

基于心电-脉搏波的心血管疾病识别研究

陈倩蓉¹, 梁永波², 赵飞骏², 朱健铭², 陈真诚²

1. 桂林电子科技大学电子工程与自动化学院, 广西 桂林 541004; 2. 桂林电子科技大学生命与环境科学学院, 广西 桂林 541004

【摘要】为实现心血管疾病的早期筛查,降低心血管疾病临床检测的成本。本研究基于上肢脉搏波传导速度(PWV)及脉搏波相关血液动力学基础理论,采集了总计51人的脉搏波与心电信号数据,提取了包括3种PWV和脉搏波特征参数总计16个特征参数,将不同的PWV与脉搏波特征组成3个样本特征数据集,分别建立了基于K近邻学习(KNN)和支持向量机(SVM)的心血管疾病识别模型。KNN模型分类准确率为66.28%,SVM模型分类准确率为84.3%,并通过对比不同PWV对模型性能的影响,确定了用于血管评估的最优脉搏波传导速度pwvm。研究表明基于SVM建立的分类模型对心血管疾病识别有一定可靠性,为低成本的心血管疾病早期筛查提供了新思路,也为穿戴式心血管系统监测提供了基础。

【关键词】心血管疾病;脉搏波;K近邻学习;支持向量机

【中图分类号】R318;R543

【文献标志码】A

【文章编号】1005-202X(2019)02-0210-05

Cardiovascular disease recognition based on electrocardiogram data and pulse wave

CHEN Qianrong¹, LIANG Yongbo², ZHAO Feijun², ZHU Jianming², CHEN Zhencheng²

1. School of Electronic Engineering and Automation, Guilin University of Electronic Technology, Guilin 541004, China; 2. School of Life and Environmental Sciences, Guilin University of Electronic Technology, Guilin 541004, China

Abstract: To achieve an early screening of cardiovascular diseases and reduce the cost of clinical detection of cardiovascular disease, a research based on pulse wave velocity (PWV) in the upper extremities and the hemodynamic theory related to pulse waves is performed. The pulse waves and electrocardiogram data of 51 volunteers were collected, and 16 feature parameters including 3 types of PWV and pulse wave features were extracted. Three sample feature data sets which are composed of different PWV and pulse wave features are used to establish two different cardiovascular diseases recognition models based on K-nearest neighbor (KNN) or support vector machine (SVM). The classification accuracy of KNN and SVM models is 66.28% and 84.3%, respectively. By comparing the effects of different PWV on the performance of models, the optimal pwvm for vascular assessment is determined. The research results show that the SVM model is reliable in the cardiovascular disease recognition, providing a new idea for the low-cost and early screening of cardiovascular diseases and providing a basis for wearable cardiovascular system monitoring.

Keywords: cardiovascular disease; pulse wave; K-nearest neighbor; support vector machine

【收稿日期】2018-08-06

【基金项目】国家自然科学基金重大科研仪器研制项目(61627807);广西自然科学基金(2017GXNSFGA198005);国家重点研发计划课题(2016YFC1305703);广西自然科学基金青年基金项目(2016GXNSFBA380145);广西自动检测技术与仪器重点实验室主任基金(YQ17118);2015年广西信息科学实验中心一般项目(YB1513);桂林电子科技大学研究生教育创新计划资助项目(2016YJCB01)

【作者简介】陈倩蓉,在读研究生,主要研究:向为生物传感与智能仪器,E-mail: chenqianro@163.com

【通信作者】陈真诚,教授,博士生导师,主要研究方向:生物传感与智能仪器,E-mail: chenzhcheng@163.com

前言

据统计,2012年我国农村居民的心血管疾病死亡人数占总死亡人数的41.09%,到2016年达到了45.01%;城市居民的占比也从41.52%增长到了42.61%^[1-2],心血管疾病死亡人数连续几年占全球死亡人数的首位。动脉硬化是诸多心血管疾病的共同病理基础和独立预测因素^[3-4],也是唯一最早出现的能够直接检测到的血管功能结构变化,它会导致动脉血管壁逐步增厚,逐渐失去弹性直至管腔狭窄、堵塞,引发心血管疾病^[5]。目前临床上检查心血管疾病的方法主要有生化分析和影像检查^[6]。生化分析是

一种间接测量的有创方法。影像检查方法准确度高,但无法检测出早期病变,当冠脉出现堵塞时才能确诊。上述两种方法不仅费用昂贵,而且不适用于普查。因此,心血管疾病的识别对心血管疾病的防控有重要意义,也是现代社会对健康监测的迫切需求。

近年来,国内外许多学者通过脉搏波传导速度(Pulse Wave Velocity, PWV)检测动脉硬化,期望通过检测动脉硬化实现心血管疾病的早期筛查。Alty等^[7]通过单点的指端脉搏波波形估计PWV,依据PWV对样本进行分类,有研究者通过指端光电容积脉搏波形态学特征评估血管生理状况^[8]。张丽娜等^[9]也研究了基于心电、脉搏波信号的动脉硬化无创检测方法,建立了基于自适应神经模糊推理系统的动脉硬化评价模型。日本有研究机构已经研制出动脉硬化仪,但由于其售价昂贵,没有得到广泛运用。

上述的研究中学者们通过动脉硬化筛查心血管疾病是可行的,但是由于动脉硬化程度随年龄的增长而加深,PWV也会随年龄的增长而加快,对不同年龄段的受试者使用同一个阈值是极不合理的。各个实验中获取PWV使用的脉搏波传导时间大相径庭,因此也就无法得知不同PWV对血管评估产生的影响。所以拥有一种低成本的心血管疾病早期筛查方法已经成为了现代社会的迫切需求。

因此,本研究提出一种低成本的适用于心血管疾病早期筛查的便捷方法。通过同时采集指端脉搏波信号与心电信号,提取3种PWV与心血管结构功能密切相关的脉搏波特征参数,基于含不同PWV的特征集分别建立心血管疾病识别模型,并通过对比PWV对血管评估的效果确定用于评估动脉血管功能结构的最优PWV。

1 PWV理论基础

目前,PWV是公认用于评估动脉硬化的指标,主要有两种测量方式获取PWV,其一是通过单点脉搏波与心电信号获取脉搏波传导时间,结合脉搏波传导路径获取PWV;其二是通过双端脉搏波获取脉搏波传导时间,结合脉搏波传导路径求得PWV。颈-股动脉PWV是评估动脉硬化的金标准,是通过上述方法二测量获取的^[10],但由于其测量区域的限制,没有得到广泛运用。刘娜等^[11]的研究表明单点PWV与冠状动脉狭窄数量密切相关,冠状动脉狭窄数量多少从侧面反映了动脉硬化程度,也证明了单点脉搏波速度与动脉硬化程度有密切联系。有研究表明动脉整体顺应性与颈-股动脉PWV的相关性很高^[12],动脉整体顺应性与动脉结构功能联系紧密,也表明身体

局部的PWV依旧可以反应身体的血管情况。

研究表明脉搏波及其微分波形上发现了与血管功能结构高相关性的变量,并发现了与动脉硬化密切联系的特征参数^[13-17],因此,本研究拟通过心电信号与单点脉搏波信号获取上肢PWV,基于脉搏波特征参数借助机器学习方法建立动脉心血管疾病识别模型,并通过模型对比确定用于心血管评估的最优PWV。

2 实验

2.1 采集系统

本研究的采集系统主要包括心电传感器HKD-10A,红外透射式脉搏波传感器HKG-07B,心电信号预处理模块、脉搏波预处理模块、控制器模块、串口与上位机的波形显示与数据存储模块。心电信号预处理模块包括光电隔离电路、放大及电平抬升电路。脉搏波信号预处理模块包括放大及电平抬升电路。通过标准I导联方法结合HKD-10A心电传感器采集心电信号,选用HKG-07B透射式红外脉搏传感器获取受试者左手食指容积脉搏波,系统框图如图1所示。脉搏波信号和心电信号采样设置为500 Hz,通过12位模数转换,采集的数据由串口发送至上位机,通过上位机中的LabVIEW界面观测波形,存储数据。

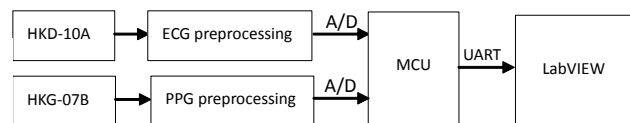


图1 采集系统

Fig.1 Acquisition system

PPG: Photoplethysmography; ECG: Electrocardiogram

2.2 采集流程

具体的数据采集流程如下:(1)控制室温在25℃左右,实验集中在春季和夏季进行,保持实验环境安静,实验前所有受试者静息20 min,并记录下受试者的年龄、身高、体质量等生理信息。(2)所有受试者都在平躺状态下同时检测左手食指的光电容积脉搏波和标准I导联下的心电信号,每个受试者都检测3组数据,每组1 min,实验过程中均保持沉默,身体保持平躺,身体各部位不得有任何位移。(3)考虑到血压计在测量过程中需要给袖带充气加压,造成血管变形,所以在第三组数据测试完成后用欧姆龙HEM-7201上臂式电子血压计在同一状态下测量3次血压。(4)将所测数据和生理信息归档保存。

2.3 数据统计

本研究共采集51人的波形数据并记录其生理信息,分为健康组与病患组两组,其中健康组有41人,病患组有10人。健康组是各项生理指标正常,没有

心脑血管疾病或者相关症状,病患组是已经被医生确诊患有心脑血管疾病如冠心病、高血压等疾病并且在服药治疗的患者。表1是受试者的基本信息。

表1 受试者基本信息统计

Tab.1 Basic statistical information of the volunteers

Physiological parameters	Health controls			Patients		
	Male	Female	Total	Male	Female	Total
Age/year	27.25±8.3	33.2±12.1	29.5±10.2	67.4±10.9	74.3±5.5	69.0±9.9
Systolic pressure/mmHg	122.7±5.6	106.8±9.3	116.3±10.9	138.2±6.6	137.3±2.5	138.0±5.6
Diastolic pressure/mmHg	72.3±7.9	65.9±6.3	69.8±8.1	76.5±4.5	71.6±10.4	75.1±6.57
Heart rate/times	69.1±8.6	70.6±10.1	69.7±9.1	71.7±7.1	82.3±2.5	74.9±7.8
Height/cm	173.7±5.8	162.5±6.2	169.4±8.1	174.1±3.1	162.1±2.1	170.5±6.4
Weight/kg	68.5±11.0	54.0±8.4	62.9±12.3	74.3±10.9	62.3±1.5	70.75±10.7

3 特征定义

如图2所示的是基于心电与容积脉搏波的特征参数定义,其中ECG为心电信号波形,PPG为脉搏波信号,VPG(Velocity Plethysmograph)为脉搏波的一次微分波形,APG(Accelerated Plethysmograph)是脉搏波的二次微分波形。图中ptto、ptta和pttm是3种脉搏波传导时间,由心电信号与脉搏波信号获取,图中的脉搏波形态学参数包括Tpp、Tc、CT、Tm_b、Tm_c、Tm_d、Tm_e、Tm_a_b、Tm_a_c、Tm_a_d、Tm_a_e、Tm_b_peak和AI,其中AI是y与x的比值,仅由上述13个脉搏波形态学参数组成的特征子集设为P,分别由ptto、ptta和pttm及受试者心脏到左手食指的距离得到pwvo、pwva和pwvm。pwvo与特征集P组成的特征集为Q_o,pwva与特征集P组成的特征集为Q_A,由pwvm与特征集P组成的特征集为Q_M。

4 数据分析

由上述特征提取部分知,每两个连续周期的脉搏波可以提取一组特征参数,每个受试者保留30组特征参数的记录,51个受试者总计有1530条特征参数记录,其中健康组1230条记录,病患组300条记录,健康组病患组数据量之比接近4:1,所以存在样本不平衡的问题,为解决样本不平衡的问题,并且避免过采样导致的过拟合问题,每次建模都在健康组随机抽取300条记录与病患组全部记录组成新数据集,基于新数据集进行建模。



图2 特征定义

Fig.2 Definition of features

ptta, pttm, ptto are 3 kinds of pulse wave transit times. Tpp, Tc, CT, Tm_b, Tm_c, Tm_d, Tm_e, Tm_a_b, Tm_a_c, Tm_a_d, Tm_a_e, Tm_b_peak, AI are pulse wave morphological parameters, and these parameters form the set p.

5 结果

本研究基于PWV和脉搏波形态学特征参数分别使用K近邻学习(K-Nearest Neighbor, KNN)与支持向量机(Support Vector Machine, SVM)两种方法对样本进行分类识别。KNN属于常用的监督学习方法,其核心思想是以某种距离度量寻找测试样本与训练样本最靠近的k个训练样本,将这k个样本中占比最

大的类别标记作为预测结果^[18]。对于KNN, k 是一个重要的参数, k 值的选择影响着分类器的精度, 所以首先要找出适合的 k 值。表2所示的是基于含不同PWV的特征集的最优 k 值对应的分类模型的准确率。

表2 基于不同特征集的KNN分类模型准确率
Tab.2 Accuracy of K-nearest neighbor model using different feature sets

Feature set	k	Accuracy
P	15	0.570 0
Q_A	17	0.574 9
Q_O	29	0.622 2
Q_M	19	0.662 8

pwvo, pwva, pwvm are 3 kinds of pulse wave velocities. P and pwvo forms the set Q_O ; P and pwva forms the set Q_A ; P and pwvm forms the set Q_M .

由表2可知加入 pwvo 和 pwvm 相比于 pwva 对分类准确率改善更明显, 一方面表明 pwvo 和 pwvm 的加入对分类器的准确性有明显改善, 另一方面说明 pwvo 和 pwvm 比 pwva 更能反映心血管生理状态, 3个PWV参数中, pwvm 是最适合用于评估心血管生理状态的PWV。

SVM分类的核心思想是在给定的训练样本空间找到一个可以把不同类别样本划分开的超平面。距离超平面最近的几个训练样本就是支持向量, 支持向量到超平面的距离称为间隔, 为了使分类器的泛化性能最好, 要找出使得间隔达到最大的超平面, 即为泛化性能最好的分类器^[18]。由于本研究的样本数据量较小, 所以选用SVM, 它对于样本量小, 非线性等问题能达到较好的分类效果, 实际问题中, 样本空间包括线性可分和线性不可分的情况, 对于线性不可分的情况, 若样本特征有限, 则可以将原样本特征映射到更高维的空间, 在映射的新空间中找合适的超平面。在这个过程中需要计算到原特征向量映射到高维空间的内积, 由于在高维空间计算非常困难, 解决方法是找到合适的核函数, 在原空间借助核函数计算在新空间中的内积。用含有 pwvm 的特征集 Q_M 在运用不同核函数时得到的分类器预测的准确率如表3所示, 由表3知, 使用不同核函数分类器的分类准确率差异不大, 当核函数为 Sigmoid 时, 分类准确率达到最高, 即 84.20%。

实际应用中查准率与查全率是评估分类器性能

表3 不同核函数的SVM分类模型的准确率
Tab.3 Accuracy of support vector machine model using different kernel functions

Kernel function	Accuracy/%
Linear	83.87
Polynomial	82.78
RBF	84.30
Sigmoid	84.20

的重要指标, 特别是在疾病识别方面的运用, 要在一定准确率的情况下尽可能提高查全率, P-R曲线以查全率为横轴, 以查准率为纵轴, 可以通过曲线直观反映出分类器的查准率与查全率, 图3是 k 为 19 时的 KNN 分类模型与核函数为 Sigmoid 的 SVM 分类模型基于不同PWV特征的P-R曲线。基于含有 pwvm 特征集 Q_M 的 SVM 分类模型性能最好, SVM 分类模型普遍比 KNN 分类模型准确率更高, pwvm 对于心血管疾病识别模型性能有更大的提升。因此选取基于含有 pwvm 的特征集 Q_M 的 SVM 分类模型作为心血管疾病识别模型。

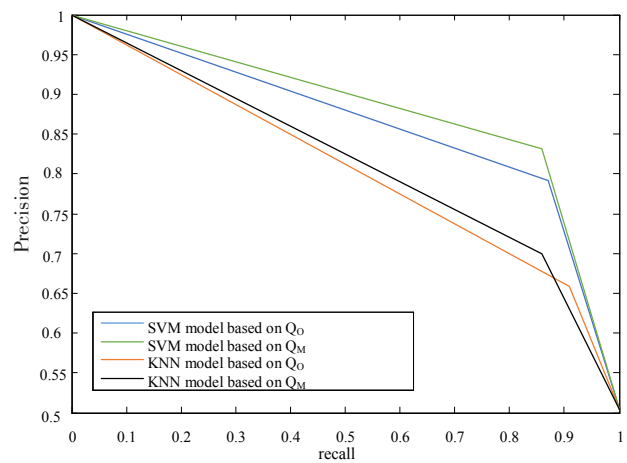


图3 两种分类模型的P-R曲线
Fig.3 P-R curve of two models

6 结论

本研究自行设计了心电-脉搏波数据采集系统, 记录了 51 个受试者的波形数据及生理信息, 提取了 16 个特征参数, 其中包括 3 种 PWV 和 13 个与心血管评估密切联系的脉搏波特征参数, 基于脉搏波特征参数及不同PWV组成的不同特征集分别用KNN和SVM建立了心血管疾病识别模型, 通过模型性能对比确定了基于特征集 Q_M 的 SVM 为最优的心血管疾

病识别模型,也确定了用于评估心血管状况的最优脉搏波波速pwvm,说明心血管疾病识别模型有一定可靠性,为心血管疾病的低成本便捷筛查提供了新思路,为实现穿戴式心血管监测提供了基础。

【参考文献】

- [1] 王文,朱曼璐,王拥军,等.中国心血管病报告2012概要[J].中国循环杂志,2013,28(6):408-412.
WANG W, ZHU M L, WANG Y J, et al. The outline of China cardiovascular disease report 2012[J]. Chinese Circulation Journal, 2013, 28(6): 408-412.
- [2] 陈伟伟,王文,隋辉,等.《中国心血管病报告2016》要点解读[J].中华高血压杂志,2017,25(7):605-608.
CHEN W W, WANG W, SUI H, et al. The key points of Chinese cardiovascular disease report 2016 [J]. Chinese Journal of Hypertension, 2017, 25(7): 605-608.
- [3] 隗玮,张存泰,吕美霞,等.无创臂踝脉搏波速度的特征及与心血管危险因素的关系[J].中华老年多器官疾病杂志,2017,16(4):274-278.
WEI W, ZHANG C T, LÜ M X, et al. Features of noninvasive brachial-ankle pulse wave velocity and their relationship with cardiovascular risk factors[J]. Chinese Journal of Multiple Organ Diseases in the Elderly, 2017, 16(4): 274-278.
- [4] CAVALCANTE J L, LIMA J A, REDHEUIL A, et al. Aortic stiffness: current understanding and future directions [J]. J Am Coll Cardiol, 2011, 57(14): 1511-1522.
- [5] 朱妙章,唐朝枢,袁文俊,等.心血管生理学基础与临床[M].北京:高等教育出版社,2011:198-199.
ZHU M Z, TANG C S, YUAN W J, et al. The basis and clinic of cardiovascular physiology[M]. Beijing: Higher Education Press, 2011: 198-199.
- [6] 汪泉,王小飞,王蕾.脉搏波传导速度应用于心血管疾病研究进展[J].中国老年学杂志,2014,35(6):1720-1722.
WANG Q, WANG X F, WANG L. Application of pulse wave velocity in cardiovascular diseases[J]. Chinese Journal of Gerontology, 2014, 35(6): 1720-1722.
- [7] ALTY S R, ANGARITA-JAIMES N, MILLASSEAU S C, et al. Predicting arterial stiffness from the digital volume pulse waveform [J]. IEEE Trans Biomed Eng, 2007, 54(12): 2268-2275.
- [8] WU H T, LIN B Y, YANG C C, et al. Assessment of vascular health with photoplethysmographic waveforms from the fingertip[J]. IEEE J Biomed Health Inform, 2017, 21(2): 382-386.
- [9] 张丽娜,周润景,武佩,等.基于心电、脉搏波信号的动脉硬化无创检测[J].生物医学工程学杂志,2016,33(4):631-638.
ZHANG L N, ZHOU R J, WU P, et al. Study on non-invasive detection of atherosclerosis based on electrocardiogram and pulse wave signals [J]. Journal of Biomedical Engineering, 2016, 33(4): 631-638.
- [10] 王开隽,张梅梅,叶继伦,等.基于指-趾脉搏波传导速度的无创动脉硬化检测系统设计[J].中国医学物理学杂志,2016,33(7):733-738.
WANG K J, ZHANG M M, YE J L, et al. Design of non-invasive atherosclerosis measurement system based on finger-toe pulse wave velocity[J]. Chinese Journal of Medical Physics, 2016, 33(7): 733-738.
- [11] 刘娜,李苗,赵静,等.单点脉搏波传导速度与冠状动脉狭窄数量的相关性研究[J].西安交通大学学报(医学版),2015,36(6):802-805.
LIU N, LI M, ZHAO J, et al. Correlation between single point pulse wave velocity and coronary artery stenosis quantity [J]. Journal of Xi'an Jiaotong University (Medical Sciences), 2015, 36(6): 802-805.
- [12] 梁宇彤,杨勇,陈鹏飞,等.动脉整体顺应性与颈-股脉搏波传播速度的相关性[J].中华医学超声杂志,2017,14(9):671-675.
LIANG Y T, YANG Y, CHEN P F, et al. Correlation of arterial total compliance and carotid-femoral pulse wave velocity: an echocardiographic study[J]. Chinese Journal of Medical Ultrasound, 2017, 14(9): 671-675.
- [13] ALLEN J. Photoplethysmography and its application in clinical physiological measurement[J]. Physiol Meas, 2007, 28(3): R1.
- [14] AHN J M. New aging index using signal features of both photoplethysmograms and acceleration plethysmograms [J]. Healthc Inform Res, 2017, 23(1): 53-59.
- [15] ELGENDI M. On the analysis of fingertip photoplethysmogram signals [J]. Curr Cardiol Rev, 2012, 8(1): 14-25.
- [16] MOHANALAKSHMI S, SIVASUBRAMANIAN A, SWARNALATHA A. Predicting arterial stiffness from physiological characteristics of photoplethysmography signals quantified through second derivative[J]. Indian J Sci Tech, 2017, 10(12): 1-8.
- [17] 周谭琪,梁永波,刘桂勇,等.基于Relief算法的心血管疾病辅助诊断研究[J].生物医学工程学杂志,2017,34(4):535-542.
ZHOU T Q, LIANG Y B, LIU G Y, et al. Study on aided diagnosis for cardiovascular diseases based on Relief algorithm [J]. Journal of Biomedical Engineering, 2017, 34(4): 535-542.
- [18] 周志华.机器学习[M].北京:清华大学出版社,2016:225-226.
ZHOU Z H. Machine learning [M]. Beijing: Tsinghua University Press, 2016: 225-226.

(编辑:薛泽玲)