

基于ResNet与Attention模块的睡眠呼吸暂停检测分析

吴培宇, 杨其宇

广东工业大学自动化学院, 广东 广州

收稿日期: 2024年12月24日; 录用日期: 2025年1月16日; 发布日期: 2025年1月30日

摘要

本文提出了一种基于ResNet架构的模型,用于睡眠呼吸暂停(SAS)的自动检测。该模型以原始心电图(ECG)信号为输入,利用多层卷积和池化结构实现高效的特征提取,省去了额外的特征工程步骤。为了进一步提升模型的性能,特别引入了注意力机制,使模型能够聚焦于信号中关键特征,有效提高了检测的准确性和鲁棒性。实验结果显示,该模型在准确度(93.84%)、灵敏度(92.00%)、特异性(94.97%)及F1分数(0.9193)等指标上均表现优异,显著优于传统网络模型。与近年来相关研究相比,本模型在SAS检测任务中展现了显著优势,突显了其在实际应用中的潜力。

关键词

SAS, 心电信号, 深度学习, ResNet, Attention

Analysis of Sleep Apnea Detection Based on ResNet and Attention Module

Peiyu Wu, Qiyu Yang

School of Automation, Guangdong University of Technology, Guangzhou Guangdong

Received: Dec. 24th, 2024; accepted: Jan. 16th, 2025; published: Jan. 30th, 2025

Abstract

This paper presents a ResNet-based model for automatic detection of sleep apnea syndrome (SAS) using raw electrocardiogram (ECG) signals. The model utilizes multi-layer convolutional and pooling structures to achieve efficient feature extraction, eliminating the need for additional feature engineering. To further enhance the model's performance, a special attention mechanism was introduced, enabling the model to focus on key features in the signal, effectively improving the detection

accuracy and robustness. Experimental results show that the model achieves excellent performance in terms of accuracy (93.84%), sensitivity (92.00%), specificity (94.97%), and F1 score (0.9193), significantly outperforming traditional network models. Compared with recent related studies, this model demonstrates significant advantages in the task of SAS detection, highlighting its potential for practical applications.

Keywords

SAS, ECG (Electrocardiogram) Signal, Deep Learning, ResNet, Attention

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

睡眠呼吸暂停综合征(SAS) [1]是一种普遍存在的慢性疾病, 全球约有超过 9 亿人受到其影响。研究表明, 男性的患病率大致在 3%至 7%之间, 而女性则在 2%至 5%之间。随着年龄的增长, SAS 的发病率呈上升趋势, 尤其在 60 岁及以上的老年人群中, 其患病率可高达 20%至 40% [2]。该疾病不仅影响睡眠质量, 还可能导致白天嗜睡、注意力不集中、工作效率下降、焦虑情绪及交通事故的风险增加。更严重的是, SAS 与心血管疾病、中风等并发症密切相关, 甚至可能危及生命。

目前, 国际上广泛认可的 SAS 诊断标准是多导睡眠图(Polysomnography, PSG) [3]。这种方法通过记录患者在夜间的脑电波(EEG)、心电图(ECG)、呼吸气流、血氧饱和度等生理参数, 来全面监测睡眠状况。患者需要佩戴多个监测设备进行睡眠, 医生则通过分析这些数据, 识别出呼吸暂停和低通气事件, 并根据计算得出的 AHI (呼吸暂停低通气指数)来确定是否患有 SAS。

尽管 PSG 被认为是最准确的诊断方法, 但其存在一些缺点。首先, 睡眠监测设备价格昂贵, 因此许多医院难以配备, 限制了广泛应用。其次, PSG 数据的解读需要专业医生, 且判读过程繁琐、耗时, 影响了诊断效率。因此, 尽管 SAS 在全球范围内普遍存在, 但在我国的确诊率却低于 10%。此外, 患者在佩戴监测传感器并在陌生环境中进行睡眠时, 往往会因为不适或紧张而导致睡眠质量下降, 从而影响数据的准确性, 不能完全反映其真实状况。

综上所述, SAS 作为一种常见的慢性疾病, 对患者的身心健康造成了显著影响。然而, 传统检测方法存在高成本、操作繁琐等问题, 导致很多患者未能及时诊断和治疗。因此, 研发一种简单、高效、低成本的自动化检测手段, 具有重要的现实意义。通过提高检测效率, 减少医生的工作负担, 可以为更多 SAS 患者提供及时、有效的诊断和治疗, 从而提升公众的整体健康水平。

Guillemineault 等人[4]在 1984 年对 400 名 SAS 患者的研究中发现, SAS 与心率变异性(HRV)相关, 表现为 SAS 时心率下降, 呼吸恢复后心率加快。Zarei 等人[5]提出了一种基于单导联 ECG 信号的自动 OSA 检测方法, 通过熵特征提取, 提高了 SAS 检测的准确性。Liu 等人[6]提出的 CNN-Transformer 模型, 使用原始 ECG 信号, 验证实验表明, 3 分钟检测窗口比 1 分钟窗口更准确, 达 88.2%。Chen 等人[7]则提出了 SE-MSCNN 网络, 利用不同长度 ECG 段的 RR 间隔和 R 波峰值, 克服噪声干扰, 并通过通道注意力机制优化特征融合, 达到了 90.64%的最佳准确度。在现有的基于 ECG 信号的 SAS 检测方法中, 尽管深度学习模型如 CNN-Transformer 和 SE-MSCNN 已取得良好效果, 但仍面临特征提取和模型泛化能力的挑战。为此, 本文提出了一种结合 ResNet 和注意力机制的混合模型, 旨在进一步提升 SAS 检测的准确性与鲁棒性。

2. 数据集

2.1. 基础数据集

本研究使用 Physionet 平台提供的 Apnea-ECG 数据集进行分析, 该数据集包含 70 份整页的单通道 ECG 记录, 数据被分为训练集和测试集, 各包含 35 份记录。每份 ECG 数据的采样率为 100 Hz, 记录时间长度在 6 小时至 8 分钟之间。每个记录不仅包含一段连续的心电图信号, 还附有专家基于同步监测的呼吸信号和其他相关生理数据制定的呼吸暂停注释, 以及由自动化算法生成的 QRS 波标记。

为了便于分析, 每个 ECG 记录被切分为 1 分钟的时间片段。如果某个片段内出现呼吸暂停事件, 则标记为“A”(呼吸暂停), 若无呼吸暂停事件, 则标记为“N”(无呼吸暂停)。在标注过程中, 低通气事件与呼吸暂停事件未做区分, 均统一为同一类事件处理。该数据库的详细信息如表 1 所示。

Table 1. Description of Physionet Apnea-ECG

表 1. Physionet Apnea-ECG 数据集描述

	训练集				测试集			
AHI 类别	AHI < 5	5 ≤ AHI < 15	15 ≤ AHI < 30	AHI ≥ 30	AHI < 5	5 ≤ AHI < 15	15 ≤ AHI < 30	AHI ≥ 30
年龄	35.5 ± 6.3	55 ± 4.2	56.75 ± 4.8	50.8 ± 6.1	32.4 ± 6.09	44.7 ± 6.66	54.9 ± 6.23	47.1 ± 10.4
样本大小	12	2	4	17	11	3	7	14
性别	7 男/5 女	2 男/0 女	4 男/0 女	17 男/0 女	5 男/6 女	3 男/0 女	7 男/0 女	12 男/2 女
AHI	0.1 ± 0.2	9.5 ± 6.4	22.2 ± 2.7	52.6 ± 17.6	0.1 ± 0.12	9.8 ± 4.7	18.9 ± 3	57.1 ± 19.3

2.2. 数据集预处理

心电信号是一种微弱的生物电信号, 典型的心电图由 P 波、QRS 波群、T 波和 U 波组成[3], 在采集过程中容易受到各种干扰影响, 包括体动、电磁干扰等, 这些因素可能导致信号质量下降。因此, 实际采集到的心电信号通常包含多种噪声, 常见的噪声类型包括工频干扰、基线漂移、肌电噪声和运动伪影等。这些噪声会降低信号的信噪比, 影响后续的分析 and 检测结果, 尤其是在睡眠呼吸暂停(SAS)检测中的准确性。

为了提高信号质量并减少噪声干扰, 本研究采取了多种预处理方法。首先, 使用独立成分分析(ICA)去除工频干扰。ICA 的目标是从一组观测信号中恢复出相互独立的源信号。假设有 X 组观测信号, 每一组信号是由若干个独立源信号通过某种线性混合模型得到的: $X = AS$, 其中, X 是观测到的信号矩阵, A 是混合矩阵, S 是源信号矩阵。ICA 的任务是估计混合矩阵 A 和恢复源信号 S, 从而使得每个源信号尽可能独立, 常见的算法包括 FastICA, 它通过迭代方法最大化信号的独立性, 并求得最优解。在使用 FastICA 进行信号分离时, 有几个核心参数需要设置: n_components (定义需要提取的独立成分的数量), max_iter (设置算法的最大迭代次数)、tol (设置算法停止的阈值), 本文分别设置为 2、200、0.0001。

其次, 为了消除基线漂移的影响, 采用了窗口大小为 0.68 秒的中值滤波器。此外, 由于肌电噪声(EMG)通常表现为高频、突变的信号, 其对心电图的干扰较为显著, 本研究使用了 0.5 Hz 到 30 Hz 的带通滤波器来去除高频肌电干扰, 进一步提升信号的清晰度。由于心电信号在不同个体之间存在明显差异, 尤其是在幅值和形态上, 为了避免个体差异对检测结果造成干扰, 本文采用零均值单位方差标准化方法对信号进行标准化处理, 以消除这种个体差异。常用的标准化方法是零均值单位方差标准化, 其公式如下, 其中 x 为原始信号, μ 为信号的均值, σ 为信号的标准差, x_{norm} 为标准化后的信号。

$$x_{norm} = \frac{x - \mu}{\sigma} \quad (1)$$

此外,在心电信号分析中,运动干扰通常会对信号造成显著的影响,导致误诊或降低分析的准确性。因此,采用小波变换作为时频分析方法,通过对信号的局部时间和频率特征进行分析,识别和提取运动干扰的特征信号。具体而言,使用了双正交小波基函数 `bior4.4`, 并进行 5 层小波变换观察不同层次的分解系数与运动干扰之间的关系,发现第 4 层的细节系数能够较为有效地反映运动干扰的起始、过程和结束,因此被选作运动干扰的主要特征信号。小波变换具体的数学公式如下:

$$W_{\Psi}(x, a) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} f(t) \Psi^* \left(\frac{t-x}{a} \right) dt \quad (2)$$

其中 $f(t)$ 是心电信号, $\Psi(t)$ 是小波基函数, a 是控制小波伸缩的尺度参数, x 是控制小波平移的参数, $\Psi^*(t)$ 是小波函数的共轭复数。

为了进一步提取和去除运动干扰,采用了移动窗口积分方法。该方法通过两步操作:首先对小波变换第 4 层的细节系数取绝对值,然后在选定的窗口(长度为 20 个采样点)内进行积分。经过观察,运动干扰部分在积分后的波形中呈现出明显的坡峰特征,这使得通过设定适当的阈值,可以有效地检测出运动干扰的存在。在本方法中,设定积分信号的阈值为 1, 即当平滑后的移动窗口积分值超过 1 时,判定为存在运动干扰。在运动干扰的标记过程中,首先使用峰值为 5 的红色方波进行初步标记。考虑到短时间的静息心电信号对结果的影响较小,对初步标记结果进行了修正,采用峰值为 10 的黄色方波进行调整。当两个运动干扰事件之间的时间间隔小于约 2.5 秒时,即对应于 100 个采样点,认为它们属于同一运动干扰过程,进行合并判断,最终识别出运动干扰的效果如图 1 所示。通过这一方法,可以消除短暂的误检或噪声干扰,确保运动干扰段的连续性和完整性,较为精确地去除心电信号中的运动干扰,提高后续分析的准确性。

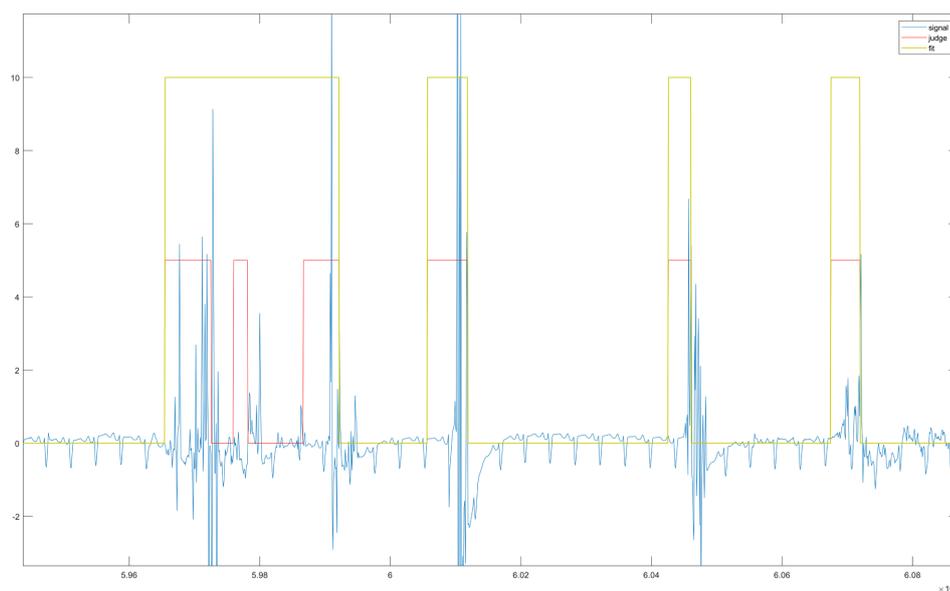


Figure 1. Motion interference recognition illustration

图 1. 运动干扰识别效果图

3. 模型构建

在模型构建中,本实验结合了 ResNet (Residual Networks) 的优势以及 Attention 机制的特性,能够有

效提取 ECG 信号中的深层特征, 捕捉信号中的时序动态变化, 并聚焦于关键时间片段, 从而大幅提升睡眠呼吸暂停事件检测的准确性和效率。这一方法在自动化诊断系统中具有重要的应用价值, 能够为 SAS 的早期发现和及时治疗提供有力支持, 模型整体架构如图 2 所示。

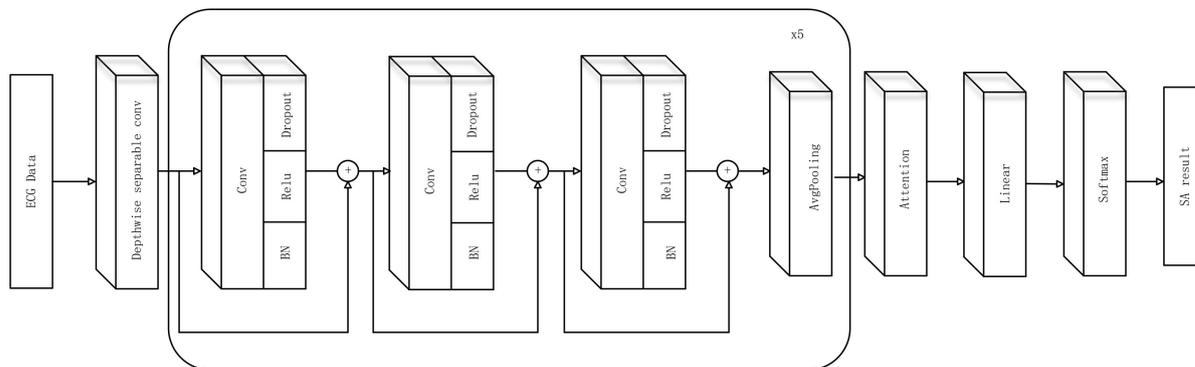


Figure 2. Model architecture
图 2. 模型架构

3.1. ResNet

ResNet (Residual Network)是由何恺明等人提出的一种深度神经网络结构, 主要解决传统深度网络在训练时因层数过多导致的梯度消失和信息丢失等问题。为了解决这些问题, ResNet 引入了残差连接(skip connections), 即将输入信号直接传递到下一层, 使得每一层的学习目标变为输入和输出之间的残差, 而不是直接学习整个映射关系。这种方法有效地减少了训练过程中的困难, 特别是在深度网络的训练上表现出明显的优势。一个典型的残差块可以通过以下公式来表示:

$$y = F(x, \{W\}) + x \quad (3)$$

其中, x 是输入, $F(x, \{W\})$ 是通过卷积层、激活函数等操作后的残差部分, y 是输出。这样的结构使得网络在梯度反向传播时, 残差连接能够帮助梯度更好地传递到更深层, 可以有效避免梯度消失问题。

3.2. Attention 模块

注意力机制(Attention Mechanism)模拟了人类视觉注意力的过程, 使得神经网络能够自动地关注输入中的关键信息。通过计算输入序列中各个元素之间的相对关系, 注意力机制能够根据需要动态调整每个元素的重要性, 从而提高模型对重要信息的敏感度。其基本原理可通过以下公式表示:

$$Attention(Q, K, V) = \text{soft max} \left(\frac{QK^T}{\sqrt{d_k}} \right) V \quad (4)$$

在其中, Q 表示查询(Query)矩阵, K 表示键(Key)矩阵, V 表示值(Value)矩阵, 而 d_k 是键的维度。通过计算查询与键之间的相似度并进行加权, 模型能够有效地聚焦于输入中最相关的信息, 从而优化任务的性能和准确度。

3.3. 损失方法

本实验采用的损失函数为交叉熵损失函数(Cross-Entropy Loss), 它通过衡量模型输出的概率分布与真实标签分布之间的差异来计算损失。在二分类问题中, 交叉熵损失函数的公式为:

$$Loss = -[y \log(p) + (1 - y) \log(1 - p)] \quad (5)$$

在其中, y 表示真实标签(0 或 1), p 是模型对类别 1 的预测概率。交叉熵损失函数通过优化过程减少预测概率与真实标签之间的差距, 从而提高模型的准确性。该损失函数通常与 softmax 激活函数结合使用, softmax 能够将模型输出的原始得分(logits)转换为概率分布, 再利用该概率分布计算损失。

4. 模型训练与评估

4.1. 训练环境

本文的实验训练环境配置如表 2 所示。

Table 2. Experimental environment

表 2. 实验环境

硬件环境		软件环境	
CPU	AMD 5600X	操作系统	Windows10
GPU	NVIDIA RTX4060	Cuda 版本	11.8
内存	32 G	Pytorch 版本	2.0

4.2. 模型训练

在模型训练前, 需要对多个超参数进行设定, 这些超参数对训练过程的效果有着重要的影响。首先, 训练集和验证集的划分比例设定为 7:3, 以确保模型在不同数据集上的泛化能力。在实验中, 为了减少结果的随机性, 我们将所有随机数种子统一设定为 28。模型训练过程中, 选用了 Adam 优化器, 并将初始学习率设置为 0.001, 以便优化训练过程中的参数更新。此外, 批次大小(batch size)设定为 50, 意味着每次训练会处理 50 个样本, 而训练的总轮数(epochs)则设置为 100 轮, 确保模型能够充分学习数据的特征。

4.3. 模型的评估指标

在评估模型的性能时, 我们采用了多种常见的评估指标, 包括混淆矩阵(Confusion Matrix)、准确率(Accuracy)、特异性(Specificity)、灵敏度(Sensitivity)和 F1 分数(F1 Score)。这些指标能够全面地反映模型在分类任务中的表现, 具体介绍如下:

4.3.1. 混淆矩阵

混淆矩阵为模型的分类性能提供了直观的可视化, 包括四个核心元素: 真正例(TP)、假正例(FP)、真负例(TN)和假负例(FN)。理想情况下, 混淆矩阵应尽量保持真正例和真负例的数量较高, 同时假正例和假负例的数量应尽可能低, 以确保模型具有较高的分类准确性。

准确率是衡量模型整体预测效果的常用指标, 它表示模型在所有预测中正确分类的比例。准确率越高, 表明模型在分类任务中表现越好。其计算公式如下:

$$Accuracy = (TP + TN) / (TP + TN + FP + FN) \quad (6)$$

4.3.2. 特异性

特异性主要用于衡量模型识别负例的能力, 即在所有实际负例中, 模型能正确识别的比例。其计算公式为:

$$Specificity = TN / (TN + FP) \quad (7)$$

高特异性意味着模型能够较少地将负例误判为正例, 从而降低了误报的风险。

4.3.3. 灵敏度

灵敏度则用来衡量模型识别正例的能力, 指的是在所有实际正例中, 模型能够正确识别的比例。计算公式为:

$$Sensitivity = TP / (TP + FN) \quad (8)$$

灵敏度值越高, 说明模型能够较好地识别出正例, 避免漏判。

4.3.4. F1 分数

F1 分数综合了精确率(Precision)和召回率(Recall)的表现, 是一个平衡精确度与召回率的指标。其计算公式如下:

$$F_1 = 2 \times TP / (2 \times TP + FP + FN) \quad (9)$$

F1 分数越高, 表明模型在精确度和召回率之间达到了良好的平衡, 既能有效识别正例, 又能减少误分类的负例。

4.4. 模型评估

为了评估模型在睡眠呼吸暂停(SAS)预测任务中的表现, 本研究采用了逐片段评估的方法。在这种评估方式中, 我们会根据每个时间片段的预测结果, 分析模型在识别每个片段中的睡眠呼吸暂停事件的能力。本研究中, 每个时间片段的长度设定为 1 分钟。评估指标不仅包括混淆矩阵, 还涉及灵敏度、准确率、特异性、F1 分数等多项指标, 以全面衡量模型的性能。

根据实验结果显示的混淆矩阵, 如图 3 所示, 在所有的评估片段中, 只有 530 个片段被误分类为睡眠呼吸暂停(SA), 而 520 个片段则被误分类为正常(Non-SA)。大多数片段的分类结果是正确的, 表明模型在分类任务中表现良好, 具备较强的识别能力。

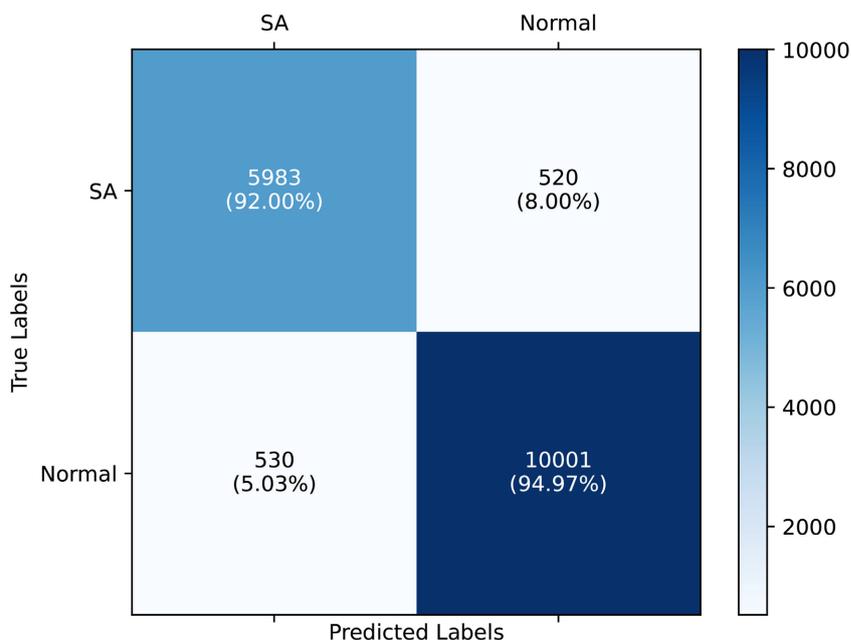


Figure 3. Confusion matrix of per-segment analysis

图 3. 逐片段分析混淆矩阵

图4展示了该模型在各项性能测试中的数值表现。具体来说,模型的准确度达到了93.84%, 敏感度为92.00%, 特异性为94.97%, F1分数为0.9193, AUC值为0.9384。相比较之下, Hu等人[8]所使用的Hybird Transformer方法的准确度为90.5%, 敏感度为86.5%, 特异性为93%, 而Li等人[9]采用的Light Transformer方法的准确度为92.12%, 敏感度为88.4%, 特异性为94.42%。这些结果表明, 所提出的模型在准确性、敏感性和特异性方面均具有显著优势, 能够有效提升心电信号分析的性能。在实际应用中, 高灵敏度意味着模型能够有效地识别真实的病例, 漏诊率较低, 因此在筛查任务中, 灵敏度是至关重要的指标。高特异性则意味着模型很少将健康个体误判为患者, 从而降低了误诊率。特异性同样在疾病的筛查和诊断过程中具有重要意义。

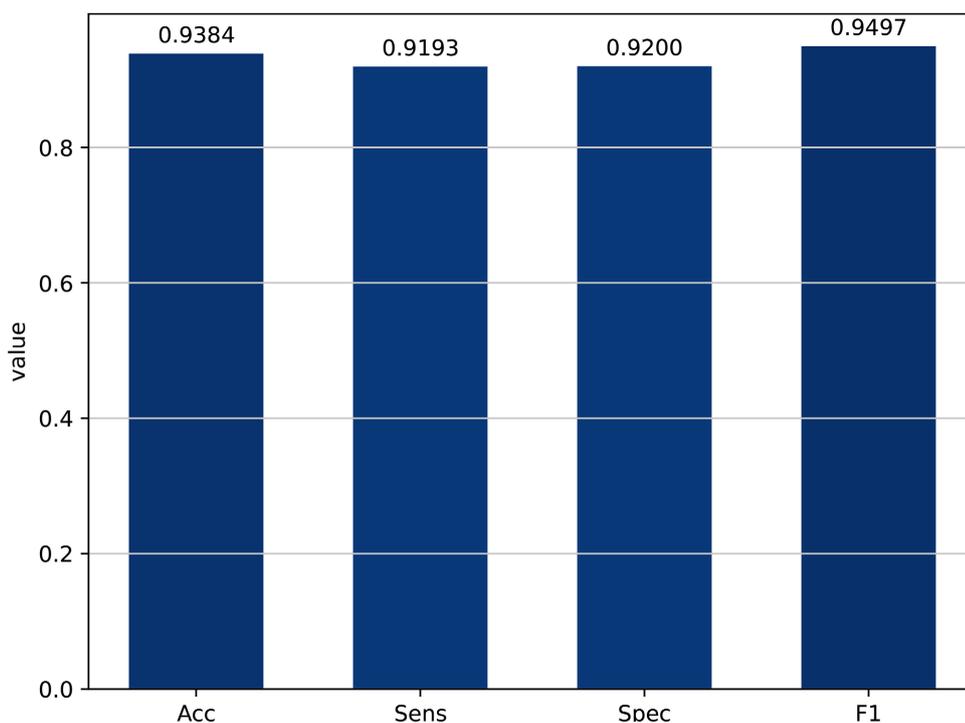


Figure 4. Bar Chart of model performance metrics on a per-segment analysis

图4. 模型逐片段分析各项指标柱状图

5. 结论

本实验所采用的模型基于 ResNet 架构, 作为特征提取的核心部分。通过多层卷积层和池化层, 模型能够从原始心电图(ECG)信号中自动提取重要特征, 无需额外的特征工程步骤, 为后续分析和识别任务奠定了坚实基础。同时, 模型引入了注意力机制, 以增强其对信号中关键特征的识别能力。通过为信号的不同部分分配不同的权重, 注意力机制使得模型能够自动聚焦于对睡眠呼吸暂停(SAS)诊断至关重要的特征, 从而进一步提升了检测的精确度和鲁棒性。

实验结果表明, 该模型在各项指标上表现优异, 准确度为93.84%, 灵敏度为92.00%, 特异性为94.97%, F1分数为0.9193。通过与近年来相关研究的检测方法进行对比, 证明该模型在 SAS 检测中具有较为出色的性能, 展现了其在实际应用中的潜力和优势。

参考文献

- [1] Benjafeld, A.V., Ayas, N.T., Eastwood, P.R., Heinzer, R., Ip, M.S.M., Morrell, M.J., *et al.* (2019) Estimation of the

-
- Global Prevalence and Burden of Obstructive Sleep Apnoea: A Literature-Based Analysis. *The Lancet Respiratory Medicine*, **7**, 687-698. [https://doi.org/10.1016/s2213-2600\(19\)30198-5](https://doi.org/10.1016/s2213-2600(19)30198-5)
- [2] Park, J.G., Ramar, K. and Olson, E.J. (2011) Updates on Definition, Consequences, and Management of Obstructive Sleep Apnea. *Mayo Clinic Proceedings*, **86**, 549-555. <https://doi.org/10.4065/mcp.2010.0810>
- [3] 李江丽. 多导睡眠监测的临床应用新进展[J]. 广东医学, 2014, 35(11): 1789-1791.
- [4] Guilleminault, C., Winkle, R., Connolly, S., Melvin, K. and Tilkian, A. (1984) Cyclical Variation of the Heart Rate in Sleep Apnoea Syndrome: Mechanisms, and Usefulness of 24 h Electrocardiography as a Screening Technique. *The Lancet*, **323**, 126-131. [https://doi.org/10.1016/s0140-6736\(84\)90062-x](https://doi.org/10.1016/s0140-6736(84)90062-x)
- [5] Zarei, A. and Asl, B.M. (2019) Automatic Detection of Obstructive Sleep Apnea Using Wavelet Transform and Entropy-Based Features from Single-Lead ECG Signal. *IEEE Journal of Biomedical and Health Informatics*, **23**, 1011-1021. <https://doi.org/10.1109/jbhi.2018.2842919>
- [6] Liu, H., Cui, S., Zhao, X. and Cong, F. (2023) Detection of Obstructive Sleep Apnea from Single-Channel ECG Signals Using a CNN-Transformer Architecture. *Biomedical Signal Processing and Control*, **82**, Article ID: 104581. <https://doi.org/10.1016/j.bspc.2023.104581>
- [7] Chen, X., Chen, Y., Ma, W., Fan, X. and Li, Y. (2021) SE-MSCNN: A Lightweight Multi-Scaled Fusion Network for Sleep Apnea Detection Using Single-Lead ECG Signals. 2021 *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 9-12 December 2021, 1276-1280. <https://doi.org/10.1109/bibm52615.2021.9669358>
- [8] Hu, S., Cai, W., Gao, T. and Wang, M. (2022) A Hybrid Transformer Model for Obstructive Sleep Apnea Detection Based on Self-Attention Mechanism Using Single-Lead ECG. *IEEE Transactions on Instrumentation and Measurement*, **71**, 1-11. <https://doi.org/10.1109/tim.2022.3193169>
- [9] Li, C., Shi, Z., You, Z., Li, N., Zhou, L., Zhang, Z., et al. (2024) Drformer: A Data Rebalancing Loss-Constrained Light Transformer for OSA Detection. *IEEE Transactions on Instrumentation and Measurement*, **73**, 1-17. <https://doi.org/10.1109/tim.2024.3440369>