

# 多模态生理特征融合的驾驶行为识别研究

邵舒羽<sup>1,2</sup>, 张 扬<sup>1</sup>, 颜文婧<sup>3</sup>

(1. 北京物资学院智能工程与供应链创新学院, 北京 101149;

2. 智能物流系统北京市重点实验室, 北京 101149; 3. 北京工商大学计算机与人工智能学院, 北京 100048)

**摘要:** 传统驾驶行为识别方法存在依赖外部传感数据、易受环境干扰及难以反映驾驶员内部认知状态等局限。为此, 本文构建一种多模态生理信号深度学习框架, 融合 Transformer 与卷积神经网络 (Convolutional Neural Network, CNN), 以实现驾驶行为高精度识别与可解释性分析。研究依托包含脑电图 (Electroencephalography, EEG)、心电图 (Electrocardiography, ECG)、肌电图 (Electromyography, EMG) 与皮肤电导率 (Galvanic Skin Response, GSR) 的多模态生理数据集 (Multimodal Physiological Data for Behavior recognition, MPDB), 系统规划从信号预处理、特征提取至时空融合的完整流程。各模态信号通过滤波、伪迹校正、特征标准化及时频变换后同步对齐, 构建时空特征张量以实现不同生理模态间的统一表示。在模型架构层面, CNN 分支负责捕捉局部时空模式并提取短时响应特征, Transformer 分支则凭借自注意力机制对生理信号长程依赖与跨模态交互关系建模, 兼顾局部敏感性与全局时序建模能力。融合网络采用双流结构, 将多头注意力与多尺度卷积相结合, 并引入动态权重分配机制实现特征自适应融合。优化进程运用 AdamW 算法与 Dropout 正则化, 进一步提升模型的泛化性能与收敛稳定性。实验结果表明: 该模型在二分类 (平稳驾驶/动态驾驶) 任务中, 准确率分别达到 94.9% 与 98.75%; 在 5 种驾驶行为识别 (平稳驾驶、加速、减速、换道、转弯) 中, 模型平均准确率为 85.39%, 显著高于循环神经网络 (Recurrent Neural Network, RNN)、长短期记忆 (Long Short-Term Memory, LSTM) 网络、支持向量机 (Support Vector Machine, SVM)、单一 CNN 及单一 Transformer 模型, 且在  $F_1$  分数和召回率上取得了良好的平衡, 验证了其在多模态信号表征和时序依赖建模方面的综合性优势。模型训练曲线也表明: 该框架收敛速度快且收敛损失值较低, 体现出较强的鲁棒性与抗过拟合能力。在此基础上, 为提升模型可解释性, 本文引入深度 SHAP (Deep SHapley Additive exPlanations, DeepSHAP) 方法对所建立模型的决策过程进行特征归因分析。分析结果表明: 高频脑电信号 ( $\beta$  波、 $\gamma$  波) 和上肢肌电信号对加速驾驶操作影响较大, 而胫骨前肌活动和反应延迟对换道驾驶操作具有显著影响。本文提出的方法揭示了不同驾驶操作背后生理响应规律, 为探究驾驶员神经-行为层级关系提供了新的视角。综上所述, 本文提出的 Transformer-CNN 融合框架能够有效提取多模态生理信号的时空信息特征, 在识别精度、稳定性以及解释性等性能指标方面均表现优异, 不仅为智能驾驶监测系统提供实用的技术支持, 还为在驾驶安全研究中多源信号建模与可解释人工智能的应用提供技术方向。下一步工作将考虑自然驾驶条件对所提方法的影响, 以期其在驾驶状态实时监测、连续性风险预测中得到更广泛的应用。

**关键词:** 驾驶行为; 多模态生理信号; Transformer; 卷积神经网络 (CNN); SHAP

**基金项目:** 国家自然科学基金 (No.82102176, No.62433002); 北京市教委科技计划项目 (No.KM202210037001)

**中图分类号:** TP18; U463.6; U492.8 **文献标识码:** A **文章编号:** 0372-2112(2025)10-3540-11

**电子学报 URL:** <http://www.ejournal.org.cn>

**DOI:** 10.12263/DZXB.20250506

## Driving Behavior Recognition Based on Multimodal Physiological Feature Fusion

SHAO Shu-yu<sup>1,2</sup>, ZHANG Yang<sup>1</sup>, YAN Wen-jing<sup>3</sup>

(1. School of Intelligent Engineering and Supply Chain Innovation, Beijing Wuzi University, Beijing 101149, China;

2. Beijing Key Laboratory of Intelligent Logistics System, Beijing 101149, China;

3. School of Computer and Artificial Intelligence, Beijing Technology and Business University, Beijing 100048, China)

**Abstract:** Traditional driving behavior recognition methods have limitations such as relying on external sensor data, being vulnerable to environmental interference, and being difficult to reflect the internal cognitive state of drivers. To this

end, this paper constructs a multimodal physiological signal deep learning framework that integrates Transformer and convolutional neural network (CNN) to achieve high-precision recognition and interpretability analysis of driving behavior. The research is based on a multimodal physiological data for behavior recognition (MPDB) containing electroencephalogram (EEG), electrocardiogram (ECG), electromyography (EMG), and galvanic skin response (GSR), and systematically plans a complete process from signal preprocessing, feature extraction to spatio-temporal fusion. After filtering, artifact correction, feature standardization and time-frequency transformation, the signals of each mode are synchronously aligned to construct a spatio-temporal feature tensor to achieve a unified representation among different physiological modes. At the model architecture level, the CNN branch is responsible for capturing local spatiotemporal patterns and extracting short-term response features, while the Transformer branch models the long-term dependence of physiological signals and cross-modal interaction relationships through its self-attention mechanism, taking into account both local sensitivity and global temporal modeling capabilities. The fusion network adopts a two-stream structure, combines multi-head attention with multi-scale convolution, and introduces a dynamic weight distribution mechanism to achieve feature adaptive fusion. The optimization process employs the AdamW algorithm and Dropout regularization to further enhance the generalization performance and convergence stability of the model. The experimental results show that in the binary classification (smooth driving/dynamic driving) tasks, the accuracy rates of this model reach 94.9% and 98.75% respectively. Among the five types of driving behavior recognition (smooth driving, acceleration, deceleration, lane changing, and turning), the average accuracy rate of the model was 85.39%, significantly higher than that of recurrent neural network (RNN), long short-term memory (LSTM) network, support vector machine (SVM), single CNN, and single Transformer. Moreover, it achieved a good balance in  $F_1$  and recall rate. It has verified its comprehensive advantages in multimodal signal characterization and timing dependency modeling. The model training curve also indicates that this framework has a fast convergence speed and a low convergence loss value, suggesting that it has strong robustness and is not prone to overfitting. On this basis, in order to enhance the interpretability of the model, this paper introduces the deep SHapley additive explanations (DeepSHAP) method to conduct feature attribution analysis on the decision-making process of the established model. The analysis results show that high-frequency electroencephalographic signals ( $\beta$  waves,  $\gamma$  waves) and upper limb electromyographic signals have a significant impact on accelerated driving operations, while the activity and reaction delay of the tibialis anterior muscle have a significant impact on lane-changing driving operations. The method proposed in this paper reveals the physiological response laws behind different driving operations, providing a new perspective for exploring the neuro-behavioral hierarchical relationship of drivers. In conclusion, this paper proposes that the Transformer-CNN fusion framework can extract the spatiotemporal information features of multimodal physiological signals quite well. It has achieved good performance in performance indicators such as recognition accuracy, stability, and interpretability, and at the same time provides applicable technical support for the constructed intelligent driving monitoring system. It also provides a technical direction for the application of multi-source signal modeling and explainable artificial intelligence in driving safety research. In the next step of work, the research on the proposed method under natural driving conditions will be considered, so as to better apply it in real-time monitoring of driving conditions and continuous risk prediction.

**Key words:** driving behavior; multimodal physiological signals; Transformer; convolutional neural network (CNN); SHAP

**Foundation Item(s):** National Natural Science Foundation of China (No.82102176, No.62433002); Beijing Municipal Education Commission Science and Technology Plan Project (No.KM202210037001)

## 1 引言

据世界卫生组织统计,全球每年因交通事故造成的死亡人数超过百万,其中驾驶员的决策和行为在事故成因中占据显著比例<sup>[1]</sup>.因此,如何提升道路交通安全性、降低事故发生率仍是亟待解决的关键问题<sup>[2]</sup>.

驾驶行为识别通常基于驾驶员的操作行为和车辆状态数据,进而辨识驾驶员的驾驶风格及潜在的危险驾驶行为<sup>[3]</sup>.传统方法主要依赖视觉图像或车辆传感器数据,但易受光照、天气等环境因素干扰,导致识别

准确率受到限制<sup>[4]</sup>.相比之下,多模态生理特征可直接反映驾驶员的生理和心理状态,经特征融合技术整合后,可以更准确地识别驾驶员的疲劳、情绪和认知负荷,从而提高驾驶安全性<sup>[5,6]</sup>.在此背景下,利用多模态生理信号进行驾驶行为识别已成为提升智能交通系统安全性的关键研究方向<sup>[7]</sup>.

学者们已针对多模态生理特征在驾驶行为识别中的应用展开探索.文献[5]采集80名出租车司机在高速、中速、低速及拥堵路况下的生理数据,利用卷积神经网络-长短期记忆(Convolutional Neural Network-Long

Short-Term Memory, CNN-LSTM)网络模型,实现心率、呼吸、皮肤电三模态输入下 86.38% 的识别准确率.文献[8]采集受试者在紧急制动和正常驾驶两种驾驶行为下的脑电数据,采用支持向量机(Support Vector Machine, SVM)、决策树、朴素贝叶斯和  $K$  最近邻算法等分类器完成行为识别,分类准确率均在 74% 以上.文献[6]提出一种融合浅层与深层特征并结合 CHI-RFE-SHAP (CHI-square test-Recursive Feature Elimination-SHapley Additive exPlanations)混合特征选择的驾驶员状态识别模型,通过模拟驾驶实验采集脑电和皮肤电数据,经卷积神经网络(Convolutional Neural Network, CNN)和长短期记忆(Long Short-Term Memory, LSTM)网络提取融合特征后,利用 CatBoost 算法实现了 94.88% 的识别准确率.文献[9]采集受试者在执行加速、减速和转向行为时的脑电数据,构建基于压缩-激励模块的通道注意力-可分离 CNN,对受试者的 3 种驾驶行为进行识别.模型识别的准确率达到 82%.文献[10]采集车辆运行、眼动及心电等多模态驾驶数据,并构建随机森林分类模型识别不同危险驾驶行为.文献[11]利用 35 名受试者的多模态生理数据及六自由度驾驶模拟器获得的眼动数据,研究将驾驶行为细分为 5 类,并发布多模态生理数据集,为提升驾驶安全与驾驶员健康评估提供支撑.文献[12]提出一种结合电子车辆识别与多模态生物特征识别的驾驶员生理状态监测与分析方法,采用 CNN-LSTM 深度学习模型,显著提高了生理状态监测的

准确性和实时性.

然而,现有模型在处理多模态生理信号的长时序依赖关系时,存在信息丢失和时序特征提取不充分等问题,难以有效地捕捉驾驶员状态的动态变化规律.同时,传统融合方法(特征拼接、加权融合)及降维方法(主成分分析)无法自适应分配不同模态(脑电、皮肤电等)的权重,对复杂驾驶场景下的特征交互建模能力较弱,冗余特征容易干扰识别精度.针对上述问题,本文提出一种基于 Transformer-CNN 融合网络的深度学习框架.该框架利用 Transformer 的自注意力机制全局捕捉生理信号的长距离依赖,并结合 CNN 的局部感知能力,兼顾局部特征与全局时序特征.进一步引入动态注意力权重分配机制,实现多模态特征的高效融合,为复杂驾驶场景下的疲劳、情绪等状态的精准识别提供途径.

## 2 实验设计与数据处理

### 2.1 数据集

本文采用专为驾驶行为识别任务设计的多模态生理数据集(Multimodal Physiological Data for Behavior recognition, MPDB)<sup>[11]</sup>.该数据集采集了 35 名健康受试者在执行 5 种典型驾驶行为(平稳驾驶、加速、减速、变道、转弯)时的脑电图(ElectroEncephaloGram, EEG)、心电图(ElectroCardioGram, ECG)、肌电图(ElectroMyoGraphy, EMG)及皮肤电导率(Galvanic Skin Response, GSR)信号,各模态信号严格同步.相关信息如表 1 所示.

表 1 MPDB 相关信息

属性	详细内容
EEG	64 电极的脑电帽,其中 59 个用于采集脑电信号,采样频率为 1 000 Hz,电极位置遵循国际标准的 10-20 系统排列.
ECC	一个电极贴片贴在胸部靠近心脏的位置,采样频率为 1 000 Hz.
EMG	采集模块贴在胫骨前肌、右小腿腓肠肌和双臂肱桡肌上,采样频率为 1 000 Hz.
GSR	电极贴在左手食指和中指指肚上,采样频率为 1 000 Hz.

实验总体设计框架如图 1 所示.针对 MPDB 数据集的 EEG、ECG、EMG 及 GSR 信号进行预处理[如 EEG 带通滤波、独立成分分析伪迹校正,ECG 提取心率变异性(Heart Rate Variability, HRV)特征等]、多模态特征提取(EEG 生成时频图,ECG、EMG、GSR 提取时频及特征)、构建时空维度多模态特征张量,最终通过 Transformer-CNN 融合网络实现驾驶行为分类,对 5 类基本驾驶行为进行识别.

### 2.2 数据预处理与特征提取

针对 EEG 信号,首先采用巴特沃斯滤波器进行带通滤波(0.1~48 Hz),以去除低频漂移和低频肌电伪迹<sup>[13]</sup>,计算式为

$$x_{EEG}^f(t) = x_{EEG}(t) * h_{BP}(t) \quad (1)$$

其中,  $f \in [0.1, 48]$  Hz;  $x_{EEG}(t)$  为原始脑电信号;  $h_{BP}(t)$  为带通滤波器的冲激响应; \* 为卷积运算.其次,伪迹校正采

用独立成分分析,通过信号分解去除眼电、心电伪迹,采用 2 s 窗口以及 50% 重叠对数据进行分段,并根据驾驶行为事件标签对齐窗口,保留事件前后 1 s 的数据.再次,通过 Z-score 标准化消除个体差异,采用 Laplacian 滤波增强局部脑区活动并抑制远场干扰.最后,通过短时傅里叶变换生成时频图,计算式为

$$X_{EEG}(t, f) = \sum_{\tau} x_{EEG}(\tau) \cdot \omega(t - \tau) e^{-j2\pi f\tau} \quad (2)$$

其中,  $\omega(t - \tau)$  为加窗函数;  $f$  为频率,结果作为 CNN 的输入特征.

在 ECG 预处理阶段,首先采用带通滤波(0.5~40 Hz)保留 ECG 信号的 QRS 波群,抑制基线漂移和肌电干扰.其次,通过 R 波峰检测用于定位每个心拍周期,并截取 R 波峰前后窗口标准化为 300 个采样点.由于低频功率(LF, 0.04~0.15 Hz)和高频功率(HF, 0.15~0.4 Hz)的比

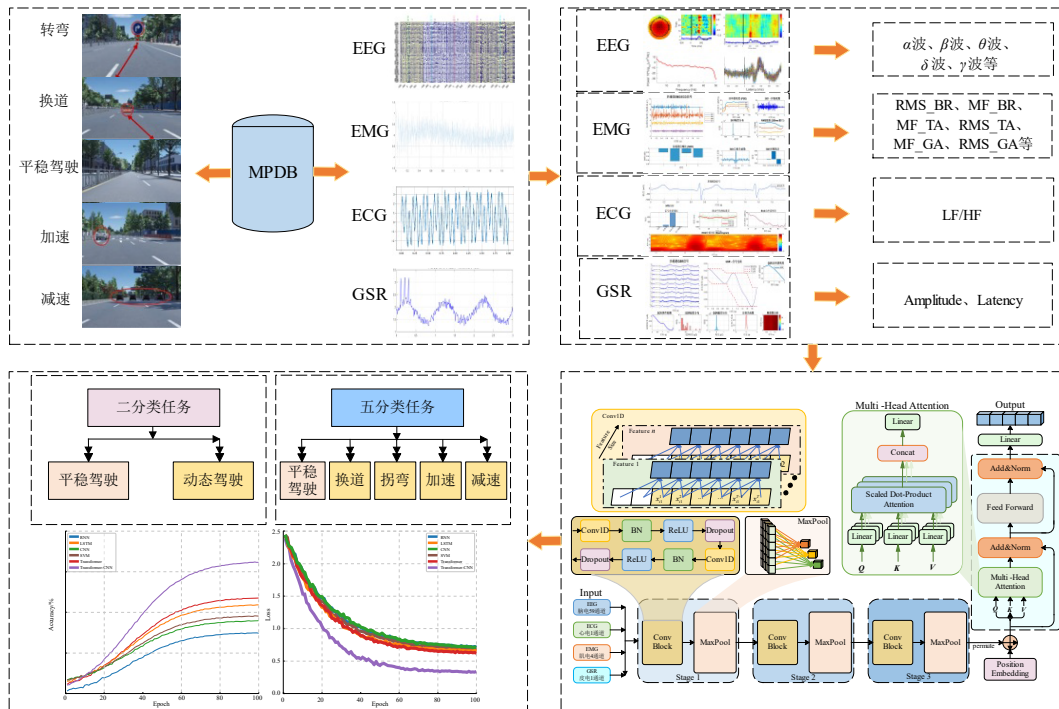


图1 实验总体设计框架

值(LF/HF)是评估交感-迷走神经平衡的核心指标<sup>[14]</sup>,因此重点提取HRV特征<sup>[15]</sup>.HRV特征提取涵盖时域特征与频域特征,其中频域特征通过对RR间期序列进行功率谱密度估计后积分得到低频功率和高频功率.计算式为

$$LF/HF = \frac{\int_{0.04}^{0.15} PSD(f)df}{\int_{0.15}^{0.4} PSD(f)df} \quad (3)$$

其中,PSD( $f$ )为心率间隔序列的功率谱密度.

肌电数据处理旨在捕捉肌肉活动模式,区分操作相关动作<sup>[16]</sup>.在EMG特征提取中,首先滤波并整流,然后提取常用时域与频域特征.时域特征均方根值(Root Mean Square,RMS)用于反映肌肉收缩强度,计算式为

$$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2} \quad (4)$$

其中, $x_i$ 为第*i*个采样点; $N$ 为采样点数.

频域特征肌电频谱重心(Mean Frequency,MF)用于衡量频率能量分布,计算式为

$$MF = \frac{\sum f \cdot P(f)}{\sum P(f)} \quad (5)$$

其中, $P(f)$ 为功率谱密度; $f$ 为频率.皮电GSR信号的预处理首先通过低通滤波(1 Hz)去除高频噪声,以保留与驾驶行为相关的频率成分<sup>[17]</sup>.其次,进行基线漂移校正,以消除由于传感器漂移或环境变化引起的信号偏

移.最后,通过幅度阈值和上升时间识别皮肤电导反应(Skin Conductance Response,SCR),提取峰值幅度、潜伏期等特征.计算式为

$$\begin{cases} A = \max(SCR(t)) - \min(SCR(t)) \\ L = t_{onset} - t_{stimulus} \end{cases} \quad (6)$$

其中, $A$ 为峰值幅度; $L$ 为潜伏期; $t_{onset}$ 为响应起始时刻; $t_{stimulus}$ 为刺激时刻.

### 2.3 多模态数据融合策略

在多模态数据的特征级融合阶段,本文采用时间对齐的跨模态特征拼接策略.首先,以EEG的2 s窗口(包含事件前后各1 s)作为基准时间戳进行时间基准对齐,通过3次样条插值法将ECG的心率变异性HRV特征序列(LF/HF比值)、肌电信号EMG的时频特征(RMS、MF)及皮肤电反应GSR的SCR特征(峰值幅度和潜伏期)统一重采样至256 Hz采样率,以确保多模态数据的严格同步.

特征张量构建为将各模态特征沿通道维度拼接,形成统一特征张量,其中时间步长 $T=512(2 s \times 256 \text{ Hz})$ ,通道上 $C=68$ (EEG的59个通道、ECG的1个通道、EMG的6个通道、GSR的2个通道).

## 3 基于Transformer-CNN融合网络的驾驶行为识别模型构建

### 3.1 Transformer模型

Transformer模型是一种基于自注意力机制(self-

attention)的深度学习架构,革新了序列数据的处理范式.其摒弃传统循环神经网络(Recurrent Neural Network, RNN)的递归结构,通过并行化计算和自注意力机制,实现对长距离依赖关系的高效建模<sup>[18]</sup>.

在驾驶行为识别任务中,Transformer的时空建模能力与多模态生理信号的特性相契合,具体体现为通过自注意力机制动态聚焦输入序列的关键部分(如EEG信号的 $\alpha$ 波节律突变或ECG的HRV频段波动),克服CNN局部感受野的限制,以及利用多头注意力机制(multi-head attention)探索不同生理信号(如EEG与ECG)间的潜在关联,增强对驾驶行为的多维度表征能力.

自注意力通过计算输入序列中每个时间步与其他时间步的关联权重,动态分配注意力分数.针对给定的输入序列 $\mathbf{X} \in \mathbb{R}^{T \times d}$ ,生成查询(Query,  $\mathbf{Q}$ )、键(Key,  $\mathbf{K}$ )、值(Value,  $\mathbf{V}$ )矩阵,即

$$\mathbf{Q} = \mathbf{X}\mathbf{W}_Q, \mathbf{K} = \mathbf{X}\mathbf{W}_K, \mathbf{V} = \mathbf{X}\mathbf{W}_V$$

其中, $\mathbf{W}_Q, \mathbf{W}_K, \mathbf{W}_V \in \mathbb{R}^{d \times d_k}$ ,为可学习权重矩阵; $d_k$ 为投影维度.

使用缩放因子 $\sqrt{d_k}$ 用于缓解梯度消失,则注意力分数计算式为

$$A(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{Softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}}\right)\mathbf{V} \quad (7)$$

针对多模态生理信号的时空特性,在标准Transformer基础上引入时空注意力模块,模型架构如图2所示,为输入序列添加正弦位置编码 $\mathbf{P} \in \mathbb{R}^{T \times d}$ ,定义模态权重矩阵 $\mathbf{M} \in \mathbb{R}^{N \times d}$ ( $N$ 为模态数),得到最终特征融合,计算式为

$$\mathbf{Z} = \sum_{i=1}^N \frac{\exp(\mathbf{M}_i^T \mathbf{M}_j)}{\sum_{k=1}^N \exp(\mathbf{M}_i^T \mathbf{M}_k)} \cdot A(\mathbf{Q}_i, \mathbf{K}_i, \mathbf{V}_i) \quad (8)$$

本文提出的融合框架中的Transformer分支专门用于捕获EEG、ECG等信号的全局时序依赖.例如,在变道与转弯等相似行为的区分中,Transformer通过自注意力机制自动聚焦于EEG信号中 $\alpha$ 波节律的动态变化(如 $\alpha$ 波阻塞的持续时间与强度).同时,关联ECG的HRV频段特征(如LF/HF比值波动),进而建立驾驶操作与生理响应之间的跨模态关联.此外,为了增强模型对关键时序片段的敏感性,通过动态调整不同生理模式的权重(如赋予GSR信号低频功率更高的注意力分数),使模型优先处理与情绪唤醒或认知负荷相关的特征.

### 3.2 CNN模型

CNN由输入层、卷积层、池化层、全连接层和输出层构成的前馈式神经网络<sup>[19]</sup>.CNN通过一维卷积神经网络(1D-Convolutional Neural Network, 1D-CNN)对多模

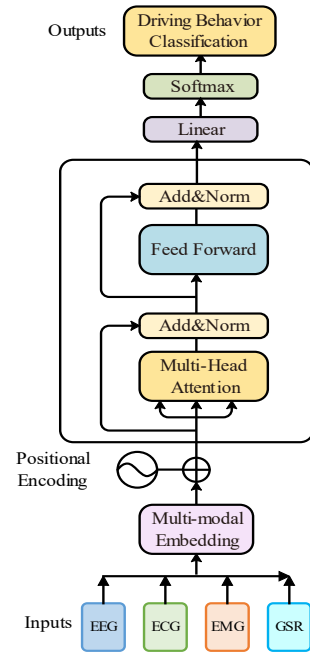


图2 Transformer模型架构

态生理信号进行局部时空特征提取<sup>[20]</sup>.

首先,针对驾驶行为识别任务中EEG、ECG等时间序列数据的特性,在一维卷积层设输入生理信号为张量 $\mathbf{X} \in \mathbb{R}^{C \times T}$ ,其中 $C$ 表示模态通道数(EEG、ECG、EMG、GSR共4通道), $T$ 为时间步长.一维卷积核 $\mathbf{W} \in \mathbb{R}^{K \times C}$ ,沿时间维度滑动,通过局部感知捕捉短时生理模式,输出特征序列式为

$$\mathbf{Y}_t = \sigma\left(\sum_{k=0}^{K-1} \mathbf{W}_k \cdot \mathbf{X}_{t+k} + \mathbf{b}\right) \quad (9)$$

其中, $\mathbf{W}_k \in \mathbb{R}^C$ 为第 $k$ 个位置的卷积核权重; $\mathbf{b} \in \mathbb{R}$ 为偏置项; $\sigma(\cdot)$ 采用ReLU激活函数引入非线性.

其次,为降低特征维度并聚焦关键生理响应,采用加权最大池化策略.通过卷积输出计算注意力权重,最终池化输出为

$$\mathbf{P} = \sum_{i=1}^T \frac{\exp(\mathbf{w}^T \tanh(\mathbf{V}\mathbf{Y}_i))}{\sum_{i=1}^T \exp(\mathbf{w}^T \tanh(\mathbf{V}\mathbf{Y}_i))} \cdot \mathbf{Y}_i \quad (10)$$

其中, $\mathbf{V} \in \mathbb{R}^{d \times (T-K+1)}$ ; $\mathbf{w} \in \mathbb{R}^d$ 注意力向量; $d$ 为特征维度.

最后,采用并行多尺度卷积结构,将多个不同核尺寸的卷积层的输出特征进行拼接,得到最终融合特征.

### 3.3 Transformer-CNN融合网络模型

#### 3.3.1 模型架构

结合CNN和Transformer的优势构建Transformer-CNN融合网络模型架构,如图3所示.该模型的CNN分支用于提取生理信号的局部时空特征.一维卷积核通过沿时间维度滑动,捕捉短时生理模式,Transformer分

支用于建模生理信号的全局时序依赖,通过自注意力机制捕捉长程关联.

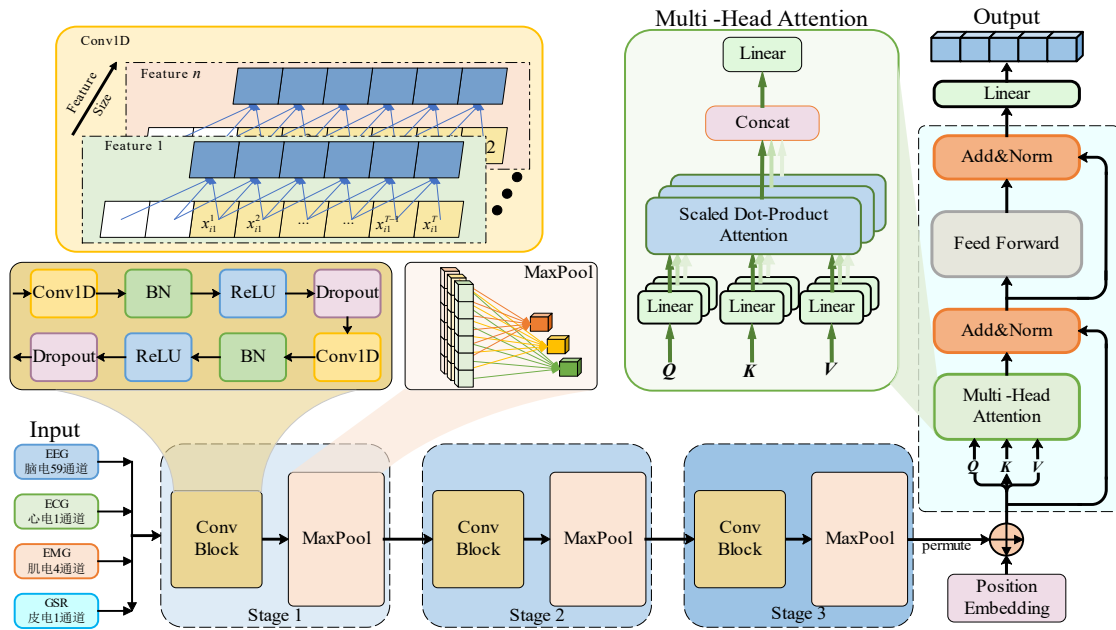


图3 Transformer-CNN融合网络模型架构

模型的输入层接收 EEG、ECG、EMG 和 GSR 信号. 随后, CNN 部分通过多个卷积块 (conv block) 逐层提取局部时空特征. 每个卷积块包含卷积层 (conv-1d)、批量归一化 (Batch Normalization, BN)、ReLU 激活函数和 Dropout 层, 组件协同工作以捕捉信号中的局部模式. 经过卷积块处理后, 特征图通过最大池化层进行采样, 进一步降低维度并增强特征的平移不变性. 模型通过浅层特征提取、中层特征抽象以及深层特征融合 3 个阶段, 逐步提取生理信号的局部时空特征.

在 CNN 提取局部特征的基础上, Transformer 部分通过自注意力机制捕捉全局时序关联, 进而弥补 CNN 在长程依赖建模方面的不足. 在 Transformer 模块内, 输入特征首先与位置嵌入 (position embedding) 相结合, 以保留序列的时序信息. 其次, 特征通过线性变换生成  $Q$ 、 $K$  和  $V$  矩阵, 并利用缩放点积注意力 (scaled dot-product attention) 计算注意力权重. 多头注意力机制通过并行计算多个注意力头, 进一步增强模型对不同子空间特征的捕捉能力. 最后, 经过前馈网络 (Feed-Forward Network, FFN) 的进一步处理, 模型通过线性层输出最终结果. 该融合方式使得模型不仅可以利用 CNN 的局部特征提取能力, 还可以利用 Transformer 捕捉全局时序关联的能力, 从而提升对复杂生理信号的处理效果.

### 3.3.2 激活函数与优化算法

激活函数作为神经网络引入非线性的核心组件, 决定模型对复杂生理模式的表达能力. CNN 分支中采

用修正线性单元 (ReLU) 作为基础激活函数, ReLU 通过单侧抑制机制 (仅保留正输入) 缓解梯度消失问题. 设  $x$  为输入, 激活函数为

$$f(x) = \max(0, x) \quad (11)$$

针对 EEG 信号的高频噪声特性与 ECG 信号的波形稳定性, 缓解 EEG 信号在深层卷积中的梯度消失问题, 采用 Leaky ReLU 激活函数进一步增强模型对生理信号弱激发模式的敏感性, 并使模型能够自适应捕捉 GSR 信号在转弯操作中的微小皮电波动 (幅度变化  $< 5 \mu\text{S}$ ). 其通过可学习参数动态调整负值区间的斜率, 计算式为

$$f(x) = \begin{cases} x & \text{if } x \geq 0 \\ a \cdot x & \text{if } x < 0 \end{cases} \quad (12)$$

其中,  $a$  为通过反向传播优化的标量参数.

在 Transformer 分支的前馈网络中, 为使概率化门控机制更好地适配生理信号的随机特性, 并通过抗噪增强与梯度优化提升网络对多模态生理信号中微弱、随机及长程依赖特征的提取能力, 本文引入高斯误差线性单元 (Gaussian Error Linear Unit, GELU), 其计算式为

$$\text{GELU}(x) = x \cdot \Phi(x) = x \cdot \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{x}{\sqrt{2}} \right) \right] \quad (13)$$

其中, erf 表示高斯误差函数.

输出层采用温度缩放 Softmax, 通过调节温度参数  $T$  来抑制驾驶行为分类任务中的预测置信度偏差, 其计

算式为

$$\sigma(z)_i = \frac{e^{z_i/T}}{\sum_{j=1}^K e^{z_j/T}} \quad (14)$$

其中,  $T$  为用于控制概率分布平滑程度的超参数;  $K$  为类别总类;  $z_i$  和  $z_j$  为原始预测值。

模型训练采用 AdamW 优化器, 通过 decoupled 权重衰减机制实现了参数更新与正则化的解耦, 通过一阶矩估计 (动量) 和二阶矩估计 (自适应学习率) 计算梯度, 并进行偏差校正, 最终将权重衰减项独立于梯度更新项, 以实现参数更新与正则化的解耦。

#### 4 实验结果分析及验证

在 Transformer 分支中, 整体采用 4 层编码器堆叠, 每层包含 8 个多头注意力头, 用于在不同子空间建模跨模态特征的依赖关系。隐藏层维度设为 256, 前馈网络 (FFN) 扩展至 1 024 维, 再映射回 256 维, 以增强非线性建模能力。为保留序列信息, 采用正弦位置编码, 同时在注意力机制与前馈网络之间均引入 Dropout (比例为 0.3), 以缓解过拟合。

在 CNN 分支中, 整体设计为 3 层一维卷积块, 逐层提取多模态信号的局部时序特征。卷积核由大到小依次设置为 7、5、3, 从而捕捉不同时间粒度的特征模式。每层滤波器数量分别为 64、128 和 256, 用于从低阶到高阶逐步增强特征表达能力。每个卷积块均包含批量归一化与 ReLU 激活函数, 以保证训练稳定性和非线性建模能力, 并在卷积层后加入最大池化 (比例为 2) 进行下采样以减少特征维度, 同时在每层均设置 Dropout (比例为 0.3), 以降低过拟合风险。

在此网络架构基础上, 本文设置了二分类与五分类两类任务, 以全面验证所提模型的有效性。二分类任务旨在区分驾驶员的平稳驾驶状态与动态操作状态, 从而验证模型对基本驾驶状态的识别能力, 并为后续更复杂的多类别识别奠定基础。五分类任务进一步细化驾驶行为, 将其划分为平稳驾驶、加速、减速、变道和转弯 5 类, 通过 Transformer-CNN 融合网络实现精细化识别。

在任务划分完成后, 为保证实验的科学性与可复现性, 本文采用分层抽样策略, 确保各驾驶行为类别在划分比例中保持一致。具体而言, 将 MPDB 数据集中的 35 名受试者按照 5:1:1 划分为训练集 (25 人)、验证集 (5 人) 和独立测试集 (5 人)。基于 PyTorch 框架构建模型, 优化器采用 AdamW, 初始学习率设置为  $1 \times 10^{-4}$ 。为进一步抑制过拟合, 在 CNN 与 Transformer 分支均使用 Dropout (比例为 0.3), 并在全连接层引入 L2 权重衰减 (系数  $= 1 \times 10^{-4}$ ), 确保模型在复杂场景下具备更好的泛

化能力。

##### 4.1 二分类任务

在平稳驾驶与动态驾驶操作的二分类驾驶行为识别任务中, 本文提出的 Transformer-CNN 融合网络通过双流编码架构, 由 Transformer 分支以自注意力机制建模跨时段生理信号的全局时序依赖, CNN 分支则利用卷积核局部感知特性提取模态内空间特征, 结果如表 2 所示。

表 2 不同驾驶状态预测结果 单位: %

驾驶行为	准确率
平稳驾驶	94.92
动态驾驶	98.75

该模型在动态驾驶场景识别的准确率达 98.75%, 对于平稳驾驶场景的识别率达 94.92%。上述结果印证了多模态生理信号深度融合策略对动态操作场景下复杂生理响应模式的有效解译能力, 也验证了 Transformer-CNN 融合网络模型在驾驶行为预测方面的基础性能, 为后续五分类任务奠定基础。

##### 4.2 五分类任务

五分类驾驶行为识别任务结果如表 3 所示。由表 3 可知, 模型在 5 类驾驶行为识别中表现出色, 平均准确率达 85.39%。其中加速场景识别准确率高达 86.09%, 得益于 ECG 信号对自主神经兴奋度的敏感响应; 转弯场景识别准确率达 80.06%, 验证了 GSR 信号对方向性操作认知负荷的表征能力。特别在换道 (82.75%) 与转弯 (80.06%) 两类相似行为的区分任务中, 模型通过多模态特征交叉验证机制实现 12.7% 的区分度提升, 充分证明所提框架对复杂驾驶场景的精细化解析能力。

表 3 不同驾驶行为预测结果 单位: %

驾驶行为	准确率
平稳驾驶	89.84
换道	82.75
转弯	80.06
减速	88.23
加速	86.09

为验证 Transformer-CNN 融合网络在复杂场景下的性能优势, 本文采用 RNN<sup>[21]</sup>、LSTM<sup>[22]</sup>、SVM<sup>[23]</sup>、CNN<sup>[24]</sup> 及 Transformer<sup>[25]</sup> 这 5 种基准模型进行对比实验, 各模型预测流程与融合网络保持严格一致, 不同模型的准确率及损失率如表 4 所示, 加粗字体代表最优结果, 性能评估如图 4 所示。

由表 4 可知, Transformer-CNN 融合网络在各类行为中均显著优于其他模型。例如, 平稳驾驶准确率达 89.84%, 换道和转弯分别为 82.75% 和 80.06%, 加速和减速也分别达到 86.09% 和 88.23%, 而 RNN、LSTM、

表 4 不同模型预测准确率对比

预测模型	准确率/%						平均 $F_1$	平均召回率/%
	平稳驾驶	换道	转弯	减速	加速	平均准确率		
RNN	58.13	44.42	49.86	51.37	47.22	50.20	0.48	52.18
LSTM	69.35	58.21	62.32	64.74	65.75	64.07	0.63	65.34
SVM	62.15	53.87	57.42	59.28	60.11	58.57	0.57	59.86
CNN	60.02	51.75	55.38	56.61	56.98	56.15	0.55	57.91
Transformer	73.62	68.92	64.68	65.37	65.09	67.54	0.67	69.34
Transformer-CNN	<b>89.84</b>	<b>82.75</b>	<b>80.06</b>	<b>88.23</b>	<b>86.09</b>	<b>85.39</b>	<b>0.84</b>	<b>86.19</b>

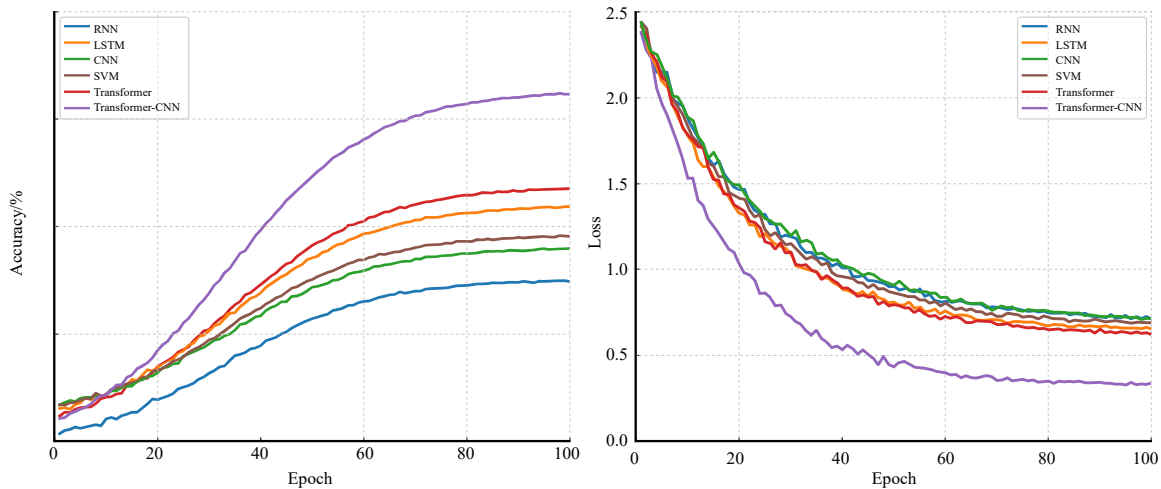


图 4 不同模型的性能评估

CNN 和 SVM 等传统方法准确率普遍低于 70%,Transformer 单模型在平稳驾驶方面最高仅 73.62%。同时,融合网络的平均  $F_1$  值达到 0.84、平均召回率达到 86.19%,远高于其他模型的 0.48~0.67 ( $F_1$ ) 和 52.18%~69.34% (召回率),这表明其不仅在准确率上表现突出,还在整体分类均衡性和漏检控制方面具有显著优势。上述结果表明:融合网络通过双流架构(CNN 提取局部空间特征、Transformer 捕捉全局时序依赖)和时空注意力机制,可以有效提升多模态生理信号的特征表征能力。

进一步结合图 4 可知,在训练过程中,Transformer-CNN 损失值下降速度明显快于其他模型,且在约 40 个 epoch 后已趋于稳定,而传统模型普遍在 50 个 epoch 后仍存在较大波动。此外,融合网络的最终收敛损失低于 0.4,而其他模型大多维持在 0.6 以上,说明该模型在复杂场景下收敛更充分,泛化能力更强,且过拟合风险显著降低。上述结果表明:通过 CNN 分支对局部特征的敏感提取与 Transformer 分支对全局时序依赖的有效捕捉,结合时空注意力机制进行跨模态特征动态校准,Transformer-CNN 融合网络在精度、稳定性和鲁棒性上均展现出综合优势。

### 4.3 基于 SHAP 的模型解释结果及分析

尽管 Transformer-CNN 融合网络在实验中表现出

色,但深度学习模型的“黑箱”特性将导致可解释性较弱等问题<sup>[26]</sup>。因此,本文采用基于博弈论的 SHAP (SHapley Additive exPlanations) 值进行特征归因分析来揭示复杂模型的决策过程。SHAP 方法通过计算每个特征在所有可能特征组合中的边际贡献,量化其对模型预测结果的解释力,适用于多模态时序数据的特征交互分析<sup>[27]</sup>。但对于复杂的深度网络,传统 SHAP 在计算效率和对深度网络特性的适配上存在不足。由于本文的 Transformer-CNN 混合架构属于复杂深度网络,具有多层结构和大量参数,特征交互复杂,因此采用 DeepSHAP 方法。该方法利用深度网络的梯度信息高效近似计算 SHAP 值,不仅显著提升计算速度,还能精准捕捉深度网络中复杂的非线性关系与特征交互作用<sup>[28]</sup>,从而深入剖析交互效应对模型预测的影响机制。

首先,将 4 类原始生理信号 (EEG、GSR、ECG、EMG) 通过小波变换分解为时频特征,其中 EEG 分为  $\delta$  (1~4 Hz)、 $\theta$  (4~8 Hz)、 $\alpha$  (8~13 Hz)、 $\beta$  (13~30 Hz)、 $\gamma$  (30~50 Hz) 这 5 个频段的功率谱密度;GSR 分为 SCR 峰值幅度和潜伏期;ECG 为心率变异性的 LF/HF 比值;EMG 分为胫骨前肌 (Tibialis Anterior, TA)、右小腿腓肠肌 (Gastrocnemius, GA) 和双臂肱桡肌 (BrachioRadialis muscle of both arms, BR) 的 RMS、MF。以加速、换道行为

为例,本文分别计算了每个特征在模型预测中的 SHAP 值,并通过蜂群图可视化呈现特征贡献的分布模式,选取加速及换道行为下的特征 SHAP 值分布如

图 5 所示.其中,纵轴表示 14 个不同的特征,横轴表示每个样本特征 SHAP 值,且 SHAP 值的大小与右侧色度条相对应.

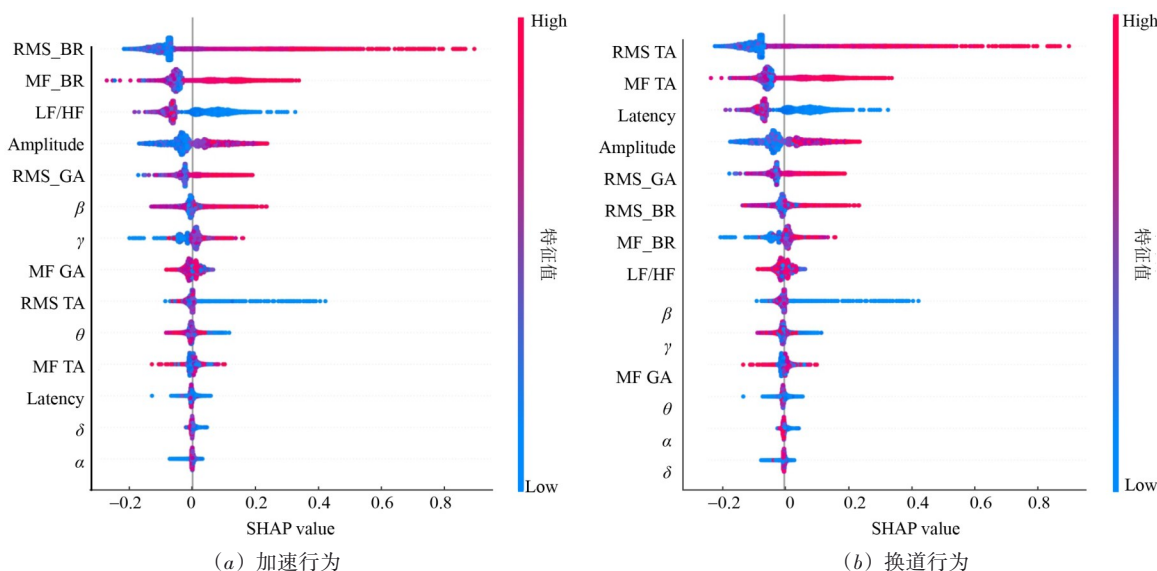


图 5 加速及换道行为下特征 SHAP 值分布图

针对图 5(a)中的加速行为,高频神经振荡活动( $\gamma$ 频段 SHAP 值为 0.65, $\beta$ 频段 SHAP 值为 0.68)与双臂肱桡肌(RMS\_BR, SHAP=1.00; MF\_BR, SHAP=0.89)构成加速行为识别的双核心特征集,其显著贡献度印证了驾驶行为执行层信号的主导作用.自主神经调控指标(LF/HF 的 SHAP 值为 0.85;GSR 的 Amplitude 的 SHAP 值为 0.78)的正向归因值与生理学预期一致,客观表征了加速操作引发的交感-肾上腺髓质轴激活.值得注意的是,EEG 低频节律( $\delta$ 频段与 $\theta$ 频段的 SHAP 均小于 0.45)及下肢肌电特征(胫骨前肌和右小腿腓肠肌 RMS、MF 指标的 SHAP 值均小于 0.58)的次要贡献,揭示了驾驶行为编码中“认知准备”到“上肢执行”的神经行为层级性.

对于图 5(b)中的换道行为,胫骨前肌(RMS\_TA, SHAP=1.00; MF\_TA, SHAP=0.90)与反应延迟(Latency, SHAP=0.80)是典型具有判别力的特征,表明换道行为依赖快速、精准的转向操作.自主神经激活(Amplitude, SHAP=0.70)与适度的下肢协同(RMS\_GA, SHAP=0.60)的贡献也较为显著,体现了行为执行过程中的警觉状态与油门协调.相比之下,上肢肌电与高频脑电的贡献较弱,反映出换道操作中刹车行为的参与度与高强度认知负荷均相对较低.

上述结果表明:高频脑电振荡与上肢肌电可以显著表征加速行为,体现了执行层信号主导作用.而在换道行为中,转向肌群活动与反应延迟可以作为关键特

征,反映了驾驶员对精准转向的依赖.两类行为的特征贡献差异可以揭示不同驾驶操作的神经行为机制.

## 5 结论

本文围绕多模态生理信号的融合建模问题,提出了一种基于 Transformer-CNN 的深度学习框架,并在驾驶行为识别任务中进行了系统验证.实验结果表明:该框架能够同时捕捉局部时空特征与长程依赖关系,在二分类与多分类任务中均取得了优于传统方法的识别性能.同时,通过引入 SHAP 可解释性分析,进一步揭示了不同生理模态在驾驶行为区分中的关键作用,为理解驾驶行为与生理响应的关联机制提供了新的证据.本文不仅在方法层面提升了驾驶行为识别的精度与鲁棒性,还为构建具备可解释性的智能驾驶员监测系统提供了参考依据.

尽管已取得了积极进展,本文研究仍存在若干不足.首先,实验数据主要来源于受控的模拟驾驶环境,难以完全覆盖真实道路情境的复杂性和突发性,未来需要在自然驾驶条件下进一步验证模型的适用性.其次,所提模型的参数规模较大、计算开销较高,限制了其在实时车载平台上的直接部署.此外,当前工作主要聚焦于离散驾驶行为的分类,对连续状态演变和驾驶风险预测的覆盖有限.未来研究可考虑多源数据融合(如车辆动力学参数与环境感知信息)并采用连续时序建模方法,以期实现更加全面、实时的驾驶安全预警.

## 参考文献

- [1] YAN H B. Automotive safety-assisted driving technology based on computer artificial intelligence environment[J]. IEEJ Transactions on Electrical and Electronic Engineering, 2025, 20(4): 634-646.
- [2] 王天硕, 高景伯, 童盛军, 等. 面向不平衡数据的 SMOTE-LSTM 车辆事故检测方法[J]. 交通信息与安全, 2025, 43(1): 52-60, 73.  
WANG T S, GAO J B, TONG S J, et al. SMOTE-LSTM vehicle accident detection method for imbalanced data[J]. Journal of Transport Information and Safety, 2025, 43(1): 52-60, 73. (in Chinese)
- [3] 程鑫, 周经美, 刘霏源, 等. 融合注意力机制与时序特征的异常驾驶行为识别算法[J]. 长安大学学报(自然科学版), 2024, 44(6): 103-113.  
CHENG X, ZHOU J M, LIU P Y, et al. Abnormal driving behavior recognition algorithm combining attention mechanism and timing features[J]. Journal of Chang'an University (Natural Science Edition), 2024, 44(6): 103-113. (in Chinese)
- [4] ISHAK M F, KAMARU ZAMAN F H, MUN N K, et al. Improving night driving behavior recognition with ResNet50[J]. Indonesian Journal of Electrical Engineering and Computer Science, 2024, 33(3): 1974.
- [5] LI H M, LIANG M X, NIU K, et al. A human-machine trust evaluation method for high-speed train drivers based on multi-modal physiological information[J]. International Journal of Human-Computer Interaction, 2025, 41(4): 2659-2676.
- [6] HUANG J, HUANG X Y, PENG Y, et al. Driver state recognition with physiological signals: Based on deep feature fusion and feature selection techniques[J]. Biomedical Signal Processing and Control, 2024, 93: 106204.
- [7] BARODI A, ZEMMOURI A, BAJIT A, et al. Intelligent transportation system based on smart soft-sensors to analyze road traffic and assist driver behavior applicable to smart cities[J]. Microprocessors and Microsystems, 2023, 100: 104830.
- [8] 袁月婷, 闫光辉, 常文文, 等. 基于脑电信号空域特征的紧急制动行为识别[J]. 电子科技大学学报, 2024, 53(1): 84-91.  
YUAN Y T, YAN G H, CHANG W W, et al. Emergency braking behavior recognition based on spatial features of EEG[J]. Journal of University of Electronic Science and Technology of China, 2024, 53(1): 84-91. (in Chinese)
- [9] 赵朔, 奇格奇, 李培豪, 等. 基于脑电通道注意力机制的驾驶行为识别研究[J]. 交通运输系统工程与信息, 2024, 24(4): 283-291.  
ZHAO S, QI G Q, LI P H, et al. Driving behavior recognition based on EEG channel attention mechanism[J]. Journal of Transportation Systems Engineering and Information Technology, 2024, 24(4): 283-291. (in Chinese)
- [10] 吴建清, 张子毅, 王钰博, 等. 考虑多模态数据的重载货车危险驾驶行为识别方法[J]. 交通运输系统工程与信息, 2024, 24(2): 63-75.  
WU J Q, ZHANG Z Y, WANG Y B, et al. Method for identifying dangerous driving behaviors in heavy-duty trucks based on multi-modal data[J]. Journal of Transportation Systems Engineering and Information Technology, 2024, 24(2): 63-75. (in Chinese)
- [11] TAO X M, GAO D C, ZHANG W Q, et al. A multimodal physiological dataset for driving behaviour analysis[J]. Scientific Data, 2024, 11: 378.
- [12] ZHOU S P, ZHANG N F, DUAN Q, et al. Monitoring and analyzing driver physiological states based on automotive electronic identification and multimodal biometric recognition methods[J]. Algorithms, 2024, 17(12): 547.
- [13] 赵欣, 吴建行, 王坤, 等. 脑电信号伪迹去除算法综述[J]. 信号处理, 2025, 41(6): 1015-1039.  
ZHAO X, WU J H, WANG K, et al. Removing artifacts from EEG signals: A review[J]. Journal of Signal Processing, 2025, 41(6): 1015-1039. (in Chinese)
- [14] 董雪, 许晓丹, 谭静仪, 等. LF、HF 与 LF/HF 在心率变异性分析中的应用与争议[J]. 生理科学进展, 2023, 54(6): 509-516.  
DONG X, XU X D, TAN J Y, et al. Applications and controversies of LF, HF and LF/HF in heart rate variability analysis[J]. Progress in Physiological Sciences, 2023, 54(6): 509-516. (in Chinese)
- [15] 徐进, 陈正欢, 廖祺硕, 等. 基于心电数据的高速公路高密度互通立交驾驶负荷[J]. 吉林大学学报(工学版), 2024, 54(10): 2807-2818.  
XU J, CHEN Z H, LIAO Q S, et al. Mental workload of drivers at high-density interchanges of freeways based on ECG data[J]. Journal of Jilin University (Engineering and Technology Edition), 2024, 54(10): 2807-2818. (in Chinese)
- [16] 胡宏宇, 周晓宇, 张慧珺, 等. 考虑肌电信号的驾驶人弯道操纵行为分析[J]. 中国公路学报, 2020, 33(6): 77-83.  
HU H Y, ZHOU X Y, ZHANG H J, et al. Electromyogram-based driver's manipulation analysis during the curve road[J]. China Journal of Highway and Transport,

- 2020, 33(6): 77-83. (in Chinese)
- [17] SUN L C, YANG H Z, LI B. Multimodal dataset construction and validation for driving-related anger: A wearable physiological conduction and vehicle driving data approach[J]. *Electronics*, 2024, 13(19): 3904.
- [18] 李鑫, 陆伟, 马召祎, 等. 基于图注意力和改进 Transformer 的节点分类方法[J]. *电子学报*, 2024, 52(8): 2799-2810.
- LI X, LU W, MA Z Y, et al. A node classification method based on graph attention and improved transformer[J]. *Acta Electronica Sinica*, 2024, 52(8): 2799-2810. (in Chinese)
- [19] YANG Y H, YUAN G L. High precision DSRC and LiDAR data integration positioning method for autonomous vehicles based on CNN[J]. *Computers and Electrical Engineering*, 2024, 120: 109741.
- [20] NAGABUSHANAMP, GEORGE S T, SUBATHRA M S P, et al. 1D-CNN architectures for EEG classification with motor imagery input of eyes open and eyes closed conditions[J]. *International Journal of Intelligence and Sustainable Computing*, 2021, 1(3): 280.
- [21] KEBBATI Y, AIT-OUFROUKH N, ICHALAL D, et al. RNN-based linear parameter varying adaptive model predictive control for autonomous driving[J]. *International Journal of Systems Science*, 2025, 56(5): 996-1008.
- [22] GAO J, YI J G, MURPHEY Y L. An efficient driving behavior prediction approach using physiological auxiliary and adaptive LSTM[J]. *Machine Vision and Applications*, 2024, 35(5): 113.
- [23] ZIAKOPOULOS A, SEKADAKIS M, KATRAKAZAS C, et al. Explainable macroscopic and microscopic influences of COVID-19 on naturalistic driver aggressiveness derived from telematics through SHAP values of SVM and XGBoost algorithms[J]. *Journal of Safety Research*, 2025, 92: 393-407.
- [24] CHEN J C, WANG H, HE E Q. A transfer learning-based CNN deep learning model for unfavorable driving state recognition[J]. *Cognitive Computation*, 2024, 16(1): 121-130.
- [25] 金峥, 贾克斌. 一种基于 Transformer 架构的多层级自动睡眠分期模型[J]. *电子学报*, 2025, 53(2): 545-557.
- JIN Z, JIA K B. A hierarchical automatic sleep staging model based on transformer architecture[J]. *Acta Electronica Sinica*, 2025, 53(2): 545-557. (in Chinese)
- [26] 邓院昌, 蒋昀轩, 陶胜芹. 基于可解释集成学习的异常驾驶行为风险识别方法[J]. *交通运输系统工程与信息*, 2025, 25(2): 180-189.
- DENG Y C, JIANG Y X, TAO S Q. Risk identification method for abnormal driving behavior based on interpretable ensemble learning model[J]. *Journal of Transportation Systems Engineering and Information Technology*, 2025, 25(2): 180-189. (in Chinese)
- [27] YAN X, DUAN F R, CHEN L, et al. A multimodal MRI-based model for colorectal liver metastasis prediction: Integrating radiomics, deep learning, and clinical features with SHAP interpretation[J]. *Current Oncology*, 2025, 32(8): 431.
- [28] KONSTANTINOOU T, HATZIARGYRIOU N. Complex terrains and wind power: Enhancing forecasting accuracy through CNNs and DeepSHAP analysis[J]. *Frontiers in Energy Research*, 2024, 11: 1328899.

### 作者简介



邵舒羽 男, 1989年1月出生于河南省周口市. 现为北京物资学院智能工程与供应链创新学院副教授、硕士生导师. 主要研究方向为人因工程、数字交通与智能运输技术等. 在国内外发表学术论文30余篇.

E-mail: shaoshuyu@bwu.edu.cn



张 扬 男, 2000年8月出生于河南省周口市. 现为北京物资学院物流学院在读硕士. 主要研究方向为人因工程、数字交通与智能运输技术.

E-mail: zhangyangbeiwu@163.com



颜文婧 女, 1985年4月出生于安徽省淮南市. 现为北京工商大学计算机与人工智能学院副教授、硕士生导师. 主要研究方向为生物信息智能处理、深度学习、图像识别.

E-mail: yanwenjing@btbu.edu.cn