

一种面向非平衡步态数据的帕金森病诊断方法

吴 玺¹, 张 永², 陈 绪², 许胜强³, 王 训⁴

WU Xi¹, ZHANG Yong², CHEN Xu², XU Shengqiang³, WANG Xun⁴

1. 合肥工业大学 计算机与信息学院, 合肥 230009

2. 合肥工业大学 工业与装备技术研究院, 合肥 230009

3. 中国科学院 合肥智能机械研究所, 合肥 230031

4. 安徽中医药大学 神经病学研究所附属医院, 合肥 230061

1. School of Computer and Information, Hefei University of Technology, Hefei 230009, China

2. Institute of Industry & Equipment Technology, Hefei University of Technology, Hefei 230009, China

3. Institute of Intelligent Machines, Chinese Academy of Sciences, Hefei 230031, China

4. Hospital Affiliated to Institute of Neurology, Anhui University of Chinese Medicine, Hefei 230061, China

WU Xi, ZHANG Yong, CHEN Xu, et al. Diagnostic approach of Parkinson's disease for imbalanced gait data. *Computer Engineering and Applications*, 2018, 54(9): 218-223.

Abstract: Dyskinesia is the significant feature of Parkinson's disease patient, therefore the data analysis on parkinsonian gait kinematic parameters can provide a strong basis for disease diagnosis and rehabilitation. The actual number of Parkinson's disease is far less than the normal human. The traditional machine learning methods are not suitable to classify the imbalanced data directly. This paper proposes a Cost Sensitive Support Vector Machine (CS-SVM) approach to distinguish between healthy human and patients with Parkinson's disease firstly. The entire subject's gait data is extracted from real U-shape electronic walkway. The classification performance is also improved by transforming the extracting features into non-dimensional forms. The experimental results indicate that the CS-SVM method can achieve the prediction accuracy of 94.16%, and the F-measure value reaches 87.08%, while it achieves a better performance comparing traditional SVM and the recognition performance is significantly improved by eliminating the influence of height, the prediction accuracy and F-measure value reach 94.81% and 88.66% respectively.

Key words: gait analysis; imbalanced data; electronic walkway system; Parkinson's disease; Cost Sensitive Support Vector Machine (CS-SVM)

摘 要: 运动障碍是帕金森病(PD)患者的重要特征, 步态信号分析可以为疾病诊断和康复治疗提供有力依据。现实中PD患者数量远小于正常人群, 传统的机器学习方法不适合对正例样本数远多于反例的非平衡数据进行分类。为了准确地区分出PD患者和健康人, 使用一种代价敏感支持向量机(CS-SVM)的方法来构建PD患者和健康人之间的步态信号分类模型。所有受试者的步态运动学特征数据是采用真实的U型电子步道系统提取的, 并将特征数据转化为无量纲的形式来消除身高对时空属性的影响。实验结果表明使用这种CS-SVM方法得到的预测准确率和F-measure值分别达到了94.16%和87.08%, 与传统的SVM方法相比性能更优。同时消除身高对时空属性的影响可以大幅提高识别性能, 预测准确率和F-measure值分别达到94.81%和88.66%。

关键词: 步态分析; 非平衡数据; 电子步道系统; 帕金森病; 代价敏感支持向量机

文献标志码: A **中图分类号:** TP391 **doi:** 10.3778/j.issn.1002-8331.1612-0225

基金项目: 国家高等学校学科创新引智计划("111")(No.B14025); (佛山市科技创新团队资助)可穿戴设备研发创新团队(No.2015IT100095); 广东省科技厅重大项目(No.2016B010108002); 安徽省科技攻关计划项目(No.1501021042, No.1301042215)。

作者简介: 吴玺, 博士, 研究方向为信号处理与系统、机器学习, E-mail: wuxi_xizi@sina.com; 张永, 硕士; 陈绪, 硕士; 许胜强, 博士; 王训, 教授。

收稿日期: 2016-12-14 **修回日期:** 2017-02-06 **文章编号:** 1002-8331(2018)09-0218-06

CNKI网络出版: 2017-05-10, <http://kns.cnki.net/kcms/detail/11.2127.TP.20170510.1116.014.html>

1 引言

帕金森病(PD)是一种锥体外系疾病,也称为运动障碍疾病,主要症状表现为运动迟缓、肌僵直、静止性震颤和姿势不稳。帕金森病造成的运动障碍往往会呈现一种特殊的步态特征,如拖步、慌张步态和步态冻结等^[1-2]。临床诊断中,医生通常利用上述步态特征作为帕金森病的诊断线索,因此对帕金森病患者的步态信号特征进行分析可以为疾病前期诊断和康复治疗提供有力的依据。

早期帕金森病诊断通常采用量表形式,如帕金森病统一评分量表(UPDRS)和Weber's分级量表^[3],这种量表通过设计针对性的动作和问题对患者的步态等特征进行直接观察,并对各项指标进行主观评分。这种诊断方式主观性强,依赖于医生经验,诊断的准确性和稳定性不高^[4]。

近年来机器学习技术的发展,为步态信号定性和定量的分析提供了重要的手段^[5-6]。如Begg等人^[7]采用支持向量机自动识别年龄与步态特征之间的变化规律,其使用了三种类型的步态变量:基本时空参数、运动属性和动力学属性。当人正常行走时,使用一种同步PEAK运动分析系统和一个平台记录并分析12位年轻人和12位老年人的步态信息。Klucken等人^[8]运用一种基于可移动的、嵌入生物传感器的人工智能步态分析(eGaIT)系统对帕金森病患者的患病阶段和运动状态进行自动识别。上述研究中分别选择帕金森病患者和老年人组与具有相似物理特征的对照组作为训练集,并且所选择的正、反两类样本大小基本相同,属于平衡样本数据分析。但是在帕金森病的监测中,帕金森病患者的数量远少于健康人群,因此实际中获得的样本为非平衡数据样本,并且将帕金森病患者判别为健康人的代价和将健康人判别为帕金森病患者的代价不同,前者导致延误治疗的最佳时机甚至危害生命,而后者以再次检查或药物副作用为代价,显然将帕金森病患者判定为健康人群的代价更大,因此使用平衡样本学习到的模型不能解决实际问题。

本文的研究目的是针对由帕金森病患者和正常人群组成的非平衡步态数据集,应用机器学习的方法构建识别模型,然后应用该识别模型来区分帕金森病患者和正常人群,从而为临床医生诊断疾病提供客观依据。研究过程是首先由U型电子步道提取出真实的帕金森病患者和健康人群的各种步态时空参数,并将上述步态时空参数和受试者的物理特征构建出非平衡特征数据样本,然后采用一种代价敏感支持向量机(CS-SVM)的方法来处理非平衡样本数据,建立自动识别帕金森病患者的分类模型。同时利用受试者的身高对相应样本的步态信号时空属性进行去量纲处理来消除身高差异对步态信号时空属性的影响,从而达到进一步提高预测模型

识别性能的目的。研究结果表明,本文构建的模型可以对帕金森病进行自动诊断,可以作为临床医生对帕金森病患者诊断的有效辅助工具。

2 材料和方法

2.1 数据获取和特征

本文使用一套由中国科学院合肥智能机械研究所运动与健康信息技术研究中心自主设计研发的电子步道,该电子步道由基于柔性阵列力敏传感器的14块压力垫(大小60 cm×60 cm,4压力点/cm²,采样频率100 Hz)、5块压力台(大小60 cm×60 cm,采样频率500 Hz)以及1块平衡测试仪组成,如图1所示。图1(a)是该U型电子步道的结构模拟图,图1(b)是U型电子步道的实物图。

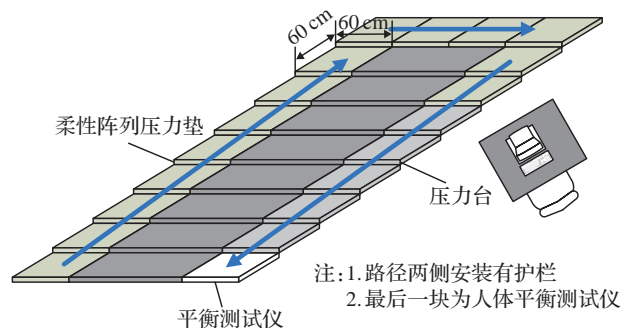


图1(a) U型电子步道模拟图



图1(b) U型电子步道实物图

所有受试的帕金森病患者是来自安徽中医药大学神经病学研究所附属医院的志愿者。对照测试人员没有未知的影响步态的损伤或异常,受试者24 h内未服用药物。所有受试者签订了知情同意书。35位患者和119位正常人被要求以正常步速行走在电子步道上,同时所有受试者的测试均在两位医生的监督下完成。受试者的物理特征如表1所示。依据由美国加利福尼亚州Rancho LosAmigos(RLA)医学中心提出的步态量化参数以及量表UPDRS-Ⅲ,利用U型电子步道系统提取直行部分的步态特征数据,如表2所示。

表1 测试者物理特征(均值±标准差)

变量	PD组(35)	NP组(119)
年龄(yrs)	64.86±9.92	56.93±8.71
性别(m:f)	17:18	38:81
体重/kg	63.80±12.74	61.90±11.22
身高/cm	160.20±7.90	161.08±8.44

注:m=male;f=female;yrs=years。

表2 提取的特征变量

特征类型(数量)	特征(单位)
时空特征(16)	步长*(cm)、步幅*(cm)、步速*(m·s ⁻¹)、步频*(step·s ⁻¹)、站立时间*(ms)、摆动时间*(ms)、摆动前期时间*(ms)、周期(ms)、双支撑时间(ms)

注:标记*的特征均分为左脚特征和右脚特征。

2.2 数据预处理

对步态特征数据的预处理过程主要分为两步:消除身高差异对时空参数的影响和特征数据的标准化。

(1) 消除身高差异对时空参数的影响

通常人行走时的步长、步幅、步速、周期等步态信号特征和身高都有一定关联。例如身高比较矮的人,其步长和步幅会相对较短,而其步频会相对比较高。因此为了消除受试者身高差异对于步态时空参数的影响,本文利用下列公式消除身高差异对时空参数的影响。

$$\hat{l} = l/l_0 \quad (1)$$

$$\hat{v} = \frac{v}{\sqrt{gl_0}} \quad (2)$$

$$\hat{f} = \frac{f}{\sqrt{g/l_0}} \quad (3)$$

$$\hat{t} = \frac{t}{\sqrt{l_0/g}} \quad (4)$$

公式(1)~(4)中, l 表示步长和步幅; t 表示时间特征,包括站立时间、摆动时间、摆动前期、周期、双支撑时间; f 表示步频; v 表示步速; l_0 表示每位受试者的身高;同时引入一个重力加速度常量 $g=9.81 \text{ m/s}^2$ 来消除身高差异对步速、步频以及时间属性的影响。

(2) 标准化

在应用分类器前,将步骤(1)处理后的特征和物理特征利用式(5)将每一特征转化为相同的统计分布。

$$\hat{x} = \frac{x - \mu}{\sigma} \quad (5)$$

式(5)中, x 表示步骤(1)获得的特征变量和物理特征, μ 和 σ 分别表示每个特征变量的均值和标准差。

2.3 分类方法

支持向量机(SVM)是由 Vapnik 提出的一种基于统计学理论的用于解决分类和回归问题的机器学习方法^[9]。对于二分类问题, SVM 分类算法的主要目标是寻找一个最优分类超平面。SVM 是通过最大化两类之间的边界距离来寻找最优超平面。假设存在一个两类的训练样本空间 R^n , 其中包含样例 $x_i \in R^n, i=1, 2, \dots, n$, 和一个与样例 x_i 一一对应的向量 $y_i \in R^n, y_i \in \{1, -1\}$ 。SVM 解决分类问题需要优化下式^[10-11]:

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i$$

$$\text{s.t. } y_i[(w^T x_i) + b] \geq 1 - \xi_i \quad i=1, 2, \dots, n \quad \xi_i \geq 0 \quad (6)$$

式(6)中 C 是误差的惩罚参数,其取值大于0。此外,还

需要选择一个核函数,将原样本中线性不可分的低维空间特征向量映射到一个高维空间中,使样本线性可分。本文选择普遍使用的径向基核函数(RBF):

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2) \quad (7)$$

公式(7)中 γ 是核函数参数。

上述传统的 SVM 算法是以分类精度为优化目标,假定每类样本的大小基本相同。但是,在实际中,数据集通常是非平衡数据。例如帕金森患者的数量是远少于正常人群。例如在 100 人的测试样本中有 99 位正常人,仅有 1 位为帕金森病患者。若分类器把所有样本都划分为正常人,此分类器能够获得 99% 的分类精度^[12]。虽然分类精度很高,但是使用这种分类模型来识别帕金森病患者是没有意义的。在诊断过程中,将帕金森病患者误诊为正常人的代价和将正常人误诊为帕金森病患者的代价是不同的。前者使帕金森病患者失去治疗机会,造成病情恶化甚至危及生命,而后者以再次诊断或药物的副作用为代价。显然,将帕金森病患者判定为正常人的误分类代价要远大于将正常人判定为帕金森病患者。因此,传统的 SVM 算法不适合处理这种非平衡数据。

本文中采集的受试者步态数据集为非平衡数据集。传统 SVM 对此类非平衡数据集学习的边界会更接近于正例(样本量较少的类),从而导致分类精度倾向于样本量较大的类,因而需要使用一种有偏的支持向量机对步态数据集进行分类,将分类边界“拖离”正例。为了准确区分帕金森病患者和正常人群,本文使用 CS-SVM 来构建帕金森病患者和正常人群的步态分类模型。采用的 CS-SVM 算法原理是对传统的 SVM 的目标函数做出修改,即正例和反例设置不同的惩罚参数,分别表示为 C_+ 和 C_- 。CS-SVM 的目标函数如下^[13-14]:

$$\min \frac{1}{2} \|w\|^2 + C_+ \sum_{i: y_i = 1} \xi_i + C_- \sum_{i: y_i = -1} \xi_i$$

$$\text{s.t. } y_i[(w^T x_i) + b] \geq 1 - \xi_i \quad i=1, 2, \dots, n \quad \xi_i \geq 0 \quad (8)$$

CS-SVM 算法对正例和反例设置不同的惩罚参数来处理非平衡数据集,这样就可以将分类超平面“拖离”正例,然后调整两类的惩罚参数和核函数参数来优化公式(8)从而获得最优的分类超平面。

2.4 评价指标

在二分类问题中,分类器标签分为正例和反例。分类器的结果可以表示为混淆矩阵或一致性表^[15],用于评价像支持向量机(SVM)、代价敏感支持向量机(CS-SVM)和 k 阶近邻(KNN)这样的监督学习算法的性能。表 3 给出的是帕金森病患者预测的混淆矩阵,其中 TP 表示真实为帕金森病患者,预测也为帕金森病患者的样例个数;FP 表示真实为健康人群,而预测为帕金森病患者的样例个数;FN 表示真实为帕金森病患者,而预测为健康

人群的样例个数;TN表示真实为健康人群,预测也为健康人群的样例个数。

表3 帕金森病预测的混淆矩阵

预测类别	真实类别	
	帕金森病患者	健康人群
帕金森病患者	TP	FP
健康人群	FN	TN

为了评估构建的帕金森病预测模型的性能,本文使用下列指标。

(1)准确率(Accuracy):

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \times 100\% \tag{9}$$

公式(9)为准确率的计算公式,表示正确分类的整体准确率。由此式本文计算的是正确分类的帕金森病患者和健康人群的样例个数与总样例个数的比值。

(2)召回率(Recall):

$$Recall = \frac{TP}{TP + FN} \times 100\% \tag{10}$$

式(10)为召回率的计算公式,表示被正确判定的正例占总体正例的比例。由此式本文计算的是正确分类的帕金森病患者样例个数与总的帕金森病患者样例个数的比值。

(3)精确度(Precision):

$$Precision = \frac{TP}{TP + FP} \times 100\% \tag{11}$$

公式(11)为精确度的计算公式,表示被分类器判定的正例中真正的正例样本所占的比例。由此式本文计算的是正确分类的帕金森病患者样例个数与全部预测为帕金森病患者样例个数的比值。

(4)F-measure值:

$$F - measure = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{12}$$

式(12)为F-measure值的计算公式,表示召回率和精确度的加权调和平均数。F-measure值表示召回率和精确度的综合评价指标,当F-measure值越高,则说明分类实验方法更加有效。

3 实验评估

3.1 实验设置

本研究实验设置为两部分:步态特征差异性检验和分类实验。在3.2节给出利用T-test检验帕金森病患者和正常人群之间步态特征的差异性分析;3.3节给出分类实验结果及评价。在分类实验中分为两组实验,实验1:未采用2.2节步骤(1)介绍的方法消除身高差异对时空参数的影响;实验2:采用2.2节步骤(1)介绍的方法来消除身高差异对时空参数影响。对于每组实验使用三折交叉验证来对比验证SVM和CS-SVM两种方法的识别性能,同时对比两组实验的分类效果。步态数据集被分为大小相同的三个子集。其中一个子集用于评估模型,其余子集用于分类器模型的训练和参数优化。然后计算2.4节给出的准确率、召回率、精确度和F-measure四种评估指标的均值和标准差来评价构建的预测模型。

在2.3节分类方法中介绍到本文使用的算法SVM算法需要设置惩罚参数C和核函数参数g,而CS-SVM算法需要设置惩罚参数C₊、C₋和核函数g。本文中SVM算法的参数为:C=4,g=0.0313;CS-SVM算法的参数为:C₊=13.6、C₋=3.4、g=0.25。

3.2 实验结果及评价

表4给出帕金森病患者和正常人群的步态参数之间的统计及差异性检验,图2给出比较两组步态参数均值的对比。

表4 步态特征统计(均值±标准差)

步态特征	帕金森病组	对照组	p-value	95%CI
步长#/cm	42.19±11.09	56.74±7.12	<0.01	[12.33,18.77]
步幅#/cm	81.91±23.56	112.76±14.15	<0.01	[24.51,37.21]
步速#/(m·s ⁻¹)	0.71±0.24	0.99±0.17	<0.01	[0.22,0.37]
步频#/(step·s ⁻¹)	0.83±0.13	0.88±0.11	0.03	[0.004,0.09]
站立时间#/ms	835.33±217.47	712.91±111.48	<0.01	[-176.58,-68.25]
摆动时间#/ms	398.48±47.54	429.71±50.24	<0.01	[12.38,50.09]
摆动前期时间#/ms	207.49±114.70	132.08±36.93	<0.01	[-99.44,-51.37]
步长*/cm	41.87±10.89	56.74±7.48	<0.01	[11.69,18.05]
步幅*/cm	83.34±22.89	113.78±14.34	<0.01	[24.12,36.76]
步速*/(m·s ⁻¹)	0.72±0.25	1.02±0.18	<0.01	[0.23,0.38]
步频*/(step·s ⁻¹)	0.86±0.12	0.92±0.10	0.03	[0.004,0.08]
站立时间*/ms	789.94±191.69	695.78±118.28	<0.01	[-146.62,-41.67]
摆动时间*/ms	400.26±51.29	420.46±40.90	0.02	[3.69,36.70]
摆动前期时间*/ms	195.47±101.67	131.18±41.64	<0.01	[-87.26,-41.31]
周期/ms	1 204.88±203.91	1 135.19±159.34	0.03	[-134.40,-4.99]
双支撑时间/ms	402.95±211.53	263.24±77.01	<0.01	[-185.62,-93.76]

注:标记为#表示为左脚时空特征;标记为*表示为右脚时空。

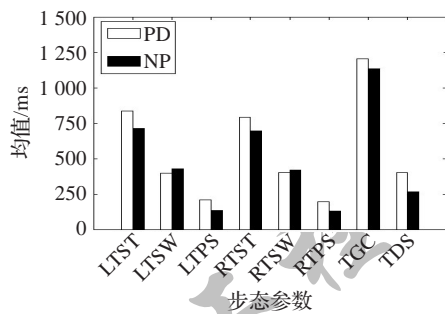


图2(a) PD和NP组的时间参数均值对比

从表4中发现帕金森病患者的步态参数:步长、步幅、步速、左脚摆动时间、摆动前期时间以及双支撑时间与正常人群的相应步态时空参数具有显著的差异(p -value<0.01),并且步频、站立时间、右脚摆动时间、周期也具有显著差异(p -value<0.05)。从图2中也可以看出提取的PD和NP组的长度和时间参数均值均具有明显的差异。因此实验中使用U型电子步道系统采集的帕金森病患者和正常人群的步态特征具有显著性差异,所以提取的步态特征是合理的,能够用于区分帕金森病患者和正常人群。图2(a)中LTST、LTSW、LTPS、RTST、RTSW、RTPS、TGC、TDS分别表示左脚站立时间、左脚摆动时间、左脚摆动前期、右脚摆动时间、右脚站立时间、右脚摆动前期、周期、双支撑时间;图2(b)中LSL、LSTL、RSL、RSTL分别表示左脚步长、左脚步幅、右脚步长、右脚步幅。

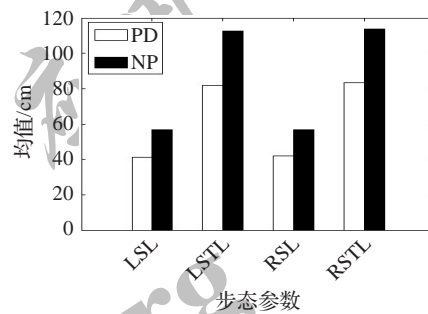


图2(b) PD和NP组的长度参数均值对比

的识别准确率相比于SVM分类器提高了约1.96%;CS-SVM的召回率达到了88.89%,比SVM分类器提高了约17.18%,因此相比SVM算法,CS-SVM算法构建的步态模型对帕金森病患者的误分类率降低了17.18%,对帕金森病患者的误判风险降低;同时CS-SVM分类器的F-measure值达到88.66%,明显高于SVM分类器,约提高6.76%。因此,综合考虑识别准确率、召回率和F-measure值,相比于SVM分类算法,CS-SVM分类器对帕金森病患者的步态特征识别性能更好。

对比表5和表6可以看出,当消除受试者身高差异对步态时空特征影响时,SVM和CS-SVM分类器的识别准确率分别为92.85%和94.81%,相比于未消除身高差异对时空属性影响时,SVM和CS-SVM算法对帕金森病患者和正常人群的步态分类准确率分别提高了约1.3%和0.65%;而且相比于未消除身高差异对步态特征影响时,针对消除受试者身高对步态特征影响的样本,SVM算法的召回率达到71.71%,提高了约3.02%,SVM算法对帕金森病患者的误分类率降低了3.02%,SVM分类器对帕金森病患者的误判风险降低;同时F-measure值达到81.90%,提高了约2.92%。CS-SVM算法的召回率达到88.89%,提高了约3.03%,CS-SVM算法对帕金森病患者步态特征的误分类率降低了3.03%,CS-SVM分类器对帕金森病患者的误判风险降低,而且CS-SVM分类器的F-measure值达到88.66%,提高了约1.58%。因此综合考虑识别准确率、召回率和F-measure值,相比于未消除身高差异对步态特征影响的样本实验,通过消除身高差异对步态特征的影响可以大幅提高帕金森病患者的识别性能。

表5 对未消除身高差异影响的样本,使用SVM和CS-SVM的预测结果(均值±标准差) %

算法	准确率	召回率	精确度	F-measure
SVM	91.55±3.01	68.69±3.50	93.33±11.55	78.98±5.81
CS-SVM	94.16±1.96	85.86±9.41	89.83±10.76	87.08±3.36

表6 对消除身高差异影响的样本,使用SVM和CS-SVM的预测结果(均值±标准差) %

算法	准确率	召回率	精确度	F-measure
SVM	92.85±4.10	71.71±12.08	95.83±7.22	81.90±10.53
CS-SVM	94.81±2.98	88.89±12.73	90.08±10.80	88.66±6.08

3.3 分类及评价

表5和表6给出实验1和实验2的分类结果。

实验1结果分析:从表5中可以看出,针对未消除身高差异对步态特征影响的样本进行识别时,SVM和CS-SVM分类器的识别准确率分别达到91.55%和94.16%,相比SVM模型,使用CS-SVM构建分类器的识别准确率更高,提高了约2.61%;召回率也更高,达到85.86%,提高了约17.17%,因此CS-SVM分类器对帕金森病患者的步态特征的识别效果更好。

实验2结果分析:从表6中可以看出,对消除身高差异对步态特征影响的样本,SVM和CS-SVM分类器的识别准确率分别达到92.85%、94.81%。CS-SVM分类器

4 结束语

本文采用自主研发的U型电子步道系统提取帕金森病患者和健康人群的步态信号时空特征,同时结合受试者的物理特征组成非平衡步态数据集,并利用T检验验证PD患者和健康人群步态特征的差异性和合理性,然后采用代价敏感支持向量机(CS-SVM)和传统支持向量机(SVM)分别来构建帕金森病患者和健康人群的步态特征分类模型,最后通过受试者的身高对步态特征进行去量纲处理,来消除身高差异对步态特征的影响,

达到了提高预测模型识别性能的目的。实验结果表明,相比于SVM算法,CS-SVM算法对于帕金森病患者的步态特征的识别性能更优且能达到94.16%的全局预测准确率,有效地降低了帕金森病的误判风险。并且消除了身高的差异性对步态特征的影响后,CS-SVM预测模型的全局识别准确率达到94.81%,相比不去量纲的预测模型,预测准确率提高了约0.65%。因此本文基于U型电子步道系统提取的非平衡步态信号数据并采用CS-SVM算法构建的预测模型可以为临床医生诊断帕金森病提供一个可靠的辅助决策工具。

参考文献:

- [1] Taktak A F G. Clinical engineering: A handbook for clinical and biomedical engineers[M]. [S.l.]: Academic Press, 2014.
- [2] Santens P, Boon P, Van Roost D, et al. The pathophysiology of motor symptoms in Parkinson's disease[J]. *Acta Neurologica Belgica*, 2003, 103(3): 129-134.
- [3] Morris T R, Cho C, Dilda V, et al. Clinical assessment of freezing of gait in Parkinson's disease from computer-generated animation[J]. *Gait & Posture*, 2013, 38(2): 326-329.
- [4] 顾磊, 吴慧中, 肖亮. 一种基于人体轮廓宽度特征的步态识别方法[J]. *计算机工程与应用*, 2007, 43(24): 4-6.
- [5] Sosnoff J J, Klaren R E, Pilutti L A, et al. Reliability of gait in multiple sclerosis over 6 months[J]. *Gait & Posture*, 2015, 41(3): 860-862.
- [6] Bae J, Kong K, Byl N, et al. A mobile gait monitoring system for gait analysis[C]//IEEE International Conference on Rehabilitation Robotics, 2009: 73-79.
- [7] Begg R, Kamruzzaman J A. A machine learning approach for automated recognition of movement patterns using basic, kinetic and kinematic gait data[J]. *Journal of Biomechanics*, 2005, 38(3): 401-408.
- [8] Klucken J, Barth J, Kugler P, et al. Unbiased and mobile gait analysis detects motor impairment in Parkinson's disease[J]. *PloS One*, 2013, 8(2): e56956.
- [9] Chang C C, Lin C J. LIBSVM: A library for support vector machines[J]. *ACM Transactions on Intelligent Systems & Technology*, 2011, 2(3).
- [10] Vapnik V. The nature of statistical learning theory[M]. [S.l.]: Springer Science & Business Media, 2013.
- [11] 张学工. 关于统计学习理论与向量机[J]. *自动化学报*, 2000, 26(1): 32-42.
- [12] 郑恩辉, 李平, 宋执环. 代价敏感支持向量机[J]. *控制与决策*, 2006, 21(4): 473-476.
- [13] Cao P, Zhao D, Zaiane O. An optimized cost-sensitive SVM for imbalanced data learning[M]//Advances in Knowledge Discovery and Data Mining. Berlin Heidelberg: Springer, 2013: 280-292.
- [14] M'hamed Abidine B, Fergani B, Oussalah M, et al. A new classification strategy for human activity recognition using cost sensitive support vector machines for imbalanced data[J]. *Kybernetes*, 2014, 43(8): 1150-1164.
- [15] Wang A G, An N, Chen G, et al. Predicting hypertension without measurement: A non-invasive, questionnaire-based approach[J]. *Expert Systems with Applications*, 2015, 42(21): 7601-7609.

(上接125页)

- [6] Zhou L G, Chen H Y. A generalization of the power aggregation operators for linguistic environment and its application in group decision making[J]. *Knowledge-Based Systems*, 2012, 26(2): 216-224.
- [7] Herrera F, Martinez L A. A 2-Tuple fuzzy linguistic representation model for computing with words[J]. *IEEE Transactions on Fuzzy Systems*, 2000, 8(6): 746-752.
- [8] 姜艳萍, 樊治平. 二元语义信息集成算子的性质分析[J]. *控制与决策*, 2003, 18(6): 754-757.
- [9] 姜艳萍, 樊治平. 基于不同粒度语言判断矩阵的群决策方法[J]. *系统工程学报*, 2006, 21(3): 249-253.
- [10] Xu Y J, Wang H M. Approaches based on 2-tuple linguistic power aggregation operators for multiple attribute group decision making under linguistic environment[J]. *Applied Soft Computing*, 2011, 11(5): 3988-3997.
- [11] Wang J Q, Wang D D, Zhang H Y, et al. Multi-criteria group decision making method based on interval 2-tuple linguistic information and Choquet integral aggregation operators[J]. *Soft Computing*, 2015, 19(2): 389-405.
- [12] Merigo J M, Gil-Lafuente A M. Induced 2-tuple linguistic generalized aggregation operators and their application in decision-making[J]. *Information Sciences*, 2013, 236(1): 1-16.
- [13] 刘金培, 林盛, 陈华友. 二元语义 Bonferroni 集成算子及其在多属性群决策中的应用[J]. *运筹与管理*, 2013, 22(5): 122-127.
- [14] Zadeh L A. The concept of a linguistic variable and its application to approximate reasoning Part I[J]. *Information Sciences*, 1975, 8(2): 99-249.
- [15] Zhang H M. The multiattribute group decision making method based on aggregation operators with interval-valued 2-tuple linguistic information[J]. *Mathematical and Computer Modelling*, 2012, 56(1): 27-35.
- [16] 林建, 兰继斌, 林耀海. 基于区间二元语义集结算子的多属性群决策方法[J]. *吉林师范大学学报*, 2009, 30(1): 5-9.
- [17] Yager R R. On generalized Bonferroni mean operators for multi-criteria aggregation[J]. *International Journal of Approximate Reasoning*, 2009, 50(8): 1279-1286.
- [18] Yager R R. On ordered weighted averaging aggregation operators in multicriteria decision making[J]. *IEEE Transactions on Systems, Man, and Cybernetics*, 1988, 18(1): 183-190.