

# 基于脑电信号多特征融合的驾驶员精神疲劳状态 多分类研究

王雨生, 杨聚芬, 刘志钢

(上海工程技术大学, 上海 201620)

**摘要:** 面向目前最普遍的L2级智能驾驶场景下, 驾驶员的疲劳状态检测预警的安全需求问题, 基于多通道的无线脑电帽采集的脑电极数据实现驾驶员的状态4分类识别。研究采用卷积循环神经网络的同时集成频域、时域和非线性特征的不同组合训练模型, 发现将非线性特征微分熵以及时域特征平均绝对值组合在一起时识别性能最好, 提出了3种集成策略来集成不同输入特征组合下的基分类器模型, 该方法能以经济的成本以及便捷的方式来满足驾驶员疲劳状态的精确多分类, 推动可穿戴式设备应用于驾驶场景, 提高驾驶安全。

**关键词:** 驾驶员疲劳检测; 脑电信号; 卷积循环神经网络; 可穿戴式设备; 驾驶安全

中图分类号: U461.91 文献标志码: A DOI: 10.3969/j.issn.2095-1469.2025.03.07

## A Study on Multi-Class Classification of Driver Mental Fatigue States Based on Multi-Feature Fusion of EEG Signals

WANG Yusheng, YANG Jufen, LIU Zhigang

(Shanghai University of Engineering Science, Shanghai 201620, China)

**Abstract:** To meet the safety requirements for driver fatigue detection and warning in the most common L2 level intelligent driving scenarios, a four-class classification of driver states is achieved using EEG signals collected by a multi-channel wireless EEG cap. A convolutional recurrent neural network is used to train models using different combinations of frequency-domain, time-domain and nonlinear features. The results show that the best recognition performance is achieved when combining the differential entropy from nonlinear features with the average absolute value from time-domain features. Furthermore, three integration strategies are proposed to fuse base classifiers trained on different feature combinations. The method achieves accurate multi-class classification of driver fatigue states in a cost-effective and user-friendly manner, and promotes the application of wearable devices in driving scenarios to improve driving safety.

**Keywords:** driver fatigue detection; EEG signal; convolutional recurrent neural network; wearable devices; driving safety

收稿日期: 2024-04-23 改稿日期: 2024-06-10 网络首发日期: 2024-06-14

基金项目: 国家自然科学基金项目(52302438)

参考文献引用格式:

王雨生, 杨聚芬, 刘志钢. 基于脑电信号多特征融合的驾驶员精神疲劳状态多分类研究[J]. 汽车工程学报, 2025, 15(3): 340-352.

WANG Yusheng, YANG Jufen, LIU Zhigang. A Study on Multi-Class Classification of Driver Mental Fatigue States Based on Multi-Feature Fusion of EEG Signals[J]. Chinese Journal of Automotive Engineering, 2025, 15(3): 340-352. (in Chinese)



驾驶疲劳的识别和检测一直是研究的热点，虽然如今的交通设施设备越来越高度自动化，但目前最普遍的是L2级的智能辅助驾驶，驾驶员仍是交通安全运营的重要组成部分<sup>[1]</sup>，所以智能辅助驾驶环境下驾驶员的驾驶疲劳不容忽略，由疲劳导致的驾驶员不安全行为十分危险<sup>[2]</sup>，驾驶期间驾驶员的全身低频振动（4~10 Hz）很容易导致嗜睡，由于智能驾驶下驾驶员的主要任务是监控<sup>[3]</sup>，难免会出现相对不易被察觉的被动疲劳<sup>[4]</sup>，当发生危险或紧急事件时，容易导致安全事故，通过可穿戴设备的传感器采集的生理数据识别和检测实际工况下驾驶员的疲劳是一项挑战，同时，随着脑机接口技术的发展，可穿戴设备在工程技术领域的应用潜能也亟待开发。

目前，大部分基于可穿戴设备的疲劳监测研究都是通过有监督的机器学习来进行的。AL-LIBAWY等<sup>[5]</sup>通过分析可穿戴设备疲劳检测方法，发现支持向量机（SVM）比人工神经网络（ANN）实现疲劳二分类的检测更准确。LIU Ninghan等<sup>[6]</sup>则基于EEG整合了人工神经网络（ANN）和支持向量机来确定司机的心理状态。结合实际工况下采集到的数据，LI Jue等<sup>[7]</sup>比较了多种监督学习算法，认为支持向量机最适合识别建筑场景中的精神疲劳。当面对数据集规模较小的情况时，HOU Zhongwei等<sup>[8]</sup>在进行疲劳监测的研究中，采用了支持向量机（SVM）级联方法。他们提出了引入长短期记忆网络（LSTM）等深度学习算法，可以提升疲劳监测模型的性能。但以上方法和研究都属于二分类，虽然简单有效但对于复杂交通路况等情况的分类效果还有待研究。

而通过深度学习技术可以有效提高监测任务的鲁棒性和可靠性<sup>[9]</sup>，HUANG Mengtao等<sup>[10]</sup>基于卷积神经网络（CNN）通过驾驶员的表情识别疲劳，通过调整卷积核大小和通道数来提高模型的准确性。为了实现情绪状态的多种分类，RONG Qian等<sup>[11]</sup>通过3D-CNN分别提取了人脸数据和红外人脸数据的时空特征。由于基于脑电信号的特征更容易获取，

并且频域特征常被选用，CHAI等<sup>[12]</sup>提出通过脑电信号的4个频段的功率谱密度（PSD）使用ICA聚类后的特征作为精神疲劳的特征。WANG Hongtao等<sup>[13]</sup>利用小波包变换（WPT）方法获得脑电信号的 $\theta$ （4~7 Hz）、 $\alpha$ （8~12 Hz）和 $\beta$ （13~30 Hz）频带频率分量，用于计算所选通道的相应PSD，并通过PSD和样本熵（SE）来计算综合指标，实现精神疲劳的预测。由此可见，时频域特征是广泛且有效的一种特征选择，非常适合用于深度学习。

除了时域特征和频域特征，空域特征在脑电信号的识别上也不容忽略。ZHAO Chunlin等<sup>[14]</sup>采用多元自回归（MVAR）模型提取额叶、中枢和枕叶的多通道脑电图（EEG）信号特征，通过不同脑区的数据进一步探索脑电信号的识别。近年来，深度学习技术在脑电信号识别方面颇有成效，卷积神经网络和循环神经网络在脑电信号研究方面开始显示出巨大的潜力和应用价值。ZAMANI等<sup>[15]</sup>通过一维卷积神经网络（CNN-1D）和递归神经网络（RNN）实现了情绪的4分类，并在RNN框架中采用门控循环单元（GRU）和长短期记忆（LSTM）来解决梯度消失的问题。CHEN Jichi等<sup>[16]</sup>通过功能脑网络（FBN）表征脑网络组织与疲劳之间的关系，通过功率谱密度（PSD）表征功率变化与疲劳之间的关系，并基于特征的融合和极限学习机（ELM）开发了疲劳检测系统。相较于其他类型的生理数据，EEG生理数据特有的空域特征也是非常重要且有效的。

以往疲劳检测的研究中明显存在着当分类情况越复杂准确度越难提高的问题，并且通过脑电信号识别的研究只考虑了时频域特征或者空域特征中的少数几种，同时集成时频域和空域特征的研究较少，并且没有考虑与时间相关的动态特性。考虑到这一特性，本研究采用了卷积循环神经网络，并结合脑电生理技术，同时融合了多种特征。这一方法旨在解决驾驶员疲劳状态发现不及时的问题，并针对L2级的驾驶场景，准确识别驾驶员所处的具体状态。

## 1 数据采集与处理

为了有效识别驾驶员的精神疲劳，使用模拟驾驶台进行仿真模拟试验，通过单调模拟驾驶和各项驾驶任务采集了多模态的原始数据集，数据集包括单调驾驶期间的脑电数据、主观量表数据以及基于 Stroop 试验范式<sup>[17]</sup>的绩效表现数据。将脑电信号分解为 4 种频段，综合考虑其时域、频域、空域上的信息，通过卷积循环神经网络 CRNN 进行识别，综合考虑不同的特征集组合和集成策略并进行比较，筛选出最准确的模型。本文的研究框架如图 1 所示。

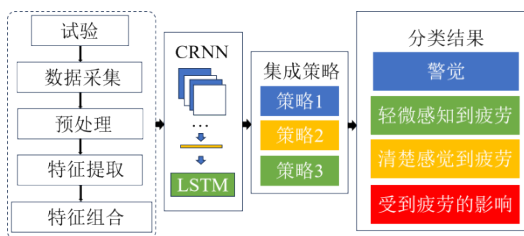


图1 研究框架

### 1.1 试验环境和试验设置

试验环境由 2 位试验员和 12 位被试者、搭载了 E-Prime 心理学试验软件的笔记本电脑、KETENG 模拟驾驶器、Neuracle 的 64 导联无线脑电帽、脑电信号接收传输模块、驾驶疲劳恢复座舱、三联显示屏组成的显示系统等构成。模拟驾驶场景选择 SILAB 软件中自带的驾驶疲劳应用场景。无线脑电帽和 E-Prime 采集到的数据将实时保存在笔记本电脑上，模拟驾驶器用于完成模拟驾驶操作，驾驶疲劳恢复座舱用于被试者在试验前保持清醒的静息状态。试验环境如图 2 和图 3 所示。

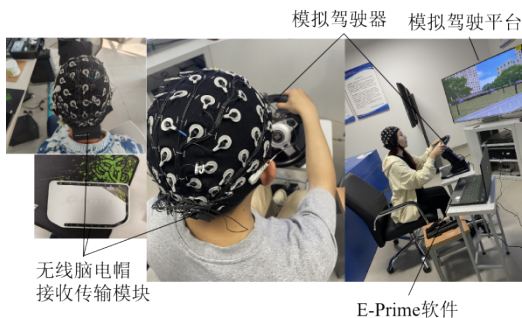


图2 试验场景



图3 驾驶疲劳恢复座舱

其中，12 位被试者年龄在 20~26 岁之间，均值为 22.25，方差为 2.69。其中，男生 7 名，女生 5 名，驾龄为 1~7 年不等，均值为 3.33，方差为 3.22。被试者身体健康，没有长期酗酒或吃药的习惯，具有正常或矫正到正常的视力，在参与试验前的 24 小时内没有饮酒或摄入咖啡因，每位被试者都至少参与了 1 次试验，最终采集到了 20 次有效驾驶数据。

面向 L2 级的驾驶任务，驾驶员无需刻意控制车距以及保持车道位置，因此，驾驶员只需要进行正常的单调驾驶操作任务，例如踩油门和刹车以及调整方向，由于驾驶员在必要的时候仍需要接管车辆，所以要持续监视路线状况和车载仪器运作情况，因此长时间驾驶更易导致被动疲劳。由于驾驶任务的本质是选择任务和辨别任务，遇到交通标志物等需消耗认知进行判别，模拟自适应巡航控制下驾驶员单调驾驶任务。为了加快被试者疲劳的产生，并且客观反映驾驶员疲劳状况，通过 E-Prime2.0 软件设置 Stroop 试验范式，实现在单调驾驶任务下确认交通信号与驾驶情况是否符合的认知任务。该任务需要被试者在 2 s 内识别一组色彩单词的颜色，而不是色彩单词本身的意思。同时，本研究还考虑了被试者主观上的感受，根据 KSS（卡罗林斯卡困倦量表）从 1（极度警觉）到 9（非常困倦）进行评估<sup>[18]</sup>，并且精神疲劳也会受到情绪调节的影响<sup>[19]</sup>，根据本文作者之前的研究，单调驾驶环境下，驾驶员的消极情绪对精神疲劳的

影响不可忽视，而根据SAM (The Self-Assessment Manikin) 可视量表的状态图可以对自身情绪状态从1~9进行评分<sup>[20]</sup>。

在开始正式的驾驶试验前，被试者会被要求戴上无线脑电帽进行10 min的模拟驾驶训练。脑电帽的设置：采样频率1 000 Hz，采用Biosemi32标准的电极，每个电极的阻抗都调整到5 k $\Omega$ 以下。然后在疲劳驾驶恢复座舱中休息10 min，以恢复到静息状态，为了确保疲劳产生，试验时长被设置得尽量长，每次试验的最长驾驶时间约为120 min。被试者除了基于Stroop试验范式设计的认知判别任务外，只需要保持正常驾驶，采取单调的驾驶操作，为了加速疲劳产生并记录被试者客观认知能力的变化，驾驶过程中被试者每5 min开始20 s的Stroop任务，同时，试验期间每5 min对被试者进行一次主观量表的采集。如果被试者连续3次的KSS量表或是SAM三维度的评级超过等级8，则停止试验。试验任务设置如图4所示。

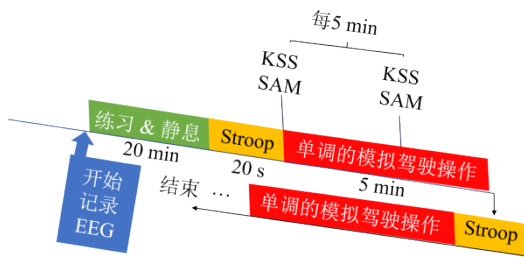


图4 试验任务设置

## 1.2 EEG信号预处理

使用Python的MNE库对原始脑电信号进行预处理，对每次驾驶试验采集到的120 min的32导联的脑电数据进行陷波滤波，去除50 Hz及其倍频的工频干扰，以消除明显的环境噪声。选择FIR (Finite Impulse Response) 线性滤波器进行30 Hz的低通滤波以及0.1 Hz的高通滤波，分别消除高频噪声和电压漂移，以有效保留脑电信号的有用数据。为了减少噪声和伪迹的影响以提高信噪比，将所有导联的平均值作为参考电极，对脑电数据进行平均重参考，然后对脑电数据进行独立主成分分析

(ICA)，ICA已经广泛用于去除脑电信号的伪影，算法假设原始信号在统计上是独立的，而观测信号则是由原始信号瞬时混合而成，可以根据一些先验知识分离观测信号，由此得到独立可靠的原始信号<sup>[21]</sup>。通过独立主成分分析找出含有眼动、肌肉、心跳等伪迹的成分，并将其去除，由于原数据的采样频率太高，不利于计算，将脑电数据的采样频率从原来的1 000 Hz降低到200 Hz，200 Hz在能提取时频域特征的同时也减少了数据量<sup>[22]</sup>，脑电信号预处理步骤如图5所示。

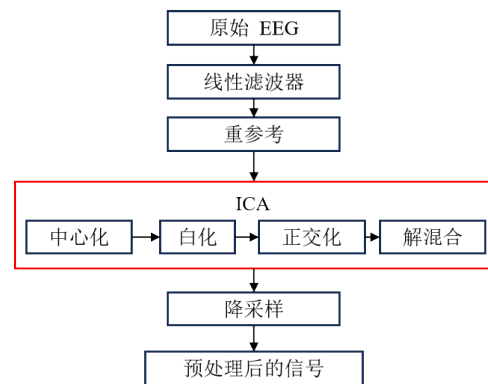


图5 脑电信号预处理步骤

## 1.3 特征提取

脑电信号的5种频段Delta波(0~4 Hz)、Theta波(4~8 Hz)、Alpha波(8~12 Hz)、Beta波(12~30 Hz)、Gamma波(30~50 Hz)都与驾驶员的表现相关。由于Delta波经常在深度睡眠中被观察到，而驾驶员的疲劳关乎安全，驾驶过程中不能睡觉，所以本研究不考虑Delta波。通过傅里叶变换将由不同频率的脑电波叠加形成的连续EEG信号从时域转换到频域，再使用Butterworth滤波器将32个导联下的脑电波分为4种节律波(Theta、Alpha、Beta、Gamma)，再分别对4种频段下的脑电信号提取时域特征和频域特征以及非线性特征进行分析。

### 1.3.1 频域特征

脑电信号的功率谱密度常被用来计算驾驶员的疲劳程度，由于脑电信号属于非平稳信号，为了减

少噪声，提高平滑度和估计精度，采用韦尔奇法计算各频段的功率谱密度，这是一种将信号分成多个重叠的窗口，并对每个窗口进行傅里叶变换，然后求平均的方法。使用了汉宁窗（Hanning Window）作为窗函数，由于窗口的大小能反映特征指标的灵敏度和准确性，选择窗口长度为 10 s，重叠度为 50%，窗口采用了与采样率相同的傅里叶变换点数，以保证频谱的分辨率。同时，对信号进行了常数去趋势处理，以消除直流分量的影响。在计算功率谱密度时，采用了密度缩放方式，以得到单位为 W/Hz 的功率谱密度。最后，对多个分段信号的功率谱密度进行了算术平均，以降低估计方差。采用韦尔奇法计算 PSD，如式(1)所示。

$$P(f) = \frac{1}{LK} \sum_{k=0}^{K-1} |X_k(f)|^2 \quad (1)$$

式中： $P(f)$ 为 PSD 的估计值； $L$ 为窗口长度； $K$ 为窗口个数； $X_k(f)$ 为第  $k$  个窗口的傅里叶变换。

### 1.3.2 时域特征

平均绝对值（MAV）是一种简单且常见的时域特征提取方法，并且 MAV 方法可以与时域和频域特征结合形成多维的特征向量，从而提高信号的分类和识别能力，对于采集到的脑电信号，在不同的时间窗口下计算信号绝对值的平均值，可以减少信号噪声，如式(2)所示。

$$M = \log\left(\frac{1}{N} \sum_{i=1}^N |x(i)|\right) \quad (2)$$

式中： $N$ 为时间窗口信号中的点数； $x(i)$ 为第  $i$  个点的值； $M$ 为对数转换值。

### 1.3.3 非线性特征

微分熵 DE 可以看作是连续化的香农熵的推广，可以提取信号的高阶统计特征，能有效反映连续随机信号的复杂度和随机性，处理脑电信号这类非平稳信号具有较好的鲁棒性和可靠性，DE 的计算如式(3)所示。

$$h(x) = -\int f(x) \log(f(x)) dx \quad (3)$$

如果随机变量服从高斯分布  $N(u, \delta^2)$ ，则 DE 可以通过式(4)进行更简单的计算。

$$h(Z) = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\delta^2}} \exp\left(-\frac{(z-u)^2}{2\delta^2}\right) \log \frac{1}{\sqrt{2\pi\delta^2}} \exp\left(-\frac{(z-u)^2}{2\delta^2}\right) dz = \frac{1}{2} \log 2\pi\epsilon\delta \quad (4)$$

## 1.4 特征组合

为了同时集成 32 个通道的时频域和空域信息，将脑电极定位位置铺设为一张  $8 \times 9$  的二维脑电图<sup>[23]</sup>，基于 1.3 节中提取的时频域特征，以 10 s 为一个时间窗口，50% 的重叠率将 4 种节律波下计算得到的 MAV、DE、PSD 分别堆叠为数个三维(3D)结构，形成  $8 \times 9 \times 4$ （高度  $h=8$ ，宽度  $w=9$ ，深度  $d=4$ ）的三维脑电图，如图 6a 所示。其中，只有 32 个格子是试验采集到的对应脑电极位置的数据，空白处的数据值设置为 0。考虑到模型信息的互补可能会提高模型识别的能力，将 MAV、DE、PSD 分别两两组合堆叠在一起，增加了输入层特征的深度，形成  $8 \times 9 \times 8$ （高度  $h=8$ ，宽度  $w=9$ ，深度  $d=8$ ）的三维脑电图，将 2 个小时的脑电数据片段拼接在一起就会分别得到 3 组  $1440 \times 8 \times 9 \times 4$  的四维脑电图数据，以及 3 组  $1440 \times 8 \times 9 \times 8$  的四维脑电图数据，如图 6b 所示，该数据同时包括脑电信号的空间位置信息、时间特征信息，以及频域特征信息。

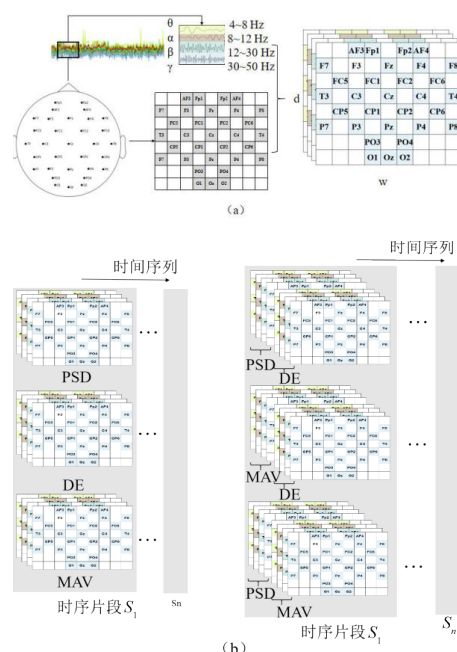


图 6 六种脑电特征组合输入模型的方案

## 2 模型搭建

### 2.1 样本集标注

除了脑电信号，试验还采集了主观的KSS疲劳量表以及SAM情绪量表数据，以及被试者在驾驶期间进行Stroop试验采集到的选择反应时间以及正确率，这些数据均用于计算与脑电数据的依赖关系从而实现疲劳程度的科学量化。

在之前的研究中，已通过TICC (Toeplitz Inverse Covariance-Based Clustering) 算法<sup>[23]</sup>进行了多模态数据的时序聚类<sup>[24]</sup>来实现疲劳的量化，该方法用于疲劳标注已得到了验证<sup>[7]</sup>。具体来说，由于脑电数据的采用频率为200 Hz，所以需要Stroop试验采集的数据和SAM以及KSS量表的数值进行时序的线性插值，并与脑电极各脑区平均电极值一并输入进行TICC时序数据聚类。聚类原理则是对于所有的输入变量定义一个独特的马尔可夫网络来表示两两变量之间的依赖关系，以此来计算分割后的时间序列中不同时间窗口数据的相似度，不同时间窗口下各变量之间的依赖网络相似则为同一类，通过计算不同分类数量下模型的贝叶斯信息准则BIC的极小值，从而确定了疲劳状态的最佳分类数量为4种。随后基于KSS量表在对应时间窗口下的值，由低到高将当前疲劳状态平均划分为4类，第1类为清醒；第2类为能感知到一点疲劳，但没有任何负面影响；第3类为能较清晰地感知到疲劳，但还在自己的掌控之中；第4类为非常明显地感受到疲劳和困倦。从而对原始脑电信号进行了标签，实现对驾驶疲劳的科学量化。数据清洗后的有效平均驾驶时长为107 min，按照前文的时间窗口划分方式，最终得到共12 840份标注好疲劳状态的样本数据，其中，1到4类数据的样本分别占41%、29%、19%、11%。

### 2.2 模型训练

为了解决如何通过多模态数据精准识别驾驶员精神疲劳这个问题，汲取了前人的研究经验，首先

搭建CNN与LSTM结合的计算框架。与完全卷积架构不同<sup>[25]</sup>的是，本文先通过CNN从数据中的不同时间窗口学习数据的时频域和空域特征，模型结构参考前人的研究<sup>[23]</sup>，只在最后一个卷积层后添加池化层。为了提高模型性能，在第1个卷积层后添加Dropout层和Batch Normalize层，而为了尽量多保留信息，具体结构如图7所示。由脑电特征数据作为模型的输入，CNN结构部分包含4个卷积层、1个Dropout层和Batch Normalize层以及最后的Maxpooling层、展平层和全连接层，应用ReLU (Rectified Linear Unit) 作为激活函数，最后输入到512个单元的全连接层，输出不同时间窗口下脑电信号的特征时序数据 $S_n = (s_1, s_2, \dots, s_T)$ 。

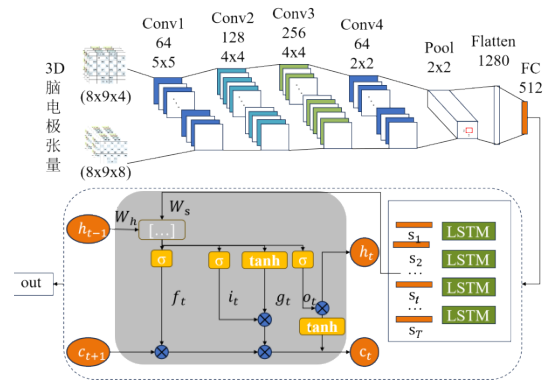


图7 模型理论框架

为了学习连续时间窗口下脑电信号的特征数据，通过带有长短期记忆(LSTM)单元的递归神经网络提取特征对时间变化的依赖性来实现准确的精神疲劳分类。在CNN输出时序数据 $S_n$ 后，作为LSTM的输入层，设置具有128个隐藏单元的LSTM层，从CNN输出的序列 $(s_1, s_2, \dots, s_T)$ 中集成时域、频域、空域的特征在时间序列上的信息，LSTM的输出由式(5)计算得到。

$$\begin{aligned}
 i_t &= \sigma(W_{si}s_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i), \\
 f_t &= \sigma(W_{sf}s_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f), \\
 c_t &= f_t c_{t-1} + i_t \tanh(W_{sc}s_t + W_{hc}h_{t-1} + b_c), \\
 o_t &= \sigma(W_{so}s_t + W_{ho}h_{t-1} + W_{co}c_t + b_o), \\
 h_t &= o_t \tanh(c_t), \\
 y_t &= W_{ho}h_t + b_o.
 \end{aligned} \quad (5)$$

LSTM的4个网络层分别由 $i$ 输入门、 $f$ 遗忘门、

$c$  细胞状态、 $o$  输出门表示,  $W$  为权重矩阵,  $b$  为偏置向量,  $\sigma$  为逻辑 Sigmoid 函数, 作为激活函数用于控制输入门、输出门、遗忘门, 数据经过遗忘门后被激活与上一个时间步的细胞状态进行 Hamadard 乘法运算并输出, 输出接近于 0 则被遗忘, 接近于 1 则被保留。 $y_n$  作为最后计算得到的特征。

采用线性变换的方法预测标签, 如式(6)所示。

$$\text{OUT} = Ay_n + b = [\text{out}_1, \text{out}_2, \dots, \text{out}_c]。 \quad (6)$$

式中:  $A$  为变换矩阵;  $b$  为偏差;  $C$  为分类的数量。

最终输出被输入到 Softmax 分类器中, 计算得到该时间窗口的片段属于某一类的概率, 通过 Softmax 分类的预测概率如式(7)所示。

$$P(c|X_n) = \max \left\{ \frac{\exp(\text{out}_j)}{\sum_{i=1}^c \exp(\text{out}_i)} \mid j = 1, \dots, C \right\}。 \quad (7)$$

由于数据维度较大, 为了避免模型过拟合, 添加 Early Stopping 方法, 监控模型在验证集上的性能指标, 如果损失函数值连续 30 个 epoch 没有恢复到最低值以下, 则立刻停止训练, 并通过交叉熵损失函数衡量模型预测结果与真实标签之间的差异。交叉熵损失值计算如式(8)所示。

$$L(x, y) = -\sum_{i=1}^c y_i \log(p_i)。 \quad (8)$$

式中:  $x$  为样本输入;  $y$  为样本标签;  $p$  为模型的预测结果;  $\log$  为自然对数函数。

交叉熵损失函数的值越小, 说明模型预测结果与真实标签越接近。

为了优化模型性能, 可以通过选择例如自适应矩估计 (Adam)、随机梯度下降 (SGD)、Adamax (Adaptive Moment Estimation with Infinity Norm) 等多种优化算法来不断地迭代调整模型的权重和偏差。由于优化器的选择并没有明确的方案, 所以为了提高模型的收敛速度, 研究匹配模型的最佳优化算法, 在调整超参数的同时对以上优化算法进行评估, 模型的理论框架如图 7 所示。

## 2.3 集成学习

对于多分类任务, 不同的特征可能具有不同的预测能力, 单个特征的模型不一定会比选择多个特征的模型性能差。为了进一步提高模型的识别准确性, 同时检验不同特征组合对于最终分类结果的影响, 采用集成学习的策略来吸收基分类器的优点, 不同的基分类器可以通过不同的特征输入、不同的数据集或是不同的算法来构建, 基于不同的基分类器的模型融合方法在实践中已被证明是有效的。以往对于脑电信号状态识别的研究往往通过不同频段下的特征构建的基分类器来构成集成策略<sup>[26]</sup>, 而本文的模型本身就包含了 4 种频段的信息, 所以数据量更加丰富, 为了充分利用信息互补, 发挥不同特征的优势, 基于不同特征组合的结果提出以下 3 种集成策略:

1) 集成分别以 MAV、DE、PSD 为输入特征的 3 种基分类模型;

2) 从以 MAV、DE、PSD 为输入特征的 3 种基分类模型中分别两两进行集成;

3) 集成堆叠输入特征组合后与另一种特征的 2 个基分类模型 (如集成 PSDDE 与 MAV 为输入的 2 种基分类模型)。

这些策略的输出结果将通过未加权平均法、动态修改权重法以及时间序列投票算法来确定。

1) 未加权平均法

对基学习器中的输出进行未加权平均是最常用的模型融合决策方法, 尤其是基于深度学习构架, 未加权平均法可以减少模型之间的方差, 从而提高集成模型的泛化性能<sup>[27]</sup>。未加权平均法如式(9)所示。

$$Y(x) = \left( \frac{1}{k} \sum_{i=1}^k y_i(x) \right)。 \quad (9)$$

式中:  $k$  为基分类器的个数;  $y$  为基分类器的分类结果。

2) 动态修改权重法

本文提出一种动态修改权重算法, 首先初始 3

个模型的权重，权重平均分配，在得到各模型的第 1 次预测结果后，之后的每一次训练迭代中如果其中一个模型的分类预测结果正确，则后续该模型的输出权重增加，然后将 3 个模型的权重归一化重新分配权重，每一次训练迭代时将 3 个模型中的权重最大者作为预测结果输出。

$$w_i = w_j / \sum w_j \quad (10)$$

$$y = \operatorname{argmax} w_i p_i(x) \quad (11)$$

式中： $w_i$  为每个模型的权重。

### 3) 时序投票算法

考虑到脑电信号是具有时间相关性的时间序列数据，时间序列投票算法通过投票将时间维度上的连续样本概率值组合起来再求平均来预测类别标签<sup>[28]</sup>，如式 (12) 所示。

$$\hat{y}_j = \operatorname{argmax} \sum_j^{t'} P_{ij} \quad (12)$$

式中： $t$  为合并样本数； $P_{ij}$  为第  $j$  个样本属于  $i$  类的概率。

## 3 模型效果

在完成数据标注后，将按时间顺序的 3D 脑图序列通过 CRNN 进行训练，最终模型的性能选择五重交叉验证方法，经过多次调试，Adam 优化算法在基于单输入特征和组合输入特征的模型下，都能提供最佳的优化精度，确定好的最佳模型参数见表 1。

表 1 模型参数

参数设置	值	
	单一特征	多特征组合
初始学习率	0.001	0.001
Patience	3 epochs	5 epochs
Dropout rate	0.2	0.6
Batch size	128 sample	256 sample
Epochs	300	300

一般来说准确率越高，分类器的性能越好。F1 分数是精确率和召回率的调和平均值。精确率 (Precision) 是指预测为正例的样本中，实际为正例的样本数占总预测为正例的样本数的比例；

Macro-Precision 则是指将每个类别的 Precision 求平均值；召回率 (Recall) 是指实际为正例的样本中，被预测为正例的样本数占总实际为正例的样本数的比例；Macro-Recall 则是指将每个类别的 Precision 求平均值；Macro-F1 综合 Macro-Precision 和 Macro-Recall 的优点。它们的计算式分别为：

$$\text{Macro-Precision} = \frac{\sum_{i=1}^l \frac{tp_i}{tp_i + fp_i}}{l} \quad (13)$$

$$\text{Macro-Recall} = \frac{\sum_{i=1}^l \frac{tp_i}{tp_i + fn_i}}{l} \quad (14)$$

$$\text{Macro-F1} = \frac{2 \times \text{Macro-Precision} \times \text{Macro-Recall}}{\text{Macro-Precision} + \text{Macro-Recall}} \quad (15)$$

式中： $l$  为分类数量； $tp_i$  为第  $i$  类的真阳性计数； $fp_i$  为第  $i$  类的假阳性计数； $fn_i$  为第  $i$  类的假阴性计数。

不同输入特征组合下模型的性能指标见表 2。

表 2 不同特征组合下模型的性能指标

模型指标	PSD	DE	MAV	PSD&DE	PSD&MAV	DE&MAV
Accuracy/%	83.33	91.28	92.85	88.46	93.10	91.89
Macro-Precision/%	76.31	88.00	86.84	86.07	85.33	87.50
Macro-Recall/%	88.88	95.34	93.02	92.13	89.65	97.95
Macro-F1/%	82.11	91.52	89.82	88.99	87.43	92.43

从准确率来看，多输入特征组合的模型准确率相对于单输入特征的模型准确率平均提高了 2.24%，但是并非多特征的性能就都比单特征的好，例如当把 PSD 和 DE 作为特征组合输入的模型准确率虽然高于 PSD 作为单输入特征时的模型准确率，但却低于 DE 作为单输入特征时的模型准确率（相对于 PSD 提高 6.15%，相对于 DE 降低 3.08%），DE 与 MAV 两输入特征组合时也是类似的情况（相对于 DE 提高 0.67%，相对于 MAV 降低 1.03%）。而当 PSD 和 MAV 作为输入特征组合时，模型的准确率又同时高于 PSD 和 MAV 分别作为单输入特征时模型的准确率（相对于 PSD 提高 11.72%，相对于 MAV 提高 0.26%），说明将多特征进行组合输入模型进行训练至少能在某种单特征相对表现不好的情

况下提高准确率。

除了当 PSD 为单输入特征时模型的 Macro-Precision 相对较低 (76.31%)，其他几类输入特征的模型 Macro-Precision 波动变化不大，平均为 86.74%，说明这些模型无论是单输入特征还是多输入特征在每种类别都能得到较高的准确率，对于疲劳状态的识别也足够可靠。PSD 为单输入特征时，模型的 Macro-Recall 与其他模型相比也相对较低 (88.88%)，而其他几类输入特征的模型 Macro-Recall 则较高，平均为 93.61%，并且模型之间的 Macro-Recall 差异较小，多输入特征的模型平均 Macro-Recall 只比单输入特征的提高了 0.89%，说明不论是多输入特征还是单输入特征仍会有一部分真阳性数据（正确的精神疲劳状态）被错误分类为

了其他状态。

由于在整个驾驶期间疲劳状态并非均匀分布，驾驶员可能会出现大部分时间处于同一种状态，而处于某种状态的时间较少，只考虑 ACC 可能会对模型的实际效果评价有失偏颇，所以可以通过 Macro-F1 对该情况下模型是否表现良好进行了评估，较高的 Macro-F1 说明模型表现良好，精确率和召回率都偏高。DE 和 MAV 组合输入模型的 Macro-F1 得分最高 (92.43%)，表明该种特征下模型对于 4 种类别的分类更加平衡，其次则是 DE 作为单输入特征时的 Macro-F1 (91.52%)。

对于分类问题，研究者们都希望尽可能地再提高一些模型的精度，表 3 为本研究集成策略的结果。

表 3 三种集成策略的结果

集成策略	方法	Accuracy/%	Macro-Precision/%	Macro-Recall/%	Macro-F1/%	
策略 1	未加权平均法	90.37	85.71	94.11	89.71	
	动态修改权重法	91.22	89.46	95.13	92.20	
	时序投票算法	90.13	88.96	93.89	91.35	
策略 2	未加权平均法	86.69	88.29	93.27	90.71	
	PSD*DE	92.59	89.14	94.7	91.83	
	DE*MAV	91.33	84.83	94.51	89.40	
	PSD*DE	87.73	84.23	94.2	88.93	
	动态修改权重法	PSD*MAV	94.65	92.81*	95.69	94.22*
	DE*MAV	91.55	91.07	96.47	93.69	
	PSD*DE	88.13	86.32	95.59	90.71	
	时序投票算法	PSD*MAV	94.27	91.05	96.83*	93.85
	DE*MAV	95.47*	92.64	94.73	93.67	
策略 3	未加权平均法	90.27	87.66	95.49	91.40	
	PSD&DE*MAV	94.09	89.46	93.22	91.30	
	MAV&DE*PSD	87.22	84.17	91.24	87.56	
	PSD&DE*MAV	93.53	87.82	92.91	90.29	
	动态修改权重法	PSD&MAV*DE	91.79	88.97	95.5	92.11
	MAV&DE*PSD	89.83	84.21	90.16	87.08	
	PSD&DE*MAV	93.37	89.37	92.88	91.09	
	时序投票算法	PSD&MAV*DE	92.34	89.49	94.11	91.74
	MAV&DE*PSD	89.02	84.98	90.43	87.62	

“\*”表示不同特征的基分类模型集成，“&”表示输入特征的组合

由表可知，采用的集成策略确实能组合各个基分类器的优势，从各方面有效地提升模型的性能，策略 1 相对于集成前各模型的平均 ACC 提升了 0.46%，策略 2 则提升了 1.35%，策略 3 提升了 1.24%。3 种方法中未加权平均法相对于集成前各模型的平均 ACC 提升了 0.23%，动态修改权重法则提升了 1.46%，时序投票算法提升了 1.84%，总体而言，采用了集成策略后模型的 ACC 平均提高了 1.17%。其中，表现最好的是当 DE 和 MAV 分别作为单输入特征的 2 个基分类器通过时序投票算法得到的最终预测结果(ACC= 95.47%)，ACC 的性能相对于原基分类器(DE 和 MAV 分别作为单输入特征的模型以及 DE 和 MAV 作为特征组合输入的模型)性能分别提高了 4.59%、2.82%、3.74%，该模型验证集和训练集的损失值如图 8 所示。基于相同的基模型结构，在选择合适的特征后模型收敛得很快，模型构建 4 分类混淆矩阵如图 9 所示。其中，数字分别对应了分类的类别(例如数字 1 对应分类 1)，该特征组合下的时序投票算法效果更好的原因可能是该方法对于状态的识别更加平滑，降低了本身十分敏感的时序数据中因为个别噪声导致的误分类可能性。根据集成学习之后的性能变化，本研究发现似乎通过集成不同特征的基分类器能补充一些组合输入特征的一些缺陷，使模型能更好地学习各个特征下的特点和规律。

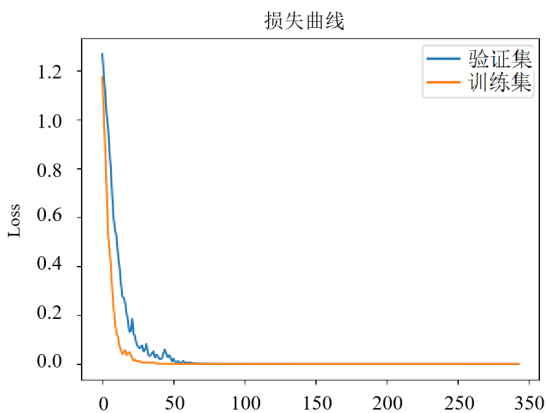


图 8 模型损失变化

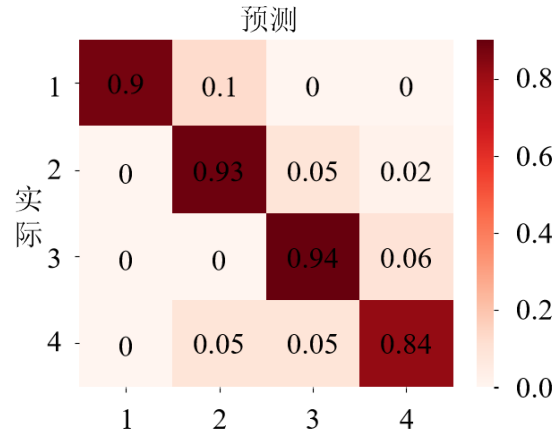


图 9 模型混淆矩阵

## 4 结果分析

### 4.1 特征分析

本研究选择了脑电信号 3 种不同类型的特征：PSD、DE、MAV。PSD 是在脑电信号的疲劳预测中最常用的特征，而本文通过探讨 3 种特征不同组合下在 CRNN 中的不同性能表现，发现 PSD 在多分类模型中的表现效果并没有那么好。以往的部分研究者也认为 DE 特征比 PSD 特征更适合基于 EEG 的驾驶疲劳检测，但当 PSD 与其他的基分类模型集成后也能贡献一些重要信息，例如当 PSD 与 MAV 两类基分类器通过动态修改权重法集成后，不仅 ACC 高于原分类器，Macro-Precision 也显著提升，所以对于 PSD 特征来说，可能其时序上的相关性或重叠性较低，对于多分类任务来说预测性能较差，更适合用于与其他模型特征数据的丰富和补充。此外，以往的研究结果和本研究的结果显示，DE 不仅在二分类任务中表现优良，在多分类的模型中也具有较好的性能，尤其是与 MAV 的组合，不仅在模型输入时将特征进行组合能提供更多有效信息，集成基分类器后的性能更是能大幅提升。之所以组合的性能较好，本文推测首先是因为当样本数据之间差异较大时，原始时序数据的大小波动就较明显，而在静息时大脑趋向于休息，活跃的脑区更集中，也更易学习到规律，所以区分静息和其他状态的分类

效果很好。由静息到唤醒再到疲劳,各脑区之间的时序差异很大,因此,MAV特征在多种状态的分类中表现较好。其次是因为在不同的时间窗口下如果数据偏差较大的话,熵值也能较易区分出波动的复杂性和规律性,所以组合得到的总体性能表现较好。而使用了PSD特征组合输入的模型,如果与DE进行组合,由于PSD特征与DE特征数据计算部分的一些重叠性和相似性,所以当DE与PSD同时进行特征组合时,其效果并不如预期的那样显著。

#### 4.2 集成策略的比较

如何通过集成策略充分利用不同特征的优势来使模型得到更好的性能,通过对3种集成策略的研究发现,总体来说,越多的特征对于基分类器模型的简单组合的识别准确性确实有所提升。在未加权平均法下,同时包含了3种特征的策略1和策略3比策略2下的模型的平均性能高0.18%~0.35%。而在其他集成方法下,多特征的优势则可以通过结合特征特点选择合适的集成方法来弥补,例如在动态修改权重法下,策略2比策略1虽然少一种特征,但模型的平均性能却要高0.09%,而采用时序投票算法后,策略2比策略1和策略3的模型平均性能则要高1.14%~2.76%。

无法确定模型中哪种特征的性能最重要时,选择未加权平均法,在尽量多的特征输入下能平均不同特征输入下基分类器的优劣,总体性能能有较小提升。而如果想要减少特征数量,则需要充分利用各特征的优势,动态修改权重法基于每个基分类器

的性能动态调整最终的模型预测结果,对于性能比较接近的基分类器,特征较少时模型性能提升较小,此时选择增加特征可能提高模型最终性能,例如在动态修改权重法下PSD与DE集成后的性能仍较差,而增加特征MAV后能提高6.61%。针对不同的模型结构,选择不同的集成方法也能有效提升模型的最终性能,由于本研究模型的LSTM模块能捕捉脑电时序数据,在不同的时间段识别驾驶员精神状态具有优越性,采用时序投票算法也能有效利用小时间窗的脑电特征的时序性特点,所以在该模型中选择时序投票算法能有效综合模型的优势。

## 5 结论

本文探讨了在精神疲劳程度不断积累的驾驶工作中,脑电信号3种不同的输入特征经过不同的组合是否能通过CRNN学习到脑电数据由于疲劳导致的时序性变化,从而有效地进行驾驶疲劳状态的多分类识别。根据特征的不同组合方式进一步提出了3种集成策略,不同的集成策略下分别通过未加权平均法、动态修改权重法、时序投票算法进行集成。结果显著提高了模型的性能,准确率最高可以达到95.47%,并且模型所需的环境简单,能达到较高性能的同时满足正式驾驶场景的数据采集时效性,实用性较强,未来可以基于此研究结果对驾驶期间的驾驶员疲劳进行实时监控,从而避免因疲劳驾驶导致的危险事故,并且基于驾驶员不同的状态设计不同的辅助功能。

## 参考文献 (References)

- [1] GAO Ting, ZHU Haiyan, LIU Zhigang, et al. Design of Real-Time Monitoring System for Physiological Status of Metro Train Drivers[C]//COTA International Conference of Transportation Professionals, July 6-8, 2019, Nanjing, China. Piscataway NJ: IEEE, c2019:392-399.
- [2] YOU Xingda, ZHU Lin, LIU Zhigang, et al. Experimental Study on the Relationship between Fatigue and Unsafe Behavior of Urban Rail Transit Drivers[J]. Transportation Research Record: Journal of the Transportation Research Board, 2021, 2675(10):1151-1160.
- [3] KÖRBER M, CINGEL A, ZIMMERMANN M, et al. Vigilance Decrement and Passive Fatigue Caused by Monotony in Automated Driving [J]. Procedia Manufacturing, 2015, 3:2403-2409.
- [4] PHILLIPS R O. A Review of Definitions of Fatigue—And a Step Towards a Whole Definition [J]. Transportation Research Part F: Traffic Psychology and Behaviour, 2015, 29:48-56.

- [5] AL-LIBAWY H, AL-ATABY A, AL-NUAIMY W, et al. HRV-Based Operator Fatigue Analysis and Classification Using Wearable Sensors [C]//2016 13th International Multi-Conference on Systems, Signals & Devices(SSD), Mar. 21–24, 2016, Leipzig, Germany. Piscataway NJ: IEEE, c2016:268–273.
- [6] LIU N H, CHIANG C Y, HSU H M. Improving Driver Alertness through Music Selection Using a Mobile EEG to Detect Brainwaves [J]. *Sensors*, 2013, 13 (7) : 8199–8221.
- [7] LI Jue, LI Heng, UMER W, et al. Identification and Classification of Construction Equipment Operators' Mental Fatigue Using Wearable Eye-Tracking Technology [J]. *Automation in Construction*, 2020, 109: 103000.1–103000.15.
- [8] HOU Zhongwei, OU Shuangjiang, XU Dengyuan. Research on Fatigue Driving Feature Detection Algorithms of Drivers Based on Machine Learning [J]. *Systems Science & Control Engineering*, 2021, 9(1) : 167–172.
- [9] LIU Fan, CHEN Delong, ZHOU Jun, et al. A Review of Driver Fatigue Detection and Its Advances on the Use of RGB-D Camera and Deep Learning [J]. *Engineering Applications of Artificial Intelligence: The International Journal of Intelligent Real-Time Automation*, 2022, 116: 105399.1–105399.19.
- [10] HUANG Mengtao, WANG Yi. Driver Fatigue Expression Recognition Research Based on Convolutional Neural Network [C]//2018 International Symposium on Computer, Consumer and Control (IS3C), Dec. 6–8, 2018, Taichung, Taiwan, China. Piscataway NJ: IEEE, c2018:6–9.
- [11] RONG Qian, DING Shuai, YUE Zijie, et al. Non-Contact Negative Mood State Detection Using Reliability-Focused Multi-Modal Fusion Model [J]. *IEEE Journal of Biomedical and Health Informatics*, 2022, 26 (9) : 4691–4701.
- [12] CHAI R, TRAN Y, LING S H, et al. Combining ICA Clustering and Power Spectral Density for Feature Extraction of Mental Fatigue of Spinal Cord Injury Patients [C]//2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), July 23–27, 2019, Berlin, Germany. Piscataway NJ: IEEE, c2019:530–533.
- [13] WANG Hongtao, DRAGOMIR A, ABBASI N I, et al. A Novel Real-Time Driving Fatigue Detection System Based on Wireless Dry EEG [J]. *Cognitive Neurodynamics*, 2018, 12(4) : 365–376.
- [14] ZHAO Chunlin, ZHENG Chongxun, ZHAO Min, et al. Multivariate Autoregressive Models and Kernel Learning Algorithms for Classifying Driving Mental Fatigue Based on Electroencephalographic [J]. *Expert Systems with Applications*, 2011, 38(3) : 1859–1865.
- [15] ZAMANI F, WULANSARI R. Emotion Classification Using 1D-CNN and RNN Based on DEAP Dataset [J]. *Natural Language Processing*, 2021 : 363–378.
- [16] CHEN Jichi, WANG Hong, HUA Chengcheng. Electroencephalography Based Fatigue Detection Using a Novel Feature Fusion and Extreme Learning Machine [J]. *Cognitive Systems Research*, 2018, 52: 715–728.
- [17] STROOP J R. Studies of Interference in Serial Verbal Reactions [J]. *Journal of Experimental Psychology General*, 1992, 121(1) : 15–23.
- [18] AKERSTEDT T, GILLBERG M. Subjective and Objective Sleepiness in the Active Individual [J]. *International Journal of Neuroscience*, 1990, 52(1–2) : 29–37.
- [19] PACE F, SCIOTTO G. The Effect of Emotional Dissonance and Mental Load on Need for Recovery and Work Engagement among Italian Fixed-Term Researchers [J]. *International Journal of Environmental Research and Public Health*, 2021, 18(1) : 18010099.1–18010099.17.
- [20] BRADLEY M M, LANG P J. Measuring Emotion: The Self-Assessment Manikin and the Semantic Differential [J]. *Journal of Behavior Therapy and Experimental Psychiatry*, 1994, 25(1) : 49–59.
- [21] ZHU Miankuan, CHEN Jiangfan, LI Haobo, et al. Vehicle Driver Drowsiness Detection Method Using Wearable EEG Based on Convolution Neural Network [J]. *Neural Computing and Applications*, 2021, 33(20) : 13965–13980.
- [22] LEE K F A, CHAN E, CAR J, et al. Lowering the Sampling Rate: Heart Rate Response During Cognitive Fatigue [J]. *Biosensors*, 2022, 12 (5) : 12050315.1–12050315.11.
- [23] SHEN Fangyao, DAI Guojun, LIN Guang, et al. EEG-Based Emotion Recognition Using 4D Convolutional Recurrent Neural Network [J]. *Cognitive Neurodynamics*, 2020, 14(6) : 815–828.
- [24] HALLAC D, VARE S, BOYD S, et al. Toeplitz Inverse Covariance-Based Clustering of Multivariate Time Series Data [C]//27th International Joint Conference on Artificial Intelligence, July 13–19, 2018, Stockholm, Sweden. 2018:5254–5258.
- [25] RUDAKOV E, LAURENT L, COUSIN V, et al. Multi-Task CNN Model for Emotion Recognition from EEG

- Brain Maps [C]//2021 4th International Conference on Bio-Engineering for Smart Technologies (BioSMART), Dec. 8–10, 2021, Paris/Creteil, France. Piscataway NJ: IEEE, c2021:1–4.
- [26] GANAIE M A, HU Minghui, MALIK A K, et al. Ensemble Deep Learning: A Review [J]. Engineering Applications of Artificial Intelligence, 2022, 115: 105151.
- [27] FAN Chaojie, PENG Yong, PENG Shuangling, et al. Detection of Train Driver Fatigue and Distraction Based on Forehead EEG: A Time-Series Ensemble Learning Method [J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(8): 13559–13569.
- [28] HUO Xueqin, ZHENG Weilong, LU Baoliang. Driving Fatigue Detection with Fusion of EEG and Forehead EOG [C]//2016 International Joint Conference on Neural Networks (IJCNN), July 24–29, 2016, Vancouver, British Columbia, Canada. Piscataway NJ: IEEE, c2016: 897–904.

#### 作者简介



王雨生 (2000–), 男, 四川绵阳人, 硕士研究生, 主要研究方向为交通安全、人因安全。  
E-mail: w\_y\_s\_710302970@163.com