

房颤智能检测中的心电特征选择和机器学习

周帆¹ 赵莉娜¹ 李钰雯¹ 李建清^{1,2} 刘澄玉^{1*}

¹ (东南大学仪器科学与工程学院, 生物电子学国家重点实验室, 南京 210096)

² (南京医科大学生物医学工程与信息学院, 南京 211166)

摘要: 房颤是最常见的心律失常疾病, 因其临床诊断不足而促进了实时自动检测算法的发展。但大多算法缺乏模型或数据库间的对比, 难以评判模型的优劣。为此考虑选择三种机器学习算法(支持向量机、随机森林、逻辑回归)构建分类模型, 在 MIT-BIH 房颤数据库上训练, 在三个不同数据库上测试并对比三种模型的表现, 并进一步分析特征选择对模型性能的影响。最终得到选择 12 个特征(3 个时域特征和 9 个非线性特征)时, 三种模型在 2018 年中国生理挑战赛训练集和可穿戴式心电设备采集数据上的灵敏度、特异度、准确性和 $F1$ 分数均达到 95% 以上, 且随机森林相较于另两种算法具有更强的稳定性和泛化能力。

关键词: 心电图; 房颤; 特征选择; 机器学习

中图分类号: R541.75; TN911.7 **文献标识码:** A **国家标准学科分类代码:** 510. 4030

ECG Feature Selection and Machine Learning in Intelligent Detection of Atrial Fibrillation

Zhou Fan¹ Zhao Lina¹ Li Yuwen¹ Li Jianqing^{1,2} Liu Chengyu^{1*}

¹ (State Key Laboratory of Bioelectronics, School of Instrument Science and Engineering, Southeast University, Nanjing 210096, Jiangsu, China)

² (School of Biomedical Engineering and Informatics, Nanjing Medical University, Nanjing 211166, Jiangsu, China)

Abstract: Atrial fibrillation (AF) is the most common cardiac arrhythmia in clinical practice, and some real-time automatic detection algorithms have been developed rapidly due to its under-detected. But mostly existing AF detection algorithms are short of comparison among kinds of models or databases, which are hard to explain the advantages and disadvantages of the model. Based on the above issues, three machine learning algorithms, including support vector machine (SVM), random forest (RF) and logistic regression (LR), were applied to build three classification models to identify AF. Three models were trained on the MIT-BIH atrial fibrillation database, and were tested and compared on three databases respectively. Furthermore, the influence of feature selection on model performance was analyzed. Finally, when 12 features (3 domain features and 9 nonlinear features) were applied, the sensitivity, specificity, accuracy and $F1$ score of the three models reached more than 95% on the CPSC2018 and the wearable ECG database. In addition, RF algorithm has stronger stability and generalization ability compared with the other two algorithms.

Key words: electrocardiogram; atrial fibrillation; feature selection; machine learning

0 引言

心房颤动 (atrial fibrillation, AF) 是临床上常见的持续性心律失常疾病, 患病率约为 0.4~1%^[1-2]。房颤不仅会增加患者中风、心力衰竭和冠状动脉等疾病的发病风险, 使患者死亡率升高, 且其发病率随着年龄的增加而增加。但是由于房颤发生时可能表现为非特异性症

基金项目: 本文受国家自然科学基金 (81871444), 江苏省自然科学基金 (SBK2020041638) 和中央高校基本科研业务费专项资金 (2242020K40140) 资助。

*通讯作者: 刘澄玉, E-mail: chengyu@seu.edu.cn

状或无症状,故房颤患者在临床中难以得到确切诊断而延误治疗^[3]。基于动态心电图的房颤自动检测算法通过实时、长程监测患者心电状况,能及时发现并预测房颤等心律失常疾病,从而弥补临床诊断不足,减轻医疗负担和经济压力,节省宝贵的医疗资源^[4-7]。

房颤发生时,心电图(electrocardiogram, ECG)波形主要表现为RR间期绝对不规则,且P波消失,代之以大小、形状各不相同的快速震荡的f波。因此,目前房颤检测的基本原理可分为基于心室响应分析和基于心房活动分析两种,即结合ECG信号RR间期,或是结合P波或f波的特征信息,然后利用机器学习等算法建立分类模型,从而实现实时自动检测房颤^[8-10]。目前,已有许多国内外学者提出各种检测算法。Andersen等^[11]从搏动间隔中提取五个时域特征,并结合支持向量机(support vector machine, SVM)提出一种房颤检测算法,最终所提算法在MIT-BIH房颤数据库(atrial fibrillation database, AFDB)上的敏感度 $\geq 95\%$ 。Zabihi等^[12]从ECG信号中提取150个特征,并采用随机森林(random forest, RF)分类器将信号分为正常、房颤、其他节律及噪声四种类型,最终在2017 PhysioNet/CinC挑战赛(The PhysioNet Computing in Cardiology Challenge 2017, CinC2017)测试集数据上得分0.83。2018年,陈志博等^[13]提出基于RR间期提取稳健变异系数、偏态参数和Lempel-Ziv复杂度三个特征,并结合SVM实现房颤的自动分类检测,最终在MIT-BIH房颤数据库数据上的灵敏度为95.81%、特异度为96.48%、准确率为96.09%,同时在MIT-BIH窦性心律数据库中实现了95.16%的特异度。

值得注意的是,尽管已有各种机器学习算法^[14]用于房颤检测,并取得了较为理想的效果,但对于不同算法之间的比较分析,及其在分类过程中的特点,所做的研究远远不够。掌握不同算法的分类能力及特点,有助于选择更适用临床数据的实时自动检测算法并进一步优化。此外在大部分房颤检测研究中多采用标准数据库(如MIT-BIH AFDB),很少研究会不同数据库上对模型加以验证,尤其缺乏在国人的临床数据集上的分析,这样的模型泛化能力差,结果缺乏说服力。针对以上问题,本文选用三种常用的机器学习算法,SVM、RF和逻辑回归(logistics regression, LR),结合ECG信号的时域、频域、形态和非线性等102个特征,在标准数据库AFDB上训练分类模型,并在CinC2017、2018中国生理挑战赛(China Physiological Signal Challenge 2018, CPSC2018)和可穿戴式心电设备采集数据三种数据库上分别测试三种模型表现,同时对特征进行筛选,从而研究不同数据库、不同算法和不同特征集这三个方面对于模型性能的影响,并对结果给予充分的验证和合理的解释。

1 方法

1.1 方法概述

本文选择RF算法对102个特征进行计算并筛选,而后在AFDB上分别训练特征选择前后的SVM、RF和LR分类器,并在CinC2017训练集数据、CPSC2018训练集数据和可穿戴式心电设备采集数据库上测试并比较模型间的分类能力。具体算法流程框图如图1所示。

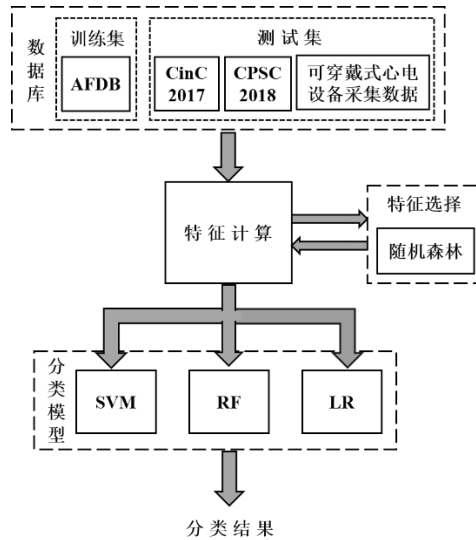


图 1 算法流程图

Fig.1 Flow chart of the proposed method

1.2 数据库

本文所研究数据来自四个数据库，分别是 AFDB、CinC2017 训练集数据、CPSC2018 训练集数据和可穿戴式心电设备采集数据，其中 AFDB 作为训练集，其他三个数据库作为测试集。具体如表 1 所示，表中数据为 30s 片段数。

表 1 各数据库的房颤和非房颤数据段数

Tab1 Numbers of AF and non-AF data for all databases

	训练集	测试集		
	AFDB	CinC2017	CPSC2018	心电设备采集数据
采样率/Hz	250	300	500	400
房颤/段	11048	855	81	11089
非房颤/段	16746	5477	66	10000
总计/段	27794	6332	147	21089

1.2.1 MIT-BIH 房颤数据库

AFDB 包括 25 例患有房颤(主要是阵发性房颤)的人体实验 ECG 记录，其中记录 00735 和 03665 在本研究中被预先排除(仅包括节律和未经审核的节拍注释文件)。每条 ECG 记录包含两个时长 10 小时的信号，采样率 250Hz。节律注释文件经过手工标注，包括 AF、AFL (atypical atrial flutter, 心房扑动)、J (房室交界区性心律)和 N (用于表示所有其他节律)四种节律注释。节拍注释文件来自于自动检测器，其中记录 05091 和 07859 经过人工标注。本文将 AFL、J 和 N 均视为非房颤信号，并将所有 ECG 记录分割为 30s/段，最终得到 27794 段 ECG 信号，其中包括 11048 条房颤数据和 16746 条非房颤数据。

1.2.2 2017 PhysioNet/CinC 挑战赛数据库

本研究采用 CinC2017 的训练集数据作为测试集 I^[15]。该数据库训练集包含 8528 条单导联 ECG 记录，其中包括 5050 条正常节律、738 条房颤、2456 条其他节律和 284 条噪声 ECG 信号。所有 ECG 信号时长从 9 秒到 60 多秒，采样频率为 300Hz，每条记录文件包括一个 ECG 信号(.mat 格式)和一个关于波形信息的注释文件(.hea 格式)。本文选用数据库

中的正常节律信号和房颤信号，并分割为 30s/段，最终选用 855 条房颤信号和 5477 条正常节律信号组成测试集 I。

1.2.3 2018 中国生理挑战赛数据库

本文采用 CPSC2018 的训练集数据作为测试集 II^[16]。该训练集是从 11 家医院收集而来，共有 6877 条的 12 导联 ECG 记录，时长为 6 秒至 60 秒，采样率是 500Hz。每条数据包含一个记录 ECG 信号采样值的.mat 文件，以及患者的性别和年龄信息。该数据集一共包含 9 种类型的 ECG 信号，本文只选用其中的 918 条窦性心律和 1098 条房颤 ECG 记录的第一导联信号，并分割为 30s/段，最终选择 81 条房颤及 66 条正常 ECG 信号段组成测试集 II。

1.2.4 可穿戴式心电设备采集数据库

为进一步验证模型在临床应用上的表现，本研究采用东南大学和联想集团^[17]联合开发的可穿戴式心电设备，从 23 名志愿者（13 名房颤患者和 10 名正常人）身上各采集 24 小时 ECG 信号，采样率为 400Hz。其中每位患者的数据对应一份节律类型人工标注文件，均由心脏疾病专家进行标注。本研究将采集到的数据分为 30s/段，经过筛选共得到 173831 条正常数据和 108830 条房颤数据。由于数据量太大导致模型运行时间过长，我们从每个人的数据中随机选出 1000 条数据，最终得到 11089 条房颤数据（其中两位志愿者只有 47 条和 42 条数据）和 10000 条正常数据组成测试集 III。

1.3 特征提取

特征提取之前，首先对所有 ECG 信号段进行 QRS 波、P 波和 T 波位置检测。本研究采用的 R 波位置检测算法^[18-19]对 P&T（Pan-Tompkins）算法做出一定的改进，应用不应期概念和回溯技术有效避免了 R 波漏检，同时保证极低振幅无变化的 ECG 信号不被检测^[20]。另外，对于检测出 RR 间期个数小于 12 的 ECG 信号段予以舍弃。而后采用 Shang 等^[21-22]提出的算法检测 T 波位置，Q 波、S 波和 P 波位置以 R 波位置为参考，采用 Datta 等^[23]编写的开源算法进行检测。

本研究根据得到的 P 波、Q 波、R 波、S 波和 T 波位置从 ECG 信号中共提取 102 个特征，可分为时域特征（27 个，编号为 F1~F27）、频域特征（13 个，编号为 F28~F40）、非线性特征（36 个，编号为 F41~F76）和形态特征（26 个，编号为 F77~F102）四类^[24-26]。具体特征指标及说明如表 2 所示。

1.3.1 时域特征

时域特征主要是根据 ECG 信号的 RR 间期得到的统计特征，包括 RR 间期的最大值（maxRR）、最小值（minRR）、概率密度估计值等特征，还有心率的中位数（medHR）、相邻 RR 间期差值（ ΔRR ）的均方根（RMSSD）及 ΔRR 间期大于 50ms 的百分比（PNN50）等。由于房颤信号的 RR 间期绝对不规则，因此这些特征在区分正常信号和房颤信号时能够发挥较好的效果。

表 2 102 个特征指标及说明

Tab2 Description of 102 features

特征编号	特征说明	特征编号	特征说明
F1	mRR, 即 RR 间期的平均值, 单位 ms	F53	Renyi 熵
F2	medRR, 即 RR 间期的中位数, 单位 ms	F54~F58	SampEn, 即嵌入维数为 1~5 时的样本熵, 计算详见附录
F3	minRR, 即 RR 间期的最小值, 单位 ms	F59	SampEn 为无限值的个数
F4	maxRR, 即 RR 间期的最大值, 单位 ms	F60	SampEn 最大值
F5	RR 间期的偏度	F61	SampEn 最小值
F6	RR 间期的峰度	F62	SampEn 最大值的维数
F7	RR 间期值的范围, 单位 ms	F63	SampEn 最小值的维数
F8	RR 间期概率密度估计 ⁽¹⁾ 值的峰度	F64~F65	最小二乘拟合多项式系数
F9	RR 间期概率密度估计值的偏度	F66	SD1, 定义为投射到垂直于 Poincare plot 中椭圆短轴的点的标准偏差
F10	RR 间期概率密度估计的局部峰值个数	F67	SD2, 定义为投射到平行于 Poincare plot 中椭圆长轴的点的标准偏差
F11	RR 间期概率密度估计的最大峰位置与其余峰距离最大值	F68~F72	ApEn, 即嵌入维数为 1~5 时的近似熵, 计算详见附录
F12	RR 间期概率密度估计的最大峰位置与其余峰距离最小值	F73	FuzzyMEn, 即模糊测量熵, 计算见附录
F13	Δ RR 间期概率密度估计值的峰度	F74	NFEn, 即归一化模糊熵 ^[30] , 计算详见附录
F14	Δ RR 间期概率密度估计值的偏度	F75	MAD, 即绝对中位差, 定义为三个相邻 RR 间期心率平均值的绝对偏差变化率的中位数
F15	R 峰幅值的概率密度估计的峰度	F76	定义为 Lorenz plot 中占 60% 面积的最小圆的半径 ^[31]
F16	R 峰幅值的概率密度估计的偏度	F77	S 波与 R 波幅值差的变异系数
F17	R 峰幅值概率密度估计的局部峰值个数	F78	S 波与 R 波幅值差的范围, 单位 mV
F18	R 峰幅值的变异系数 ⁽²⁾	F79	S 波与 R 波幅值差的中位数, 单位 mV
F19	相邻 R 峰间所有采样点幅值的变异系数	F80	ST 段斜率的变异系数
F20	pNN20, 定义为 Δ RR 间期大于 20ms 的百分比	F81	ST 段斜率小于 0 的个数与 R 波个数的比值
F21	pNN50, 定义为 Δ RR 间期大于 50ms 的百分比	F82	S 波幅值小于 Q 波幅值的个数与 S 波个数的比值
F22	SDSD, 即 Δ RR 间期的标准差, 单位 ms	F83	ST 段幅值大于前一个 TQ 段幅值中位数的位置
F23	medHR, 即心率的中位数, 单位次/分钟	F84	T 波与 R 波幅值比的中位数
F24	SDNN, 即 RR 间期的标准差, 单位 ms	F85	P 波个数与 R 波个数的比值
F25	RMSSD, 即 Δ RR 间期的均方根, 单位 ms	F86	QRS 波宽度的中位数, 单位 ms
F26	RR 间期的变异系数	F87	QRS 波宽度的方差, 单位 ms ²
F27	Δ RR 间期的变异系数	F88	QR 间期的中位数, 单位 ms
F28	谱质心, 反映信号频谱各个子带主要能量集中的频率范围	F89	QR 间期的方差, 单位 ms ²
F29	窗口内信号的峰度	F90	Q 波与 R 波幅值比的变异系数
F30	细节小波系数, 反映信号高频分量信息的小波变换结果	F91	Q 波与 R 波幅值比的范围
F31~F36	第 2、4、7、9、10、11 个线性预测 ⁽³⁾ 系数	F92	Q 波与 R 波幅值比的中位数
F37	质心频率, 定义为给定子带的加权平均频率, 权值为子带中各频率分量的归一化能量	F93	Q 波和 S 波幅值差与 Q 波和 R 波幅值差之比
F38	归一化低频功率 LF_n , 定义为 $LF_n = LF / (LF + HF)$	F94	Q 波和 S 波幅值差与 Q 波和 R 波幅值差比值的变异系数
F39	归一化高频功率 HF_n , 定义为 $HF_n = HF / (LF + HF)$	F95	QR 斜率的变异系数
F40	LF/HF, 定义为低频功率 LF (0.04Hz~0.15Hz) 与高频功率 HF (0.15Hz~0.4Hz) 之比	F96	S 波斜率的变异系数
F41~F46	Lorenz plot 相关参数, 详见附录	F97	Bazett 公式 ⁽⁵⁾ 计算 QT 间期的中位数, 单位 ms
F47	Poincare plot ⁽⁴⁾ 中 RR 间隔的平均步进增量	F98	Bazett 公式计算 QT 间期的变异系数
F48	Poincare plot 中对角线周围点的分布	F99	Fridericia 公式 ⁽⁶⁾ 计算 QT 间期的中位数, 单位 ms
F49	Poincare plot 中集群的数量	F100	Fridericia 公式计算 QT 间期的变异系数
F50	Shannon 熵, 量化信息量的大小	F101	Sagie 公式 ⁽⁷⁾ 计算 QT 间期的中位数, 单位 ms
F51	Hjorth 参数移动性..	F102	Sagie 公式计算 QT 间期的变异系数
F52	Hjorth 参数复杂度		

(1)概率密度估计: 已知某类别 ω_i 的样本 $X_i (i=1,2,\dots,N)$, 采用某种规则估计出样本所属类的概率函数 $P(X|\omega_i)$ 。

(2)变异系数: $CV = \sigma / \mu$, σ 为标准差, μ 为均值。

(3)线性预测: 根据随机信号过去的 p 个已知抽样值序列 $S_{n-1}, S_{n-2}, \dots, S_{n-p}$, 预测当前值 S_n 的估计值的方法。

(4) Poincare plot: 以当前心拍间隔 RR_i 为横轴, 下一个心拍间隔 RR_{i+1} 为纵轴得到的散点图。

(5) Bazett 公式: $QTcB = QT / RR^{0.5}$ (6) Fridericia 公式: $QTcF = QT / RR^{0.33}$ (7) Sagie 公式: $QTcS = QT + 0.154 * (1 - RR)$

1.3.2 频域特征

频域特征主要是对 ECG 信号加窗, 计算窗口信号的谱参数, 包括谱质心、质心频率等, 同时还计算了小波变换系数、归一化低频功率 (normalized low-frequency power, LF_n) 和归

一化高频功率（normalized high-frequency power, HF_n）等特征。

1.3.3 非线性特征

非线性特征主要计算 Lorenz plot 相关参数、Poincare plot 参数^[27]、Hjorth 参数^[28]等，和绝对中位差（median absolute deviation, MAD），即三个相邻 RR 间期心率平均值的绝对偏差变化率的中位数，同时包括样本熵^[29]（sample entropy, SampEn）、近似熵（approximate entropy, ApEn）、模糊测量熵（fuzzy entropy, FuzzyEn）及归一化模糊熵（normalized fuzzy entropy, NFEn）^[30]等熵值。详细计算过程参看附录。

1.3.4 形态特征

形态特征主要是根据 P 波、Q 波、R 波、S 波和 T 波位置及幅值，判断 ECG 波形的变形情况，主要计算 S 波和 Q 波相对 R 波的深度、ST 段斜率、QT 间期及 QRS 波群宽度等。

根据以上分析和计算，提取 102 个特征值后，便可进一步做特征选择和模型建立。

1.4 特征选择

特征选择在模型建立过程中发挥着重要作用。过多的特征不仅会增大计算量，降低程序运行速度，不适合的特征还可能误导模型结果，使分类准确度下降。因此，需要在大量特征中进行合理筛选，得到一组既能全面反映信号重要特征信息，又能有效区分不同类被测信号的特征指标。

本研究采用 RF 特征选择算法，并按其重要性得分对特征进行排序^[31]。L.Breiman 将 RF 定义为一组“树状结构分类器 $h(X, \theta_k), k=1, 2, \dots, K$ 的集合”，其中 θ_k 是独立同分布随机向量， K 为树的个数。从原始训练数据中随机抽取每棵树的样本集，并在每次节点分裂时随机选择特征集，生成 RF 分类模型。对于输入 X ，通过每棵树投票选择票数最多的作为最后的分类结果。

RF 模型进行分类时，每次均选择最优特征进行节点分裂。最优特征的选择主要根据两种常规的特征重要性度量方法，即根据 Gini 指数和 OOB 数据分类正确率进行度量。本文采用基于 Gini 指数的特征重要性 IMP_i^{Gini} ，即第 i 个特征 x_i 在节点 n 上的数据划分到其左右子节点 n_l 和 n_r 前后 Gini 指数的变化量，如式(1)：

$$IMP_{in}^{Gini} = I_G(n) - I_G(n_l) - I_G(n_r) \quad (1)$$

其中， $I_G = 1 - \sum_{c=1}^2 p_c^2$ 为节点 n 的 Gini 指数。

若特征 x_i 在第 k 棵决策树中作为节点分割属性出现的集合为 N ，则该特征在这棵决策树上的重要性为：

$$IMP_{i-k}^{Gini} = \sum_{n \in N} IMP_{in}^{Gini} \quad (2)$$

进一步计算特征 x_i 在整个 RF 中的重要性，如式(3)：

$$IMP_i^{Gini} = \frac{1}{K} \sum_{k=1}^K IMP_{i-k}^{Gini} \quad (3)$$

通过上述过程，最终得到了 102 个特征的重要性得分，并对得分作归一化处理，得分越高，特征越重要。本文对所有特征按照降序排列，分别选择 12 个特征和 9 个特征用于后续

的分类模型训练。

1.5 机器学习分类模型

1.5.1 支持向量机

对一组训练样本 $\{(\mathbf{x}_i, \mathbf{y}_i), i=1, \dots, N\}$ ，其中 $\mathbf{x}_i \in \mathbf{R}^{N_f}$ (N_f 是输入空间维度)，分类标签为 $\mathbf{y}_i = \{+1, -1\}$ 。SVM 用于分类的核心思想是最优超平面 $\boldsymbol{\omega}^T \mathbf{x} + b = 0$ 的构建^[10]，即使得向量 \mathbf{x} 与超平面之间的最小距离最大，可表示为：

$$\max_{\boldsymbol{\omega}, b} \min \tilde{\gamma} = \max_{\boldsymbol{\omega}, b} \min \left(\mathbf{y} \frac{\boldsymbol{\omega}^T \mathbf{x} + b}{\|\boldsymbol{\omega}\|} \right) \quad (4)$$

其中， $\tilde{\gamma}$ 为向量 \mathbf{x} 与超平面间的几何距离。为方便计算，定义几何间隔 $\tilde{\gamma} = 1/\|\boldsymbol{\omega}\|$ ，边界宽度等于 $2/\|\boldsymbol{\omega}\|$ 。问题(4)可重新表述为：

$$\min_{\boldsymbol{\omega}, b} \tilde{\gamma}(\boldsymbol{\omega}) = \frac{1}{2} \|\boldsymbol{\omega}\|^2 \quad \text{s.t.} \quad \mathbf{y}_i (\boldsymbol{\omega}^T \mathbf{x}_i + b) \geq 1, i=1, \dots, N \quad (5)$$

目标优化问题(5)可根据拉格朗日对偶性，将原始问题转化为其对偶问题进行求解。为将分类中的非线性关系转化为线性关系，需引入核函数概念，本文选择的是高斯核函数。

1.5.2 随机森林

根据 1.4 节中对 RF 算法的定义、相关概念和基于 RF 的特征选择算法的介绍，可知 RF 算法应用于分类问题时，最终分类结果由所有树节点投票众数决定。对于分类结果的可靠性，采用余量函数 (margin function) 进行度量，定义为：

$$mg(\mathbf{X}, \mathbf{Y}) = av_K I(h_K(\mathbf{X}) = \mathbf{Y}) - \max_{j \neq \mathbf{Y}} av_K I(h_K(\mathbf{X}) = j) \quad (6)$$

其中， $h_K(\mathbf{X})$ 表示第 K 个树分类器的预测结果， \mathbf{Y} 表示分类标签， $I(\bullet)$ 表示示性函数， $av_K(\bullet)$ 表示取平均。由余量函数可得预测结果的泛化误差为：

$$PE^* = P_{\mathbf{X}, \mathbf{Y}}(mg(\mathbf{X}, \mathbf{Y}) < 0) \quad (7)$$

其中，下标 \mathbf{X}, \mathbf{Y} 表示概率 P 覆盖 \mathbf{X}, \mathbf{Y} 的空间。当 RF 中的树分类器足够多时， $h_K(\mathbf{X}) = h(\mathbf{X}, \theta_K)$ 服从强大数定律，即随着决策树个数的增多，对于所有序列 $\theta_1, \theta_2, \dots, \theta_k$ ， PE^* 几乎处处收敛于：

$$P_{\mathbf{X}, \mathbf{Y}}\{P_\theta(h(\mathbf{X}, \theta) = \mathbf{Y}) - \max_{j \neq \mathbf{Y}} P_\theta(h(\mathbf{X}, \theta) = j) < 0\} \quad (8)$$

公式(8)表明，随着决策树个数的增加，RF 不会出现过拟合问题，但可能会产生一定范围内的泛化误差。

1.5.3 逻辑回归

LR 是一种概率分类统计模型，该模型通过将 Sigmoid 函数 $g(z) = 1/(1 + e^{-z})$ 与线性回归函数 $y = g(\boldsymbol{\omega}^T \mathbf{x}) = \boldsymbol{\omega}^T \mathbf{x} + b$ 融合，很好地解决了回归函数易受离群值影响的问题，并将分类器的输出转换为 $[0, 1]$ 之间的有界值^[10]。本文用 $y = 0$ 或 1 代表分类结果，且各类概率之和为 1，则可得：

$$p(y = 0 | \mathbf{x}) = \exp(\boldsymbol{\omega}^T \mathbf{x}) / (1 + \exp(\boldsymbol{\omega}^T \mathbf{x})) = \phi(z) \quad (9)$$

$$p(y = 1 | \mathbf{x}) = 1 / (1 + \exp(\boldsymbol{\omega}^T \mathbf{x})) = 1 - \phi(z) \quad (10)$$

其中， ω 是回归系数向量， x 是待预测变量， $p(y=j|x)$ 表示 x 属于类 j 的条件概率。为使函数模型预测正确率最高，根据最大似然估计提出了 LR 模型的交叉熵损失函数：

$$J(\omega) = -l(\omega) = \sum_{i=1}^n [-y_i \ln(\phi(z_i)) - (1-y_i) \ln(1-\phi(z_i))] \quad (11)$$

本文采用梯度下降法求解交叉熵损失函数，求得 ω 后便得到了 LR 的函数模型，然后设定合适的阈值，若预测值大于阈值，则分类结果 $y=1$ ，反之 $y=0$ 。

1.6 模型训练及测试

由 1.5.1-1.5.3 节中对 SVM、RF 及 LR 三种算法的介绍，可在训练集上建立相应的分类模型。在此之前，首先需要对训练集和测试集的所有特征值进行中心化和标准化处理，即： $x_i' = (x_i - \mu_i) / \sigma_i$ ，其中 μ_i 和 σ_i 分别是训练集特征 x_i 的均值和标准差，对测试集使用同样的参数 μ_i 和 σ_i 进行处理。

数据预处理后，分别选择 102 个、12 个和 9 个特征在 AFDB 上建立 SVM、RF、LR 分类模型，并在 CinC2017、CPSC2018 和可穿戴式心电设备采集数据上测试三个模型的分类能力，进而对不同特征集建立的各个模型在多个数据库上的表现进行分析比较和总结。

2 结果

2.1 特征选择结果

特征选择得到了 102 个特征的重要性排序，图 2 只列出前 51 个特征的重要性得分。由图 2 可得，特征重要性得分在 0.04 以上的有 9 个，得分总和约占 76%；得分在 0.02 以上的有 12 个，得分总和约占 85%。因此本文最终从 102 个特征中分别选择 12 个和 9 个特征，以比较两种特征选择结果的差别和所建立模型的性能。9 个特征分别为 3 个时域特征 F27： ΔRR 间期的变异系数、F24：SDNN、F20：PNN20，和 6 个非线性特征 F74：NFE_n、F41：Lorenz plot 证据特征（atrial fibrillation evidence, AFE_v）、F47：Poincare plot 中 RR 间期的平均步进增量、F75：MAD、F55：SampEn（嵌入维数 $m=2$ ）、F66：SD1。12 个特征是除了这 9 个特征外，还包括 3 个非线性特征 F43：不规则度证据特征（irregularity evidence, IrrEv）、F54：SampEn（嵌入维数 $m=1$ ）和 F62：SampEn 最大时的维数。在所有房颤数据和非房颤数据上分别计算 12 个特征的平均值及标准差，如表 3 所示。

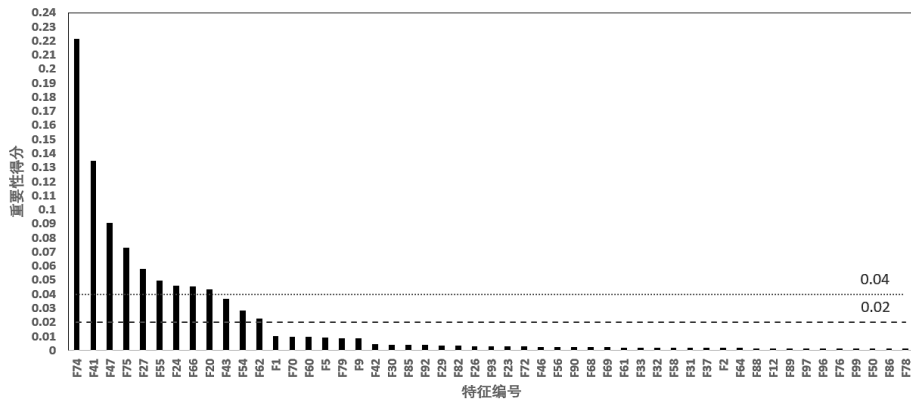


图 2 前 51 个特征的重要性得分

Fig.2 Importance scores for the first 51 features

表 3 不同数据库上房颤与非房颤的特征值（平均值±标准差）比较

Tab.3 The mean±standard deviation values of nine features for comparison between the AF and Non-AF groups of different databases

特征编号	AFDB		CinC2017		CPSC2018		心电设备采集数据	
	房颤	非房颤	房颤	非房颤	房颤	非房颤	房颤	非房颤
信号个数	11048	16746	855	5477	81	66	11089	10000
F74	0.94±0.43	-0.62±0.75	0.77±0.62	-0.54±0.64	0.90±0.47	-0.51±0.38	1.00±0.44	-0.67±0.83
F41	1.08±0.32	-0.71±0.57	0.92±0.57	-0.61±0.52	1.08±0.42	-0.49±0.48	1.01±0.29	-0.67±0.48
F47	0.87±0.62	-0.57±0.76	0.85±0.79	-0.71±0.49	1.08±1.10	-0.83±0.13	1.02±0.54	-0.55±0.80
F75	0.87±0.72	-0.57±0.70	0.80±0.87	-0.63±0.48	0.90±0.77	-0.72±0.13	0.75±0.68	-0.64±0.54
F27	0.70±0.80	-0.46±0.83	0.72±0.79	-0.60±0.60	0.90±1.26	-0.80±0.12	0.78±0.50	-0.47±0.76
F55	0.07±0.34	0.05±1.26	-0.11±0.06	-0.06±0.82	-0.08±0.05	-0.09±0.04	-0.02±0.68	0.04±1.26
F24	0.50±0.89	-0.33±0.93	0.50±0.90	-0.28±0.89	0.75±1.94	-0.55±0.14	0.71±0.61	-0.15±0.81
F66	0.49±0.90	-0.32±0.93	0.53±0.96	-0.42±0.75	0.75±1.86	-0.63±0.10	0.72±0.65	-0.23±0.82
F20	0.78±0.88	-0.51±0.70	0.73±1.04	-0.67±0.41	0.90±0.80	-0.84±0.05	1.16±0.93	-0.44±0.86
F43	1.09±0.50	-0.72±0.44	0.93±0.76	-0.78±0.29	1.11±0.63	-0.85±0.21	0.94±0.45	-0.80±0.29
F54	0.21±0.46	-0.14±1.21	-0.15±0.78	-0.03±1.06	0.16±0.51	0.47±0.38	0.31±0.35	-0.05±1.12
F62	0.33±0.81	-0.22±1.05	0.01±1.07	-0.20±1.10	0.41±0.80	0.36±0.70	0.34±0.62	-0.25±1.03

2.2 模型测试结果

2.2.1 评价指标

为了比较三种模型对房颤信号的分类表现，我们采用如下常用的评价指标来衡量，即：

$$\text{灵敏度: } Se = \frac{TP}{TP + FN} \times 100\% \quad (12)$$

$$\text{特异度: } Sp = \frac{TN}{TN + FP} \times 100\% \quad (13)$$

$$\text{准确度: } Acc = \frac{TP + TN}{TP + FP + FN + TN} \times 100\% \quad (14)$$

$$F1 \text{ 分数: } F1 = 2 \cdot \frac{S \cdot Sp}{Se + Sp} \times 100\% \quad (15)$$

其中， TP 为预测结果中的真阳性数， TN 为真阴性数， FP 为假阳性数， FN 为假阴性数。

2.2.2 SVM、RF 和 LR 模型测试结果

分别采用 102 个特征、12 个特征和 9 个特征训练 SVM、RF 和 LR 模型，并在 CinC2017、CPSC2018 和可穿戴式心电设备采集数据上测试三种模型性能，测试结果如表 4 所示。

由表可得，特征选择前后，SVM 模型和 LR 模型的各项性能指标均显著提高，尤其是 Se 和 $F1$ 分数。在 CinC2017 数据库上，SVM 模型的 Se 从 37.78% 提高到 77.54%， $F1$ 分数从 0.5484 提高到 0.8140；LR 模型的 Se 由 34.74% 提高到 80.12%， $F1$ 分数由 0.5152 提高到 0.8333；同样在心电设备采集数据上，LR 模型的 Se 由 0.01% 陡然提高到 99.31%， $F1$ 分数由特征选择前的仅 0.0002 上升到 0.9909，同时 Acc 也由 47.42% 上升到 99.04%。此外可以观察到，选择 12 个特征和选择 9 个特征的模型性能几乎无差。对于 CinC2017 数据库，选择 9 个特征时 SVM 模型表现稍好，选择 12 个特征时 LR 模型略优；而在 CPSC2018 数据库，选择 12 个特征的 SVM 模型较好，LR 模型则在两种特征选择上性能相同；在心电设备采集数

据上，选择 12 个特征的 SVM 模型和 LR 模型表现更好。因此，相较之下，选择 12 个特征训练两种模型，最终得到在 CinC2017 数据库上的 SVM 模型和 LR 模型 Se 分别为 77.54%、80.12%， Sp 为 97.97%、98.10%， Acc 为 95.21%、95.67%， $F1$ 为 0.8140、0.8333；在 CPSC2018 数据库上两种模型的 Se 分别为 96.30%、95.06%， Sp 分别为 100%、98.48%， Acc 为 97.96%、96.60%， $F1$ 为 0.9811、0.9686；在心电设备采集数据上的 Se 分别为 99.35%、99.31%， Sp 为 98.79%、98.75%， Acc 为 99.08%、99.04%， $F1$ 为 0.9913、0.9909。

表 4 SVM、RF 和 LR 模型的分类结果
Tab.4 Classification results of SVM, RF and LR models

数据库	特征个数	SVM				RF				LR			
		$Se/\%$	$Sp/\%$	$Acc/\%$	$F1$	$Se/\%$	$Sp/\%$	$Acc/\%$	$F1$	$Se/\%$	$Sp/\%$	$Acc/\%$	$F1$
CinC2017	102	37.78	100.00	91.60	0.5484	79.30	98.78	96.15	0.8475	34.74	99.98	91.17	0.5152
	12	77.54	97.97	95.21	0.8140	77.89	99.03	96.18	0.8463	80.12	98.10	95.67	0.8333
	9	79.65	97.74	95.29	0.8205	77.31	98.30	95.47	0.8216	80.35	98.01	95.63	0.8322
CPSC2018	102	88.89	98.48	93.20	0.9351	96.30	100.00	97.96	0.9811	97.53	78.79	89.12	0.9080
	12	96.30	100.00	97.96	0.9811	95.06	100.00	97.28	0.9747	95.06	98.48	96.60	0.9686
	9	93.83	100.00	96.60	0.9682	96.30	100.00	97.96	0.9811	95.06	98.48	96.60	0.9686
心电设备 采集数据	102	80.20	100.00	89.59	0.8901	99.40	99.62	99.50	0.9953	0.01	100.00	47.42	0.0002
	12	99.35	98.79	99.08	0.9913	98.69	99.39	99.02	0.9907	99.31	98.75	99.04	0.9909
	9	99.22	97.97	98.62	0.9870	98.76	99.23	98.99	0.9903	99.27	98.12	98.72	0.9879

与特征选择前后的 SVM 模型和 LR 模型不同，RF 模型性能不仅没有太大变化，甚至在三个数据库上的 Se 和 $F1$ 分数略微下降，但始终呈现优秀的分类能力。因此特征选择对 RF 模型性能影响不大，但极大减少了运算复杂度，节省时间成本。同时，选择 12 个特征训练得到的 RF 模型在 CinC2017 和心电设备采集数据库上比选择 9 个特征时表现略好，故选择 12 个特征训练 RF 模型最为合适。最终得到 RF 模型在 CinC2017 数据库上的 Se 为 77.89%， Sp 为 99.03%， Acc 为 96.18%， $F1$ 为 0.8463；在 CPSC2018 数据库上的 Se 为 95.06%， Sp 为 100%， Acc 为 97.28%， $F1$ 为 0.9747；在心电设备采集数据上的 Se 为 98.69%， Sp 为 99.39%， Acc 为 99.02%， $F1$ 为 0.9907。

3 讨论

比较三种模型在 CinC2017、CPSC2018 和心电设备采集数据上的分类结果可以得到：当选择合适的特征及特征个数时，每种算法都能得到较为理想的分类模型，因此特征选择是建立优秀分类模型中的关键步骤；同时亦可说明，本文所选用的特征选择算法效果较为理想。再由 2.2.2 节的分析可以得到，选择 12 个特征与 9 个特征对于三种模型的性能影响不大，说明这 9 个综合了时域和非线性信息的特征已足以反映房颤和正常 ECG 信号的不同点，特征 F43、F54 和 F62 对分类结果并没有决定性作用。但是考虑到两种选择计算量相差不大且选用 12 个特征时模型性能略好，最终选择 12 个特征构建三种分类模型。

值得注意的是，即便经过特征选择后模型的性能大大提升，三种模型在 CinC2017 数据库上的表现相比于另外两个测试集，仍是不尽人意。选择 12 个特征训练模型时，三种分类器在 CPSC2018 和实际采集数据上的 Se 、 Sp 、 Acc 和 $F1$ 分数都达到了 95% 以上，而在

CinC2017 数据库上 Se 和 $F1$ 分数都在 85% 以下, 这说明三种模型对于 CinC2017 数据库中的房颤信号段检测能力不足。因此本文分别统计了三种模型在 CinC2017 中 855 条房颤数据上检测出错的数据, 其中, SVM 模型检测错误 192 条, RF 模型检测错误 189 条, LR 模型检测错误 170 条。通过对这些检错数据进行手动标注, 发现其中约有 110 条并不属于房颤, 而是典型的早搏信号, 如图 3 所示, 图中数据分别来自于 CinC2017 数据库中的 A00680.mat、A03391.mat 和 A00624.mat。

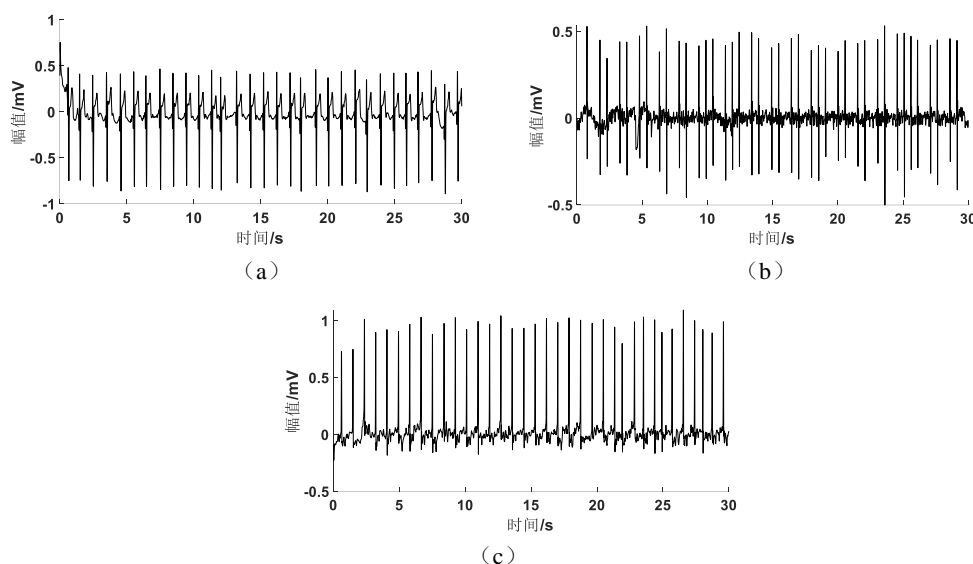


图 3 CinC2017 数据库中分类标签错误的数据。(a) A00680.mat; (b) A03391.mat; (c) A00624.mat

Fig.3 Datas with incorrect label in CinC2017 database. (a) A00680.mat; (b) A03391.mat; (c) A00624.mat

由此可知, 三种模型在 CinC2017 数据库上表现不理想, 主要是由于 CinC2017 数据库中存在部分数据的分类标签错误, 与模型本身关系不大, 若去除这些分类错误的数据, 三种分类器的 Se 和 $F1$ 分数至少可达到 90%。以上分析说明本研究所建立的三种分类模型具有很强的泛化能力, 可以适应不同的数据库, 而且模型在可穿戴式心电设备采集的临床数据上表现优秀, 证明它们具有可期的临床应用价值。

尽管如此, 通过对比三种算法在相同条件下的分类结果, 仍能发现它们之间存在差别。从表 4 中可以看出, 特征选择前, SVM 模型和 LR 模型在 CinC2017 和 CPSC2018 数据库上性能相差不大, 但在可穿戴式心电设备采集的临床数据集上, LR 模型表现很差, 而 SVM 模型表现相对稳定。这反映出 SVM 模型在未知数据集上可能具有更理想的表现, 而 LR 模型更为脆弱, 泛化能力更差, 更加依赖于优质特征的筛选才能得到理想的模型。再观察特征选择后, SVM 模型和 LR 模型的性能均显著提高, 而 RF 模型表现却无明显变化。这说明 RF 算法对各种特征的包容性更好, 其分类能力不易被效果差的特征影响, 而且特征个数越多, RF 接收到的信号信息越全面, 分类模型的表现越好。由此可得, 三种分类算法的性能关系: $RF > SVM > LR$ 。

之所以出现这样的情况, 与三种算法的分类原理密不可分。RF 是由一组决策树分类器组成, 分类时, 不仅每棵树的样本集是从训练数据中随机抽取的, 每棵树的特征集合也是从

原始特征集合中随机抽取的，因此得到的分类器具有更强的泛化能力。更关键的是，RF 能够对输入特征作重要性度量，每次都选择较为重要的特征参与训练，这突出了信息量较大的优质特征在分类中的作用，而弱化无用特征的影响，从而保证分类准确度。所以 RF 算法不但操作简单，易于实现，且无需考虑特征间的复杂关系，便能得到性能优秀的分类模型，非常适用于临床初步构建房颤自动检测模型。

相较而言，SVM 模型和 LR 模型都是通过确定分类边界完成分类，一个是最优超平面的建立，另一个是分界函数模型的建立。这两种算法在分类时，都易受边界值的影响。当特征集中存在大量无用特征时，SVM 和 LR 算法并不能区分这些特征的重要度，而将所有特征同等对待，这将会干扰真正的分类边界，使得建立的模型在新数据集上表现很差。所以在使用 SVM 和 LR 算法训练分类模型时，一定要进行特征选择。

除此之外，观察表 4，还可发现特征选择前后，SVM 模型和 LR 模型在 CPSC2018 数据库上的分类能力，与其他两个测试集相比变化幅度很小。这可能是由于该数据库只有 147 条数据，与其余测试集相比，数据量太少，不能全面反映三种模型的分类能力及特征选择的效果，参考意义不大，但能证明模型的泛化能力。

同时从表 4 可以观察到，当分类模型在 CinC2017 测试集上性能较差时， Se 在 40% 以下， Acc 却在 90% 以上，此时 $F1$ 分数更加符合模型真实性能。这种现象是由 CinC2017 数据库的数据分布不均衡导致的，该数据库有 855 条房颤数据，5477 条非房颤数据。因此在数据分布不平衡的数据库上作模型评价时，计算 $F1$ 分数相较 Acc 更为可靠。

4 结论

针对目前大部分房颤检测算法的通病，即缺乏不同模型或不同数据集，尤其是在临床数据集上的性能对比，无法可靠评价它们在临床应用上的表现，本文提出在两个公开数据库和一个自主研发的可穿戴式心电设备采集数据集上，分别测试并对比分析三种机器学习算法构建得到模型的分类能力，并通过特征选择得到较优特征集。结果表明特征选择后得到的三种模型分类准确度高且泛化能力强，具有较大的临床应用价值。此外三种模型分类能力关系为： $RF > SVM > LR$ 。下一步可以丰富数据库，选用不同的特征选择算法，比较更多分类算法得到的模型，深入了解不同特征集、不同方法对于分类的作用和影响，并提出更为针对性的优化，以期得到更适用于动态心电图实时、自动且准确检测房颤的特征集及分类算法。

参考文献

- [1] LIP G, FAUCHIER L, FREEDMAN S, et al. Atrial fibrillation [J]. Nat Rev Dis Primers, 2016, 2(1): 16016(1)-16016(26).
- [2] KIRCHHOF P, BENUSSI S, KOTTECHA D, et al. 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS [J]. European Heart Journal, 2016, 37(38): 2893-2962.
- [3] 黄从新, 张澍, 黄德嘉, 华伟. 心房颤动:目前的认识和治疗的建议-2018 [J]. 中国心脏起搏与心电生理杂志, 2018, 32(04): 315-368.
- HUANG C X, ZHANG S, HUANG D J, HUA W. Current knowledge and management recommendations of atrial fibrillation-2018 [J]. Chinese Journal of Cardiac Pacing and Electrophysiology, 2018, 32(04): 315-368.
- [4] 张丹, 隋文涛, 梁钊, 王峰. 基于 VMD 和 KNN 的心电信号分类算法 [J]. 电子测量与仪器学报, 2019, 33(04): 140-145.
- ZHANG D, SUI W T, LIANG Z, WANG F. ECG signal classification algorithm based on VMD and KNN [J]. Journal of Electronic Measurement and Instrumentation, 2019, 33(04): 140-145.
- [5] 李梁, 文笃石. 基于 ARM 的心电监护仪的设计与实现 [J]. 国外电子测量技术, 2017, 36(03): 70-73.
- LI L, WEN D S. Design and implementation of ECG monitor based on ARM [J]. Foreign Electronic Measurement Technology, 2017, 36(03): 70-73.
- [6] 霍东风, 谭励夫, 胡学亮, 郭伟桥. 一种提取呼吸信号的可穿戴无线心电监测系统[J]. 国外电子测量技术, 2018, 37(07): 98-102.
- HUO D F, TAN L F, HU X L, GUO W Q. Wearable wireless ECG monitoring system for extracting respiratory signals [J]. Foreign Electronic Measurement Technology, 2018, 37(07): 98-102.
- [7] KARUNADAS C P, MATHEW C. Comparison of arrhythmia detection by conventional Holter and a novel ambulatory ECG system using patch and android App, over 24h period [J]. Indian Pacing and Electrophysiology Journal, 2020, 20(2): 49-53.
- [8] 黄超. 动态心电图中心房颤自动检测算法研究及其临床应用 [D]. 浙江: 浙江大学, 2013.
- HUANG C. Research of automatic detection algorithm of atrial fibrillation in ambulatory ECG and its clinical application [D]. ZheJiang University, 2013.
- [9] HAGIWARA Y, FUJITA H, OH S L, et al. Computer-aided diagnosis of atrial fibrillation based on ECG signals: a review [J]. Information Sciences, 2018, 467: 99-114.
- [10] COLLOCA R. Implementation and testing of atrial fibrillation detectors for a mobile phone application [D]. Italy: Politecnico di Milano, 2013.
- [11] ANDERSEN R S, POULSEN E S, PUTHUSSERYPADY S, et al. A novel approach for automatic detection of atrial fibrillation based on inter beat intervals and support vector machine [C]//International Conference of the IEEE Engineering in Medicine & Biology Society. Conf Proc IEEE Eng Med Biol Soc, 2017: 2039-2042.
- [12] ZABIHI M, BAHRAMI RAD A, KATSAGGELOS A K, et al. Detection of atrial fibrillation in ECG hand-held devices using a random forest classifier [C]//Computing in Cardiology, 2017, 44: 1-4.
- [13] 陈志博, 李健, 李智, 彭韵陶, 高兴姣. 基于 RR 间期和多特征值的房颤自动检测分类 [J]. 生物医学工程学杂志, 2018, 35(04): 550-556.
- CHEN Z B, LI J, LI Z, PENG Y T, GAO X J. Automatic detection and classification of atrial fibrillation using RR intervals and multi-eigenvalue [J]. Journal of Biomedical Engineering, 2018, 35(04): 550-556.
- [14] 孙振. 基于机器学习的网络流量特征选择 [J]. 电子测量技术, 2017, 40(07): 131-136.
- SUN Z. Research of network flow feature selection based on machine learning [J]. Electronic Measurement Technology, 2017, 40(07): 131-136.
- [15] CLIFFORD G D, LIU C Y, MOODY B, et al. AF classification from a short single lead ECG recording: The PhysioNet/Computing in Cardiology Challenge 2017 [C]//2017 Computing in Cardiology (CinC). IEEE, 2017: 1-4.
- [16] LIU F F, LIU C Y, ZHAO L N, et al. An open access database for evaluating the algorithms of ECG rhythm and morphology abnormal detection [J]. Journal of Medical Imaging and Health Informatics, 2018, 8(7): 1368-1373.

- [17] LIU C Y, ZHANG X Y, ZHAO L N, et al. Signal quality assessment and lightweight QRS detection for wearable ECG SmartVest system [J]. IEEE Internet of Things Journal, 2019, 6(2): 1363-1374.
- [18] BEHAR J, JOHNSON A, CLIFFORD G D, et al. A comparison of single channel fetal ECG extraction methods [J]. Annals of Biomedical Engineering, 2014, 42(6): 1340-1353.
- [19] BEHAR J, OSTER J, CLIFFORD G D. Combining and benchmarking methods of foetal ECG extraction without maternal or scalp electrode data [J]. Physiological Measurement, 2014, 35(8): 1569-1589.
- [20] SHARMA T, SHARMA K K. QRS complex detection in ECG signals using locally adaptive weighted total variation denoising [J]. Computers in Biology and Medicine, 2017, 87: 187-199.
- [21] SHANG H X, WEI S S, LIU F F, et al. An improved sliding window area method for T wave detection [J]. Computational and Mathematical Methods in Medicine, 2019: 1-11.
- [22] SUÁREZ-LEÓN A A, VARON C, WILLEMS R, et al. T-wave end detection using neural networks and support vector machines [J]. Computers in Biology and Medicine, 2018, 96: 116-127.
- [23] DATTA S, PURI C, MUKHERJEE A, et al. Identifying normal, AF and other abnormal ECG rhythms using a cascaded binary classifier [C]// Computing in Cardiology, 2017.
- [24] ALCARAZ R, RIETA J J. Applications of nonlinear methods to atrial fibrillation [M]//Barbieri R, Scilingo E P, Valenza G. Complexity and Nonlinearity in Cardiovascular Signals. Cham; Springer International Publishing. 2017: 387-426.
- [25] LI Q, LIU C Y, OSTER J, et al. Signal processing and feature selection preprocessing for classification in noisy healthcare data [M]. Machine Learning for Healthcare Technologies, 2016.
- [26] SIECIŃSKI S, KOSTKA P S, TKACZ E J. Comparison of atrial fibrillation detection performance using decision trees, SVM and artificial neural network [C]// Rocha Á, Ferrás C, Paredes M. Information Technology and Systems. Springer Cham, 2019: 693-701.
- [27] TUBOLY G, KOZMANN G, KISS O, et al. Atrial fibrillation detection based on poincaré plot and P wave analysis [C]// 2019 12th International Conference on Measurement, Smolenice: IEEE, 2019: 17-20.
- [28] OH S H, LEE Y R, KIM H N. A novel EEG feature extraction method using Hjorth parameter [J]. 2014, 2: 106-110.
- [29] 田曼, 杨风雷, 张艺. 基于熵的脑电特征选择情绪识别研究 [J]. 电子测量技术, 2018, 41(19): 106-111.
TIAN M, YANG F L, ZHANG Y. Research on emotion recognition with EEG signal feature selection based on entropy [J]. Electronic Measurement Technology, 2018, 41(19): 106-111.
- [30] LIU C Y, JULIEN O, ERIK R, et al. A comparison of entropy approaches for AF discrimination [J]. Physiological Measurement, 2018, 39(7): 74002(1)-74002(18).
- [31] BIN G Y, SHAO M G, BIN G H, et al. Detection of atrial fibrillation using decision tree ensemble [C]//Computing in Cardiology, 2017, 44.
- [32] 施文, 陆宁云, 姜斌, 支有冉, 许志兴. 数据驱动的地铁车门微小故障智能诊断方法 [J]. 仪器仪表学报, 2019, 40(06): 192-201.
SHI W, LU N Y, JIANG B, ZHI Y R, XU Z X. Data-driven intelligent incipient fault diagnosis for subway vehicle door system [J]. Chinese Journal of Scientific Instrument, 2019, 40(06): 192-201.
- [33] GREGORUTTI B, MICHEL B, SAINT-PIERRE P. Correlation and variable importance in random forests [J]. Statistics and Computing, 2017, 27(3): 659-678.

附录 特征计算方法

F41: Lorenz plot 特征^{错误!未找到引用源。} (AFEv), Lorenz plot 如图 0 所示:

$$AFEvidence = IrregularityEvidence - OriginCount - 2 \times PACEvidence \quad (A-1)$$

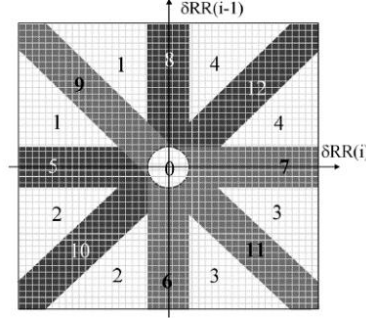


图 A-1 二维直方图 $\{\Delta RR(i), \Delta RR(i-1)\}$, 即 Lorenz plot 的数学表示^{错误!未找到引用源。}
Fig.A-1 The 2-D histogram $\{\Delta RR(i), \Delta RR(i-1)\}$, that's a numeric representation of a Lorenz plot

F42: Lorenz plot 中包含原点的被 $\{\Delta RR(i), \Delta RR(i-1)\}$ 值填充的方格数 (OriginCount)

F43: 测量 Lorenz plot 中点分布的稀疏性 (IrrEv)

$$IrregularityEvidence = \sum_{n=1}^{12} BinCount_n \quad (A-2)$$

式中 $BinCount_n$ 计算直方图的区段 n 中至少被填充过一次的方格数。

F44: 测量了补偿性停顿的特征^{错误!未找到引用源。} (PACEv)

$$PACEvidence = \sum_{n=1}^4 (PointCount_n - BinCount_n) + \sum_{n=5,6,10} (PointCount_n - inCount_n) - \sum_{n=7,8,12} (PointCount_n - BinCount_n) \quad (A-3)$$

式中 $PointCount_n$ 计算直方图区段 n 中被填充的点数。

F45: 测量点聚集的密度^{错误!未找到引用源。} (DensityEvidence)

$$DensityEvidence = \sum_{n=5}^{12} (PointCount_n - BinCount_n) \quad (A-4)$$

F46: 测量点分布的方向^{错误!未找到引用源。} (AnisotropyEvidence)

$$AnisotropyEvidence = \left| \sum_{n=9,11} PointCount_n - \sum_{n=10,12} PointCount_n \right| + \left| \sum_{n=6,7} PointCount_n - \sum_{n=5,8} PointCount_n \right| \quad (A-5)$$

F54~F58: 嵌入维数 $m=1 \sim 5$ 时的 SampEn, 计算过程^[30]如下:

对一 RR 时间序列 $\mathbf{x}(i) (1 \leq i \leq N)$, 给定嵌入维数 m 和距离阈值 r , 定义:

$$\mathbf{X}_i^m = \{\mathbf{x}(i), \mathbf{x}(i+1), \dots, \mathbf{x}(i+m-1)\} \quad (A-6)$$

向量 $\mathbf{X}_i^m (1 \leq i \leq N-m)$ 表示 m 个连续的向量 $\mathbf{x}(i)$ 的值。然后计算 \mathbf{X}_i^m 和 \mathbf{X}_j^m 之间的距离, 即最大绝对差为:

$$d_{i,j}^m = \max_{0 \leq k \leq m-1} |\mathbf{x}(i+k) - \mathbf{x}(j+k)| \quad (A-7)$$

对每个 \mathbf{X}_i^m , $(1 \leq i \leq N-m)$, 定义 $B_i^m(r)$ 为满足 $d_{i,j}^m \leq r$ 的 \mathbf{X}_j^m 的数量的 $1/(N-m)$ 。类

似地，定义 $A_i^{m+1}(r)$ 为满足 $d_{i,j}^{m+1} \ll r$ 的 \mathbf{X}_j^{m+1} 的数量的 $1/(N-m)$ 。在维数 $m+1$ 和 m 处的相似向量的总数计算为：

$$A^{m+1}(r) = \sum_{i=1}^{N-m} A_i^{m+1}(r), B^m(r) = \sum_{i=1}^{N-m} B_i^m(r) \quad (\text{A-8})$$

SampEn 定义为：

$$\text{SampEn} = -\ln\left(\frac{A^{m+1}(r)}{B^m(r)}\right) \quad (\text{A-9})$$

F68~F72：嵌入维数为 1~5 时的 ApEn，计算过程如下：

对于 N 点原始时间序列 $\mathbf{x}(1), \mathbf{x}(2), \dots, \mathbf{x}(N)$ ，以时间延迟 τ 重建 m 维相空间，得到一组 m 维矢量：

$$\mathbf{X}_i^m = \{\mathbf{x}(i), \mathbf{x}(i+\tau), \dots, \mathbf{x}(i+(m-1)\tau)\} \quad (\text{A-10})$$

其中 $(1 \ll i \ll N-m)$ ，定义 \mathbf{X}_i^m 和 \mathbf{X}_j^m 之间的距离 $d_{i,j}^m$ 为两者对应元素中差值最大的一个，即：

$$d_{i,j}^m = \max_{0 \leq k \leq m-1} |\mathbf{x}(i+k) - \mathbf{x}(j+k)| \quad (\text{A-11})$$

统计对每个 i 值的 $d_{i,j}^m$ 小于 r 的数目，并求其与距离总数 $N-m$ 的比值，记作：

$$C_i^m(r) = \frac{\text{num}\{d_{i,j}^m < r\}}{N-m} \quad (\text{A-12})$$

则

$$\phi^m(r) = (N-m+1)^{-1} \sum_{i=1}^{N-m+1} \ln[C_i^m(r)] \quad (\text{A-13})$$

ApEn 定义为：

$$\text{ApEn}(m, r, N) = \phi^m(r) - \phi^{(m+1)}(r) \quad (\text{A-14})$$

F73：FuzzyMEn，计算过程^[30]如下：

对一时间序列 $\mathbf{x}(i) (1 \ll i \ll N)$ ，分别计算局部矢量序列 \mathbf{L}_i^m 和全局向量序列 \mathbf{G}_i^m ，其中 $(1 \ll i \ll N-m)$ ，则

$$\mathbf{L}_i^m = \{\mathbf{x}(i), \mathbf{x}(i+1), \dots, \mathbf{x}(i+m-1)\} - \bar{\mathbf{x}}(i) \quad (\text{A-15})$$

$$\mathbf{G}_i^m = \{\mathbf{x}(i), \mathbf{x}(i+1), \dots, \mathbf{x}(i+m-1)\} - \bar{\mathbf{x}} \quad (\text{A-16})$$

矢量 \mathbf{L}_i^m 表示 m 个去掉局部平均值的连续时间序列 $\mathbf{x}(i)$ ，其中局部平均值 $\bar{\mathbf{x}}(i) = \frac{1}{m} \sum_{k=0}^{m-1} \mathbf{x}(i+k)$ ；矢量 \mathbf{G}_i^m 代表 m 个去掉全局平均值的连续时间序列 $\mathbf{x}(i)$ ，其中全局平均值 $\bar{\mathbf{x}} = \frac{1}{N} \sum_{i=1}^N \mathbf{x}(i)$ 。局部矢量序列 \mathbf{L}_i^m 和 \mathbf{L}_j^m 之间的距离以及全局矢量序列 \mathbf{G}_i^m 和 \mathbf{G}_j^m 之间的距离分别定义为：

$$dL_{i,j}^m = \max_{0 \leq k \leq m-1} \left| \left(\mathbf{x}(i+k) - \bar{\mathbf{x}}(i) \right) - \left(\mathbf{x}(j+k) - \bar{\mathbf{x}}(j) \right) \right| \quad (\text{A-17})$$

$$dG_{i,j}^m = \max_{0 \leq k \leq m-1} \left| \left(\mathbf{x}(i+k) - \bar{\mathbf{x}} \right) - \left(\mathbf{x}(j+k) - \bar{\mathbf{x}} \right) \right| \quad (\text{A-18})$$

给定以下参数：局部相似性权重 n_L ，局部容忍阈值 r_L ，全局相似性权重 n_G 和全局容忍阈值 r_G ，然后通过模糊函数计算局部向量 \mathbf{L}_i^m 和 \mathbf{L}_j^m 之间的相似度 $DL_{i,j}^m(n_L, r_L)$ 和全局矢量 \mathbf{G}_i^m 和 \mathbf{G}_j^m 之间的相似度 $DG_{i,j}^m(n_G, r_G)$ ：

$$DL_{i,j}^m(n_L, r_L) = \exp\left(-\frac{(dL_{i,j}^m)n_L}{r_L}\right) \quad (\text{A-19})$$

$$DG_{i,j}^m(n_G, r_G) = \exp\left(-\frac{(dG_{i,j}^m)n_G}{r_G}\right) \quad (\text{A-20})$$

函数 $BL^m(n_L, r_L)$ 和 $BG^m(n_G, r_G)$ 定义为：

$$BL^m(n_L, r_L) = \frac{1}{N-m} \sum_{i=1}^{N-m} \left(\frac{1}{N-m} \sum_{j=1}^{N-m} DL_{i,j}^m(n_L, r_L) \right) \quad (\text{A-21})$$

$$BG^m(n_G, r_G) = \frac{1}{N-m} \sum_{i=1}^{N-m} \left(\frac{1}{N-m} \sum_{j=1}^{N-m} DG_{i,j}^m(n_G, r_G) \right) \quad (\text{A-22})$$

$BL^m(n_L, r_L)$ 和 $BG^m(n_G, r_G)$ 分别测量 m 维局部向量和全局向量的平均相似度。同样定义 $m+1$ 维平均相似度的函数为 $AL^{m+1}(n_L, r_L)$ 和 $AG^{m+1}(n_G, r_G)$ ，计算公式分别为：

$$AL^{m+1}(n_L, r_L) = \frac{1}{N-m} \sum_{i=1}^{N-m} \left(\frac{1}{N-m} \sum_{j=1}^{N-m} DL_{i,j}^{m+1}(n_L, r_L) \right) \quad (\text{A-23})$$

$$AG^{m+1}(n_G, r_G) = \frac{1}{N-m} \sum_{i=1}^{N-m} \left(\frac{1}{N-m} \sum_{j=1}^{N-m} DG_{i,j}^{m+1}(n_G, r_G) \right) \quad (\text{A-24})$$

然后，二次模糊局部测量熵 (H_L^θ) 和二次模糊全局测量熵 (H_G^θ) 分别定义为：

$$H_L^\theta = -\ln\left(\frac{AL^{m+1}(n_L, r_L)}{BL^m(n_L, r_L)}\right) \quad (\text{A-25})$$

$$H_G^\theta = -\ln\left(\frac{AG^{m+1}(n_G, r_G)}{BG^m(n_G, r_G)}\right) \quad (\text{A-26})$$

最终 FuzzyEn 计算为：

$$FuzzyMEN = H_L^\theta(m, n_L, r_L, N) + H_G^\theta(m, n_G, r_G, N) \quad (\text{A-27})$$

F74: NFE_n^[30]

首先定义嵌入维数 m 和距离阈值 r 。与使用基于概率方法的 FuzzyMEN 不同，NFE_n 使用基于密度的方法来生成二次模糊局部测量熵 (H_L^θ) 和二次模糊全局测量熵 (H_G^θ)，即采用每个匹配区段的体积如 $(2r)^m$ ：

$$H_L^\theta = -\ln\left(\frac{AL^{m+1}(n_L, r_L)/(2r_L)^{m+1}}{BL^m(n_L, r_L)/(2r_L)^m}\right) = -\ln\left(\frac{AL^{m+1}(n_L, r_L)}{BL^m(n_L, r_L)}\right) + \ln(2r_L) \quad (\text{A-28})$$

$$H_G^\theta = -\ln\left(\frac{AG^{m+1}(n_G, r_G)/(2r_G)^{m+1}}{BG^m(n_G, r_G)/(2r_G)^m}\right) = -\ln\left(\frac{AG^{m+1}(n_G, r_G)}{BG^m(n_G, r_G)}\right) + \ln(2r_G) \quad (\text{A-29})$$

从 H_L^θ 和 H_G^θ 中减去平均 RR 间隔的自然对数，如下所示：

$$H_L^\theta = -\ln\left(\frac{AL^{m+1}(n_L, r_L)}{BL^m(n_L, r_L)}\right) + \ln(2r_L) - \ln(mRR) \quad (\text{A-30})$$

$$H_G^\theta = -\ln\left(\frac{AG^{m+1}(n_G, r_G)}{BG^m(n_G, r_G)}\right) + \ln(2r_G) - \ln(mRR) \quad (\text{A-31})$$

最终 NFE_n 计算为：

$$\begin{aligned} NFE_n &= H_L^\theta + H_G^\theta \\ &= -\ln\left(\frac{AL^{m+1}(n_L, r_L)}{BL^m(n_L, r_L)}\right) - \ln\left(\frac{AG^{m+1}(n_G, r_G)}{BG^m(n_G, r_G)}\right) + \ln(2r_L) + \ln(2r_G) - 2 \times \ln(mRR) \end{aligned} \quad (\text{A-32})$$

其中 mRR 是当前窗口中 RR 间隔的平均值， r_L 、 r_G 和 mRR 以秒为单位。

作者简介:



周帆, 2018 年于西南交通大学获得学士学位, 现为东南大学硕士在读, 主要研究方向为房颤信号检测。

E-mail: fanz1996@seu.edu.cn

Zhou Fan received her B.Sc. degree in 2018 from Southwest Jiaotong University, now she is a M. Sc. candidate in Southeast University. Her main research interest includes atrial fibrillation detection.



刘澄玉, 2005 年于山东大学获得学士学位, 2010 年于山东大学获得博士学位, 现为东南大学教授, 主要研究方向为健康智能监测、医学大数据处理与机器学习、计算机辅助诊断的早期检测、计算机辅助诊断设备的发展、睡眠与情绪监测。

E-mail: chengyu@seu.edu.cn

Liu Chengyu (Corresponding author) received his B.Sc. degree in 2005 from Shandong University, received his Ph.D. degree in 2010 from Shandong University, now he is a professor in Southeast University. His main research interests include mHealth and intelligent monitoring, machine learning and big data processing for physiological signals, early detection of CADs, device development for CADs, sleep and emotion monitoring.