与PD3组进行分类,以AUC值为结果进行评估,评估结果如表1所示。由表1可知,特征融合的效果比单一特征效果好,在ALS与正常组中,使用NB分类器的效果最佳,其AUC值为0.8750;在HD与正常组中,使用NB分类器效果最佳,AUC值为0.9406;在PD与正常组中,3种分类器效果均高于平均水平,其中,NB及RF分类器的效果较好,分别为0.9583与0.9580。

在重新扩展的数据集上,采用相同的拓扑特征以及分类器对模型进行验证,并以 AUC 值为

结果进行评估,实验结果如表 2 所示。由表 2 可知,在ALS 与正常组中,采用 RF 分类器效果最佳,AUC 值达到 91.46%,明显优于基于原始步态数据的实验结果;在 HD 与正常组中,NB 和 SVM 的结果要明显弱于 RF 分类器,最佳 AUC 值为 96.23%;在 PD 与正常组中,NB 分类器没有良好的鲁棒性,其最佳分类结果是由 RF 分类器产生的,AUC 值为 96.14%。

### 4 讨 论

目前,关于步态节律以及半周期的研究都经 历了较长的发展期,研究者们采用不同的特征 来区分健康受试者与 ALS、HD 和 PD 受试者, 以及其他异常步态者。传统的平均值、标准差

表1	基于原始数据的评估结果	

方法	特征	ALS 与正常				HD 与正常		PD 与正常		
		NB (%)	SVM (%)	RF (%)	NB (%)	SVM (%)	RF (%)	NB (%)	SVM (%)	RF (%)
单个阶段	左脚步幅间隔	75.71	81.73	73.56	90.94	92.19	90.63	87.71	83.75	91.67
	右脚步幅间隔	81.73	86.06	84.62	96.25	89.69	91.25	87.08	88.75	86.67
	左脚摆动间隔	78.85	70.67	75.96	87.19	83.12	85.63	88.33	93.33	93.33
	右脚摆动间隔	80.77	69.23	68.27	92.50	85.63	89.37	92.92	90.00	92.08
	左脚站立间隔	83.17	72.12	75.24	91.88	85.94	88.75	86.46	79.58	82.92
	右脚站立间隔	88.70	91.34	86.06	91.88	87.19	84.38	88.96	81.67	79.58
多个阶段	特征融合	87.50	83.17	82.08	94.06	90.63	92.81	95.83	93.33	95.80

 Table 1
 Experimental results based on raw data

#### 表 2 基于重新扩建数据集的实验结果

 Table 2
 Experimental results based on the re-expanded data set

方法	特征	ALS 与正常				HD 与正常	÷ I		PD 与正常		
		NB (%)	SVM (%)	RF (%)	NB (%)	SVM (%)	RF (%)	NB (%)	SVM (%)	RF (%)	
	左脚步幅间隔	74.95	75.27	80.99	85.79	85.43	89.23	76.26	77.37	81.24	
单个阶段	右脚步幅间隔	72.78	76.34	82.02	84.86	86.02	91.01	83.49	84.88	85.90	
	左脚摆动间隔	71.90	76.49	85.84	85.02	83.77	90.49	86.19	87.02	91.80	
	右脚摆动间隔	73.95	77.89	84.96	87.31	85.10	90.62	85.41	86.19	90.11	
	左脚站立间隔	72.18	71.96	77.37	79.97	81.00	84.44	78.10	80.58	85.79	
	右脚站立间隔	76.70	78.92	80.91	86.63	85.37	92.09	82.81	80.38	84.51	
多个阶段	特征融合	80.20	82.90	91.46	89.33	91.01	96.23	89.92	93.11	96.14	

等统计参数特征提取方法只考虑了有限的一维 数据特征,忽略了步态节律在相位空间中的非线 性动力学特征。在特征提取的过程中,大部分 主要针对步幅以及步幅变化进行研究,对左右脚 步幅间隔、摆动间隔以及站立间隔相结合的研究 较少,在特征信息的提取过程中,会导致部分 信息遗漏。在同一数据集上,将本文方法与其 他方法进行对比<sup>[26]</sup>,对比结果如表 3 所示。由 表 3 可知, 在 ALS 患者与正常受试者步态的检 测中,效果最好的是近似熵与模糊熵,但其分类 结果只有 71.42%, 明显低于本文方法 87.50% 的 识别准确率。在 HD 患者与正常受试者步态的识 别中,模糊熵的识别准确率最高,但也只达到了 77.78%,比本文方法识别的准确率还要低 16%。 在 PD 患者与正常受试者步态的检测中, 样本 熵、模糊熵的非线性特征检测结果较好,然而仅 达到 75.00%, 与本文方法 95.83% 的识别率仍有 较大差距。由近似熵、样本熵、模糊熵 3 种非线 性特征的对比可知, 在异常步态识别中, 模糊熵

的表现最好,但仍低于本文所用的 PL 非线性特征。综上所述,在基于非线性特征的异常步态识别中,本文方法具有较为明显的优势。

本文提出一种全新的方法用于异常步态的识别,该方法研究了步态相空间的拓扑特征,由于 拓扑特征具有随着疾病的干扰而改变的特点,所 以这些特征能够作为 NDDs 潜在的生物标志之 一。此外,步态节律也会受到信号的非线性以及 非平稳性的影响,因此,采用拓扑非线性动态建 模的方法能够很好地保留步态特征的空间特性以 及非线性特征。由表 1 可知,在异常步态识别的 应用中,进一步扩展数据集,拓扑非线性动态建 模的方法仍可取得较好的结果,使得该方法的优 良性能得到了进一步验证。

综上所述,本文提出的基于拓扑非线性动力 学的异常步态检测的优点如下:

(1)基于步态周期的特征提取,考虑了不同 阶段的步态特征及左右脚之间的协同信息;

(2)提取特征时,采用一种新的拓扑特征,保 留了步态变化过程中非线性动力学的特点;

(3)在 ALS、HD 以及 PD 分类中,本文方法 均获得较高的准确率,明显优于其他非线性动力 学特征。

# 5 结论与展望

本文提出了一种基于拓扑特征分析的异常步态非线性分析方法,探讨了拓扑特征在 NDDs 方面的应用,研究了 ALS、HD、PD 患者的步态拓扑特征差异,构造了基于步态拓扑特征信息的机器学习分类模型,实现了 NDDs 异常步态的精准分类。在对 ALS、HD 以及 PD 患者步态的检测 任务中,分别获得了 91.46%、96.23%、96.14%的准确率,验证了拓扑非线性特征在步态分析中显著的表征能力,及良好的异常识别能力,为临床或康复中的步态信息分析提供了一种新的思路。

表 3 相同数据集的非线性特征结果对比

Table 3 Comparison of nonlinear characteristic results of the same data set

林东东	ALS 与正常				HD 与正常			PD 与正常			
1寸1匹	NB (%)	SVM (%)	RF (%)	NB (%)	SVM (%)	RF (%)	NB (%)	SVM (%)	RF (%)		
近似熵	71.42	71.42	57.14	44.44	66.67	66.67	62.50	50.00	62.50		
样本熵	42.85	57.14	28.57	66.67	66.60	33.33	37.50	75.00	37.50		
模糊熵	71.42	71.42	57.14	77.78	77.78	62.50	75.00	62.50	62.50		
本文方法	87.50	83.17	82.08	94.06	90.63	92.81	95.83	93.33	95.80		

#### 参考文献

- Whittle MW. Clinical gait analysis: a review [J]. Human Movement Science, 1996, 15(3): 369-387.
- [2] Perry J, Davids JR. Gait analysis: normal and pathological function [J]. Journal of Pediatric Orthopaedics, 1992, 12(6): 815.
- [3] Kiernan MC, Vucic S, Cheah BC, et al. Amyotrophic lateral sclerosis [J]. The Lancet, 2011, 377(9769): 942-955.
- [4] Roos RAC. Huntington's disease: a clinical review[J]. Orphanet Journal of Rare Diseases, 2010, 5(1): 1-8.
- [5] Lang AE, Lozano AM. Parkinson's disease. First of two parts [J]. New England Journal of Medicine, 1998, 339(16): 1130-1143.
- [6] Niyogi SA, Adelson EH. Analyzing gait with spatiotemporal surfaces [C] // Proceedings of the IEEE Workshop on Motion of Non-rigid & Articulated Objects, 1994: 64-69.
- [7] Tang YQ, Li ZR, Tian HW, et al. Detecting toe-off events utilizing a vision-based method [J]. Entropy, 2019, 21(4): 329.
- [8] 徐俞弢, 王伟杰, 徐腾, 等. Vicon 运动捕捉系统在 人行走过程中左右踝关节运动特征的研究 [J]. 现 代信息科技, 2018, 2(6): 7-9.

Xu YT, Wang WJ, Xu T, et al. Study on the characteristics of left and right ankle joints in Vicon motion capture system during human walking [J]. Modern Information Technology, 2018, 2(6): 7-9.

- [9] 赵雨佳. 基于稀疏学习的神经退行性疾病早期诊断方法研究 [D]. 深圳: 深圳大学, 2019.
   Zhao YJ. Sparse feature learning for neurodegenerative disease early diagnosis [D]. Shenzhen: Shenzhen University, 2019.
- [10] Anderson FC, Pandy MG. Dynamic optimization

of human walking [J]. Journal of Biomechanical Engineering, 2001, 123(5): 381-390.

- [11] Becker S, Berger J, Backfisch M, et al. Effects of a 6-week strength training of the neck flexors and extensors on the head acceleration during headers in soccer [J]. Journal of Sports Science & Medicine, 2019, 18(4): 729-737.
- [12] Zawawi TNST, Abdullah AR, Sudriman R, et al. Classification of EMG signal for health screening task for musculoskeletal disorder [J]. International Journal of Engineering and Technology, 2019, 8(1.7): 219-226.
- [13] Mileti I, Germanotta M, Alcaro S, et al. Gait partitioning methods in Parkinson's disease patients with motor fluctuations: a comparative analysis
  [C] // Proceedings of the 2017 IEEE International Symposium on Medical Measurements and Applications, 2017: 402-407.
- [14] Camps J, Samà A, Martín M, et al. Deep learning for freezing of gait detection in Parkinson's disease patients in their homes using a waist-worn inertial measurement unit [J]. Knowledge-Based Systems, 2018, 139: 119-131.
- [15] Lee J, Park S, Shin H. Detection of hemiplegic walking using a wearable inertia sensing device [J]. Sensors, 2018, 18(6): 1736.
- [16] Hausdorff JM. Gait variability: methods, modeling and meaning [J]. Journal of Neuroengineering and Rehabilitation, 2005, 2(1): 1-9.
- [17] Wu Y, Krishnan S. Statistical analysis of gait rhythm in patients with Parkinson's disease
  [J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2009, 18(2): 150-158.
- [18] Wu YF, Shi L. Analysis of altered gait cycle duration in amyotrophic lateral sclerosis based on nonparametric probability density function

estimation [J]. Medical Engineering & Physics, 2011, 33(3): 347-355.

- [19] Xia Y, Gao QW, Ye Q. Classification of gait rhythm signals between patients with neuro-degenerative diseases and normal subjects: experiments with statistical features and different classification models [J]. Biomedical Signal Processing and Control, 2015, 18: 254-262.
- [20] Wahid F, Begg RK, Hass CJ, et al. Classification of Parkinson's disease gait using spatial-temporal gait features [J]. IEEE Journal of Biomedical and Health Informatics, 2015, 19(6): 1794-1802.
- [21] Zhao A, Qi L, Li J, et al. A hybrid spatiotemporal model for detection and severity rating of Parkinson's disease from gait data [J]. Neurocomputing, 2018, 315: 1-8.
- [22] Baratin E, Sugavaneswaran L, Umapathy K, et al. Wavelet-based characterization of gait signal for neurological abnormalities [J]. Gait & Posture, 2015, 41(2): 634-639.
- [23] Ye Q, Xia Y, Yao ZM. Classification of gait patterns in patients with neurodegenerative disease using adaptive neuro-fuzzy inference system [J]. Computational and Mathematical Methods in Medicine, 2018, 2018: 1-8.
- [24] Zeng W, Wang C. Classification of neurodegenerative diseases using gait dynamics via deterministic learning [J]. Information Sciences, 2015, 317: 246-258.
- [25] Ren P, Hu SA, Han ZF, et al. Movement symmetry assessment by bilateral motion data fusion [J]. IEEE Transactions on Biomedical Engineering, 2018, 66(1): 225-236.
- [26] 江俊敏. 基于步态信号非线性动力学分析的神经 退行性疾病智能分类方法 [D]. 杭州: 杭州电子科 技大学, 2020.

Jiang JM. Nonlinear dynamic analysis based on gait signals intelligent classification of neurodegenerative diseases [D]. Hangzhou: Hangzhou Dianzi University, 2020.

- [27] Ren P, Zhao WH, Zhao ZY, et al. Analysis of gait rhythm fluctuations for neurodegenerative diseases by phase synchronization and conditional entropy
   [J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2015, 24(2): 291-299.
- [28] Ren P, Tang SJ, Fang F, et al. Gait rhythm fluctuation analys is for neurodegenerative diseases by empirical mode decomposition [J]. IEEE Transactions on Biomedical Engineering, 2016, 64(1): 52-60.
- [29] Gupta K, Khajuria A, Chatterjee N, et al. Rule based classification of neurodegenerative diseases using data driven gait features [J]. Health and Technology, 2019, 9(4): 547-560.
- [30] Ghaderyan P, Beyrami SMG. Neurodegenerative diseases detection using distance metrics and sparse coding: a new perspective on gait symmetric features [J]. Computers in Biology and Medicine, 2020, 120: 103736.
- [31] Dutta S, Ghosh D, Samanta S. Nonlinear approach to study the dynamics of neurodegenerative diseases by Multifractal Detrended Crosscorrelation Analysis—a quantitative assessment on gait disease [J]. Physica A: Statistical Mechanics and its Applications, 2016, 448: 181-195.
- [32] Pham PT. Texture classification and visualization of time series of gait dynamics in patients with neuro-degenerative diseases [J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2017, 26(1): 188-196.
- [33] Goldberger AL, Amaral LAN, Glass L, et al. PhysioBank, PhysioToolkit, and PhysioNet:

components of a new research resource for Complex Physiologic Signals [J]. Circulation, 2000, 101(23): e215-e220.

- [34] Hausdorff JM, Cudkowicz ME, Firtion R, et al. Gait variability and basal ganglia disorders: stride-tostride variations of gait cycle timing in Parkinson's disease and Huntington's disease [J]. Movement Disorders, 1998, 13(3): 428-437.
- [35] Kennel MB, Brown R, Abarbanel HDI.
   Determining embedding dimension for phase-space reconstruction using a geometrical construction [J].
   Physical Review A, 1992, 45(6): 3403.
- [36] Fraser AM, Swinney HL. Independent coordinates for strange attractors from mutual information [J]. Physical Review A, 1986, 33(2): 1134.
- [37] Liebert W, Schuster HG. Proper choice of the time delay for the analysis of chaotic time series [J]. Physics Letters A, 1989, 142(2-3): 107-111.
- [38] Grassberger P, Procaccia I. Measuring the

strangeness of strange attractors [M]. The theory of chaotic attractors. Springer, New York, 2004: 170-189.

- [39] Rosenstein MT, Collins JJ, De Luca CJ. Reconstruction expansion as a geometry-based framework for choosing proper delay times [J]. Physica D: Nonlinear Phenomena, 1994, 73(1-2): 82-98.
- [40] Pecora LM, Moniz L, Nichols J, et al. A unified approach to attractor reconstruction [J]. Chaos: an Interdisciplinary Journal of Nonlinear Science, 2007, 17(1): 013110.
- [41] Rieck B. Persistent homology in multivariate data visualization [D]. Heidelberg: Heidelberg University, 2017.
- [42] Emrani S, Gentimis T, Krim H. Persistent homology of delay embeddings and its application to wheeze detection [J]. IEEE Signal Processing Letters, 2014, 21(4): 459-463.