

DOI: 10.3969/j.issn.1671-024x.2025.04.007

## 基于心率变异性与机器学习的养老护理人员疲劳分类

张欣<sup>1</sup>, 马帅<sup>2</sup>, 欧宗锟<sup>2</sup>, 彭成<sup>3</sup>, 韦然<sup>4</sup>

(1. 民政职业大学 老年福祉学院, 北京 102600; 2. 天津工业大学 电子与信息工程学院, 天津 300387; 3. 天津工业大学 生命科学学院, 天津 300387; 4. 民政职业大学 康复工程学院, 北京 102600)

**摘要:** 针对养老护理人员疲劳评估中主观量表法的局限性, 提出基于心率变异性(HRV)与机器学习的养老护理人员疲劳分类方法。通过可穿戴设备采集78名护理人员736 h心电信号, 使用巴特沃斯滤波与Pan-Tompkins算法进行预处理, 系统提取了心率变异性的时域、频域及非线性域共42个特征参数, 并采用皮尔逊相关系数法筛选18个关键特征, 构建基于XGBoost的疲劳分类模型。结果表明: 通过特征降维显著提升了模型性能, 分类准确率从0.78提升至0.91, 其中重度疲劳分类准确率达到0.99; 与SVM、KNN等传统模型相比, XGBoost在特征精简后展现出最优的非线性拟合能力; 2 min短时窗心电数据的分类性能最优, 准确率达0.91, 显著优于4~10 min心电数据准确率(0.61~0.69)。该方法实现了护理疲劳的高效精准监测, 为预防职业性心脑血管疾病及优化养老服务管理提供了可穿戴智能技术支持。

**关键词:** 心率变异性; 机器学习; 心电信号; 特征降维; XGBoost; 疲劳分类

中图分类号: TP391.4

文献标志码: A

文章编号: 1671-024X(2025)04-0044-08

### Fatigue classification of service staff in elderly care based on heart rate variability and machine learning

ZHANG Xin<sup>1</sup>, MA Shuai<sup>2</sup>, OU Zongkun<sup>2</sup>, PENG Cheng<sup>3</sup>, WEI Ran<sup>4</sup>

(1. College of Elderly Welfare, China Civil Affairs University, Beijing 102600, China; 2. School of Electronic and Information Engineering, Tiangong University, Tianjin 300387, China; 3. School of Life Sciences, Tiangong University, Tianjin 300387, China; 4. College of Rehabilitation Technology, China Civil Affairs University, Beijing 102600, China)

**Abstract:** Aiming at the limitations of the subjective scale method in the fatigue assessment of elderly caregivers, this study proposes a fatigue classification method for elderly caregivers based on heart rate variability (HRV) and machine learning. We collected 736 h of ECG data from 78 caregivers through wearable devices, used Butterworth filtering and Pan-Tompkins algorithm for preprocessing, systematically extracted a total of 42 feature parameters in the time domain, frequency domain, and nonlinear domain of HRV, and used Pearson correlation coefficient to filter 18 key features to construct a fatigue classification model based on XGBoost. The experimental results show that the model performance is significantly improved by feature dimensionality reduction, and the classification accuracy is increased from 0.78 to 0.91, among which the accuracy of heavy fatigue classification reaches 0.99; compared with the traditional models such as SVM, KNN, etc., XGBoost demonstrates the optimal nonlinear fitting ability after feature streamlining. Further studies showed that the classification performance of 2 min short-time window ECG data was optimal with an accuracy of 0.91, which was significantly better than the 4~10 min ECG data accuracy of 0.61~0.69, respectively. This method achieves efficient and accurate monitoring of nursing fatigue, and provides wearable smart technology support for the prevention of occupational cardiovascular and cerebrovascular diseases and the optimization of elderly service management.

**Key words:** heart rate variability (HRV); machine learning; electrocardiogram (ECG); feature dimension reduction; XGBoost; fatigue classification

收稿日期: 2025-03-22

基金项目: 中央高校基本科研业务费资助项目(JBKYJJCX2024-10)

第一作者: 张欣(1981—), 女, 硕士, 副教授, 主要研究方向为医养结合和老年护理。E-mail: zhangxin@bcsa.edu.cn

通信作者: 韦然(1982—), 男, 博士, 副教授, 主要研究方向为医学电子诊疗技术。E-mail: wrccau@163.com

随着我国老龄化进程的加速,养老需求日益增长,使得养老机构护理人员面临着日益繁重的工作压力,其身心健康状况不容乐观<sup>[1]</sup>。疲劳作为一种普遍存在的生理和心理状态,不仅反映了身体和心理机能的变化,还可能对个人的工作效率、学习能力以及日常生活质量产生显著的负面影响<sup>[2]</sup>。研究表明,长期处于高强度工作状态的护理人员普遍存在显著的疲劳症状<sup>[3]</sup>。这种职业性疲劳不仅会导致工作效率下降和服务质量降低,更可能引发自主神经功能紊乱、心律失常等一系列心血管健康问题<sup>[4]</sup>。因此,建立科学有效的疲劳分类评估体系,对于保障护理人员职业健康、提升养老服务品质具有重要的现实意义。

传统疲劳评估方法多采用主观量表<sup>[5-6]</sup>,存在响应延迟和个体偏差等固有缺陷,难以准确反映护理人员的真实疲劳状态。近年来,心电信号(Electrocardiogram, ECG)<sup>[7-8]</sup>逐渐被用于疲劳监测领域。通过分析心电信号的特征,可以揭示心脏在疲劳状态下的电生理活动变化,为预防和管理疲劳相关的心脏疾病提供重要的信息<sup>[9]</sup>。研究表明,疲劳状态通常伴随着心率恢复时间延长,且疲劳状态可能增加心律失常的风险。研究与疲劳相关的心率变异性(heart rate variability, HRV)的改变是理解疲劳进程的一个重要方面<sup>[10-11]</sup>。HRV是描述心率变化的自主活动的量度,其功率谱分析提供了心率自主调节的测量。在疲劳状态下,身体对自主神经系统的调控能力不足,从而造成交感神经系统和副交感神经系统失衡。这种平衡失调会反映在HRV的低频成分中,从而反映出交感神经和副交感神经的共同活动水平,因此,将HRV作为研究主体,可有效判断被试者是否处于疲劳状态<sup>[12-14]</sup>。然而,HRV特征与疲劳状态的关联具有高度非线性和个体差异性,传统统计方法(如线性回归)难以充分捕捉其复杂模式。极限梯度提升(extreme gradient boosting, XGBoost)<sup>[15]</sup>是一种高效的机器学习优化模型,在处理高维HRV特征时表现出色。

基于上述背景,本研究以养老机构护理人员为研究对象,采用可穿戴设备实时采集其心电信号,从中提取HRV特征;为避免特征之间的高度相关性对模型造成不良影响,采用皮尔逊相关系数法对特征进行筛选,并结合主观疲劳量表,利用XGBoost模型对疲劳进行分类识别。XGBoost通过迭代训练一系列决策树并整合其结果,有效拟合HRV多维度特征间的复杂非线性关系,精准提取疲劳状态的生理特征。研究结果表明,该方法不仅提高了护理人员的工作效率,

减少因疲劳导致的操作失误风险,还能有效预防因过度疲劳引发的心脑血管疾病甚至猝死事件,从而改善养老机构护理人员的健康状况,促进养老服务业的可持续发展。

## 1 基于HRV的疲劳特征提取

### 1.1 研究对象

本研究采用横断面调查设计,通过自行开发的一套可穿戴远程心电监护设备,连续采集高质量的ECG信号。设备采用五导联进行采集,采样率为500 Hz,同时本设备对各模块进行了集成封装,质量仅为77.5 g,大小8 cm,便于携带。采集某三甲医院养老护理人员的工作时心电数据。研究纳入标准:①持有效执业证书的在岗护理人员;②连续工作时间 $\geq 1$ 年;③自愿参与研究并签署知情同意书。排除标准:①患有严重心血管疾病者;②近30 d内服用影响自主神经功能药物者;③数据采集不全者(有效时长 $< 8$  h/d)。最终纳入78名护理人员,其中普通病区56名,重症病区22名;男性14名,女性64名;年龄( $30.78 \pm 3.58$ )岁(范围:25~42岁)。共获取有效心电数据147段,每段时长 $\geq 30$  min,总监测时长达736 h。

### 1.2 心电数据预处理

为了更好地进行心电信号的特征提取,首先采用一阶巴特沃斯低通滤波器(截止频率2 Hz)进行零相位滤波,提取信号低频成分;随后,通过原始信号减去低频成分实现高通滤波,保留0.05~100.00 Hz有效频带信号,心电信号预处理前后对比图如图1所示。

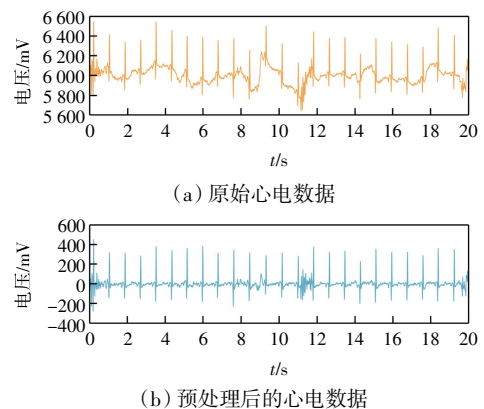


图1 心电信号预处理前后对比

Fig.1 Comparison of ECG signal before and after preprocessing

由图1可以看出,去噪后的心电信号波形变得平滑,具有心电信号的所有特征,心电信号的质量明显

增强,提高了波形检测的准确率。

为确保 HRV 分析的准确性,本文采用 Pan-Tompkins 算法进行 R 波检测,并对异常搏动(如室性早搏、噪声干扰等)进行人工校正<sup>[6]</sup>。为满足频域特征与非线性特征的计算要求,将预处理后的数据以 2 min 为时间窗进行分段,剔除不完整片段(时长 < 120 s),最终获得 6 120 组有效数据。根据护理人员工作状态(基于 NASA-TLX 量表评分),将数据划分为轻度疲劳(0)、中度疲劳(1)和重度疲劳(2)3 类并建立相应标签。

### 1.3 HRV 特征提取

本文以 HRV 的时域、频域和非线性域作为疲劳度的主要观察指标,共选取关键参数,包括平均心率(HR)、RR 间期标准差(SDNN)、相邻 RR 间期差值的均方根(RMSSD)、大于 50 ms 的 RR 间期差值百分比(pNN50)、归一化低频功率(LFn)、归一化高频功率(HFn)以及归一化低频与高频功率比值(LFn/HFn)。这些参数能够全面反映自主神经系统的活动状态,为疲劳度的评估提供科学依据<sup>[7]</sup>。本文首先使用 Pan\_tompkin 算法寻找心电信号的 R 峰,之后通过 R 峰的序列得到 HRV 参数,共计算出 42 个 HRV 参数。它们的分布情况如表 1 所示。

表 1 所使用的所有 HRV 特征  
Tab.1 All HRV features used

特征种类	特征数	特征简写
时域	9	meanHR、MEAN、SDNN、RMSSD、TINN、SDNNindex、pNN50、SDSD、HrvIndex
频域	7	HF、LF、VLF、TotalPower、LFndHFn、HFn、LFn
非线性域	26	SD1、SD2、index、VLI、VAI、ED(1,1)、ED(1,2)、ED(1,3)、ED(1,4)、Rtf、Rpf、Rnf、SP_EN、DB_EN、AP_EN、GI、PI、EI、S、CCM_1、GI_MSTI、PI_MSTI、DI_HD、TI、CTM_SD1、CTM_SD2、CTM_120

#### 1.3.1 时域特征

本研究提取了 9 个时域特征。其中包括 meanHR、MEAN、SDNN、RMSSD、TINN、SDNNindex、pNN50、SDSD、HrvIndex。meanHR 为平均心率,反映整体心脏活动水平;MEAN 为 RR 间期均值,表征心脏节律的平均状态;SDNN 为 RR 间期标准差,评估整体 HRV 变化幅度;RMSSD 为相邻 RR 间期差值的均方根,反映副交感神经活性;TINN 表示 RR 间期直方图三角插值,描述 HRV 分布范围;SDNNindex 为 5 min 时段 SDNN 的均值,评估短期 HRV 稳定性;pNN50 表示相邻 RR 间期差值大于 50 ms 的比例,量化心率快速调节能力;SDSD 为相邻 RR 间期差值的标准差,进一步刻画短期 HRV 波动;HrvIndex 为 HRV 指数,基于 RR 间期直方图的总面积与峰值之比,综合反映 HRV 复杂度。

#### 1.3.2 频域特征

本文提取了 7 个频域特征,用于分析 HRV 的频率成分及其与自主神经系统的关联。具体特征及其生理意义如下:HF(高频功率,0.15~0.40 Hz)反映副交感神经活性;LF(低频功率,0.04~0.15 Hz)表征交感神经与副交感神经的共同调节;VLF(极低频功率,0.003 3~0.040 0 Hz)与体温调节和激素分泌相关;TotalPower(总功率,0~0.40 Hz)评估 HRV 的整体能量水平;LFn(归一化低频功率,LF/(TotalPower-VLF) × 100)量化交感神经相对活性;HFn(归一化高频功率,HF/(TotalPower-VLF) × 100)反映副交感神经相对活性;LFndHFn(LF/HF 比值)表征交感-副交感神经平衡状态。研究表明,LF/HF 比值与疲劳程度显著相关(p < 0.01),能够有效区分不同疲劳状态。这些频域特征为深入理解自主神经系统在疲劳状态下的调节机制提供了重要的量化指标<sup>[8]</sup>。

#### 1.3.3 非线性域特征

本文提取了 26 个非线性域特征,用于全面刻画 HRV 的复杂性和非线性动力学特性。具体特征及其生理意义如下:SD1 和 SD2(Poincaré 图的短轴和长轴标准差)分别反映短期和长期 HRV 波动;index(散点图指数)量化 RR 间期分布的非线性特征;VLI 和 VAI(向量长度指数和向量角度指数)描述 RR 间期变化的几何特性;ED(1,1)至 ED(1,4)(经验分布函数)评估 RR 间期分布的统计特性;Rtf、Rpf 和 Rnf(时间反转、相位反转和噪声反转特征)刻画 HRV 的时间不对称性;SP\_EN、DB\_EN 和 AP\_EN(Shannon、分布基和近似熵)评估 HRV 的复杂度;GI、PI 和 EI(全局、局部和熵指数)反映 HRV 的多尺度特性;S(样本熵)量化 HRV 的规则性;CCM\_1(交叉条件熵)评估 RR 间期序列的依赖性;GI\_MSTI 和 PI\_MSTI(多尺度全局和局部指数)刻画 HRV 的多尺度非线性特性;DI\_HD(高阶差分指数)评估 RR 间期的高阶统计特性;TI(趋势指数)描述 HRV 的长期趋势;CTM\_SD1、CTM\_SD2 和 CTM\_120(中心趋势测度)分别量化 Poincaré 图和 120 s 窗口内 HRV 的中心趋势。研究表明,样本熵(S)和近似熵(AP\_EN)与疲劳程度显著相关(p < 0.05),能够有效区分不同疲劳状态。这些非线性域特征为深入理解 HRV 的复杂动力学特性及其在疲劳状态下的变化规律提供了多维度的量化指标<sup>[9]</sup>。

### 1.4 基于皮尔逊相关系数的心电特征选择

在基于心电信号(ECG)的疲劳分类研究中,初始特征集通常包含大量 HRV 指标及其他相关特征。然而,这些特征之间可能存在高度相关性,即多重共线性问题,这不仅增加了模型的复杂度,还可能导致回

归模型权重失真,降低模型的泛化能力和预测性能。为解决这一问题,本文采用皮尔逊相关系数法对初始42个特征进行降维优化,以剔除冗余特征,保留独立性强、信息量高的特征,从而提升模型的效率和鲁棒性。

皮尔逊相关系数(Pearson correlation coefficient, PCC)是一种用于衡量2个连续变量之间线性相关性的统计指标,其值域范围<sup>[20]</sup>为[-1, 1]。当PCC的绝对值越接近1时,表明变量之间的线性相关性越强;而当其值接近0时,则表明变量之间无显著线性相关性<sup>[21]</sup>。在本研究中,PCC被用于量化每个HRV特征与疲劳等级标签之间的相关性,从而筛选出对分类任务最具判别力的特征子集。

对于给定的HRV特征集  $X=\{x_1, x_2, \dots, x_n\}$  和疲劳等级标签  $Y$ ,特征  $x_i$  与标签  $Y$  的皮尔逊相关系数  $r_i$  计算公式为:

$$r_i = \frac{\text{cov}(x_i, Y)}{\sigma_{x_i} \cdot \sigma_Y} \quad (1)$$

式中: $\text{cov}(x_i, Y)$ 为  $x_i$  与  $Y$  的协方差; $\sigma_{x_i}$  和  $\sigma_Y$  分别为  $x_i$  和  $Y$  的标准差。为了消除不同HRV特征之间量纲的影响,本文对所有HRV特征进行Z-score标准化处理,其计算公式为:

$$x'_i = \frac{x_i - \mu_{x_i}}{\sigma_{x_i}} \quad (2)$$

式中: $\mu_{x_i}$  为  $x_i$  特征的均值。通过计算每个标准化后的HRV特征  $x_i$  与疲劳等级标签  $Y$  的PCC值  $r_i$ 。根据预设阈值  $\theta$ (本文取  $\theta = 0.3$ ),保留  $|r_i| \geq \theta$  的特征,剔除低相关性特征。将筛选后的特征按  $|r_i|$  值从大到小排序,构建特征重要性列表。

本文首先对初始42个特征进行两两皮尔逊相关系数分析,结果如图2所示。图中横纵轴分别表示特征名称,数值代表两个特征之间的相关程度。

以0.9为阈值,若2个特征的相关系数  $\geq 0.9$ ,则认为它们高度相关,需剔除其中之一以减少冗余。基于上述方法,本文共剔除了24个高度相关特征(序号:0, 2, 3, 4, 5, 6, 7, 8, 10, 12, 13, 14, 15, 20, 23, 26, 27, 31, 32, 35, 38, 40, 41),最终保留了18个独立性强、信息冗余低的特征。对剩余特征重新编号(0-17)后,再次计算相关系数,结果如图3所示。

通过验证,所有特征之间的相关系数均  $< 0.9$ ,满足模型输入要求。特征数量从42降至18,减少了57%的数据维度,显著降低了模型训练时间与内存占用,提升了运算效率。高度相关特征易导致回归模型权重失真,剔除这些特征后,参数估计的稳定性得到显著增强。此外,冗余特征的移除减少了噪声干扰,避免了模型过拟合训练数据,从而提高了模型在高压新场景下的预测鲁棒性。

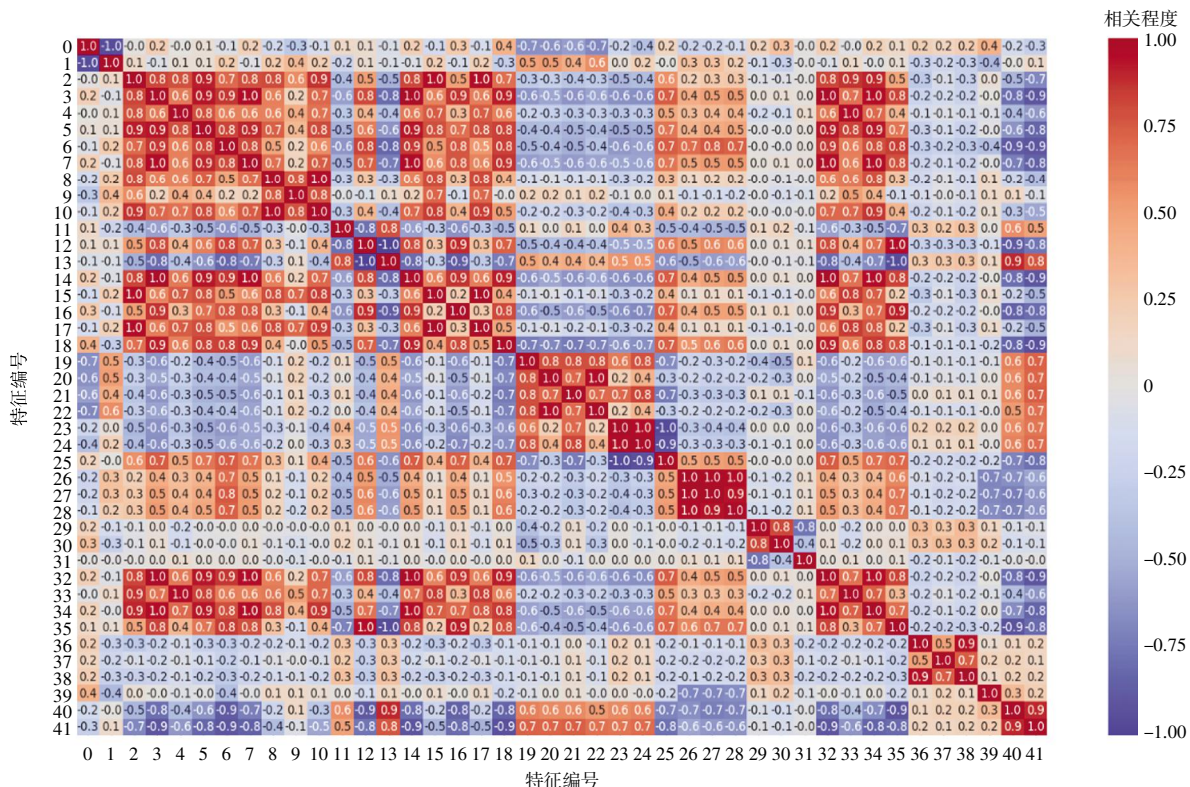


图2 42个特征的皮尔逊相关系数  
Fig.2 Pearson correlation coefficient plot of 42 features

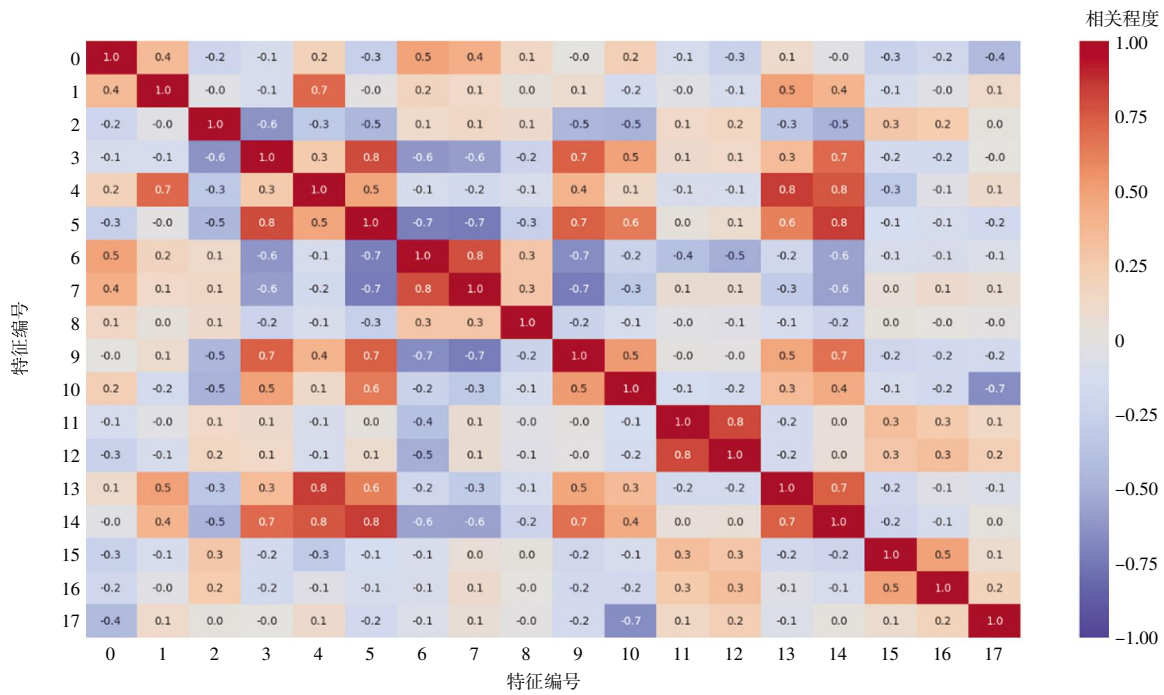


图 3 18 个特征的皮尔逊相关系数

Fig.3 Pearson correlation coefficient plot of 18 features

## 2 基于 HRV 的疲劳分类

### 2.1 评价指标

为了评估模型分类的性能, 本文通过准确率、精确率(Precision)、召回率(Recall)和 F1 分数 4 个指标对模型进行评估。各指标计算公式为:

$$\text{准确率} = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

$$\text{精确率} = \frac{TP}{TP + FP} \quad (4)$$

$$\text{召回率} = \frac{TP}{TP + FN} \quad (5)$$

$$\text{F1 分数} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (6)$$

式中:  $TP$  为模型成功预测为正类别的正样本;  $TN$  为模型成功预测为负类别的负样本;  $FP$  为模型错误地将负样本预测为正样本;  $FN$  为模型错误地将正样本预测为负样本。

### 2.2 不同特征维度的疲劳分类结果对比

#### 2.2.1 基于 42 个特征的 XGBoost 模型疲劳分类

为评估疲劳分类模型的性能, 本文采用分层抽样方法将数据集划分为训练集和测试集。其中, 训练集包含 5 304 例样本, 测试集包含 816 例样本。每例样本由 42 个 HRV 特征值及其对应的疲劳等级标签组成, 疲劳等级分为 3 类: 0(轻度疲劳)、1(中度疲劳)和 2(重度疲劳)。XGBoost 模型通过迭代构建多个决策树,

优化加权对数损失函数以最小化分类误差, 并结合 L1 和 L2 正则化防止过拟合。在多分类任务中, XGBoost 采用“softmax”目标函数, 通过加权投票确定最终类别<sup>[22-23]</sup>。本研究基于 XGBoost 分类模型对疲劳进行三分类, 其中模型参数中的树的最大深度为 18, 树数量为 300, 学习率为 0.01。在 42 个特征的全量数据集上, XGBoost 模型在测试集上的分类结果如表 2 所示。

表 2 42 个特征 XGBoost 分类结果

Tab.2 Classification results of XGBoost with 42 features

类别	准确率	精确率	召回率	F1 分数
轻度疲劳	0.882 5	0.887 9	0.882 5	0.881 2
中度疲劳	0.713 5	0.569 5	0.713 5	0.633 4
重度疲劳	0.699 3	0.823 1	0.699 3	0.756 2

由表 2 可以看出, 模型对不同疲劳类别的分类性能存在显著差异。轻度疲劳的分类效果最佳, 其准确率、精确率、召回率和 F1 分数分别为 0.882 5、0.887 9、0.882 5 和 0.881 2, 表明模型对该类疲劳的识别能力较强。中度疲劳的分类性能相对较低, 其准确率和 F1 分数分别为 0.713 5 和 0.633 4, 表明模型在区分中度疲劳时存在一定困难。重度疲劳的分类性能介于轻度疲劳和重度疲劳之间, 其准确率和 F1 分数分别为 0.699 3 和 0.756 2, 虽然优于对中度疲劳的区分效果, 但仍有较大的提升空间。

#### 2.2.2 基于 18 个特征的 XGBoost 模型疲劳分类

在 18 个特征的优化数据集上, XGBoost 模型在测试集上的分类结果如表 3 所示。

表 3 18 个特征 XGBoost 分类结果

Tab.3 Classification results of XGBoost with 18 features

类别	准确率	精确率	召回率	F1 分数
轻度疲劳	0.891 6	0.964 2	0.891 6	0.926 4
中度疲劳	0.797 8	0.904 5	0.797 8	0.847 8
重度疲劳	0.990 2	0.860 8	0.990 2	0.921 0

由表 3 可以看出,模型对不同疲劳类别的分类性能显著提升,尤其是在中度疲劳和重度疲劳类别上表现尤为突出。轻度疲劳的分类效果依然优异,其准确率、精确率、召回率和 F1 分数分别为 0.891 6、0.964 2、0.891 6 和 0.926 4,表明模型对该类别的识别能力进一步增强。中度疲劳的分类性能显著改善,其准确率和 F1 分数分别提升至 0.797 8 和 0.847 8,表明特征降维有效增强了模型对中度疲劳关键特征的捕捉能力。重度疲劳的分类性能接近完美,其准确率和 F1 分数分别达到 0.990 2 和 0.921 0,表明保留的 18 个特征中可能包含区分该类别的决定性生理信号参数。总体而言,特征降维不仅显著提升了模型在中度疲劳和重度疲劳类别上的分类性能,还保持了轻度疲劳分类效果的稳定性。

2.2.3 特征降维前后的 XGBoost 模型分类性能对比

为评估特征降维对模型分类性能的影响,本文分别绘制了特征降维前后的混淆矩阵,如图 4 所示。图 4 中,对角线上的数值表示预测正确的样本数量,而非对角线上的数值则表示分类错误的样本数量。

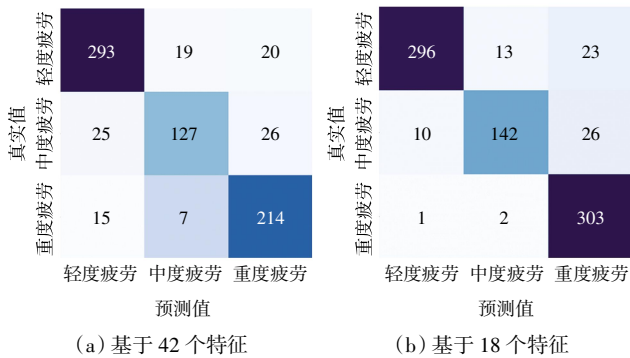


图 4 特征降维前后的混淆矩阵

Fig.4 Confusion matrices before and after feature dimensionality reduction

为进一步量化特征降维对模型性能的提升效果,本文还绘制了特征降维前后的 ROC 曲线,如图 5 所示。图 5 中曲线下面积(area under curve, AUC)是衡量模型性能的重要指标,AUC 值越大,表明模型的性能越好。由图 5 可以发现,特征降维后的 ROC 曲线更接近左上角,且 AUC 值显著高于降维前。由此表明,特征降维不仅提升了模型在各类别上的分类性能,还

显著增强了模型的整体区分能力。

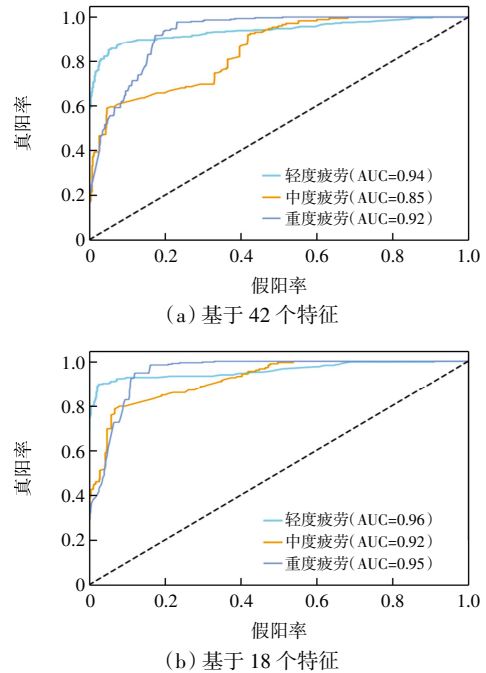


图 5 特征降维前后的 ROC 曲线

Fig.5 ROC curves before and after feature dimensionality reduction

2.2.4 特征降维前后不同模型的疲劳分类结果对比

为全面评估特征降维对模型性能的影响,本文选取了支持向量机(SVM)、K 近邻算法(KNN)、决策树(DT)、逻辑回归(LR)、随机森林(RF)以及 XGBoost 等 6 种机器学习模型进行对比分析。对比指标包括准确率、加权平均下的精确率、召回率及 F1 分数。表 4 和表 5 分别展示了 42 个特征和 18 个特征下不同模型的性能对比结果。

表 4 42 个特征不同模型对比结果

Tab.4 Classification results of different models with 42 features

模型	准确率	精确率	召回率	F1 分数
SVM	0.41	0.36	0.41	0.37
KNN	0.55	0.60	0.55	0.56
DT	0.57	0.63	0.57	0.59
LR	0.53	0.53	0.53	0.53
RF	0.74	0.74	0.74	0.73
XGBoost	0.78	0.79	0.78	0.78

表 5 18 个特征不同模型对比结果

Tab.5 Classification results of different models with 18 features

模型	准确率	精确率	召回率	F1 分数
SVM	0.47	0.41	0.43	0.39
KNN	0.58	0.64	0.58	0.59
DT	0.75	0.79	0.75	0.76
LR	0.52	0.52	0.52	0.52
RF	0.83	0.84	0.83	0.83
XGBoost	0.91	0.91	0.91	0.91

由表 4 和表 5 对比分析可知,特征降维显著提升了多数模型的性能。XGBoost 在 18 个特征下表现最佳,相比于 42 个特征,分类准确率和 F1 分数均从 0.78 提升至 0.91,表明其在处理精简特征时具有极强的非线性拟合能力;随机森林次之,准确率为 0.83,适合需要解释特征重要性的场景;决策树在降维后准确率从 0.58 大幅提升至 0.75,证明降维有效缓解了过拟合问题。相比之下,SVM 和 KNN 在 2 种特征集下均表现较差,可能因数据分布或算法特性不匹配;逻辑回归因线性模型特性无法充分捕捉数据的非线性关系,表现不佳。综上所述,本文所使用的特征降维通过剔除冗余特征,增强了模型对关键信息的捕捉能力,显著提升了分类性能,为疲劳监测研究提供了重要的方法学参考。

### 2.3 不同时长心电计算 HRV 的分类结果对比

HRV 是评估自主神经系统功能的重要指标,其计算通常依赖于心电数据的时长。为探究不同时长心电数据对 HRV 分类模型性能的影响,本研究选取了 2、4、6、8、10 min 这 5 种时长的的心电数据,基于 XGBoost 模型进行疲劳分类实验。通过对比不同时长下的分类结果,旨在为实际应用中选取最优心电数据时长提供科学依据。不同时长心电数据下的分类结果如表 6 所示。

表 6 不同时长心电计算 HRV 的分类结果

Tab.6 Classification results of HRV calculated from electrocardiograms of different durations

时长/min	准确率	精确率	召回率	F1 分数
2	0.91	0.91	0.91	0.91
4	0.62	0.64	0.62	0.62
6	0.66	0.68	0.66	0.66
8	0.61	0.63	0.61	0.61
10	0.69	0.71	0.69	0.69

由表 6 可以看出,2 min 数据在疲劳分类任务中表现最优,准确率达 0.91,显著优于其他时长数据。这一结果表明,短时长心电数据在实际应用中具有较高的信息密度和计算效率,能够有效支持 HRV 分类模型的构建。这一结果为心电数据时长的选取提供了科学依据,同时也为未来基于 HRV 的疲劳监测研究提供了重要的参考。

## 3 结 论

本文基于可穿戴设备采集养老护理人员心电数据,使用巴特沃斯滤波与 Pan-Tompkins 算法进行预处理,系统提取了 HRV 的时域、频域及非线性域共 42

个特征参数,并采用皮尔逊相关系数法筛选 18 个关键特征,构建基于 XGBoost 的疲劳分类模型,实现了对疲劳状态的三分类识别。研究结果表明:

(1) 基于 XGBoost 模型的分方法在特征降维后表现优异,显著优于其他模型。特征降维通过剔除冗余特征,增强了模型对关键信息的捕捉能力,有效提升了分类性能,尤其是在中度疲劳和重度疲劳类别上表现尤为突出。

(2) 2 min 心电数据在 HRV 分类任务中表现最优,准确率达 0.91,显著优于 4~10 min 心电数据准确率(0.61~0.69),表明短时长数据在实际应用中具有较高的信息密度和计算效率。

通过将心电信号分析与机器学习技术相结合,本研究不仅为疲劳监测提供了更为客观、精准的方法,也为护理人员的健康管理提供了科学化的解决方案。未来研究可进一步探索:其他特征选择方法(如互信息法、主成分分析等)和机器学习模型(如深度学习模型)在疲劳分类中的应用,并结合更多类型的生理信号数据(如脑电图、肌电图等)以进一步提高分类性能;特征选择与模型优化的实时应用,以及不同心电数据时长在更广泛场景下的适用性。通过不断优化算法和数据采集方法,有望为护理人员的疲劳监测与健康管理服务提供更加精准、高效的技术支持,从而促进养老服务业的可持续发展。

### 参考文献:

- [1] 贾春阳. 养老环境下基于深度神经网络的类人认知推理模型研究[D]. 沈阳:沈阳工业大学,2024.  
JIA C Y. Research on a human-like cognitive reasoning model based on deep neural networks in aged care environment[D]. Shenyang: Shenyang University of Technology, 2024 (in Chinese).
- [2] BEHRENS M, GUBE M, CHAABENE H, et al. Fatigue and human performance: An updated framework[J]. Sports Medicine, 2023, 53(1): 7-31.
- [3] RATHNAYAKE S, DASANAYAKE D, MAITHREEPALA S D, et al. Nurses' perspectives of taking care of patients with Coronavirus disease 2019: A phenomenological study[J]. PLoS One, 2021, 16(9): e0257064.
- [4] MAGHSOUD F, REZAEI M, ASGARIAN F S, et al. Workload and quality of nursing care: The mediating role of implicit rationing of nursing care, job satisfaction and emotional exhaustion by using structural equations modeling approach[J]. BMC Nursing, 2022, 21(1): 273.
- [5] 陈彦龙, 毛万丽, 刘鸿优. 主观疲劳量表(RPE)评估足球运动员训练负荷的实证研究[J]. 体育科研, 2019, 40(1): 57-63.

- CHEN Y L, MAO W L, LIU H Y. An empirical study on the assessment of football players' training load by ratings of perceived exertion[J]. *Sport Science Research*, 2019, 40(1): 57-63(in Chinese).
- [6] 吕静,杨君,张尚荣. 军队驾驶员行车前后疲劳状态与情绪变化的主观评定研究[J]. *职业卫生与病伤*, 2015, 30(1): 48-51.
- LYU J, YANG J, ZHANG S R. Subjective evaluation research on fatigue state and emotional changes of military drivers before and after driving[J]. *Journal of Occupational Health and Damage*, 2015, 30(1): 48-51(in Chinese).
- [7] GE Z Y, CHENG H Q, TONG Z, et al. ECG-MAKE: An ECG signal delineation approach based on medical attribute knowledge extraction[J]. *Information Sciences*, 2023, 637: 118978.
- [8] MOUSAVI S, AFGHAH F, KHADEM F, et al. ECG language processing (ELP): A new technique to analyze ECG signals[J]. *Computer Methods and Programs in Biomedicine*, 2021, 202: 105959.
- [9] GAO R T, YAN H C, DUAN J L, et al. Study on the non-fatigue and fatigue states of orchard workers based on electrocardiogram signal analysis[J]. *Scientific Reports*, 2022, 12(1): 4858.
- [10] PHAM T, LAU Z J, CHEN S H A, et al. Heart rate variability in psychology: A review of HRV indices and an analysis tutorial[J]. *Sensors*, 2021, 21(12): 3998.
- [11] MU S Q, LIAO S W, TAO K, et al. Intelligent fatigue detection based on hierarchical multi-scale ECG representations and HRV measures[J]. *Biomedical Signal Processing and Control*, 2024, 92: 106127.
- [12] FENG W B, ZENG K L, ZENG X M, et al. Predicting physical fatigue in athletes in rope skipping training using ECG signals[J]. *Biomedical Signal Processing and Control*, 2023, 83: 104663.
- [13] HUANG S T, LI J, ZHANG P Z, et al. Detection of mental fatigue state with wearable ECG devices[J]. *International Journal of Medical Informatics*, 2018, 119: 39-46.
- [14] NI Z Q, SUN F M, LI Y. Heart rate variability-based subjective physical fatigue assessment[J]. *Sensors*, 2022, 22(9): 3199.
- [15] CHEN T Q, GUESTRIN C. XGBoost: A scalable tree boosting system[J]. *ArXiv e-Prints*, 2016: arXiv: 1603.02754.
- [16] LIU F, WEI S, LI Y, et al. The accuracy on the common Pan-Tompkins based QRS detection methods through low-quality electrocardiogram database[J]. *Journal of Medical Imaging and Health Informatics*, 2017, 7(5): 1039-1043.
- [17] BUTKEVIČIŪTĖ E, ERINŠ M, BIKULČIENĖ L. An adaptable human fatigue evaluation system[J]. *Procedia Computer Science*, 2021, 192: 1274-1284.
- [18] ANWER S, LI H, UMER W, et al. Identification and classification of physical fatigue in construction workers using linear and nonlinear heart rate variability measurements[J]. *Journal of Construction Engineering and Management*, 2023, 149(7): 04023057.
- [19] NAYAK S K, PRADHAN B, MOHANTY B, et al. A review of methods and applications for a heart rate variability analysis[J]. *Algorithms*, 2023, 16(9): 433.
- [20] CIRIC D G, PERIC Z H, MILENKOVIC M, et al. Evaluating similarity of spectrogram-like images of DC motor sounds by Pearson correlation coefficient[J]. *Elektronika Ir Elektrotechnika*, 2022, 28(3): 37-44.
- [21] LI G Z, ZHANG A N, ZHANG Q Z, et al. Pearson correlation coefficient-based performance enhancement of broad learning system for stock price prediction[J]. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 2022, 69(5): 2413-2417.
- [22] XU Y G, JIAN Q Z, ZHU K S, et al. Identifying fatigue of climbing workers using physiological data based on the XGBoost algorithm[J]. *Frontiers in Public Health*, 2024, 12: 1462675.
- [23] PUNITHA N, DIVYA BHARATHI K, MANUSKANDAN S R, et al. Analysis of muscle fatigue progression using geometric features of surface electromyography signals and explainable XGBoost classifier[J]. *Journal of Medical and Biological Engineering*, 2024, 44(2): 191-197.

#### 本文引文格式:

张欣,马帅,欧宗锬,等. 基于心率变异性与机器学习的养老护理人员疲劳分类[J]. *天津工业大学学报*, 2025, 44(4): 44-51.

ZHANG X, MA S, OU Z K, et al. Fatigue classification of service staff in elderly care based on heart rate variability and machine learning[J]. *Journal of Tiangong University*, 2025, 44(4): 44-51(in Chinese).