【交通物流 / Transportation Logistics】

基于多模态生理信号的驾驶人脑力负荷评估

张奇良，杨坤华，曲行达，陶达

深圳大学人因工程研究所，广东深圳 518060

摘　要：准确评估驾驶人脑力负荷状态对降低因驾驶人脑力负荷过载导致的交通事故具有重要意义. 基于典型驾驶场景，结合*N*-back认知负荷次任务，设计不同难度的驾驶任务实验，研究驾驶人脑力负荷. 实验收集驾驶人在任务完成过程中的多种模态生理信号(脑电、心电和皮电信号)及美国航空航天局任务负荷指数量表主观脑力负荷数据，提出基于多模态生理信号特征分析和模式识别的驾驶人脑力负荷分类模型，并比较不同模态生理信号及其组合在3种典型机器学习算法(随机森林、决策树和k最近邻模型)中的脑力负荷分类识别效果. 研究表明，基于不同模态生理信号组合的脑力负荷分类模型具有不同的分类准确率. 单一模态生理信号的分类模型中，基于皮电、心电和脑电信号的分类模型准确率依次增加；基于多模态生理信号的分类模型准确率普遍优于单一模态分类模型；基于脑电、心电及皮电3模态生理信号的随机森林分类模型具有最高的分类准确率.

关键词：交通信息工程及控制；脑力负荷；机器学习；多模态；脑电信号；心电信号；皮电信号

中图分类号：U491；TN911.72　　**文献标志码：**A　　**doi：**10.3724/SP.J.1249.2022.03278

Evaluation of drivers′ mental workload based on multi-modal physiological signals

ZHANG Qiliang, YANG Kunhua, QU Xingda, and TAO Da

Institute of Human Factors and Ergonomics, Shenzhen University, Shenzhen 518060, Guangdong Province, P. R. China

Abstract: Accurately assessing the driver′s mental workload is of great significance to reduce the traffic accidents caused by the driver′s mental overload. This study aims to evaluate drivers′ mental workload in simulated typical driving scenarios, with *N*-back cognitive tasks used to manipulate varied levels of task difficulty. We collect data on multi-modal physiological signals including electroencephalogram (EEG), electrocardiogram (ECG), and electrodermal activity (EDA) signals, and subjective mental load of the National Aeronautics and Space Administration task load index (NASA_TLX) during the task completion process of the driver in the experiment, and propose a series of mental workload classification models based on feature analysis and pattern recognition of the multi-modal physiological signals. These classification models are verified by machine learning algorithms of random forest, decision tree and *k*-nearest neighbor models. The results show that the accuracy of classification models varies with different modalities of physiological signals. EEG-based classification models yield the highest accuracy among single-modal classification models, followed by EDA-based and ECG-based models. Multi-modal-based classification models generally perform better than single-modal classification models. The random forest classification algorithm based on three-modal physiological signals of EEG, ECG and EDA has the highest classification accuracy.

Received: 2021-06-15; **Revised:** 2021-11-31; **Accepted:** 2021-12-30; **Online (CNKI):** 2022-03-01

Foundation: Shenzhen Basic Research Foundation for General Projects (20200813225029002); National Natural Science Foundation of China (72101161)

Corresponding author: Assistant professor TAO Da. E-mail: taoda@szu.edu.cn

Citation: ZHANG Qiliang, YANG Kunhua, QU Xingda, et al. Evaluation of drivers′ mental workload based on multi-modal physiological signals [J]. Journal of Shenzhen University Science and Engineering, 2022, 39(3): 278-286. (in Chinese)

http://journal.szu.edu.cn

Key words: traffic information engineering and control; mental workload; machine learning; multi-modal; electroencephalographic signal; electrocardiographic signal; electrodermal activity signal

车辆驾驶系统是由驾驶人、车辆及交通环境构成的高度耦合的人机交互系统[1]，其中，驾驶人需对复杂路况及车辆状态信息进行连续不断的认知处理．因此，驾驶人需付出大量的脑力活动以维持良好的驾驶表现．随着时间增加，驾驶人的脑力负荷(mental workload)不断增大，易导致操作失误，进而引发交通事故．事实上，驾驶人脑力负荷过载导致的驾驶安全问题一直是社会发展中面临的一个严峻安全问题．美国国家安全委员会报告表明，因驾驶人脑力负荷过载所导致的交通事故率占比超过21%[2]．因此，对驾驶人脑力负荷进行准确评估对于解决脑力负荷过载问题，以及降低交通事故发生率至关重要．

脑力负荷不同于体力劳动负荷和任务负荷，其强调由任务需求引起的压力[3]．常见的脑力负荷测量指标包括：主观问卷指标、生理指标、绩效指标以及这些指标的组合[4]．与主观测量和绩效测量相比，生理指标具有敏感性、非侵入性和客观性等优势．

目前，可用来反映脑力负荷的生理指标包括脑电指标、心电指标、皮电指标、眼动指标及呼吸指标等[5-7]．脑电指标中，KOSTI等[8]在软件工程师的代码错误检查任务实验中，发现θ功率与脑力负荷具有较强的相关性，而大脑额叶区信号中的θ波与脑力负荷高度相关[9]．心电指标中，心率(heart rate, HR)对不同脑力负荷水平的任务较为敏感，DE等[10]在飞行任务实验中发现，HR随着任务难度增加而显著增加，TAO等[11]认为，心动间隔(interbeat interval, IBI)均值是报告最多的时域指标，其次是相邻R峰值间隔差值平方的均方根(the square root of the mean of the sum of the squares of difference between successive R intervals differences, RMSSD)、心动间隔的标准差(standard deviation of heart beat interval, SDNN)及低频(low frequency, LF)功率等；在频域指标中，低频功率与高频功率的比率（the ratio of high frequency to low frequency, LF/HF)的应用最广泛[11]，HWANG等[12]在模拟核电厂停堆程序实验中发现，LF/HF随脑力负荷的增大而增加．皮电指标中，FOY等[13]在模拟驾驶实验中发现，Tonic渐变信号均值随着任务难度的增加而增大，DING等[14]研究发现，皮肤电导(skin conductance, SC)均值在随着任务难度的增加而变大．

然而，以上指标并不是在任何情境下都能有效反映脑力负荷．TAO等[11]认为，几乎所有生理指标在评估脑力负荷的敏感性和有效性时，都会随情境的变化(如任务、人员和环境的不同)而有所不同．一些指标在某些特定情境下能有效评估脑力负荷，但在另外一些情境下则变得无效，表明单一指标可能存在可靠性不足、不确定性高以及难以全面反映个体差异的局限[15]．因此，如何充分利用不同来源的生理指标，开发更加稳健的评估方法来有效评估脑力负荷仍是一项难题．

有应用机器学习等算法融合多维指标进行脑力负荷的评估与预测的研究[16-18]．与传统使用单一指标进行评估不同，机器学习算法可以构建智能分类器来融合多维、多模态指标数据，并能从这些多维指标数据所反映的内在模式中进行自动学习，从而给出更准确的评估与预测信息[19]．决策树、k最近邻、随机森林和人工神经网络等是目前在脑力负荷评估中常用的机器学习算法，虽已取得较好的评估预测效果[1, 14, 20]，但仍存在一定局限性．首先，已有研究大多仅采用脑电、心电或皮电等单一模态的多维生理指标进行模型构建与评估，不能全面反映脑力负荷所引起的人体多方面生理指标变化，容易导致脑力负荷分类模型的分类准确率不高．如KOEING等[20]基于心电信号，利用人工神经网络模型判别驾驶人的二分类脑力负荷，其准确率仅为68%．其次，尽管已有研究采用多模态组合的机器学习模型来评估脑力负荷，但对于各种模态及其组合在脑力负荷评估中的有效性比较研究仍较为缺乏，如YIN等[21]基于脑电和心电双模态指标数据，利用支持向量机构建脑力负荷评估模型；YAN等[22]通过融合眼动和心电信号双模态指标数据，建立一种驾驶人脑力负荷预测模型，以提高驾驶安全；郑玲等[1]基于心电和皮电指标数据，并结合车辆信息特征，利用随机森林算法，构建具有较好分类性能的驾驶人脑力负荷分类模型．然而，这些研究仍难以说明各个模态的生理指标及其组合，在各种机器学习算法评估和预测脑力负荷时的效果优劣如何．最后，目前驾驶人脑力负荷评估研究中大都以构建二分类模型(如仅区分脑力负荷的高低)这类最简单的评估模型为主，而脑力负荷的变化实际是

http://journal.szu.edu.cn

一个较为连续的过程，因此，多分类模型对脑力负荷的准确评估和预测更有意义.

因此，本研究基于典型驾驶任务场景开展实证研究，采用多模态生理指标数据，应用随机森林、k最近邻及决策树算法，构建不同模态生理指标及其组合的驾驶人脑力负荷三分类评估模型，系统评估各模态生理指标在脑力负荷评估中的有效性.

1 研究方法

1.1 受试者

本实验数据来自18名视力正常且无认知障碍的在校学生(平均年龄23.2岁，标准方差为1.6).受试者的选取标准为：短发男性(短发受试者能降低脑电信号采集时的头皮表面阻抗)；具有1 a以上的有效中国驾照，且有实际上路驾驶经验；无3维眩晕反应；无脑疾病史或心脏病史.本实验的每个参与者都知悉、填写了知情同意书，并会获得相应的实验参与报酬.

1.2 实验方案

采用组内实验设计方式，组内因子为任务难度，分为低难度、中难度和高难度.实验以主任务和次任务的形式展开，主任务为驾驶任务，受试者需在模拟驾驶环境中以大约80 km/h的速度驾驶车辆保持在中间车道行驶(图1)，实验场景中设计了12辆/(km·车道)的稳定交通流，场景中周围环境车辆的速度为70~85 km/h，并不均匀分布了20个障碍物，驾驶人若遇到障碍物可以变道超车，完成超车后需回到中间车道.次任务为N-back记忆任务，用来调节实验任务难度，即低难度(仅驾驶任务，无次任务)、中难度(驾驶任务和1-back次任务)和高难度(驾驶任务和2-back次任务).



图1 实验及数据采集照片

Fig. 1 (Color online) Picture of experiment and data collection

N-back任务通过让受试者在驾驶过程中对语音播报的随机数字做出相应反应(按下方向盘上的机械按键)实现，N-back任务与驾驶实验同步进行.其中，1-back任务中，受试者听到2个相同的相邻数字时，需按下按键；2-back任务中，受试者听到的某数字若与该数字之前间隔1个的数字相同时，需要按下按键.图2为N-back任务示意图(听到图2中带下划线的数字时需要按下按钮).

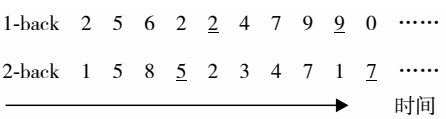


图2 N-back任务示意图

Fig. 2 Schematic diagram of N-back task

生理数据潜在提供了客观、连续和非侵入性的脑力负荷测量[23]，在实验过程中分别采集了驾驶人的脑电、心电及皮电3种模态的生理指标数据及美国航空航天局任务负荷指数(National Aeronautics and Space Administration task load index，NASA_TLX)量表主观脑力负荷数据，其中，大脑额叶区信号中的θ波被认为与脑力负荷高度相关[9].心电(electrocardiogram，ECG)信号中的心率变异性(heart rate variability，HRV)反映了心跳周期差异的变化情况，常见的HRV时域指标包括Mean IBI、Mean HR、RMSSD及SDNN，频域指标包括LF、HF及LF/HF[24].皮电(electrodermal activity，EDA)信号中的SC、Tonic、Phasic(突变信号)常被用来评估脑力负荷[25].本研究采集的各生理信号指标如表1.

1.3 实验设备及流程

本实验采用基于FORUM8开发的驾驶模拟平台，该平台由视觉仿真单元、音频仿真单元及运动单元组成，并配备罗技力反馈方向盘踏板套装.脑电信号通过BitBrain 32导脑电仪收集，采样频率为256 Hz；心电和皮电信号通过北京津发生理采集系统收集，采样频率分别为256 Hz和64 Hz.

考虑到驾驶疲劳对实验的影响，本研究将安排受试者在连续3 d的相同时间段分别完成1种难度的驾驶任务，实验任务难度顺序随机选择.实验开始前，受试者需要进行设备仪器佩戴.脑电仪根据10~20系统布置于受试者头皮表层，耳垂电极作为脑电帽的参考电极，实验前要求受试者清洗头皮并保持干燥以降低头皮表层阻抗.心电图原始数据通

http://journal.szu.edu.cn

表1　所选取的各模态生理信号指标

Table 1　The selected physiological signal indicators for each of the modalities

模态类型	特征名称	单位
EEG	θ：脑电波能量在4~8 Hz频段分布的均值	dB
ECG	mean IBI：心动间隔的均值	ms
	mean HR：平均心率	次/min
	RMSSD：相邻R峰值间隔差值的平均值	ms
	SDNN：心动间隔的标准差	ms
	HF data power：高频功率	W
	LF data power：低频功率	W
	LF/HF ratio：低频功率与高频功率的比值	
EDA	SC data mean：皮肤电导均值	μs
	Tonic data mean：渐变信号均值	μs
	Phasic data mean：突变信号均值	μs

过将3个电极分别放置在左(负)、右(接地)锁骨及左肋软骨(正)下方进行记录．皮电原始数据采集自左手食指和中指，实验前通过酒精棉签擦洗皮肤表面以降低皮肤阻抗．受试者在佩戴好生理数据采集设备后，进入驾驶模拟平台进行实验练习任务，以熟悉驾驶模拟器操作和*N*-back任务．正式实验开始后，受试者在选定的任务难度下进行连续1 h的驾驶实验，实验结束后填写NASA_TLX量表．实验流程如图3．

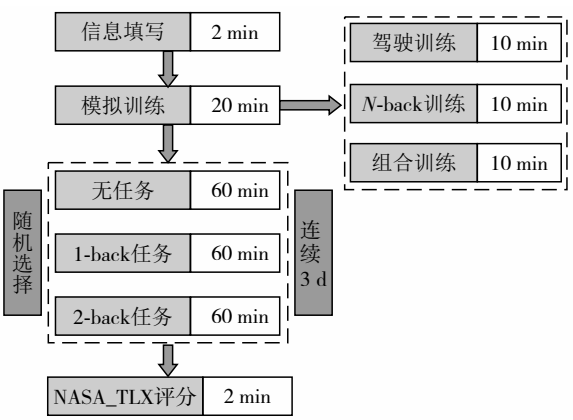


图3　实验流程

Fig. 3　Schematic diagram of the experimental process

1.4　生理数据提取与数据分析方法

在对脑电数据进行50 Hz的带阻滤波降低工频干扰后，使用0.1 Hz的高通滤波和30.0 Hz的低通滤波滤除不需要的频段脑电数据，并使用Welch法进行功率谱分析提取θ频段的功率值．心电信号采用中等强度的小波降噪来消除白噪声，并以1~30 Hz的频率对数据进行带通滤波，其中，R峰值提取的最大心率为120 次/min．皮电信号采用1~500 Hz的带通滤波处理原始数据后，使用窗口大小为78 ms的高斯滤波降噪．皮肤电导反应（skin conductance response，SCR）提取选择的是中等大小的峰值检测灵敏度，限定最大上升时间为4 s，最大衰减时间为4 s，最小的SCR振幅被限定为0.03 μs．

单因素方差用来分析不同任务难度对驾驶人主观脑力负荷的影响，独立样本t检验用来辨别次任务难度对驾驶人任务绩效的影响．Pearson相关性分析用来剔除存在强相关性的生理数据特征以避免数据冗余．脑力负荷评估模型分别应用随机森林、决策树和k最近邻算法模型，通过准确率、宏平均召回率和宏平均精确率等指标评估模型性能，其中，宏平均召回率和宏平均精确率分别是指在对每一类别的召回率和精确率进行统计后，计算所有类别的算术平均值．

2　结果及分析

2.1　主观脑力负荷评价与次任务绩效分析

NASA_TLX主观评分与次任务错误次数均值见表2．单因素方差分析表明，次任务难度对主观评分结果具有显著影响，随着任务难度的增加，得分显著增加（$F_{(2,\,69)}=9.692$，$P<0.001$），NASA_TLX评分说明驾驶人主观认知负荷随着次任务难度的增加而增加，且次任务难度设置较为合理，驾驶人能明显感受到次任务的难度变化．事后检验结果显示任何2个不同难度任务之间存在显著差异（$P<0.05$）．次任务绩效数据结果表明，随着任务难度的增加，1-back和2-back的错误次数和反应时间显著增加．独立样本t检验分析结果表明，次任务难度对次任务错误次数和次任务反应时间具有显著影响，其中，1-back和2-back任务间的绩效数据差异显著（错误次数为$t=-3.202$，$P=0.002$；反应时间为$t=-5.949$，$P=0.000$）．

2.2　生理信号数据分析结果

各指标方差分析结果显示，大脑额叶区θ波各脑电导在无任务状态和简单次任务状态下的特征值具有明显递增趋势，但简单次任务和复杂次任务间的特征值趋势相同．单因素方差分析显示，额叶区的电极导Fp1、Fp2、F3、F4、F7、F8及Fpz的θ波

http://journal.szu.edu.cn

表2 三种次任务下NASA_TLX评分与次任务错误次数
（平均值±标准差）

Table 2 NASA_TLX scores and the number of sub-task errors under three sub-tasks
（mean ± standard deviation）

名称	无次任务	1-back	2-back
NASA_TLX	36.97±14.59	46.89±18.22	57.09±14.99
次任务错误次数	—	18.08±16.09	36.46±23.06
次任务反应时间	—	7.36±0.71	9.02±1.16

能量值在3种任务难度下差异显著($P < 0.005$)，而Fz在不同任务难度间没有显著差异($F_{(2,69)} = 2.607$，$P = 0.081$)．心电特征中的IBI在不同任务难度下呈递减趋势，且差异显著($F_{(2,69)} = 10.638$，$P < 0.001$)，其他时域特征HR、SDNN、RMSSD与任务难度间成正比，在3种不同任务难度间均差异显著($P < 0.005$)，而心电信号中的频域特征LF随着任务难度递增下变化明显，HF尽管在均值中随着任务难度增加呈现递增趋势，但方差分析显示差异不显著($P = 0.058$)，而LF/HF差异显著($P = 0.002$)．皮电信号中的SC、Tonic及Phasic在不同任务难度下递增趋势显著($P < 0.050$)，可作为评价驾驶人脑力负荷的有效指标．方差分析结果请扫描论文末页右下角二维码查看表S1．

3 脑力负荷分类模型

3.1 建模数据预处理

为避免不同特征之间数据量级相差过大导致的量纲问题，本研究对所有数据进行归一化处理．通过对数据进行切片以增加样本量，片段选定时间长度为10 min．

3.2 特征提取

通过数据处理提取了7种EEG特征、6种ECG特征及3种EDA特征．为减小强相关性特征造成的数据冗余，提高分类器精度，采用Pearson相关分析去除强相关特征，以降低数据维度．若相关性系数绝对值大于0.6(代表特征之间高度线性相关)，则保留其中1个特征[1]．生理特征Pearson相关性分析系数结果显示，脑电信号的Fp1与Fpz、Fp2与F7、F3与F8电极导的θ波能量值存在强相关性；心电信号特征中的IBI与HR、SDNN与LF、RMSSD与LF/HF存在强相关性；皮电信号特征中的各指标无强相关性．因此，降维后的EEG四种特征为Fp1、Fp2、F3及F4电极导的θ波能量值，3种HRV特征为LF/HF、IBI及SDNN，3种EDA特征为SC、Tonic及Phasic．Pearson相关性分析系数结果详情请扫描论文末页右下角二维码查看表S2．

3.3 脑力负荷评估模型

本研究系统比较不同模态生理信号特征及其组合在随机森林、k-最近邻和决策树算法中的分类识别效果．各模态及其组合分别为：① 单模态：脑电信号、心电信号及皮电信号；② 双模态：脑电+心电信号、脑电+皮电信号、皮电+心电信号；③ 三模态：脑电+心电+皮电信号．各模型中每一个特征指标的输入共有324个数据(18名参与者×3种任务难度×6个时间节点)．将不同任务难度作为分类器的输出标签．本研究随机选择80%的数据样本作为模型训练集，剩余20%的数据作为测试集，基于10折交叉验证以观测模型的泛化能力．模型构建过程如图4(以三模态为例)．

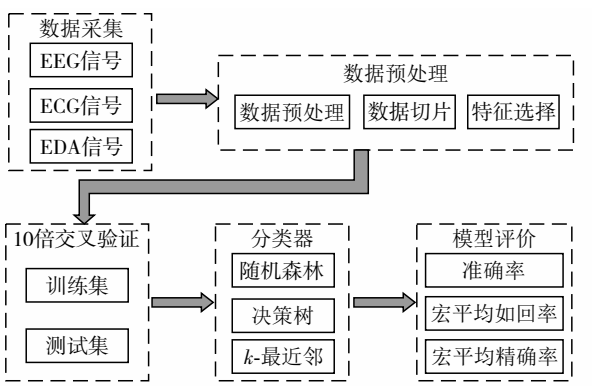


图4 分类模型框图

Fig. 4 Block diagram of classification model

表3列出以不同模态的生理数据作为模型输入时的分类器及其性能指标．可见，① 以单模态生理数据作为模型输入时，基于皮电、心电和脑电信号的模型分类准确率依次增加，其中，以脑电信号作为输入的模型准确率显著高于以心电或皮电作为输入的模型，随机森林在单模态生理数据中表现普遍最佳，基于脑电信号的模型准确率达到65.43%，3个脑力负荷等级的宏平均召回率和宏平均精确率分别为67.00%和67.33%．② 以双模态生理数据作为模型输入时，发现脑电和心电的组合特征在k最近邻($k = 3$)模型中表现优异，准确率可达79.01%，宏平均召回率和宏平均精确率分别达到80.00%和79.00%．而皮电和心电的组合特征在3种模型中的准确率均较低，在表现最好的随机森林模型中的准确率也仅为66.26%．③ 当使用脑电、

http://journal.szu.edu.cn

心电及皮电三模态生理信号组合的形式作为分类模型输入时，发现随机森林的分类准确率可达81.54%，宏平均召回率和宏平均精确率分别达到81.07%和82.17%，相比之下，k最近邻的分类准确率也可达76.54%，相应的宏平均召回率和宏平均精确率分别达到77.67%和76.76%．因此，在单一模态生理信号的分类模型中，基于脑电信号的随机森林模型表现最佳．在双模态生理信号的分类模型中，基于脑电和心电信号的k最近邻分类模型的准确率最高．在多模态生理信号的分类模型中，基于脑电、心电及皮电三模态生理信号的随机森林分类模型具有最高的分类准确率．

表3　以不同模态生理数据作为模型输入的分类结果汇总

Table 3　Summary of classification results with different modal physiological data as model inputs　%

模态	模型	准确率	宏平均召回率	宏平均精确率
脑电	随机森林	65.43	67.00	67.33
	k最近邻	62.96	63.33	63.33
	决策树	56.79	57.00	57.00
心电	随机森林	59.26	61.00	60.33
	k最近邻	56.79	57.00	57.67
	决策树	55.55	55.33	54.67
皮电	随机森林	49.38	51.66	52.33
	k最近邻	43.21	43.33	44.33
	决策树	45.68	52.00	53.67
脑电+心电	随机森林	72.83	73.00	72.00
	k最近邻	79.01	80.00	79.00
	决策树	64.20	65.33	63.67
脑电+皮电	随机森林	75.31	76.33	74.67
	k最近邻	70.37	70.67	71.00
	决策树	62.96	56.00	56.00
皮电+心电	随机森林	66.26	76.33	74.67
	k最近邻	54.32	55.67	54.67
	决策树	56.79	58.00	57.00
脑电+心电+皮电	随机森林	81.54	81.07	82.17
	k最近邻	76.54	77.67	76.67
	决策树	67.90	66.33	66.00

为了提供三模态生理数据组合作为模型输入的更多分类结果细节，本研究绘制了混淆矩阵可视化随机森林在3种脑力负荷等级下的测试集性能，如图5．图5中每一列为真实类中该类别的样本数量，每一行为预测类中该类别的样本数量，对角线上的元素表示不同类别中正确预测的实例数量，对角线之外的元素表示错误预测实例的类别和数量．最右边的列和最下面的行分别表示相应类别的精确率和召回率，右下角元素表示整体的分类准确率．由图5可见，最大的分类错误率出现在低负荷和中负荷之间，随机森林模型将低脑力负荷错误地预测为中脑力负荷，占所有实例的7.69%．有4.62%的实例错误地将中脑力负荷预测为高脑力负荷．总体来说，随机森林模型较好预测了不同等级的脑力负

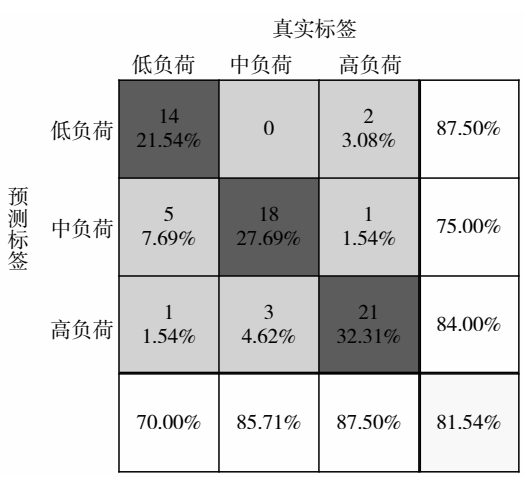


图5　基于随机森林的三级脑力负荷分类的混淆矩阵

Fig. 5　(Color online) Confusion matrix of three-level brain workload classification based on random forest

http://journal.szu.edu.cn

荷，总体准确率达到81.54%.

4 讨 论

本研究使用单模态生理数据作为模型输入时发现，基于脑电信号随机森林模型的分类准确率高于心电与皮电信号. 尽管随机森林模型本身无偏估计导致的模型泛化能力强的优点[26]可能是原因之一，但在其他两种机器学习模型中采用脑电信号作为模型输入的分类准确率亦高于心电和皮电信号，表明相比于心电与皮电信号，脑电信号能更好反映真实的脑力负荷水平，这可能是一些研究仅采用脑电信号作为脑力负荷评估指标的原因之一[5, 27]. 本研究也发现基于皮电信号的3种机器学习模型的准确率低于脑电与心电信号，在分类准确率最高的随机森林模型中仅为49.38%. 类似地，郑玲等[1]比较了基于单模态信号在人工神经网络模型中的准确率，发现基于皮电的脑力负荷评估模型准确率仅为40.9%，低于心电信号. 但不同的是，DING等[14]在基于单模态生理信号评估脑力负荷时发现，皮电信号作为反向传播网络模型输入时的准确率为47.4%，高于心电或呼吸信号作为该模型输入，表明皮电信号表征脑力负荷可能受到模型类型与具体任务场景的影响.

由基于双模态生理数据的脑力负荷评估模型结果发现，基于脑电和心电指标组合的分类模型准确率最高，其在k最近邻模型中获得了79.01%的准确率，而脑电和皮电组合(70.37%)、心电和皮电组合(54.32%)在k最近邻分类模型中的准确率稍低. 基于脑电和皮电的生理指标组合在随机森林模型中的分类准确率高于脑电和心电指标组合，表明在比较双模态生理指标组合类型时要考虑算法模型本身的特性，如k最近邻模型适合判别存在异常点的数据[28]，FAN等[29]基于脑电和心电信号的双模态数据，在支持向量机模型中得到80%以上的准确率. 本研究还发现，在相同脑力负荷分类模型中，基于双模态生理信号的模型分类准确率高于基于单模态生理数据的分类准确率. 类似地，郑玲等[1]比较了基于心电、皮电信号的单、双模态生理信号在人工神经网络(artificial neural network，ANN)模型中的识别准确率，发现基于心电和皮电的双模态生理信号组合在ANN模型中的准确率达到65.9%，高于基于心电(63.8%)、皮电(40.9%)的单模态生理信号的准确率. DING等[14]在基于心电、皮电和呼吸信号两两组合的双模态作为模型输入时，发现基于心电和皮电的双模态生理信号在反向传播神经网络(back propagation neural network，BPNN)模型中得到58.50%的准确率，高于基于心电(46.00%)、皮电(47.40%)及呼吸(43.90%)等单模态生理数据的模型准确率. 因此，可以认为基于双模态生理数据组合的脑力负荷分类模型准确率高于单模态生理信号模型.

基于脑电、心电和皮电的三模态生理信号组合作为模型输入时发现，在随机森林模型中获得了81.54%的准确率，高于基于单、双模态生理信号时的模型准确率，表明三模态生理信号可以显著提高脑力负荷评估模型的准确率. 类似地，DING等[14]基于心电、皮电和呼吸信号，对比不同模态生理信号组合作为BPNN模型的输入，尽管这3种生理信号在BPNN模型中的准确率仅为78.30%，三模态生理信号组合的准确率仍然高于单、双模态生理信号组合，充分说明三模态生理信号在评估脑力负荷中的有效性要高于单、双模态. FENG等[25]基于眼动、心电和皮电信号构建logistic回归模型，得到84.85%的分类准确率，表明基于三模态生理信号的脑力负荷评估模型准确率较高.

本研究得到的结论对于准确评估驾驶人脑力负荷状态，并降低因驾驶人脑力负荷过载导致的交通事故发生具有重要意义. 本研究亦存在一些局限. 首先，考虑到实验数据收集的便利性，本研究数据均采集自男性，未来研究可纳入女性对象以评估研究结果在性别因素上的稳健性；其次，由于本研究聚焦于生理指标模态，而前人研究表明驾驶行为数据与脑力负荷也存在相关性[30]，因此，未来研究可进一步融合驾驶行为数据来提高脑力负荷评估模型的有效性.

结 语

本研究通过模拟驾驶场景下的驾驶任务，利用N-back作为次任务调控驾驶人脑力负荷，以采集不同脑力负荷状态下驾驶人的生理信号(脑电、心电、皮电信号)，比较了不同生理指标模态及其组合在随机森林、决策树及k最近邻等模型中评估多等级脑力负荷的有效性. 研究结果可为后续的驾驶人脑力负荷实时检测系统开发提供指导性建议. ① 研

http://journal.szu.edu.cn

究发现在单模态生理信号的分类模型中，基于皮电、心电和脑电信号的分类模型准确率依次增加．该结论可适用于仅采集单一生理信号以检测驾驶人脑力负荷的便携式系统的开发．尽管本研究是多分类的脑力负荷评估模型，但基于单模态生理信号的分类结果仍不够理想．② 本研究比较了双模态生理信号多种分类模型的准确率，发现基于双模态（脑电、心电信号组合）生理信号的k最近邻分类模型的准确率显著高于其他双模态生理信号组合．该结论提示驾驶人脑力负荷检测系统的信号采集应紧紧围绕于脑电、心电信号开展．③ 本研究发现基于脑电、心电及皮电三模态生理信号的随机森林分类模型具有最高的分类准确率，可为准确率要求较高的大型驾驶人脑力负荷评估系统提供明确的指导意见，即在优先考虑识别准确性的前提下，脑力负荷评估系统应该基于脑电、心电及皮电三模态生理信号构建．本研究结论为脑力负荷实时检测系统的识别准确率提供有力理论支撑，检测系统准确率的提高可显著提高现实世界中驾驶人的健康和交通安全水平．

基金项目：深圳市基础研究面上计划资助项目（20200813225029002）；国家自然科学基金资助项目（72101161）
作者简介：张奇良（1996—），深圳大学硕士研究生．研究方向：人因工程．E-mail: 747826028@qq.com
引　　文：张奇良，杨坤华，曲行达，等．基于多模态生理信号的驾驶人脑力负荷评估[J]．深圳大学学报理工版，2022，39(3)：278-286.

参考文献 / References：

[1] 郑玲，乔旭强，倪涛，等．基于多维信息特征分析的驾驶人认知负荷研究[J]．中国公路学报，2021，34(4)：240-250.
ZHENG Ling, QIAO Xuqiang, NI Tao, et al. Driver cognitive loads based on multi-dimensional information feature analysis [J]. China Journal of Highway and Transport, 2021, 34(4): 240-250. (in Chinese)

[2] National Safety Council. Understanding the distracted brain: why driving while using hands-free cell phones is risky behavior [R]. Spring Lake: National Safety Council, 2012.

[3] GALY E. Consideration of several mental workload categories: perspectives for elaboration of new ergonomic recommendations concerning shiftwork [J]. Theoretical Issues in Ergonomics Science, 2018, 19(4): 483-497.

[4] DONG Yanchao, HU Zhencheng, UCHIMURA K, et al. Driver inattention monitoring system for intelligent vehicles: a review [J]. IEEE Transactions on Intelligent Transportation Systems, 2010, 12(2): 596-614.

[5] DI FLUMERI G, BORGHINI G, ARICÒ P, et al. EEG-based mental workload neurometric to evaluate the impact of different traffic and road conditions in real driving settings [J]. Frontiers in Human Neuroscience, 2018, 12: 509.

[6] JIMENEZ-MOLINA A, RETAMAL C, LIRA H. Using psychophysiological sensors to assess mental workload in web browsing [J]. Sensors, 2017, 18(2): 458.

[7] OMURTAG A, ROY R N, DEHAIS F, et al. Tracking team mental workload by multimodal measurements in the operating room [M/OL]. Neuroergonomics-The Brain at Work and in Everyday Life. New York, USA: 2019: 99-103.

[8] KOSTI M A, GEORGIADIS K, ADAMOS D A, et al. Towards an affordable brain computer interface for the assessment of programmers′ mental workload [J]. International Journal of Human Computer Studies, 2018, 115: 52-66.

[9] DIAZ-PIEDRA C, SEBASTIÁN M V, DI STASI L L. EEG theta power activity reflects workload among army combat drivers: an experimental study [J]. Brain Sciences, 2020, 10(4): 199.

[10] DE RIVECOURT M, KUPERUS M, POST W, et al. Cardiovascular and eye activity measures as indices for momentary changes in mental effort during simulated flight [J]. Ergonomics, 2008, 51(9): 1295-1319.

[11] TAO Da, TAN Haibo, WANG Hailiang, et al. A systematic review of physiological measures of mental workload [J]. International Journal of Environmental Research and Public Health, 2019, 16(15): 2716.

[12] HWANG S L, YAU Y J, LIN Y T, et al. Predicting work performance in nuclear power plants [J]. Safety Science, 2008, 46(7): 1115-1124.

[13] FOY H J, CHAPMAN P J A E. Mental workload is reflected in driver behaviour, physiology, eye movements and prefrontal cortex activation [J]. Applied Ergonomics, 2018, 73: 90-99.

[14] DING Yi, CAO Yaqin, DUFFY V G, et al. Measurement and identification of mental workload during simulated computer tasks with multimodal methods and machine learning [J]. Ergonomics, 2020, 63(7): 896-908.

[15] DEBIE E, ROJAS R F, FIDOCK J, et al. Multimodal fusion for objective assessment of cognitive workload: a

http://journal.szu.edu.cn

review [J]. IEEE Transactions on Cybernetics, 2019, 51 (3): 1542-1555.

[16] CINAZ B, ARNRICH B, LA MARCA R, et al. Monitoring of mental workload levels during an everyday life office-work scenario [J]. Personal Ubiquitous Computing, 2013, 17(2): 229-239.

[17] MATTHEWS G, REINERMAN-JONES L E, BARBER D J, et al. The psychometrics of mental workload: multiple measures are sensitive but divergent [J]. Human Factors, 2015, 57(1): 125-143.

[18] ZHAO Guozhen, LIU Yongjin, SHI Yuanchun. Real-time assessment of the cross-task mental workload using physiological measures during anomaly detection [J]. IEEE Transactions on Human-Machine Systems, 2018, 48 (2): 149-160.

[19] HOLZINGER A, KIESEBERG P, WEIPPL E, et al. Current advances, trends and challenges of machine learning and knowledge extraction: from machine learning to explainable AI [C]// International Cross-Domain Conference for Machine Learning and Knowledge Extraction. Hamburg, Germany: Springer, 2018: 1-8.

[20] KOENIG A, REHG T, RASSHOFER R. Statistical sensor fusion of ECG data using automotive-grade sensors [J]. Advances in Radio Science, 2015, 13: 197-202.

[21] YIN Zhong, ZHANG Jianhua. Operator functional state classification using least-square support vector machine based recursive feature elimination technique [J]. Computer Methods & Programs in Biomedicine, 2014, 113(1): 101-115.

[22] YAN Shengyuan, TRAN C C, WEI Yingying, et al. Driver′s mental workload prediction model based on physiological indices [J]. International Journal of Occupational Safety & Ergonomics, 2017, 25(2): 1-37.

[23] DAMOS D. Multiple-task performance [M]. Hants, UK: Taylor and Francis, 1991.

[24] HEINE T, LENIS G, REICHENSPERGER P, et al. Electrocardiographic features for the measurement of drivers′ mental workload [J]. Applied Ergonomics, 2017, 61: 31-43.

[25] FENG Chuanyan, WANYAN Xiaoru, YANG Kun, et al. A comprehensive prediction and evaluation method of pilot workload [J]. Technology and Health Care, 2018, 26 (Suppl. 1): 65-78.

[26] BIAU G, SCORNET E. A random forest guided tour [J]. Test, 2016, 25(2): 197-227.

[27] SO W K, WONG S W, MAK J N, et al. An evaluation of mental workload with frontal EEG [J]. PloS One, 2017, 12 (4): e0174949.

[28] ZHANG Minling, ZHOU Zhihua. ML-KNN: a lazy learning approach to multi-label learning [J]. Pattern Recognition, 2007, 40(7): 2038-2048.

[29] FAN Xiaoli, ZHAO Chaoyi, ZHANG Xin, et al. Assessment of mental workload based on multi-physiological signals [J]. Technology Health Care, 2020, 28(Suppl. 1): 67-80.

[30] ISLAM M R, BARUA S, AHMED M U, et al. A novel mutual information based feature set for drivers′ mental workload evaluation using machine learning [J]. Brain Sciences, 2020, 10(8): 551.

【中文责编：方圆；英文责编：淡紫】

补充材料

http://journal.szu.edu.cn