

# 面向未来状态预警的多模态疲劳识别与连续预测研究

黄冬梅<sup>1\*</sup>, 蔡洁云<sup>1</sup>, 崔悦佟<sup>2</sup>

<sup>1</sup>西安电子科技大学数学与统计学院, 陕西 西安

<sup>2</sup>西安电子科技大学人工智能学院, 陕西 西安

收稿日期: 2025年11月3日; 录用日期: 2025年12月3日; 发布日期: 2025年12月11日

## 摘要

为解决传统疲劳检测方法主观性强、实时性不足的问题, 本研究提出了一种面向未来状态预警的多模态疲劳识别与连续预测框架。该框架融合了基于光电容积描记(PPG)信号的生理指标与基于面部图像的行为特征, 旨在实现对个体疲劳状态的精准动态监测与预测。研究首先利用集成学习构建了PPG信号三分类(清醒、中度疲劳、高度疲劳)模型, 准确率达到98%。同时, 设计了结合时序卷积(TimeDistributed CNN)与长短期记忆网络(LSTM)的视觉模型, 以捕捉动态面部特征。在此基础上, 本文提出一种基于标签对齐的后期融合策略, 通过逻辑回归模型自适应加权各模态贡献, 有效提升了分类鲁棒性。为实现状态预警, 进一步构建了多变量LSTM联合预测模型, 以预测关键生理指标的未来趋势。实验结果表明, 该多模态框架能够准确识别不同疲劳等级, 并为实现从“事后识别”到“事前干预”的智能疲劳管理提供了有效的方法论支持。

## 关键词

疲劳识别, 多模态融合, 状态预警, 深度学习, 长短期记忆网络(LSTM)

# Research on Multimodal Fatigue Recognition and Continuous Prediction for Future State Early Warning

Dongmei Huang<sup>1\*</sup>, Jieyun Cai<sup>1</sup>, Yuetong Cui<sup>2</sup>

<sup>1</sup>School of Mathematics and Statistics, Xidian University, Xi'an Shaanxi

<sup>2</sup>School of Artificial Intelligence, Xidian University, Xi'an Shaanxi

Received: November 3, 2025; accepted: December 3, 2025; published: December 11, 2025

\*通讯作者。

文章引用: 黄冬梅, 蔡洁云, 崔悦佟. 面向未来状态预警的多模态疲劳识别与连续预测研究[J]. 计算机科学与应用, 2025, 15(12): 187-198. DOI: 10.12677/csa.2025.1512334

## Abstract

To address the subjectivity and real-time limitations of traditional fatigue detection methods, this study proposes a multimodal fatigue recognition and continuous prediction framework for future state early warning. The framework integrates physiological indicators from photoplethysmography (PPG) signals with behavioral features from facial images, aiming to enable accurate dynamic monitoring and prediction of individual fatigue states. Initially, an ensemble learning model was developed to classify PPG signals into three states (awake, moderate fatigue, and severe fatigue), achieving 98% accuracy. Concurrently, a visual model combining a TimeDistributed Convolutional Neural Network (CNN) and a Long Short-Term Memory (LSTM) network was designed to capture dynamic facial features. Building on this, a late fusion strategy based on label alignment is introduced, which uses a logistic regression model to adaptively weight the contributions of each modality, effectively enhancing classification robustness. To enable early warning, a multivariate LSTM model was further constructed to forecast the future trends of key physiological indicators. Experimental results demonstrate that the proposed multimodal framework can accurately identify different levels of fatigue and provides effective methodological support for transitioning from post-hoc recognition to proactive intervention in intelligent fatigue management.

## Keywords

**Fatigue Recognition, Multimodal Fusion, Early Warning, Deep Learning, Long Short-Term Memory (LSTM)**

---

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

随着社会节奏加快与工作强度提升，疲劳问题在交通运输、航空航天、医疗护理等高风险领域日益突出[1]。疲劳不仅削弱个体的注意力与反应能力，还显著降低操作安全与工作效率。研究表明，长期或短时的过度疲劳是导致交通事故[2]、工业设备误操作及飞行失误[3]的重要诱因。因此，如何准确检测个体疲劳状态并预测其发展趋势，已成为保障安全与提升人机系统可靠性的关键课题。

传统疲劳检测方法多依赖主观量表(如 KSS、SSS)，虽简便但实时性差、结果易受主观因素影响。近年来，随着传感技术与人工智能的发展，基于客观生理信号与行为特征的自动化检测成为主流[4]。心率变异性(HRV)、脑电(EEG)、光电容积描记(PPG)等生理信号能够反映自主神经系统在疲劳状态下的动态变化[5]；同时，眨眼频率、眼睑闭合度、头部姿态等视觉特征在计算机视觉模型的支持下亦展现出良好的检测能力。多模态融合方法通过结合生理与视觉信息，显著提升了疲劳识别的准确性与鲁棒性。然而，现有研究多聚焦于“事后识别”，缺乏对疲劳趋势的前瞻性分析。疲劳状态具有明显的累积性与时间依赖性，仅依赖瞬时特征难以及时干预[6]。为此，基于时间序列的预测模型(如 LSTM、GRU、Transformer)被引入，用于捕捉个体生理与行为特征的演化规律[7]，实现未来疲劳水平的动态预测，为个体化休息调度与任务优化提供数据支撑。

尽管已有诸多探索，疲劳检测与预测仍面临个体差异、数据噪声、环境干扰及隐私保护等挑战。未来研究亟需在个性化建模、小样本学习、跨模态融合与可解释性设计方面取得突破，以实现高精度、低

延迟、强泛化的实际应用。

基于此，本文提出一种融合生理信号与视觉信息的智能疲劳检测与预测系统。该系统通过多变量 LSTM 联合建模、残差注意力 GRU 结构及分类 - 预测联动机制[8]，实现对疲劳状态的实时识别与未来趋势预测。研究旨在推动疲劳分析从静态识别迈向动态预测，从单一模态感知走向多模态融合，为智能交通、办公健康管理及人因安全提供可落地的预测式干预方案。

## 2. 多模态融合的疲劳分类模型

### 2.1. 基于 PPG 生物信号的疲劳分类模型

#### 2.1.1. PPG 数据处理

光电容积描记法(Photoplethysmography, PPG)是一种非侵入式生理信号采集方法，通过检测心脏搏动引起的血液容积变化导致的光吸收率差异，获取心率、血氧等信息。本研究利用红外或红光照射皮肤，经光电探测器接收反射信号形成 PPG 曲线。使用的数据集是由 Canaria Technologies 采集[9]，旨在研究视频游戏玩家在长时间游戏过程中的疲劳变化。该数据集的采集协议源于一项历时 24 小时的电脑游戏实验，研究人员在为期 24 小时的游戏周末活动中，对 5 名游戏玩家进行了观察和数据采集。数据采集采用 PPG 技术，通过参与者佩戴的生物识别可穿戴设备实时收集生理信号，设备上的传感器借助红光和红外 LED 灯照射参与者耳垂以捕捉信号。数据集样本规模为 5 名游戏玩家，记录了他们为期 24 小时的完整原始 PPG 时间序列数据，这些数据可用于心率、心率变异性和平滑度的分析，进而研究由电脑游戏引起认知疲劳。

为了消除运动伪差、低频漂移与高频噪声，首先对原始 PPG 信号采用带通滤波器进行预处理。本研究使用 Butterworth 带通滤波器，截止频率设置为 0.5 Hz 至 4 Hz，保留与心率变化相关的频率成分。之后通过局部极值检测算法，提取 PPG 信号中的心跳峰值，从而构建心搏间期序列(RR 间期)。在此基础上，提取每个时间窗口内的多个生理特征。滑动窗口设置如下：每 5 分钟为一个分析窗口，步长为 1 分钟，对每个窗口分别计算以下指标作为模型输入。

**平均心率(MeanHR)**，先通过 PPG 信号检测出每个心跳峰值位置，之后计算相邻峰值之间的时间间隔即心搏间期  $RR_i$ ，之后再根据下面的公式计算每个窗口的平均心率：

$$MeanHR = \frac{60}{\overline{RR}} = \frac{60}{\frac{1}{N} \sum_{i=1}^N RR_i} \quad (1)$$

其中  $RR_i$  为相邻心跳峰之间的时间间隔(单位：秒)， $N$  为该窗口的心跳对数。

**心率变异性标准差(SDNN)**表示正常心搏间期(NN 间期)的标准差，衡量心率整体波动幅度，其具体计算公式为

$$SDNN = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (RR_i - \overline{RR})^2} \quad (2)$$

**心率变异性短时波动(RMSSD)**衡量相邻心搏间期之间差值的均方根，常用于评估短时心率变异性即副交感神经活动水平，其具体计算公式为

$$RMSSD = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N-1} (RR_{i+1} - RR_i)^2} \quad (3)$$

**心率突变比例(PNN50)**表示连续相邻  $RR_i$  间期差值大于 50 ms 的比率，是反映心率短时不稳定性或副交感神经活动强度的指标：

$$PNN50 = \frac{1}{N-1} \sum_{i=1}^{N-1} \delta(|RR_{i+1} - RR_i| > 0.05) \times 100\% \quad (4)$$

其中  $\delta(\cdot)$  为指示函数，当条件满足时为 1，否则取 0。

**呼吸频率(Respiratory Rate)**由于呼吸活动对 PPG 波形存在调制效应，呼吸频率可通过快速傅里叶变换(FFT)在低频段提取：

$$\text{Respiratory Rate} = f_{peak} \times 60 \quad (5)$$

其中  $f_{peak}$  为 PPG 功率谱密度在 0.1 Hz 至 0.4 Hz 之间的最大频率分量。

### 2.1.2. 基于 PPG 生物信号的疲劳分类模型

所使用的数据集是每小时记录一次，包括 5 位受试者在实验期间的 SSS 评分(1~7 级)和注意力反应时变化。其中 SSS (Stanford Sleepiness Scale，斯坦福嗜睡量表) [6] 是一种用于评估主观嗜睡程度的量表，通常分为 7 个等级，每个等级对应一个主观感受描述。常见分类如下表 1：

**Table 1.** Stanford sleepiness scale

**表 1.** 斯坦福嗜睡量表

等级	描述	简化分类
1	Feeling active, vital, alert, or wide awake	清醒
2	Functioning at a high level, but not at peak; able to concentrate	正常
3	Relaxed; awake, but not fully alert	轻度疲劳
4	A little foggy; not at peak; let down	疲劳
5	Fogginess; beginning to lose interest in remaining awake;	高度疲劳
6	Sleepy, woozy, fighting sleep; prefer to lie down	极度疲劳
7	No longer fighting sleep; sleep onset soon;	睡眠边缘

在本文研究中将 SSS 分值进行了三级离散化处理：SSS  $\leq 2$  判为清醒状态，SSS 为 3~4 判为中度疲劳，SSS 为 5~7 判为高度疲劳，最终构建三分类问题。

为提升 PPG 生理信号疲劳识别模型的分类性能与泛化能力，本文构建了一个基于集成学习的三分类预测框架。该框架整合了三种主流树模型：XGBoost、LightGBM 与随机森林，并通过软投票(Soft Voting)策略进行融合。该结构不仅兼顾了不同模型的决策机制，还增强了模型在不同数据分布下的鲁棒性。

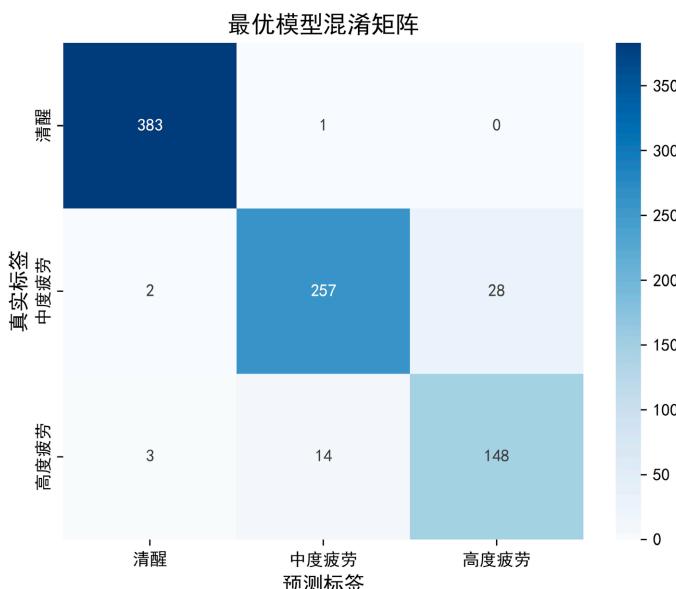
在本研究中，PPG 信号统计特征维度较低但可能存在噪声干扰，同时类别分布存在不均衡。XGBoost 的结构正好能适应此类结构化小样本数据场景，其自动处理缺失值与精细的控制参数设置使其在疲劳等级分类任务中具备出色的建模能力。尤其在多分类情形下，XGBoost 提供的 Softmax 结构天然支持主观疲劳等级(如 SSS)的区分，是不可或缺的基线模型之一。LightGBM 每次选择全局最优的叶子节点进行分裂，这种机制更容易捕捉 PPGRR、SDNN、RMSSD 等疲劳生理信号中的微妙变化。此外，LightGBM 对少量数据场景下的信息增益优化能力较强，能够快速定位“中度疲劳”与“高度疲劳”之间的临界边界，提升模型的辨别能力。在 PPG 生理信号建模中，信号中的微小波动、偶发异常和采集误差是常见挑战。随机森林通过“多树平均”的方式天然对这些问题具备鲁棒性，尤其适合本研究中统计型特征(如 PNN50 和呼吸率)波动性较大的情况。它无需复杂的超参调优，也不易过拟合，是基础模型中稳定性最强的选择之一，在本研究中，它作为 Voting 模型中的“抗方差”模型，与提升树模型形成互补。选择 RF 的另一个重要原因是：它提供了良好的特征重要性评估机制，可以协助我们理解“哪些生理特征对疲劳识别最关键”，对后续的模型解释性分析与特征优化具有现实意义。

为了进一步提高分类性能，本文采用 VotingClassifier 实现模型融合。该方法属于堆叠式集成方法的基本形式，通过将多个模型输出的类别概率进行加权平均，再选择概率最高者作为最终分类结果。VotingClassifier 使用如下融合公式：

$$y = \arg \max_k \left( \frac{1}{N} \sum_{i=1}^N p_k^{(i)} \right) \quad (6)$$

其中  $p_k^{(i)}$  为第  $i$  个模型预测为类别  $k$  的概率， $N = 3$  为模型数量。融合模型综合考虑各分类器在不同样本点上的表现，提升了模型在中度与高度疲劳状态下的识别能力。

本研究基于 PPG 信号构建的疲劳检测模型采用 Voting 集成分类器对三类疲劳状态(清醒、中度疲劳、高度疲劳)进行识别。图 1 所示为模型在测试集上的混淆矩阵，表 2 则展示了各类别的分类。



**Figure 1.** Confusion matrix for the PPG model  
**图 1.** PPG 模型的混淆矩阵

**Table 2.** PPG model results  
**表 2.** PPG 模型结果

类别	准确率	召回率	F1 分数	支持度
清醒	0.98	0.99	0.99	615
中度疲劳	0.97	0.97	0.97	389
高度疲劳	0.97	0.97	0.97	121

从混淆矩阵可见，模型在“清醒”类别上的识别能力最强，共有 615 个样本中正确识别 609 个，仅误判 6 个为“中度疲劳”，未出现将“清醒”样本错误分类为“高度疲劳”的情况。对应的精确率与召回率分别为 0.98 与 0.99，F1 分数达到 0.99，表明模型对清醒状态具有极高的识别能力。

对于“中度疲劳”类别，共有 389 个样本，其中 376 个被正确分类，仅 10 个被误判为“清醒”，3 个被误判为“高度疲劳”。该类别的精确率、召回率和 F1 分数均为 0.97，显示出模型对该状态也具备较为稳健的识别能力。值得注意的是，中度疲劳处于疲劳等级的中间区间，面部表情与生理特征往往与清醒状态存在部分重叠，例如眼睑开合频率、头部姿态或心率波动的变化幅度均较为细微。这导致模型在

特征空间上难以形成明显的区分边界，因而出现少量将“中度疲劳”误判为“清醒”的情况。这种现象提示模型在捕捉轻微疲劳特征时的判别能力仍有待提升，可考虑在后续工作中引入更加细粒度的时序特征或注意力机制，以增强对过渡性疲劳状态的敏感性。

在“高度疲劳”类别中，121个样本中有117个被正确识别，仅4个被误判为“中度疲劳”，无一被误判为“清醒”，体现出模型对高度疲劳状态识别的准确性与安全性。其精确率与召回率均为0.97，F1分数为0.97，进一步验证了模型在检测极端疲劳状态下的可靠性。

综合在整个数据集上，该模型的准确率达到98%，宏平均(macro avg)F1分数为0.98，说明模型在各类别之间具有较好的均衡性。同时，权重平均(weighted avg)精确率、召回率与F1分数均为0.98，进一步印证了模型在面对不同分布的疲劳样本时仍能保持高性能表现，但在中间过渡区域仍存在一定模糊性，这也揭示了疲劳特征连续分布下边界样本的识别难点。

## 2.2. 基于图像识别的疲劳分类模型

### 2.2.1. 图像标签处理

本研究采用的视觉数据源于公开的驾驶员疲劳数据集(DDD) [10]，其原始标签为“清醒”(Non Drowsy)与“疲劳”(Drowsy)二分类。采集协议为：从真实驾驶场景下的驾驶员视频中，通过 VLC 软件提取图像帧，再利用 Viola-Jones 算法提取面部关键区域，数据按疲劳(d\_x.mp4 标注)和非疲劳(n\_x.mp4 标注)两类标签划分，采集过程未刻意引导驾驶员行为；数据由驾驶员在自身车辆中用手机自主录制，能反映真实驾驶中的自然状态；样本规模包含超 41,790 张 RGB 图像(尺寸统一为 227 × 227 像素)，其中疲劳类图像约 22,300 份、非疲劳类图像约 19,400 份，文件总大小 2.32 GB。为提升模型对疲劳程度的辨识粒度，并与生理信号的三分类标签对齐，我们对“疲劳”类别进行了进一步细分[11]。具体流程如下：

#### (1) 图像预处理

在加载图像数据前，统一将图像缩放为 227 × 227 大小，并通过以下公式对每个像素值进行归一化处理：

$$x_{\text{norm}} = \frac{x - \mu}{\sigma} \quad (7)$$

其中， $\mu = [0.485, 0.456, 0.406]$ ， $\sigma = [0.229, 0.224, 0.225]$ ，对应 ImageNet 数据集的通道均值与标准差，以适配预训练 ResNet 模型。

#### (2) 特征提取

采用预训练的 ResNet50 模型作为特征提取器，并将其最后一层输出修改为适应二分类任务。通过去除模型最后的 Softmax 层，提取每张图像在特征空间中的深度表征。之后使用 KMeans 将特征聚类为 2 类：轻度疲劳和严重疲劳。最后得到即“清醒”、“中度疲劳”和“高度疲劳”。需要指出的是，聚类结果可能受特征分布及初始参数的影响，因此所得标签仅反映特征空间中的相对分布特征，而非严格的真实标注，可在后续研究中通过人工验证或更复杂的半监督方法进一步提升标注可靠性。

### 2.2.2. 基于图像识别的疲劳分类模型

为有效捕捉由疲劳引起的动态面部特征变化(如哈欠、点头、持续眨眼等)，本研究构建了一个结合时序卷积网络(TimeDistributed CNN)与长短期记忆网络(LSTM)的深度学习模型[12]。

该模型以连续 5 帧图像序列作为输入。首先，一个三层的 TimeDistributed CNN 模块负责逐帧提取图像的空间特征，有效识别如眼睛开合度、嘴部形态等关键信息。随后，将 CNN 提取的序列化特征输入到 LSTM 模块中，由其对时间维度上的信息进行聚合与建模，捕捉疲劳状态的动态演变过程。最终，通过两层全连接层及 Softmax 激活函数输出三种疲劳状态的预测概率。

在测试集上，该模型表现出优异的性能，总准确率达到了 95.19%。其中，“清醒”、“中度疲劳”和“高度疲劳”三类的 F1-score 分别为 93.2%、98.7% 和 89.2%，表明模型对不同疲劳等级均具备强大的识别能力，尤其对界定模糊的中度疲劳状态识别效果最佳。图 2 为模型在测试集上的混淆矩阵。

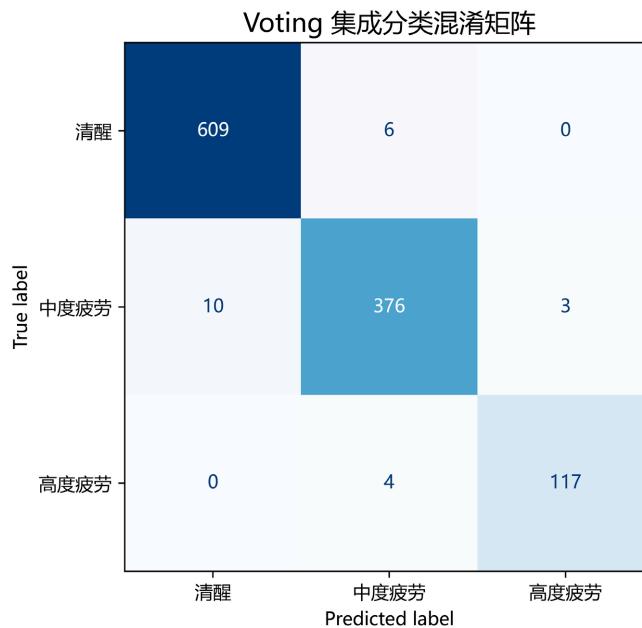


Figure 2. Confusion matrix of the best-performing image model

图 2. 图像模型最优混淆矩阵

### 2.3. 多模态融合模型

为充分结合生理信号在深度疲劳状态下的敏感性与视觉特征在清醒状态下的直观性，我们设计了一种基于标签对齐的后期融合(Late Fusion)机制[13]。该机制的核心思想是利用一个简单的分类器来学习并自适应调整来自不同模态的决策权重。

首先对输入数据进行处理，对于同一疲劳状态的样本，已独立训练好的 PPG 分类模型和图像分类模型分别输出其预测为“清醒”、“中度疲劳”、“高度疲劳”的概率向量  $p = [p_1, p_2, p_3]$  和  $q = [q_1, q_2, q_3]$ ，然后将两个概率向量拼接为一个 6 维的融合特征向量  $X = [p_1, p_2, p_3, q_1, q_2, q_3]$ 。

之后采用逻辑回归(Logistic Regression)作为最终的融合分类器。逻辑回归模型能够学习不同输入特征(即不同模态在不同类别下的预测概率)对最终分类结果的贡献权重。

权重分析表明，该融合机制具备良好的可解释性与自适应能力：在判别“清醒”状态时，模型赋予图像模态更高的权重；而在判别“高度疲劳”状态时，则更依赖于 PPG 生理信号的变化。这种动态加权机制使得模型能够博采众长，在不同疲劳区间内自动信赖更可靠的信息源，从而显著提升了整体分类的准确性与鲁棒性。

## 3. 多变量的疲劳预测模型

为实现对主观疲劳程度的精准预测，本文从生理信号出发，设计了一种多模态分布式建模框架，分别对不同特征变量进行针对性建模处理，最后用于疲劳等级或评分预测。模型的核心主要由三部分组成：多变量生理特征的联合建模，关键突变变量 S2P 的独立序列建模，波动敏感变量 PNN50 的残差注意力优化建模。

### 3.1. 多变量 LSTM 联合预测模型

为实现对疲劳状态的事前干预, 本研究从生理信号出发, 构建了一个多变量长短期记忆网络(LSTM)联合预测模型[14], 其核心目标是提前预测未来一段时间内个体的生理状态变化趋势。

#### 模型设计

该模型专注于预测四项与疲劳状态高度相关的核心生理指标: 平均心率(MeanHR)、心率变异性标准差(SDNN)、均方根差(RMSSD)以及呼吸频率(Respiratory Rate)。模型组成如下:

在输入层中每个时间步输入为一个 4 维向量, 分别对应 MeanHR、SDNN、RMSSD、Respiratory\_Rate 四个生理指标;

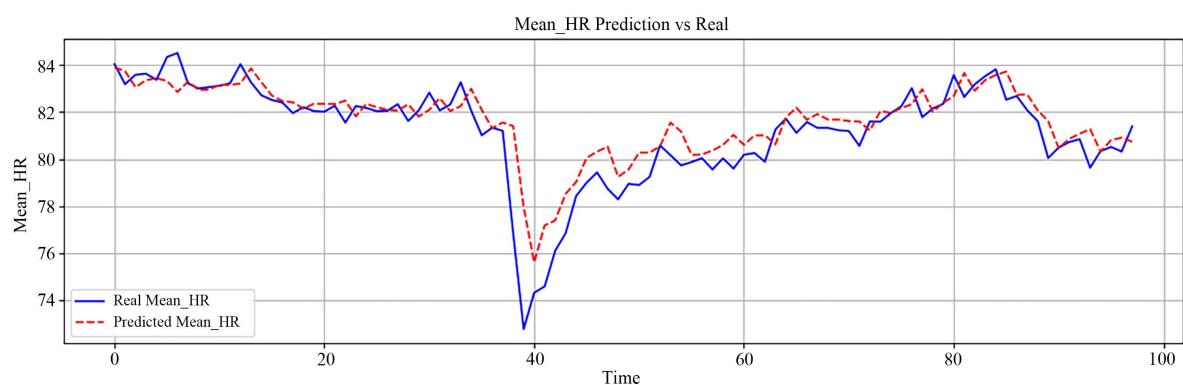
之后我们采用单层 LSTM 结构[15], 隐藏单元数量设置为 64, 负责提取输入序列中的时序依赖与状态变化特征;

在输出层我们将 LSTM 最后一个时间步的隐藏状态映射为对应变量的预测值(共 4 维), 并支持多个预测步长, 其中输出层默认使用线性激活, 适配连续回归任务。该模型结构如公式所示:

$$h_t = \text{LSTM}(x_t, h_{t-1}) \quad (8)$$

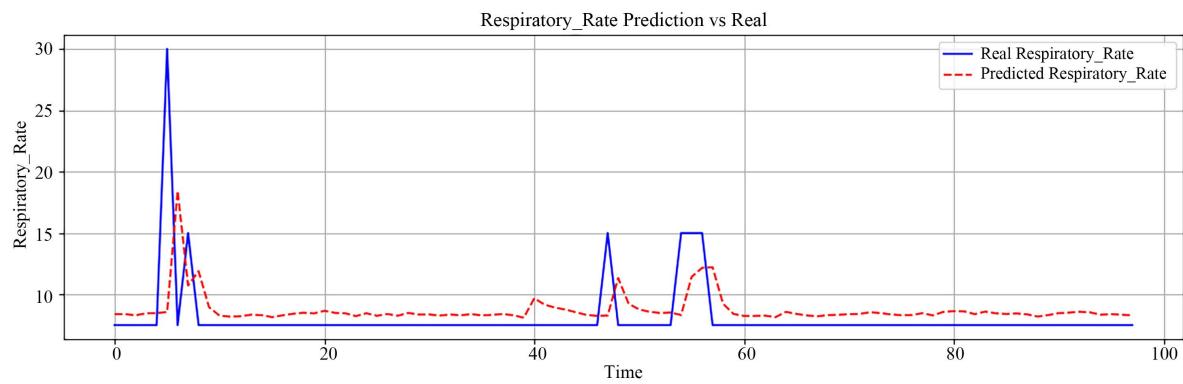
$$y_{t+1} = W_{\text{out}} \cdot h_t + b_{\text{out}} \quad (9)$$

其中,  $x_t \in \mathbb{R}^4$  表示输入时间步,  $h_t \in \mathbb{R}^{64}$  表示 LSTM 隐藏状态,  $W_{\text{out}} \in \mathbb{R}^{4 \times 64}$  是输出映射权重。



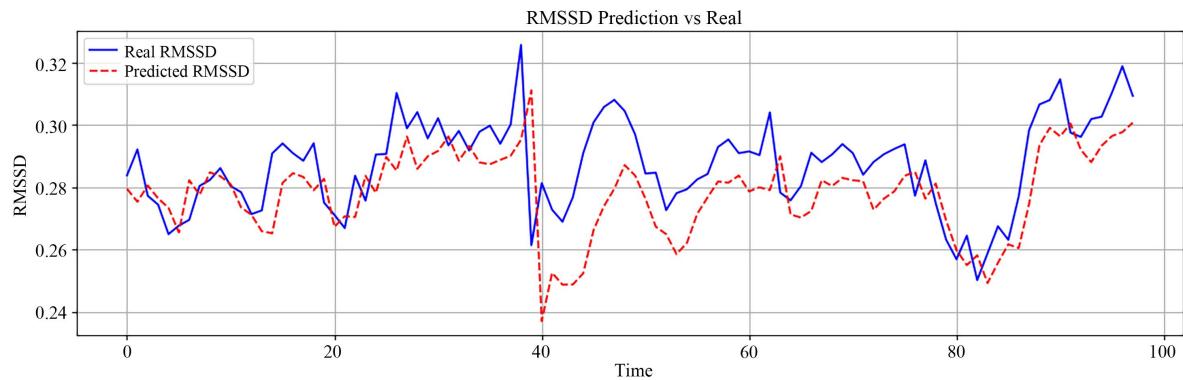
**Figure 3.** Comparison of real and predicted mean heart rate

**图 3.** 平均心率真实与预测值对比图

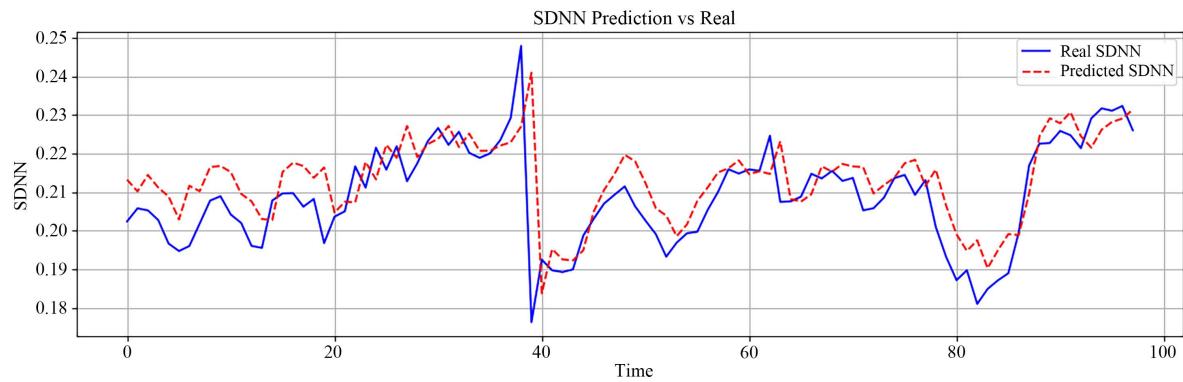


**Figure 4.** Comparison of real and predicted respiratory rate

**图 4.** 呼吸频率真实与预测对比图



**Figure 5.** Comparison of real and predicted RMSSD  
**图 5.** RMSSD 真实与预测对比图



**Figure 6.** Real vs. predicted SDNN  
**图 6.** SDNN 真实与预测对比图

输入数据通过滑动窗口方式构建，并采用 Min-Max 归一化策略以避免变量间的尺度差异影响模型学习。实验结果(见图 3、图 4、图 5、图 6)显示，该模型能有效跟踪并拟合各项平稳生理信号的变化趋势，为判断个体是否即将进入疲劳状态提供了可靠的量化依据，是实现从状态识别到状态预警的关键环节。

### 3.2. S2P 和 PNN50 指标建模

S2P (Sleep-to-Peak Reaction Time) 具有延迟性强、变化突发等特点[16]，常规 LSTM 难以捕捉其关键转变。本文单独设计基于 Encoder-Decoder 结构的 GRU 模型进行建模。该框架的设定主要是在 Encoder 部分通过输入过去连续 64 个时间步的 S2P 值，输出压缩后的隐藏状态向量  $h_{enc}$ ；在 Decoder 部分输入上一个时间步的真实值，输出目标值序列  $\hat{y}_{t+1:t+k}$ ，支持未来 1 个及多个时间点的预测任务。模型的主要公式如下：

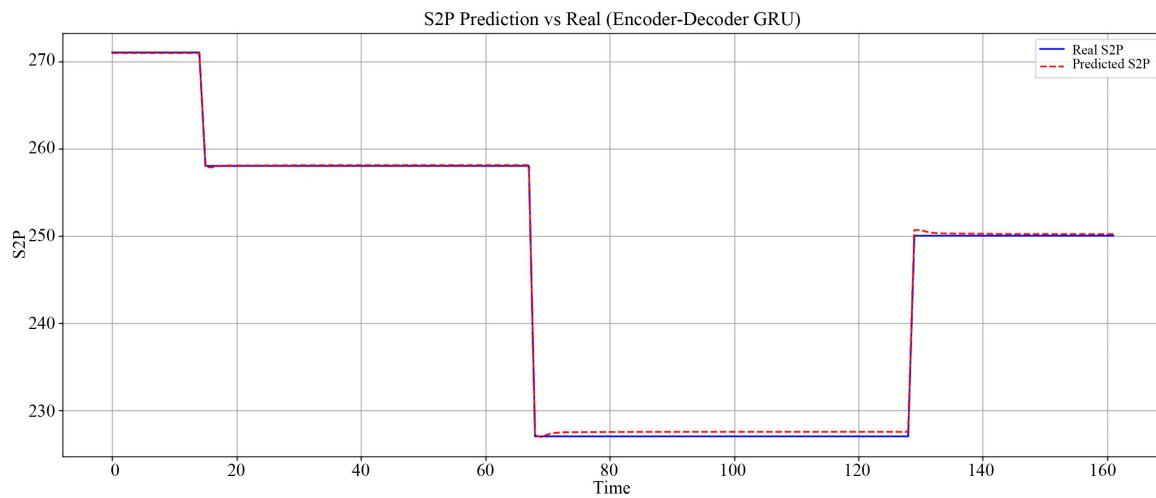
$$h_{enc} = GRU_{encoder}(x_1, \dots, x_T) \quad (10)$$

$$\hat{y}_{t+1} = GRU_{decoder}(h_{enc}, y_t) \quad (11)$$

进一步考虑绘制其真实值与预测值的对比图(见图 7、图 8)，从图中可以清晰看出，该模型能够紧密跟踪 PNN50 序列的高频波动，体现出优异的建模与拟合能力。

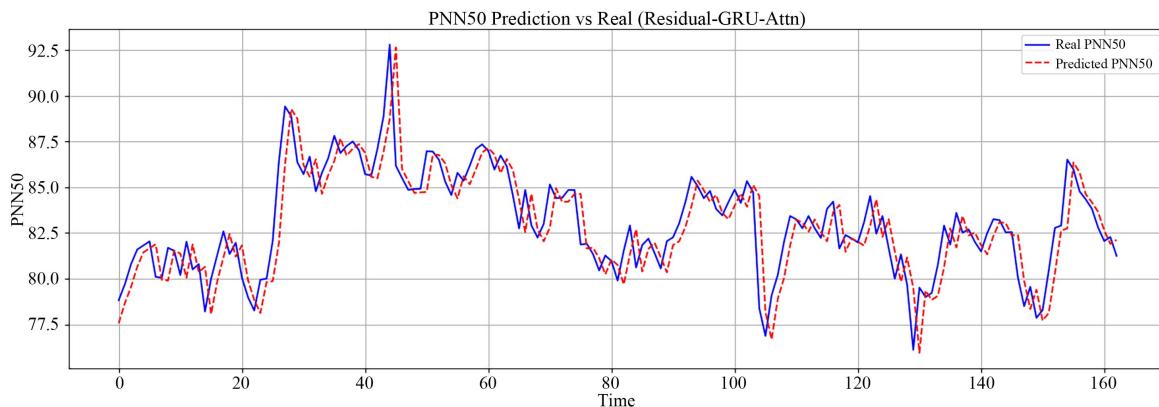
PNN50 是一个反映心率短时波动性的变量[17]，易受高频干扰与局部突变影响。为了更好地建模其

快速起伏和突变行为，本文提出一种基于残差连接与注意力机制结合的 GRU 模型。使用 GRU 层提取原始序列特征，在 Attention 模块，对 GRU 输出进行时间维度加权，聚焦关键片段；残差路径是将输入序列最后一值直接与输出相加，保留局部趋势，最后线性映射得到最终预测值。最后的预测效果可以从图 7 看出红色预测曲线几乎全程贴合蓝色真实值，并且模型能准确预测三个主要突变点( $t \approx 15$ 、 $t \approx 68$ 、 $t \approx 130$ )，且无明显延迟或偏移，在跳变后的平稳期，预测值呈现微小抖动但整体走势正确，这些均表明 Encoder-Decoder GRU 模型的强大拟合能力。



**Figure 7.** Comparison of real and predicted S2P

图 7. S2P 真实与预测对比图



**Figure 8.** Comparison of real and predicted PNN50

图 8. PNN50 真实与预测对比图

为从量化角度更系统、更精确地评估各个预测模型的性能，我们计算了所有生理指标在测试集上的预测值与真实值之间的平均绝对误差(MAE)和均方根误差(RMSE)，见表 3。

表 3 的量化评估证实了所提系列模型的有效性。其中，多变量 LSTM 模型对 SDNN、RMSSD 等平稳指标的预测达到了极高的精度，专用的 Encoder-Decoder GRU 模型也对具有强延迟和突变特性的 S2P 指标取得了 0.2716 的优异 MAE。虽然对于波动性更强的呼吸频率和 PNN50 指标，其预测误差因其固有的高频特性及易受外界干扰而相对偏高，但模型的整体性能依然稳健。总体而言，这些结果共同表明，本研究构建的模型体系能够可靠地预测关键生理指标的未来趋势，为实现从“事后识别”到“事前预

警”的转变提供了坚实的方法论支持。

**Table 3.** Quantitative performance evaluation of the prediction models  
**表 3.** 各预测模型的量化性能评估

预测指标	预测模型	MAE	RMSE
MeanHR	多变量 LSTM	0.7885	1.1419
SDNN	多变量 LSTM	0.0073	0.0103
RMSSD	多变量 LSTM	0.0111	0.0143
RespRate	多变量 LSTM	1.8875	3.0339
PNN50	Residual-GRU-Attn	1.1730	1.5538
S2P	Encoder-Decoder GRU	0.2716	0.3379

在此基础上，本研究结合第二节提出的疲劳分类模型，最终实现了对未来疲劳状态的连续预测：当系统判断个体当前处于非疲劳状态时，便可调用上述各预测模型，输出下一时间点的六个关键生理指标，并将其输入至预先训练好的疲劳分类器中，从而判断个体在未来是否会进入疲劳状态。通过这种迭代机制，系统能够实现对未来多个时间点的连续预警。

#### 4. 结语

本研究的核心价值在于，实现了一个集“模态融合识别”与“时序动态预测”于一体的前瞻性系统。通过创新的模型架构与后期融合策略，系统能够精准捕捉疲劳的生理与行为表征，并基于此预测未来的疲劳状态转折点，为主动干预提供了关键时间窗口。这一成果为实施个性化的劳动强度管理与高危行业的疲劳风险防控奠定了坚实的技术基础。未来的工作将聚焦于技术的实用化与普适化，通过迁移学习与边缘计算等手段，推动该系统从研究走向广泛的实际应用。

#### 参考文献

- [1] 中华人民共和国应急管理部. 疲劳驾驶安全事故典型案例分析[R]. 北京: 应急管理部, 2021.
- [2] 世界卫生组织. 道路交通事故预防报告[R]. 日内瓦: 世界卫生组织, 2023.
- [3] 刘炯坤. 基于多模态信号融合的飞行员疲劳状态监测技术研究[D]: [硕士学位论文]. 成都: 电子科技大学, 2023.
- [4] 董占勋, 孙守迁, 吴群, 徐娟芳. 心率变异性与驾驶疲劳相关性研究[J]. 浙江大学学报(工学版), 2010, 44(1): 46-50.
- [5] 李壮. 基于脑电信号的疲劳状态检测方法研究[D]: [硕士学位论文]. 秦皇岛: 燕山大学, 2020.
- [6] 郭永彩, 李文涛, 高潮. 基于 PERCLOS 的驾驶员疲劳检测算法[J]. 计算机系统应用, 2009, 18(8): 54-57.
- [7] 马召宾. 融合眼部特征及头部姿态的实时疲劳驾驶检测技术研究[D]: [硕士学位论文]. 济南: 山东大学, 2016.
- [8] 杨硕, 丁建清, 王磊, 等. 卷积神经网络在脑疲劳检测中的研究[J]. 信号处理, 2019, 35(4): 704-711.
- [9] Canaria, R.F., et al. (2018) PPG Heart Beat for Cognitive Fatigue Prediction. Kaggle.
- [10] Nasri, I., Karrouchi, M., Snoussi, H., Kassmi, K. and Messaoudi, A. (2021) Detection and Prediction of Driver Drowsiness for the Prevention of Road Accidents Using Deep Neural Networks Techniques. In: *Lecture Notes in Electrical Engineering*, Springer Singapore, 57-64. [https://doi.org/10.1007/978-981-33-6893-4\\_6](https://doi.org/10.1007/978-981-33-6893-4_6)
- [11] 高绮煌, 谢凯, 贺正方, 等. 复杂环境下多模态特征融合的疲劳驾驶检测[J]. 电子测量技术, 2023, 46(6): 106-115.
- [12] 徐敬一, 田瑾, 刘翔, 等. 基于注意力机制改进的疲劳驾驶检测方法[J]. 传感器与微系统, 2024, 43(4): 115-118.
- [13] 方亚南, 汪晓红, 丁一. 基于残差神经网络和表面肌电信号的肌肉疲劳监测研究[J]. 沈阳工程学院学报(自然科学版), 2025, 21(1): 71-77.

- [14] 丁晓彬. 贝叶斯网络分类器的参数学习算法研究与实现[D]: [硕士学位论文]. 南京: 南京航空航天大学, 2020.
- [15] 马鑫鑫, 苏新华, 葛焕敏. 基于混合 LSTM-Transformer 的运动疲劳检测[C]//中国体育科学学会. 第五届全民健身科学大会论文摘要集——专题报告(三). 北京: 北京体育大学, 2024: 303-304.
- [16] 张利伟. 基于视觉的驾驶疲劳实时检测系统研究[D]: [硕士学位论文]. 长沙: 中南大学, 2025.
- [17] 牛寅, 高扬, 李涛, 等. 基于网格搜索优化的 LSTM 多轴疲劳寿命预测方法[J]. 失效分析与预防, 2024, 19(4): 242-249.