



## 多模态融合网络的睡眠分期研究

陈玲玲, 毕晓君

引用本文:

陈玲玲, 毕晓君. 多模态融合网络的睡眠分期研究[J]. 智能系统学报, 2022, 17(6): 1194–1200.

CHEN Lingling, BI Xiaojun. Sleep staging model based on multimodal fusion[J]. *CAAI Transactions on Intelligent Systems*, 2022, 17(6): 1194–1200.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202202018>

## 您可能感兴趣的其他文章

### 基于双向LSTM卷积网络与注意力机制的自动睡眠分期模型

Automatic sleep staging model based on the bi-directional LSTM convolutional network and attention mechanism

智能系统学报. 2022, 17(3): 523–530 <https://dx.doi.org/10.11992/tis.202103013>

### 多尺度特征融合网络的视网膜OCT图像分类

Retinal optical coherence tomography image classification based on multiscale feature fusion

智能系统学报. 2022, 17(2): 360–367 <https://dx.doi.org/10.11992/tis.202111024>

### 面向科技学术会议的命名实体识别研究

Research on named entity recognition for scientific and technological conferences

智能系统学报. 2022, 17(1): 50–58 <https://dx.doi.org/10.11992/tis.202107010>

### 基于分割注意力机制残差网络的城市区域客流量预测

Passenger flow prediction in urban areas based on residual networks with split attention mechanism

智能系统学报. 2022, 17(4): 839–848 <https://dx.doi.org/10.11992/tis.202202014>

### 用于关系抽取的注意力图长短时记忆神经网络

Attention graph long short-term memory neural network for relation extraction

智能系统学报. 2021, 16(3): 518–527 <https://dx.doi.org/10.11992/tis.202008036>

DOI: 10.11992/tis.202202018

网络出版地址: <https://kns.cnki.net/kcms/detail/23.1538.tp.20220825.1115.002.html>

# 多模态融合网络的睡眠分期研究

陈玲玲<sup>1</sup>, 毕晓君<sup>2</sup>

(1. 哈尔滨工程大学 信息与通信工程学院, 黑龙江 哈尔滨 150001; 2. 中央民族大学 信息工程学院, 北京 100081)

**摘 要:** 针对睡眠多导图各模态信息在睡眠各阶段存在差异性, 而导致特征利用不充分的问题, 本文提出了一种基于通道注意力机制和多模态门控机制的睡眠分期模型。首先利用残差收缩网络设计各模态特征提取网络用于提取各模态特征, 并在通道维度上进行拼接融合, 利用通道注意力机制进一步对融合特征进行重标定得到睡眠多导图的时不变特征; 之后提出了一种基于自适应门控机制的多模态门控模块, 对各模态特征及时不变特征按照重要程度进行加权融合, 实现特征融合; 最后利用双向长短时记忆网络提取睡眠多导图的时序特征。实验结果表明, 本文提出的睡眠分期模型在欧洲数据格式睡眠数据集 (sleep-European data format, sleep-EDF) 上准确率为 87.6%,  $M_{F1}$  为 82.0%, 取得了目前最好的分期效果。

**关键词:** 深度学习; 睡眠分期; 通道注意力机制; 睡眠多导图; 多模态; 双向长短时记忆网络; 欧洲数据格式睡眠数据集; 残差收缩网络

**中图分类号:** TP18; TH79    **文献标志码:** A    **文章编号:** 1673-4785(2022)06-1194-07

中文引用格式: 陈玲玲, 毕晓君. 多模态融合网络的睡眠分期研究 [J]. 智能系统学报, 2022, 17(6): 1194-1200.

英文引用格式: CHEN Lingling, BI Xiaojun. Sleep staging model based on multimodal fusion[J]. CAAI transactions on intelligent systems, 2022, 17(6): 1194-1200.

## Sleep staging model based on multimodal fusion

CHEN Lingling<sup>1</sup>, BI Xiaojun<sup>2</sup>

(1. College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China; 2. School of Information Engineering, Minzu University of China, Beijing 100081, China)

**Abstract:** To solve the problem of insufficient feature use due to the differences in modal information in sleep polygraphs at different stages, this paper proposes a sleep staging method based on the channel attention mechanism and multimodal gating module. First, a residual shrinkage network is used to design each modal feature extraction network to extract the features of all modules, which are spliced and fused in the channel dimension, and the fused features are further recalibrated using the channel attention mechanism to obtain time-invariant polysomnography features. Then, a multimodal gating module based on an adaptive gating mechanism is proposed, which weighs and fuses all modal and time-invariant features according to their importance, thereby realizing the organic fusion of features. Finally, the time sequence characteristics of polysomnography are obtained using a bidirectional long short-term memory network. The experimental result shows that the accuracy of the sleep staging model proposed in this paper is 87.6% on the Sleep-EDF dataset, with an  $M_{F1}$  of 82.0%, thus achieving the best staging effect to date.

**Keywords:** deep learning; sleep stage; channel attention mechanism; polysomnogram; multimodal; bidirectional long short-term memory; sleep-European data format dataset; residual shrinkage network

睡眠分期是睡眠监测领域的一项基础研究, 通过对睡眠时期进行划分, 可以帮助医学专家更好地认识睡眠结构, 并为失眠、嗜睡等相关睡眠

疾病的诊断与治疗提供依据。AASM 标准<sup>[1]</sup>指出, 睡眠可分为清醒期 (wake, W)、非快速动眼期 (non-rapid eye movements, NREM) 和快速动眼期 (rapid eye movements, REM)。NREM 又被细分为 N1、N2、N3 期<sup>[2]</sup>。而睡眠多导图 (polysomnogram, PSG) 被称为睡眠分期的“黄金标准”, 其包含睡眠

收稿日期: 2022-02-23. 网络出版日期: 2022-08-26.

基金项目: 国家社会科学基金重大项目 (20&ZD279).

通信作者: 毕晓君. E-mail: [bixiaojun@hrbeu.edu.cn](mailto:bixiaojun@hrbeu.edu.cn).

过程中脑电 (electroencephalogram, EEG)、眼电 (electrooculogram, EOG)、肌电 (electromyography, EMG) 等信号的记录。通过综合分析睡眠多导图中的各模态信息, 能够实现睡眠阶段的划分<sup>[1]</sup>, 但耗费了睡眠专家大量时间和精力, 因此睡眠多导图的分期任务成为了热点问题<sup>[3]</sup>。

通过非线性时域、频域分析等操作可实现睡眠 PSG 特征的提取<sup>[4]</sup>, 利用支持向量机等技术进行分类, 但易出现错分, 且模型泛化能力不高<sup>[5]</sup>。近年来深度学习在图像领域取得突破性进展, 结合深度学习利用睡眠 PSG 进行睡眠分期研究逐渐成为研究的热点。一些在图像领域取得显著效果的模型如 VGG<sup>[6]</sup>、ResNet<sup>[7]</sup>、膨胀卷积<sup>[8]</sup>、深度可分离卷积<sup>[9]</sup> 等都被应用到睡眠分期中, 取得了较好的效果。但是, 由于输入模态的多样性与差异性, 如何实现多模态信息的有机融合成为亟待解决的问题。目前睡眠分期领域最常用的模态融合方法分别是数据<sup>[10-14]</sup>和特征<sup>[15]</sup>两个角度进行融合。Phan 等<sup>[10]</sup>从数据融合的角度, 将 PSG 经短时傅里叶变换为含 EEG、EOG 信息的多通道时频图, 然后利用 2 维卷积神经网络提取时不变特征。除此之外, 小波变化<sup>[11]</sup>、希尔伯特变换<sup>[12]</sup>等方法也被应用到了睡眠分期中。这种方法虽然综合考虑了多模态的信息, 但是转换为时频图时会丢失关键的时域信息<sup>[13]</sup>。随后, 从特征角度进行融合的方法成为主流, 其中代表性的有 Duan 等<sup>[15]</sup>利用卷积神经网络提取各个模态的时不变特征, 利用多模态联合网络实现睡眠分期。

虽然从特征角度进行模态融合的实验效果优

于从数据角度融合, 但是没有充分考虑各模态在不同时期的差异性, 而主要依据了 EEG 特征, 例如在 REM 期会出现明显的眼球运动, EOG 信号更为重要。

本文针对睡眠 PSG 中各模态特征存在差异性导致的特征融合不充分的问题, 主要进行了如下工作:

1) 设计了一种基于通道注意力机制的睡眠 PSG 的不变提取特征网络, 根据各通道特征的重要性对模态特征进行重标定, 得到睡眠多导图的时不变特征;

2) 设计了一种基于自适应门控机制的多模态门控模块, 对各模态特征及时不变特征按照重要性程度进行加权融合;

3) 利用睡眠分期最常用的 Sleep-EDF 数据集的多导图数据设计 20 折交叉验证实验, 通过与对比算法比较, 证明了算法具有一定的有效性和先进性。

## 1 模型框架

基于通道注意力机制和多模态门控机制的睡眠分期模型如图 1 所示, 首先利用各模态特征提取网络分别提取不同模态的高维特征, 并利用通道注意力机制对在通道维度上拼接的各模态高维特征进行重标定, 得到睡眠多导图的时不变特征, 然后设计多模态门控模块实现不同模态特征的融合, 得到能充分表征各模态信息的融合特征, 最后, 将其输入到时序特征提取网络中, 提取信号的时序特征。下面将详细介绍本文所设计的模型框架。

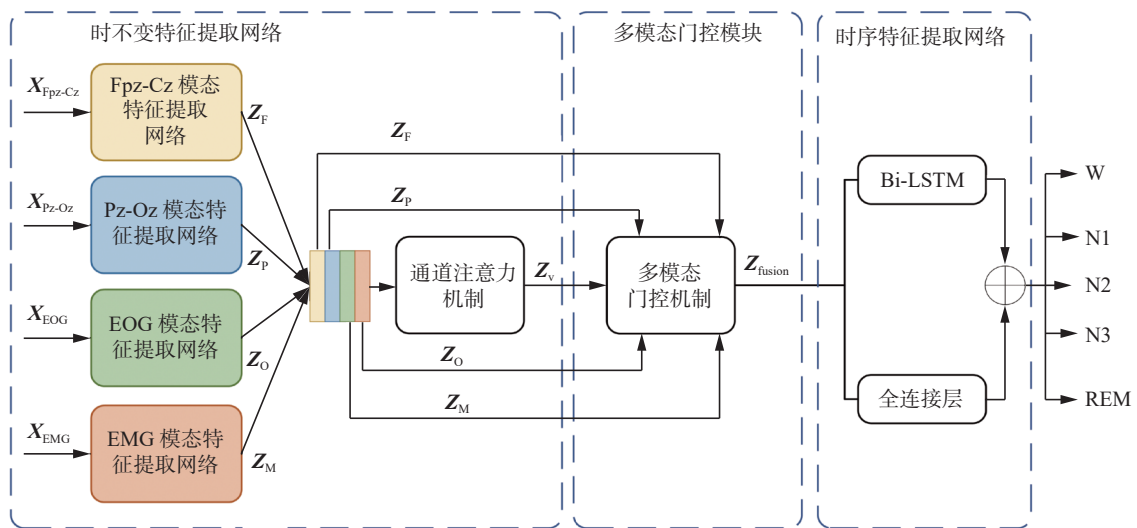


图 1 基于多模态融合的睡眠分期网络结构

Fig. 1 Network structure of sleep staging based on multimodal fusion

### 1.1 时不变特征提取网络

针对 PSG 中各模态信号中含有噪声的问题,

本文首先利用残差收缩网络设计各个模态的特征提取网络, 然后利用通道注意力机制学习不同模

态特征间的相关性,进而得到睡眠多导图的时不变特征,本文采用的时不变特征提取网络如图 2 所示。

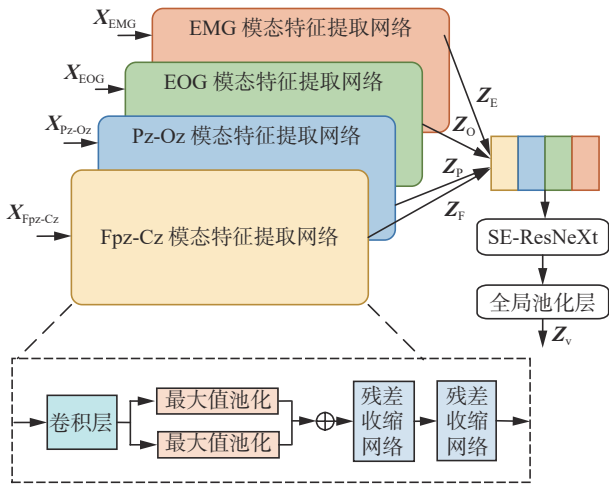


图 2 时不变特征提取网络

Fig. 2 Architecture of time-invariant feature extraction

首先利用各模态特征提取网络分别提取不同模态的高维特征,分别将每个模态信息输入到卷积尺度为 30,步长为 6 的 1 维卷积提取原始信号的低维特征。然后利用 2 个并行的最大池化层实现下采样,提取不同尺度的关键信息。

利用残差收缩网络去除融合特征的冗余信息<sup>[16]</sup>,残差收缩网络结构如图 3 所示,进而得到每个模态信息的高维抽象特征。

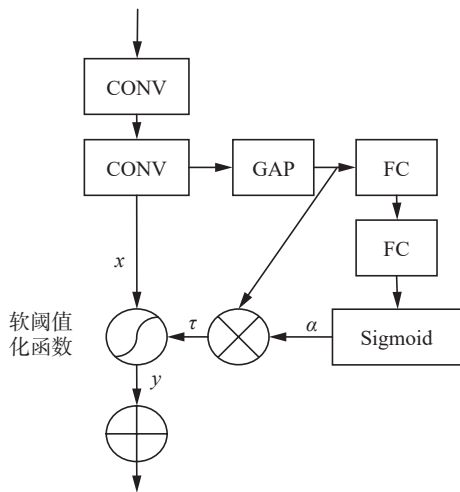


图 3 残差收缩网络

Fig. 3 Residual shrinkage network

由图 3 可知,残差收缩网络由 ResNet 网络(residual neural network)、阈值学习网络和软阈值化函数组成。利用 ResNet 学习抽象特征表示,然后通过阈值学习网络学习反映信息重要程度的阈值,其中阈值学习网络的表达式为

$$\tau = \alpha \cdot \text{average}|x_{i,j,c}| \quad (1)$$

式中  $\alpha$  表示由全连接层和 Sigmoid 函数得到的权

值。通过阈值学习网络可以确保  $\tau$  是一个很小的正实数。然后利用软阈值化函数将特征中不重要的信息置为 0,其中软阈值化函数是小波降噪算法中的核心步骤,其表达式为

$$y = \begin{cases} x - \tau, & x > \tau \\ 0, & -\tau \leq x \leq \tau \\ x + \tau, & x < -\tau \end{cases} \quad (2)$$

通过软阈值化函数的引入,可以有效地增强有效的信息,抑制冗余的信息,进而消除信号中噪声的影响,得到高效的各模态抽象特征表示。

将各模态的抽象特征在通道维度上进行拼接,如式(3)所示,得到初步融合特征。

$$Z = \text{Concat}(Z_F, Z_P, Z_O, Z_M) \quad (3)$$

式中:  $Z_F, Z_P, Z_O, Z_M \in \mathbf{R}^{l \times C}$ , 代表各个模态特征,  $l$  为特征图的尺度,  $C$  为输出滤波器的个数;  $Z$  代表拼接特征,  $Z \in \mathbf{R}^{l \times nC}$ ,  $n$  代表模态的个数。

随后利用通道注意力机制学习不同通道特征间的相关性。本文利用 SE-ResNeXt 模块<sup>[17]</sup> 根据不同通道特征重要程度对特征进行整合。其中, SE-ResNeXt 由 ResNeXt 和 SENet (squeeze and excitation network) 两部分组成,如图 4 所示。残差网络通过增加模型深度和宽度来提高特征提取的质量,但会造成模型参数量增加、计算开销增加的问题, ResNet 在此基础上引入了基数的超参数,即卷积支路的个数。通过增加基数,在不增加网络参数量的前提下有效提高了模型的特征提取能力。

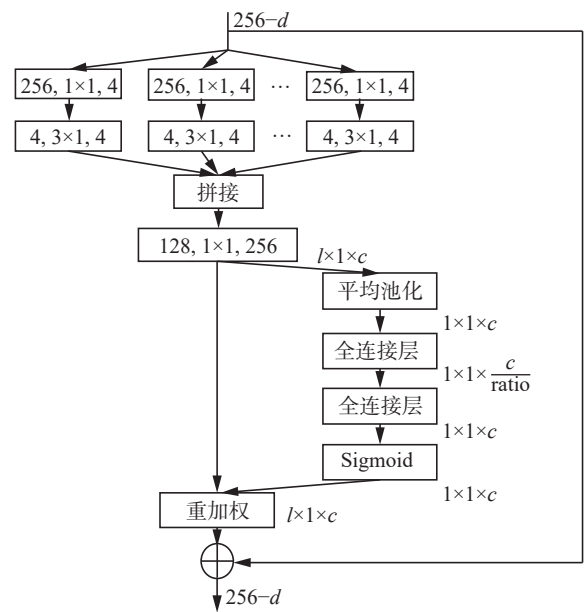


图 4 SE-ResNeXt 结构

Fig. 4 Architecture of SE-ResNeXt

SENet 通过自适应获得各个通道特征的重要程度,对输出特征进行重标定,进而提升有用的通道特征,抑制重要程度低的特征。图中  $l$  为特征图尺度,  $c$  为通道数, ratio 为特征维度缩放比。

SENet 分别经过压缩和激励变换, 建模各通道特征的重要权重, 最后各通道特征逐点相乘的操作将归一化的权重加权到输入特征中, 得到含通道重要程度的特征向量。式(4)~(6)分别给出了压缩、激励、加权的公式。

$$z_c = F_{sq}(u_c) = \frac{1}{l} \sum_{i=1}^l u_c(i) \quad (4)$$

$$s = F_{ex}(z, W) = \sigma(W_2 \delta(W_1 z)) \quad (5)$$

$$\tilde{x}_c = F_{scale}(u_c, s_c) = s_c \cdot u_c \quad (6)$$

式中:  $W_1$ 和 $W_2$ 分别代表两个全连接层的权重;  $\delta$ 代表 ReLu 函数;  $\sigma$ 代表 Sigmoid 函数;  $F_{sq}$ 、 $F_{ex}$ 、 $F_{scale}$ 分别代表压缩、激励和尺度变换的效果。

## 1.2 多模态门控模块

多模态学习中模态融合对模型性能有很大的影响, 为了充分发挥不同模态间的差异性<sup>[18]</sup>, 本文模拟人工判读以关键信号为主, 其他信号为辅的判读模式, 引入了多模态门控模块, 建模各个模态特征对于睡眠分期任务的重要程度, 通过自适应门控机制<sup>[19]</sup>对不同模态特征按照重要程度进行加权, 根据不同模态特征的组合自适应地找到更适合睡眠分期任务的特征表示。设计的多模态门控模块如图 5 所示。图中 $Z_v$ 为睡眠多导图的时不变特征,  $Z_F$ 、 $Z_P$ 、 $Z_O$ 、 $Z_M$ 分别为各模态高维抽象特征表示。

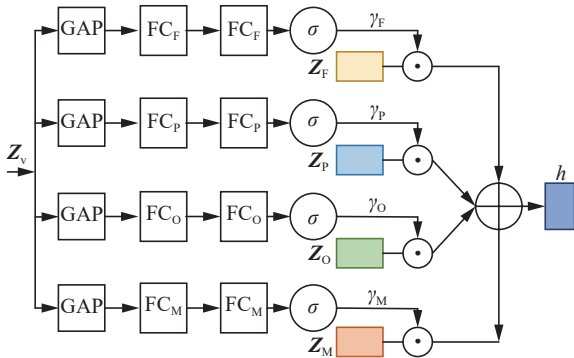


图 5 多模态门控模块

Fig. 5 Multimodal gating module

将睡眠多导图的时不变特征分别首先经过全局平均池化层得到全局信息, 然后经过两层全连接层和 Sigmoid 函数增加特征的非线性, 通过式(7)~(10)分别学习各个模态特征的权重, 然后通过式(11)得到输出特征的最优组合, 寻找对于输出贡献度最大的各个模态的特征组合, 进而得到更丰富的融合特征表示。

$$\gamma_F = \text{Sigmoid}(W_F^1(W_F^1(\text{GAP}(Z_v)))) \quad (7)$$

$$\gamma_P = \text{Sigmoid}(W_P^1(W_P^1(\text{GAP}(Z_v)))) \quad (8)$$

$$\gamma_O = \text{Sigmoid}(W_O^1(W_O^1(\text{GAP}(Z_v)))) \quad (9)$$

$$\gamma_M = \text{Sigmoid}(W_M^1(W_M^1(\text{GAP}(Z_v)))) \quad (10)$$

$$h = \gamma_F \cdot Z_F + \gamma_P \cdot Z_P + \gamma_O \cdot Z_O + \gamma_M \cdot Z_M \quad (11)$$

式中:  $W_F^1$ 、 $W_P^1$ 、 $W_O^1$ 、 $W_M^1$ 、 $W_F^2$ 、 $W_P^2$ 、 $W_O^2$ 、 $W_M^2$ 分别代表相应全连接层的参数向量;  $\gamma_F$ 、 $\gamma_P$ 、 $\gamma_O$ 、 $\gamma_M$ 分别代表各模态的重要性权值;  $Z_v$ 代表时不变特征;  $h$ 代表融合特征。

## 1.3 时序特征提取网络

睡眠 PSG 是一种时序信号, 相邻样本间具有时序相关性。在对  $t$  时刻的样本进行分期时, 可以利用  $t-1$  和  $t+1$  时刻的特征或分期结果辅助进行  $t$  时刻样本阶段的分类。

本文利用双向长短时记忆网络 (bi-directional long short-term memory, Bi-LSTM)<sup>[20]</sup> 设计时序特征提取网络。与长短时记忆网络只能学习序列前向传播规律不同的是 Bi-LSTM 分别由一个前向传播和一个反向传播的 LSTM 组成, 保证了时序特征提取网络既可以学习前向传播规律, 还可以学习序列的反向传播规律。最后将 Bi-LSTM 学习到的时序特征与多模态特征进行融合得到睡眠 PSG 的时序特征, 得到的时序特征表达式为

$$O_t = g(w_F h'_F + w_B h'_B) + w_{FC} h_t \quad (12)$$

式中:  $h_t$ 、 $O_t$ 分别代表  $t$  时刻的输入和输出;  $w_F$ 、 $w_B$ 、 $w_{FC}$ 分别代表 LSTM<sub>F</sub>、LSTM<sub>B</sub> 的输出权重和全连接的权重;  $h'_F$ 、 $h'_B$ 分别代表 LSTM<sub>F</sub> 和 LSTM<sub>B</sub> 的输出;  $g$  代表拼接操作。

## 2 实验结果及分析

为了验证本文创新工作的有效性和先进性, 进行了 3 个方面的实验: 1) 分别利用 PSG 中不同的模态组合设置相关实验, 进而验证本文利用多模态信号进行睡眠分期的有效性, 并且得到最优的模态组合; 2) 设计不同初步融合方式的对比实验, 验证本文采用 SE-ResNeXt 进行特征重标定的有效性; 3) 将本文提出的 MSN 算法与 4 种先进的方法进行对比, 验证本文算法的先进性。

### 2.1 数据集

利用 Sleep-EDF 数据集<sup>[21]</sup> 设计相关的实验共包括 39 个完整夜晚的 PSG 记录。每个 PSG 包含个两模态的 EEG, 分别是 Fpz-Cz 和 Pz-Cz, 一个模态的 EOG 和一个模态的 EMG<sup>[22]</sup>。每个样本都已由睡眠专家标记为 W、N1、N2、N3、N4、REM, 根据 AASM 标准将 N3 和 N4 合并为 N3。最后, W、N1、N2、N3、REM 各个类别的样本量分别为 7927、2804、17799、5703、7717, 共 41950 个样本。

### 2.2 实验条件

本文所有实验均是在 64 位的 Ubuntu16.04 系统中进行, 采用的深度学习框架为 tensorflow 1.14.0, 在训练过程中利用 GeForce RTX 980Ti 显卡进行加速, 并且采用 python3.5 编程语言完成代码的编写。

优化过程中采用 Adam 函数, 学习率设为 0.001, 序列学习过程中 batch size 为 25, 子序列长度为 10, ResNeXt 中的基数设置为 8, ratio 设置为 4。利用丢失率为 0.5 dropout 层来防止过拟合。进行 20 折交叉验证, 按 19:1 划分训练集和验证集。

### 2.3 评价指标

采用  $M_{F_1}$  分数 (macro-averaging  $F_1$ -score,  $M_{F_1}$ )、整体准确率 (overall-accuracy,  $A_{CC}$ ) 作为评价指标, 具体公式为

$$A_{CC} = \frac{\sum_{i=1}^C T_{P_i}}{N} \times 100\% \quad (13)$$

$$M_{F_1} = \frac{\sum_{i=1}^C F_{1_i}}{C} \times 100\% \quad (14)$$

式中:  $N$  代表样本总数;  $C$  代表类别总数;  $F_{1_i}$  代表类别  $i$  的  $F_1$  分数, 即

$$F_{1_i} = \frac{2P_iR_i}{P_i+R_i} \quad (15)$$

式中:  $P_i = \frac{T_{P_i}}{T_{P_i}+F_{P_i}}$ ;  $R_i = \frac{T_{P_i}}{T_{P_i}+F_{N_i}}$ ;  $T_{P_i}$  代表正确分数为  $i$  类的样本数;  $F_{P_i}$  代表被错分为  $i$  类的样本数;  $F_{N_i}$  代表被错分为其他类别的样本数。

### 2.4 多模态信息有效性验证

由于 EEG 是睡眠分期的主要判读依据, 在验证利用多模态信息有效性时选择在 EEG 的基础上加入不同模态的数据, 验证不同模态数据对睡眠分期的影响。得到的结果如表 1 所示, 表中 1 代表两个模态的 EEG 组合, 2 代表两个模态 EEG 加 EOG 的组合, 3 代表两个模态 EEG 加 EOG 加 EMG 的组合。

表 1 多模态信息有效性验证  
Table 1 Verification of the multimodal information %

| 模态 | $A_{CC}$    | $M_{F_1}$   | 每个类别 $F_1$  |             |             |             |             |
|----|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
|    |             |             | W           | N1          | N2          | N3          | REM         |
| 1  | 86.1        | 80.3        | 89.7        | 48.3        | 88.7        | <b>89.2</b> | 85.8        |
| 2  | 87.3        | 81.4        | 90.7        | 49.9        | <b>90.0</b> | 89.1        | 87.6        |
| 3  | <b>87.6</b> | <b>82.0</b> | <b>91.9</b> | <b>51.3</b> | 89.0        | 87.3        | <b>87.9</b> |

从表 1 中可以看出, 仅用两模态 EEG 进行睡

眠研究时平均准确率为 86.1%,  $M_{F_1}$  为 80.3%, 但是当加入眼电信号后平均准确率达到 87.3%, 相较于仅利用脑电信号提高了 1.4%,  $M_{F_1}$  达到了 81.4%, 提高了 1.4%。并且由于快速眼动是 REM 期的典型特征, 在加入眼电信号后, REM 期的  $F_1$  有了较大的提升, 提升了 2.0%, 同时对于 W 期、N1 期、N2 期也有了一定的提升, 说明眼电的加入可以有效地丰富信息特征, 增强模型对于睡眠各阶段的识别能力。在脑电、眼电信号的基础上加入肌电后, 平均准确率达到 87.6%, 与利用脑电和眼电相比提升了 0.3%,  $M_{F_1}$  达到了 82%, 提升了 0.7%, 说明肌电信号的加入进一步丰富了信息特征, 可以让网络模型学习到更丰富的分期特征, 特别是肌肉运动明显的时期如 W 和 N1 期的识别精度有了明显的提升, 其中 W 期的  $F_1$  提升了 1.3%, N1 期的  $F_1$  提升了 2.8%, 说明肌电信号的加入有效地增大了 W、N1 期与其他类别的阶段差异性, 进而提高了网络模型的分期性能。综上证明了同时利用 EEG、EOG、EMG 可以得到更好的睡眠分期结果。

### 2.5 特征融合方式验证实验

本文将各模态特征在通道维度上进行拼接, 然后利用通道注意力机制对拼接特征进行重标定的方式实现各模态的初步融合, 为了进一步验证该方法的有效性, 分别与逐点相加、通道拼接的特征融合方式进行对比, 表 2 分别给出了不同融合方式得到的分期结果。从表 2 中可以看出, 采用各通道特征逐点相加进行特征融合的整体准确率为 86.4%,  $M_{F_1}$  为 80.4%, 但是将特征融合的方式变为在通道维度上拼接后, 准确率提高了 0.5%,  $M_{F_1}$  提高了 1.1%, 证明利用在通道维度上拼接的方式能够更好地利用不同特征之间的差异性, 进而得到更充分的初步融合特征。随后加入通道注意力机制后准确率提高了 0.9%,  $M_{F_1}$  提高了 0.8%, 并且各个类别的  $F_1$  都有了一定的提高, 进而证明了本文采用 SE-ResNeXt 对特征按照通道重要性权重进行重标定, 能够让特征融合得更充分, 进而验证了本文采用通道维度拼接和 SE-ResNeXt 对特征进行初步融合的有效性。

表 2 特征融合方式有效性验证  
Table 2 Verification of modal fusion %

| 融合方法       | $A_{CC}$    | $M_{F_1}$   | 每个类别 $F_1$  |             |             |             |             |
|------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
|            |             |             | W           | N1          | N2          | N3          | REM         |
| 逐点相加       | 86.4        | 80.4        | 91.9        | 48.3        | 88.8        | 87.4        | 85.7        |
| 通道拼接       | 86.8        | 81.3        | 91.9        | 50.0        | 89.3        | 88.3        | 87.7        |
| 通道拼接+通道注意力 | <b>87.6</b> | <b>82.0</b> | <b>91.9</b> | <b>51.3</b> | <b>90.0</b> | <b>89.0</b> | <b>87.9</b> |

## 2.6 算法先进性验证

为了验证本文算法的先进性与目前 4 种先进的算法进行对比, 表 3 给出了对比结果。

表 3 算法先进性验证  
Table 3 Verification of algorithm advancement %

| 方法     | $A_{CC}$    | $M_{F1}$    | 每个类别 $F_1$  |             |             |             |             |
|--------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
|        |             |             | W           | N1          | N2          | N3          | REM         |
| 文献[10] | 82.3        | 75.0        | —           | —           | —           | —           | —           |
| 文献[23] | 84.9        | 77.3        | 91.4        | 38.8        | 87.4        | 83.5        | 85.1        |
| 文献[24] | 85.2        | 79.6        | 85.4        | 50.9        | 88.8        | 86.4        | 86.5        |
| 文献[25] | 86.0        | 82.0        | 89.0        | <b>57.0</b> | 89.0        | 88.0        | 87.0        |
| 本文     | <b>87.6</b> | <b>82.0</b> | <b>91.9</b> | 51.3        | <b>90.0</b> | <b>89.0</b> | <b>87.9</b> |

从表 3 中可以看出, 本文提出的模型的平均准确率为 87.6%,  $M_{F1}$  为 82.0%, 均高于目前先进的对比算法, 并且除 N1 期外, 各个类别的  $F_1$  分数均取得了最优的水平, 证明本文算法能够充分学习不同睡眠阶段的特征, 取得了优异的睡眠分期性能, 进而证明了本文算法的先进性。

采用混淆矩阵对本文提出的算法进行可视化操作。图 6 给出了本文算法利用两个模态 EEG、一个模态 EOG 和一个模态 EMG 达到的最好分类效果时的混淆矩阵。图中每一行代表实际标签为该类别的样本数, 每一列代表预测为该类别的样本数, 并且颜色越深代表预测为该类别的样本数越多。从混淆矩阵中可以看出, 本文利用所提出的多模态融合的方法进行睡眠分期研究, 大多数样本都能被正确分类, 并且主要的错分出现在相邻阶段的样本中, 这主要是因为转换阶段的样本特征中存在一定的中和, 甚至睡眠专家也无法对转换期的样本进行正确分类, 导致模型学到的转换期特征边界较模糊, 进而造成了错分主要出现在相邻阶段转换期的样本。由于睡眠本身的特征, N1 期存在时间短于其他时期, 样本量较小, 所以 N1 期的  $F_1$  低于其他类别。

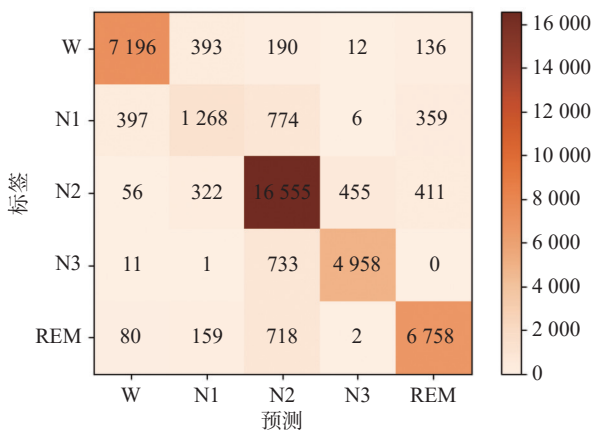


图 6 混淆矩阵

Fig. 6 Confusion matrix

## 3 结束语

本文为了充分利用睡眠 PSG 中各模态信息, 利用通道注意力机制和多模态门控机制设计了基于多模态融合的睡眠分期模型。首先利用残差收缩网络设计各模态特征提取网络提取各个模态的高维特征, 并将各个模态的高维特征在通道维度上进行融合得到初步融合特征, 然后利用 SE-ResNeXt 根据各通道信息的重要性程度对特征进行重标定, 得到睡眠多导图的时不变特征。然后输入到多模态门控模块中, 利用全连接层和非线性函数, 学习各个模态信息的重要性程度, 结合乘法门对各模态特征进行加权, 进而找到能充分表示各模态信息的最优组合。随后, 利用 Bi-LSTM 提取相邻样本间的时序相关性。本文利用 Sleep-EDF 中的 Fpz-Cz、Pz-Oz、EOG、EMG 模态信息进行了二十折交叉验证实验, 并通过与目前先进的 4 种对比算法进行比较, 本文提出的算法无论是整体准确率还是  $M_{F1}$  均高于对比算法, 进而证明了本文算法的有效性和先进性。

## 参考文献:

- [1] BERRY R B, BUDHIRAJA R, GOTTLIEB D J, et al. Rules for scoring respiratory events in sleep: update of the 2007 AASM manual for the scoring of sleep and associated events[J]. *Journal of clinical sleep medicine*, 2012, 8(5): 597–619.
- [2] 王茜侨. 有关美国睡眠医学学会睡眠分期的最新判读标准指南解析 [J]. *诊断学理论与实践*, 2009, 8(6): 575–578.  
WANG Hanqiao. Analysis of the latest guidelines for sleep staging of the american academy of sleep medicine[J]. *Journal of diagnostics concepts & practice*, 2009, 8(6): 575–578.
- [3] ZHANG Linda, FABBRI D, UPENDER R, et al. Automated sleep stage scoring of the sleep heart health study using deep neural networks[J]. *Sleep*, 2019, 42(11): zsz159.
- [4] ZHAO Dechun, WANG Yi, WANG Qiangqiang, et al. Comparative analysis of different characteristics of automatic sleep stages[J]. *Computer methods and programs in biomedicine*, 2019, 175: 53–72.
- [5] FAUST O, RAZAGHI H, BARIKA R, et al. A review of automated sleep stage scoring based on physiological signals for the new millennia[J]. *Computer methods and programs in biomedicine*, 2019, 176: 81–91.
- [6] VILAMALA A, MADSEN K H, HANSEN L K. Deep convolutional neural networks for interpretable analysis of EEG sleep stage scoring[C]//2017 IEEE 27th International Workshop on Machine Learning for Signal Processing. Tokyo: IEEE, 2017: 1–6.

- [7] HUMAYUN A I, SUSHMIT A S, HASAN T, et al. End-to-end sleep staging with raw single channel EEG using deep residual ConvNets[C]//2019 IEEE EMBS International Conference on Biomedical & Health Informatics. Chicago: IEEE, 2019: 1–5.
- [8] 贾子钰, 林友芳, 张宏钧, 等. 基于深度卷积神经网络的睡眠分期模型[J]. 浙江大学学报(工学版), 2020, 54(10): 1899–1905.  
JIA Ziyu, LIN Youfang, ZHANG Hongjun, et al. Sleep stage classification model based on deep convolutional neural network[J]. Journal of Zhejiang University (engineering science edition), 2020, 54(10): 1899–1905.
- [9] FERNANDEZ-BLANCO E, RIVERO D, PAZOS A. EEG signal processing with separable convolutional neural network for automatic scoring of sleeping stage[J]. *Neurocomputing*, 2020, 410: 220–228.
- [10] PHAN H, ANDREOTTI F, COORAY N, et al. Joint classification and prediction CNN framework for automatic sleep stage classification[J]. *IEEE transactions on bio-medical engineering*, 2019, 66(5): 1285–1296.
- [11] JADHAV P, RAJGURU G, DATTA D, et al. Automatic sleep stage classification using time-frequency images of CWT and transfer learning using convolution neural network[J]. *Biocybernetics and biomedical engineering*, 2020, 40(1): 494–504.
- [12] YÜCELBAŞ C, YÜCELBAŞ Ş, ÖZŞEN S, et al. Effect of the Hilbert-Huang transform method on sleep staging[C]//2017 25th Signal Processing and Communications Applications Conference (SIU). Antalya: IEEE, 2017: 1–4.
- [13] LOH H W, OOI C P, VICNESH J, et al. Automated detection of sleep stages using deep learning techniques: a systematic review of the last decade (2010–2020)[J]. *Applied sciences*, 2020, 10(24): 8963.
- [14] SOKOLOVSKY M, GUERRERO F, PAISARNSRI-SOMSUK S, et al. Deep learning for automated feature discovery and classification of sleep stages[J]. *IEEE/ACM transactions on computational biology and bioinformatics*, 2020, 17(6): 1835–1845.
- [15] DUAN Lijuan, LI Mengying, WANG Changming, et al. A novel sleep staging network based on data adaptation and multimodal fusion[J]. *Frontiers in human neuroscience*, 2021, 15: 727139.
- [16] ZHAO Minghang, ZHONG Shisheng, FU Xuyun, et al. Deep residual shrinkage networks for fault diagnosis[J]. *IEEE transactions on industrial informatics*, 2020, 16(7): 4681–4690.
- [17] HU Jie, SHEN Li, SUN Gang. Squeeze-and-excitation networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 7132–7141.
- [18] 张峰, 李希城, 董春茹, 等. 基于深度情感唤醒网络的多模态情感分析与情绪识别[J]. 控制与决策, 2022, 37(11): 2984–2992.
- ZHANG Feng, LI Xicheng, DONG Chunru, et al. Deep emotional arousal network for multimodal sentiment analysis and emotion recognition[J]. *Control and decision*, 2022, 37(11): 2984–2992.
- [19] WANG Yansen, SHEN Ying, LIU Zhun, et al. Words can shift: dynamically adjusting word representations using nonverbal behaviors[C]//Proceedings of the AAAI Conference on Artificial Intelligence. Honolulu: Assoc Advancement Artificial Intelligence, 2019: 7216–7223.
- [20] ABDELHAMEED A M, DAOUD H G, BAYOUMI M. Deep convolutional bidirectional LSTM recurrent neural network for epileptic seizure detection[C]//2018 16th IEEE International New Circuits and Systems Conference. Montreal: IEEE, 2018: 139–143.
- [21] KEMP B, ZWINDERMAN A H, TUK B, et al. Analysis of a sleep-dependent neuronal feedback loop: the slow-wave microcontinuity of the EEG[J]. *IEEE transactions on bio-medical engineering*, 2000, 47(9): 1185–1194.
- [22] GOLDBERGER A L, AMARAL L A, GLASS L, et al. PhysioBank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals[J]. *Circulation*, 2000, 101(23): E215–E220.
- [23] ZHANG Lei, CHEN Dan, CHEN Peilu, et al. Dual-CNN based multi-modal sleep scoring with temporal correlation driven fine-tuning[J]. *Neurocomputing*, 2021, 420: 317–328.
- [24] PHAN H, CHÉN O Y, KOCH P, et al. Towards more accurate automatic sleep staging via deep transfer learning[J]. *IEEE transactions on biomedical engineering*, 2021, 68(6): 1787–1798.
- [25] YAN Rui, LI Fan, ZHOU Dong dong, et al. Automatic sleep scoring: a deep learning architecture for multi-modality time series[J]. *Journal of neuroscience methods*, 2021, 348: 108971.

#### 作者简介:



陈玲玲, 硕士研究生, 主要研究方向为生物医学信号处理、深度学习。



毕晓君, 教授, 博士生导师, 主要研究方向为智能信息处理技术、数字图像处理及机器学习。主持国家重点研发项目、国家社科基金重大项目等国家级、省部级项目多项。获得高等学科科学进步技术一等奖 1 项, 省部级科学技术奖 7 项。发表学术论文 170 余篇。