

9-15-2024

## A Multimodal Residual Spatial-temporal Fusion Model Based on Automatic Sleep Classification

Yecai Guo

*School of Electronic and Information Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, China; School of Electronics and Information Engineering, Wuxi University, Wuxi 214105, China*

Shuang Tong

*School of Electronic and Information Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, China*

Follow this and additional works at: <https://dc-china-simulation.researchcommons.org/journal>



Part of the Artificial Intelligence and Robotics Commons, Computer Engineering Commons, Numerical Analysis and Scientific Computing Commons, Operations Research, Systems Engineering and Industrial Engineering Commons, and the Systems Science Commons

---

This Paper is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation. For more information, please contact [xtfzxb@126.com](mailto:xtfzxb@126.com).

---

# A Multimodal Residual Spatial-temporal Fusion Model Based on Automatic Sleep Classification

## Abstract

**Abstract:** Highly accurate sleep staging plays a crucial role in correctly assessing sleep conditions. Aiming at the problem that the existing convolutional network cannot obtain the topological characteristics of physiological signals, a sleep staging algorithm based on multi-modal residual spatio-temporal fusion is proposed. Time-frequency images and spatio-temporal images are obtained using short-time Fourier transform and adaptive map convolution, which are converted into high-dimensional feature vectors; lightweight interaction of feature information flow is realized through time-frequency feature and spatiotemporal feature extraction modules; the feature enhancement fusion module fuses feature information to outputs sleep staging results. The results show that the model has a high accuracy. On the ISRUC-S3 data set, the overall accuracy is 85.3%, the F1 score is 83.8%, Cohen's kappa is 81%, and the N1 stage accuracy reaches 69.81%. Experiments on the ISRUC-S1 dataset demonstrate the generality of the model.

## Keywords

sleep staging, multi-view fusion, graph convolutional network, deep learning, electroencephalogram

## Recommended Citation

Guo Yecai, Tong Shuang. A Multimodal Residual Spatial-temporal Fusion Model Based on Automatic Sleep Classification[J]. Journal of System Simulation, 2024, 36(9): 2065-2074.

# 基于自动睡眠分期的多模态残差时空融合模型

郭业才<sup>1,2</sup>, 全爽<sup>1</sup>

(1. 南京信息工程大学 电子与信息工程学院, 江苏 南京 210044; 2. 无锡学院 电子信息工程学院, 江苏 无锡 214105)

**摘要:** 高精度的睡眠分期对于正确评定睡眠情况起到了至关重要的作用。针对现有的卷积网络无法获取生理信号拓扑特征的问题, 提出了一种基于多模态残差时空融合的睡眠分期算法。利用短时傅里叶变换和自适应图卷积获取时频图像和时空图像, 将其转换为高维的特征向量; 通过时频特征和时空特征提取模块实现特征信息流的轻量化交互; 使用特征增强融合模块融合特征信息, 输出睡眠分期结果。结果表明: 该模型具有较高的准确率, 在ISRUC-S3数据集上整体准确率为85.3%, F1分数为83.8%, Cohen's kappa为81%, N1阶段准确率达到69.81%。ISRUC-S1数据集上的实验证明了模型的普遍性。

**关键词:** 睡眠分期; 多视图融合; 图卷积网络; 深度学习; 脑电信号

中图分类号: TP391

文献标志码: A

文章编号: 1004-731X(2024)09-2065-10

DOI: 10.16182/j.issn1004731x.joss.23-0588

**引用格式:** 郭业才, 全爽. 基于自动睡眠分期的多模态残差时空融合模型[J]. 系统仿真学报, 2024, 36(9): 2065-2074.

**Reference format:** Guo Yecai, Tong Shuang. A Multimodal Residual Spatial-temporal Fusion Model Based on Automatic Sleep Classification[J]. Journal of System Simulation, 2024, 36(9): 2065-2074.

## A Multimodal Residual Spatial-temporal Fusion Model Based on Automatic Sleep Classification

Guo Yecai<sup>1,2</sup>, Tong Shuang<sup>1</sup>

(1. School of Electronic and Information Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, China;  
2. School of Electronics and Information Engineering, Wuxi University, Wuxi 214105, China)

**Abstract:** Highly accurate sleep staging plays a crucial role in correctly assessing sleep conditions. Aiming at the problem that the existing convolutional network cannot obtain the topological characteristics of physiological signals, a sleep staging algorithm based on multi-modal residual spatio-temporal fusion is proposed. Time-frequency images and spatio-temporal images are obtained using short-time Fourier transform and adaptive map convolution, which are converted into high-dimensional feature vectors; lightweight interaction of feature information flow is realized through time-frequency feature and spatio-temporal feature extraction modules; the feature enhancement fusion module fuses feature information to outputs sleep staging results. The results show that the model has a high accuracy. On the ISRUC-S3 data set, the overall accuracy is 85.3%, the F1 score is 83.8%, Cohen's kappa is 81%, and the N1 stage accuracy reaches 69.81%. Experiments on the ISRUC-S1 dataset demonstrate the generality of the model.

**Keywords:** sleep staging; multi-view fusion; graph convolutional network; deep learning; electroencephalogram

收稿日期: 2023-05-18 修回日期: 2023-08-24

基金项目: 国家自然科学基金(61673222)

第一作者: 郭业才(1962-), 男, 教授, 博士, 研究方向为通信信号处理、水声信号处理等。

## 0 引言

睡眠质量的好坏在一定程度上反映了人体的健康状况<sup>[1]</sup>。了解睡眠阶段转换的周期性,对客观、准确地评估睡眠状态具有重要意义。利用脑电信号辅助医生进行睡眠诊断已成为一种新趋势<sup>[2]</sup>。多导睡眠图(polysomnography, PSG)记录来自身体各个部位的电信号,包括脑电图(electroencephalogram, EEG)、肌电图(electromyography, EMG)、眼电图(electrooculogram, EOG)和心电图(electrocardiogram, ECG)。研究人员通过睡眠分期规则对PSG信号进行划分,PSG被分成以30 s为1个周期的信号片段,根据睡眠分期标准将睡眠信号片段分为不同的阶段。例如,R&K规则指南<sup>[3]</sup>将睡眠分为3个阶段:清醒(W)、快速眼动(REM)和非快速眼动(NREM)。其中,NREM包括N1、N2、N3、N4共4个阶段。美国睡眠医学会(AASM)睡眠分类标准<sup>[4]</sup>将NREM中N3和N4阶段合并为N3阶段。

通过观察睡眠阶段,可以有效地检验身体的状况,但手动划分睡眠阶段存在时间长、专业要求高、分类结果具有高度主观性等弊端,导致该领域发展缓慢。传统的机器学习方法,如随机下采样增强<sup>[5]</sup>、可调Q小波变换<sup>[6]</sup>和支持向量机<sup>[7]</sup>已被用于识别睡眠阶段。然而,机器学习方法的性能很大程度上取决于特征工程的选择是否合理。相比之下,深度学习可以通过神经网络处理原始数据,获取丰富的信息,并直接输出分类结果,实现了更高的分类准确率和更好的泛化能力<sup>[8-10]</sup>。

睡眠分期按照波形的频率划分为不同的睡眠阶段,借助CNN对频率变化的敏感性和RNN对前后序列的捕获能力,取得了良好的分类结果。DeepSleep<sup>[11]</sup>使用Bi-LSTM和残差连接在ISRUC-S3数据集上实现了78.8%的准确率。U-Sleep<sup>[12]</sup>使用对称的编码结构,将图像处理的方式用于解决睡眠分期任务,在ISRUC-S3数据集上实现了77%

的准确率。U<sup>2</sup>-Sleep<sup>[13]</sup>使用嵌套的U型结构,自适应地捕获对睡眠分期起关键作用的特征,在ISRUC-S3数据集上实现了79.9%的准确率。

CNN对于处理非欧几里德数据有不足,单纯的使用CNN网络,无法利用大脑不同区域之间的联系<sup>[14]</sup>。相比之下,GCN<sup>[15-16]</sup>通过探索生物信号的多通道空间特征,为睡眠分类提供了新的解决思路。例如,STGCN<sup>[17]</sup>通过GCN从原始数据中提取特征;MSTGCN<sup>[18]</sup>使用域泛化提高模型的泛化能力;JKSTGCN<sup>[19]</sup>借助跳跃融合图卷积提取特征。这类模型在训练过程中由于没有卷积核作为共享参数,只能进行简单的特征汇总,无法获取深层信息,模型精度受限于输入数据的质量。此外,单纯的使用GCN网络忽略了多模态生理信号中丰富的时频信息。

综上所述,睡眠分期任务存在以下问题:①现有方法大多仅从时域或频谱域学习特征,忽略了多模态生理信号中空间域的特征信息;②相邻时期的生物信号通道之间没有联系,无法利用信号过渡时期的转换规则;③复杂的网络结构带来精确度增加的同时也加大了模型训练的难度。为此,本文设计了一种用于自动睡眠分期的多模态残差融合图卷积模型。本文使用短时傅里叶变换(short time fourier transform, STFT)<sup>[20]</sup>和AlexNet构建生理信号的时频图,使用残差融合结构(residual fusion structure, RFS)和自适应图卷积(adaptive graph convolution, AGC)<sup>[21]</sup>构建生理信号的时空图;设计了时频特征提取模块(time-frequency feature extraction, TFF),通过门控循环注意力(gated recurrent unit attention, GRU\_ATT)单元和扩张卷积捕获时频序列特征;设计了时空特征提取模块(spatial-temporal feature extraction, SFE),将特征解耦为通道映射和维度映射,通过残差连接建立前后序列的联系;设计了特征增强融合模块(feature enhancement fusion, FEF)合并特征并完成睡眠分期任务。

# 1 模型结构

PSG 为非欧几里德数据, 与 CNN 和 RNN 擅长处理的数据类型不匹配, 无法利用 PSG 信号中的时空信息<sup>[22-23]</sup>。为了捕获通道之间的信息, 本文使用 GCN 和 CNN 捕获拓扑信息, 模型结构如图 1 所示。利用 STFT、AlexNet、AGC 和 RF 将输入信号转换为含有时频信息和时空信息的高维特征向量; 借助 TFF 和 SFE, 通过扩大感受野, 增加前后序列的联系, 对高维特征进行重标定; 通过 FEF 将各模态信号进行融合, 提取信号的时序特征, 完成睡眠分期任务。

## 1.1 多模态关系表示

在时频特征提取阶段, STFT 通过滑动窗口函数对非平稳信号进行分割, 在时域进行积分。STFT 的定义为

$$S(t, f) = \int_{-\infty}^{\infty} h(t - \tau)x(\tau)e^{-j2\pi f\tau} d\tau \quad (1)$$

式中:  $h(t - \tau)$  为以  $t$  为中心的窗函数;  $x(t)$  为要处理的信号。

矩形窗是捕获相近频率信号的有效方法。本文采用矩形窗函数作为 STFT 的基本函数。信号经 STFT 处理后变为能够表示时间频率的二维图像。利用传统的算法将时频图中蕴含的特征转换为高

维向量, 方便后续任务的进行。AlexNet 在设计之初, 因计算机算力的限制将模型分为上下两部分, 分别在两块 GPU 上训练。考虑到数据集本身特点, 本文使用 AlexNet 网络的单分支作为 STFT 特征的提取结构。

在时空特征提取阶段, 可以采用 AGC 提取特征。空间信息可以表示为  $A_{mn}^{TS}$ , 采用 ReLU 激活函数保证  $A_{mn}^{TS}$  始终为正:

$$A_{mn}^{TS} = g(x_m, x_n) = \frac{\exp(\text{ReLU}(\mathbf{w}^T|x_m - x_n|))}{\sum_{n=1}^N \exp(\text{ReLU}(\mathbf{w}^T|x_m - x_n|))} \quad (2)$$

式中:  $x_n$  和  $x_m$  为不同的节点;  $\mathbf{w}$  为权重。每个节点提取的空间信息通过 RF 进行融合:

$$X_G = \text{ReLU}(g_\theta * Gx_{(m-1)}) + \text{Sigmoid}(\text{ReLU}(g_{\theta'} * G'x_{(m-2)})) \quad (3)$$

式中:  $g_\theta$  和  $g_{\theta'}$  为不同尺度的卷积核;  $*G$  和  $*G'$  为不考虑时间信息和考虑时间信息的自适应图卷积;  $x_{(m-1)}$  和  $x_{(m-2)}$  为  $m-1$  和  $m-2$  层的输入。

## 1.2 多视图特征提取

文献[24-25]表明, 从多个域中提取特征可以改善睡眠阶段分类的准确性。每个域都有其独特的特征, 例如, 时频域的信号变换关系和时空域的空间连接。因此, MJ-Sleep 采用不同的特征结构来提取多模态特征。

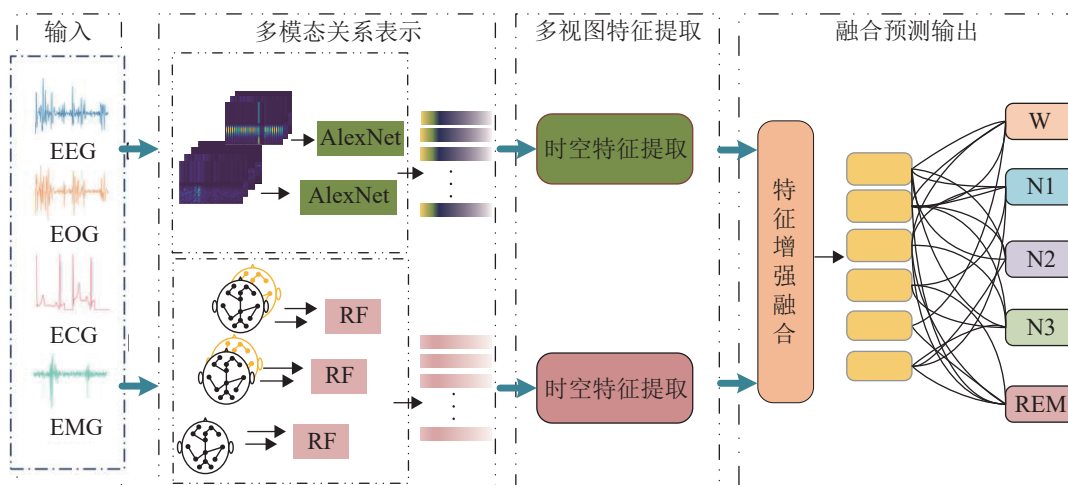


图 1 MJ-Sleep 网络的总体框架  
Fig. 1 Overall framework of MJ-Sleep network



睡眠数据是连续时间内采集的信号样本，前后序列间存在着转换关系。通过前后序列所处的阶段预测当前序列的阶段，是一种常见的判断方法。DC扩展感受野范围以获取更丰富的信息，RNN可以利用信号先后的关联性获取信号的过渡关系。从理论上讲，这2种方法并不矛盾。本文将2种方法结合起来以获得更丰富的时间信息，并将设计的模块命名为TFF模块，如图2所示，为了减小计算量，本文采用的卷积核的大小为 $3 \times 3$ ，步长为1。将经过GRU\_ATT处理过的信号分别输送到扩张因子不同的扩张卷积中，提取信号的短、中、长3个信号特征，通过拼接操作得到多尺度时频特征。

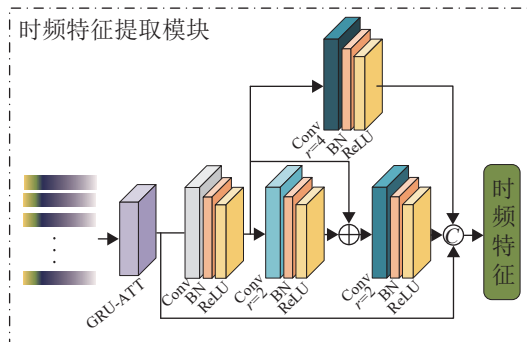


图2 时频特征提取模块

Fig. 2 Time-frequency feature extraction module

信号在神经网络传输的过程，会将信号中的某一部分增强或减弱，本质上是将信息进行重组。信号的重组可能会丢失一些有用的信息。为了解决这个问题，设计了空间特征提取模块SFE。使用 $1 \times 1$ 的卷积核进行恒等映射，加强前后序列的联系，利用ReLU6激活函数防止维度转换过程中的信息丢失或损坏，模型结构如图3所示，DSCConv表示深度可分离卷积。此外，在设计阶段使用深度可分离卷积(depthwise separable convolution, DSC)来减少训练时间并提高模型的实用性。

### 1.3 多视图特征融合

融合模块有利于充分利用特征信息，提高睡眠分期任务的准确性。本文模型利用FEF模块结合时频和时空特征，实现了特征融合，如图4所

示。FEF模块通过式(4)融合时频特征和时空特征，使用式(5)加强融合特征图的重要通道。

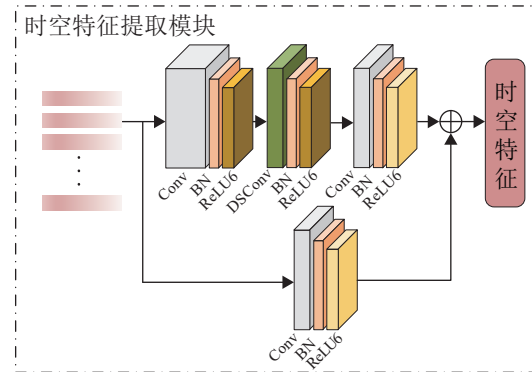


图3 时空特征提取模块

Fig. 3 Spatiotemporal feature extraction module

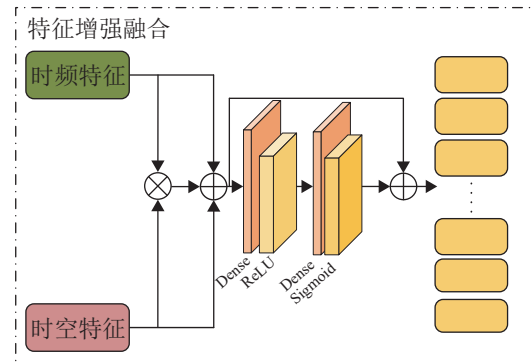


图4 特征增强融合模块

Fig. 4 Feature enhancement fusion module

将经过特征增强融合模块处理的信息 $X_{out}$ 输入到softmax激活层，得到对应的预测类别。

$$\mathbf{X} = \mathbf{X}_{TFF} \oplus \mathbf{X}_{SFE} \oplus (\mathbf{X}_{TFF} \otimes \mathbf{X}_{SFE}) \quad (4)$$

$$\mathbf{X}_{out} = \mathbf{X} \oplus \text{Sigmoid}(F_2(\text{ReLU}(F_1(\mathbf{X})))) \quad (5)$$

式中： $\mathbf{X}_{TFF}$ 为经过时频特征提取模块处理后的特征； $\mathbf{X}_{SFE}$ 为经过时空特征提取模块处理后的特征； $F_1$ 、 $F_2$ 为第1个和第2个全连接层。

多模态残差融合图卷积模型的伪代码如下所示，其中， $I$ 表示训练次数，在本模型中 $I=60$ 。

Model: MJ-Sleep

Input: 脑电信号  $x$

Output: 睡眠阶段的概率分布  $\hat{y}$

Initialization:  $i=0$ , 训练 epoch  $i$  和伪节点标签  $y_i$

While  $i \leq I$  do

for 脑电信号  $x_{epoch} \in x$  do

$X_G \leftarrow \text{ReLU}(g_\theta * Gx_{(m-1)}) + \text{Sigmoid}(\text{ReLU}(g_\theta * G'x_{(m-2)}))$  //建立不同通道间的联系, 获取空间域信息  
 $X_{\text{SFE}} \leftarrow \text{SFE}(X_G)$  //获取空间域特征信息  
 $X_{\text{TFF}} \leftarrow \text{TFF}(\text{AlexNet}(\text{STFT}(x)))$  //获取到电信号的时域特征  
 end  
 $X \leftarrow X_{\text{TFF}} \oplus X_{\text{SFE}} \oplus (X_{\text{TFF}} \odot X_{\text{SFE}})$  //融合特征信号  
 $X_{\text{out}} \leftarrow X \oplus \text{Sigmoid}(F_2(\text{ReLU}(F_1(X))))$  //建立残差连接  
 $\hat{y} = \text{softmax}(X_{\text{out}})$  //计算睡眠阶段概率分布  
 $i = i + 1$   
 end

## 1.4 损失函数

损失函数的设计对于确定模型训练的方向起着至关重要的作用。本文模型考虑了时间信息对频域和空间域的影响。

基于自适应图学习的损失函数:

$$L_{\text{graph\_learning}} = \sum_{m,n=1}^N \|x_m - x_n\|_2^2 A_{mn}^{\text{TS}} + \lambda \|\text{avg}(X_G \cdot W)\|_F^2 \quad (6)$$

式中:  $\lambda$  为大于0的正则化参数;  $\text{avg}$  为计算相邻矩阵时间步长  $t-d$  到  $t+d$  的平均值;  $X_G$  为RF融合后空间信息;  $W$  为可训练参数。

交叉熵损失函数:

$$L_{\text{cross\_entropy}} = -\frac{1}{C} \sum_{c=1}^C \sum_{r=1}^R y_{c,r} \ln \hat{y}_{c,r} \quad (7)$$

式中:  $C$  为样本数;  $R$  为类别标签数;  $y$  为真实标签;  $\hat{y}$  为预测标签。

总的损失函数:

$$L = L_{\text{cross\_entropy}} + L_{\text{graph\_learning}} \quad (8)$$

## 2 实验结果与讨论

### 2.1 实验数据集与设置

本文在开放的 ISRUC-Sleep 数据集<sup>[26]</sup>上评估了

模型性能, 该数据集包括3个子集: 健康受试者、睡眠障碍受试者和服用安眠药的受试者。数据使用标准 EDF+数据格式存储, 扩展名为 .rec。每段睡眠信号由2位睡眠专家使用 AASM 准则进行阶段划分。ISRUC-Sleep 中提供的具有不同特征的受试者数据有利于验证模型的准确性和泛化性。在 ISRUC-Sleep 数据集中, 2位专家根据 AASM 指南将数据分为5个睡眠阶段, 将未知片段标注为 U。标注为 U 的片段极少, 因此在统计的过程中常将此忽略。表1为 ISRUC-S3 和 ISRUC-S1 数据集中每个阶段的分布。

表1 ISRUC-S3 和 ISRUC-S1 数据分类及各期比例  
Table 1 ISRUC-S3 and ISRUC-S1 data classification and proportion of each period

数据集	阶段	阶段数	比例/%	总数
S1	W	20 098	23.05	87 187
	N1	11 062	12.69	
	N2	27 511	31.55	
	N3	17 251	19.79	
	REM	11 265	12.92	
S3	W	1 651	19.31	8 549
	N1	1 215	14.21	
	N2	2 609	30.52	
	N3	2 014	23.56	
	REM	1 060	12.40	

本文所有的实验使用 GeForce RTX 2060 显卡加速, 在深度学习框架 Tensorflow1.15.0 中使用 Python3.7 编译代码, 根据表2设置的超参数进行训练。

表2 训练参数设置  
Table 2 Training parameter settings

超参数	值
训练次数	60
训练批次	32
优化器	Adam
学习率	0.000 02
下采样率	0.5
自适应图卷积正则化参数	0.000 5
切比雪夫多项式阶数	9
卷积核大小	3×3

## 2.2 性能指标

为了可靠地评估模型的性能,使用10折交叉验证来分析ISRUC-S3和ISRUC-S1上的精确度( $P$ )、召回率( $R$ )、F1分数( $F_1$ )、准确率( $A$ )和Cohen's kappa( $K$ )值。这些指标可以更深入地了解模型,正确识别真阳性( $T_p$ )、假阳性( $F_p$ )、真阴性( $T_n$ )和假阴性( $F_n$ ),并提供对其整体有效性的综合衡量。

$$\begin{aligned} A &= (T_p + T_n) / (T_p + F_n + F_p + T_n) \\ F_1 &= (2 \times R \times P) / (R + P) \\ R &= T_p / (T_p + F_n) \\ P &= T_p / (T_p + F_p) \\ K &= (p_0 - p_e) / (1 - p_e) \end{aligned} \quad (9)$$

式中: $p_0$ 为模型的总体精度; $p_e$ 为随机同意的假设概率。

## 2.3 实验结果

### 2.3.1 模型分类结果

本文将MJ-Sleep模型的睡眠分期结果与睡眠专家手动评分的结果进行可视化比较,如图5所示。MJ-Sleep与专家的差异主要体现在过渡阶段。造成这一结果的原图是在睡眠阶段转换期间,样本存在一些重叠的特征,导致过渡阶段的样本信号特征复杂,人为的区分使模型学习到的转换周期特征变得模糊,导致模型在过渡时期的错误增加,例如,在后330个epoch,人进入多梦时期,过渡时期具有多个睡眠阶段的混淆特征,错误的判断大多将REM阶段误判为N2,或是N1阶段误判为N2,此类错误在其他模型中也广泛存在。

尽管如此,从表3显示的MJ-Sleep在ISRUC-S3数据集的分类效果可知,模型在W、N3和REM睡眠阶段均达到了90%左右的准确度。N1阶段达到了69.81%的准确率,模型的整体性能令人满意。

### 2.3.2 模型先进性验证

将MJ-Sleep模型与相关研究进行了比较,结

果如表4所示。根据表4可知,模型在分类方面优于其他模型。传统的深度学习方法无法利用脑电信号的空间位置,提取特征不足导致模型性能较低,现有的图卷积网络虽然充分提取了脑电信号的空间信息,但无法建立通道间的联系,导致提取的特征与真实信号特征有一定的区别。

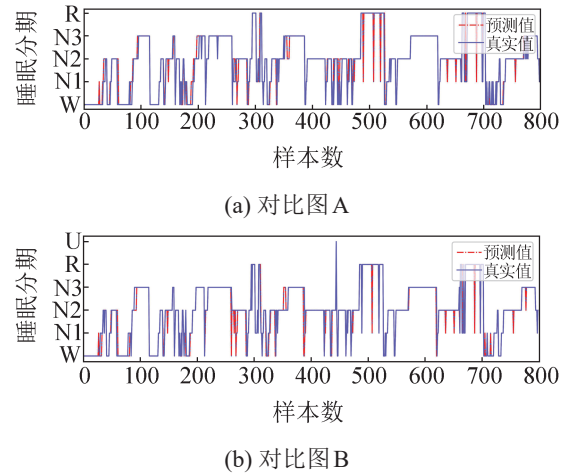


图5 睡眠图结果对比

Fig. 5 Comparison of sleep chart results

表3 MJ-Sleep在ISRUC-S3数据集上的分期结果  
Table 3 Staging results of MJ-Sleep on ISRUC-S3 dataset

睡眠阶段	预测值					性能指标/%	
	W	N1	N2	N3	REM	Pre	Re
W	<b>1 518</b>	99	24	4	6	89.08	91.94
N1	140	<b>733</b>	215	2	125	69.81	60.33
N2	31	159	<b>2 260</b>	118	38	82.97	86.62
N3	1	1	189	<b>1 822</b>	1	93.58	90.47
REM	8	61	36	1	<b>954</b>	84.88	90.00

注:主对角线加粗的值为预测正确的结果。

MJ-Sleep模型充分提取多种特征信息,建立不同通道的联系,在各个睡眠阶段均达到最优的结果。与文献[19]模型相比,ISRUC-S3数据集上N1阶段的F1分数提升了4.9%,在N2阶段提升了2.2%。而N2阶段的数据量是N1阶段的2.15倍,表明MJ-Sleep能够更好地缓解类别不平衡问题。与MJ-Sleep模型在ISRUC-S1数据集上的出色分类结果证实了其能够准确执行睡眠分期任务,同时具有一定程度的泛化能力。



表4 先进成果与MJ-Sleep性能比较  
Table 4 Performance comparison between advanced results and MJ-Sleep %

数据集	模型	总体性能指标			各阶段F1分数				
		A	F <sub>1</sub>	K	W	N1	N2	N3	REM
ISRUC-S3	DeepSleep <sup>[11]</sup>	78.8	77.9	73.0	88.7	<u>60.2</u>	74.6	85.8	80.8
	U-Sleep <sup>[12]</sup>	77.0	76.4	—	<u>90.0</u>	55.0	78.0	74.0	<u>85.0</u>
	SalientSleepNet <sup>[13]</sup>	79.9	78.6	74.2	86.0	58.9	79.3	88.6	80.2
	GraphSleepNet <sup>[17]</sup>	79.9	78.7	74.1	87.8	57.4	77.6	86.4	84.1
	MSTGCN <sup>[18]</sup>	82.1	80.8	76.9	89.4	59.6	80.6	89.0	85.6
	JK-STGCN <sup>[19]</sup>	<u>83.1</u>	<u>81.4</u>	<u>78.2</u>	<u>90.0</u>	59.8	<u>82.6</u>	<u>90.1</u>	84.7
	MJ-Sleep	<b>85.3</b>	<b>83.8</b>	<b>81.0</b>	<b>90.5</b>	<b>64.7</b>	<b>84.8</b>	<b>92.0</b>	<b>87.4</b>
ISRUC-S1	DeepSleep <sup>[11]</sup>	71.7	69.1	63.8	82.3	46.6	73.8	80.9	62.1
	U-Sleep <sup>[12]</sup>	77.0	77.0	—	89.0	52.0	79.0	77.0	<b>88.0</b>
	SalientSleepNet <sup>[13]</sup>	81.5	<u>80.1</u>	76.2	<u>89.9</u>	<u>57.0</u>	80.0	87.8	85.7
	GraphSleepNet <sup>[17]</sup>	78.6	75.4	72.3	88.4	43.7	77.5	83.8	83.5
	MSTGCN <sup>[18]</sup>	80.4	78.5	74.8	88.7	54.5	79.1	87.2	83.2
	JK-STGCN <sup>[19]</sup>	<u>82.0</u>	79.8	<u>76.7</u>	89.5	55.0	<u>81.1</u>	<u>88.3</u>	<u>85.0</u>
	MJ-Sleep	<b>83.2</b>	<b>81.9</b>	<b>78.4</b>	<b>90.1</b>	<b>61.5</b>	<b>82.8</b>	<b>90.3</b>	<u>84.8</u>

注: 粗体表示算法的最佳结果, 下划线表示次优结果。

## 2.4 消融实验

### 2.4.1 模块有效性检验

为了进一步验证模块的有效性, 进行了广泛的消融实验, 具体描述如下:

1、变体A(基本模型): 基本模型是一个独立的自适应图学习STGCN模型。

2、变体B(+RF): 在变体A的基础上添加RF模块, 融合时空信息。

3、变体C(+时频分支): 在变体B的基础上添加时频分支, 丰富时频信息。

4、变体D(+SFE模块): 在变体C的基础上添加时空特征提取模块, 丰富时空特征。

5、变体E(+TFF模块): 在变体D的基础上添加时频特征提取模块, 整合不同维度的时频特征。

6、变体F(+FFM模块): 在变体E的基础上添加了特征融合模块, 融合时频、时空特征。

图6为消融实验的结果, 证明了模型在提高睡眠分期阶段的有效性。将RF模块添加到基线模型可以获取信号不同通道之间的联系, 捕捉信号的过渡特征, 有利于混淆阶段的准确分类, 使N1

阶段分类的准确率得到了提升。添加SFE和TFF模块可以获取更深层的特征信息, 有利于模型区分各个阶段的特征, 总体上提高了模型分类准确度。结合FEF模块进一步提高了模型的性能。归因于融合模块为睡眠分期任务结合了更多的活跃信息。

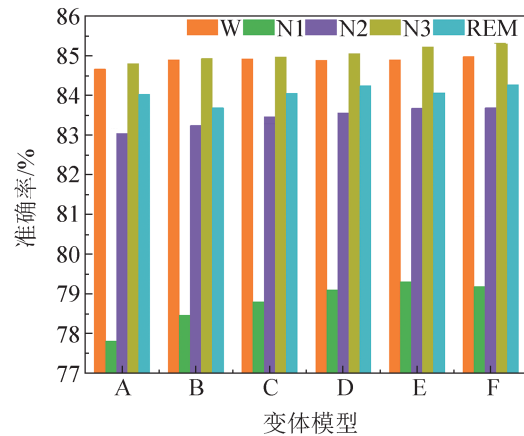


图6 模块有效性验证

Fig. 6 Module validity verification

### 2.4.2 窗函数

为了更好地捕捉信号特征, 对STFT使用的窗函数进行讨论。一般使用第一旁瓣衰减大, 旁瓣峰值衰减快的窗函数可以减小分割过程中产生的

频谱泄露问题，但具有这种特征的窗函数往往主瓣宽度较大，会带来一定的副作用。为此，本文使用了多种窗函数，如布莱克曼窗函数、布莱克曼-哈里斯窗函数、汉明窗函数和矩形窗函数。混淆矩阵的结果如图7所示。结果表明，不同的窗函数均可取得良好的效果。矩形窗函数更擅长提取频率信息，与睡眠阶段的划分依据相吻合。N1阶段和REM阶段在睡眠的680~700个epoch中有多次转换。矩形窗根据频率变换的敏感性，在REM阶段的分类精度明显优于其他窗函数。

### 2.4.3 时频特征提取

特征提取的效果在一定程度上决定着睡眠分期结果的上限。由于数据集存在类别不平衡、参数量小等特点，本文尝试使用不同的基本框架来处理睡眠数据集，结果如图8所示。AlexNet采用

的简单结构将脑电信号中的时频特征捕获，将计算的特征权重信息用于模型训练，取得了比其他方法更高的准确率，原因在于PSG信号是反映神经活动的电信号，内部结构简单，不需要复杂的特征提取结构。表明了睡眠分期任务有希望在小型的移动端部署。

### 2.4.4 时空特征提取

DC借助较大的感受野获取前后信号特征，RNN利用信号先后的关联性获取信号的过渡关系。两者在原理上并不矛盾，本文将2种方法结合使用，获取更丰富的特征信息。由于处理时序信号的基本单元有很多，例如，双向门循环单元(Bi-GRU)、门循环单元(GRU)和GRU\_ATT。为了获得更精确的分类结果，本文通过大量的实验来寻找最佳的TFF模块集成方法。

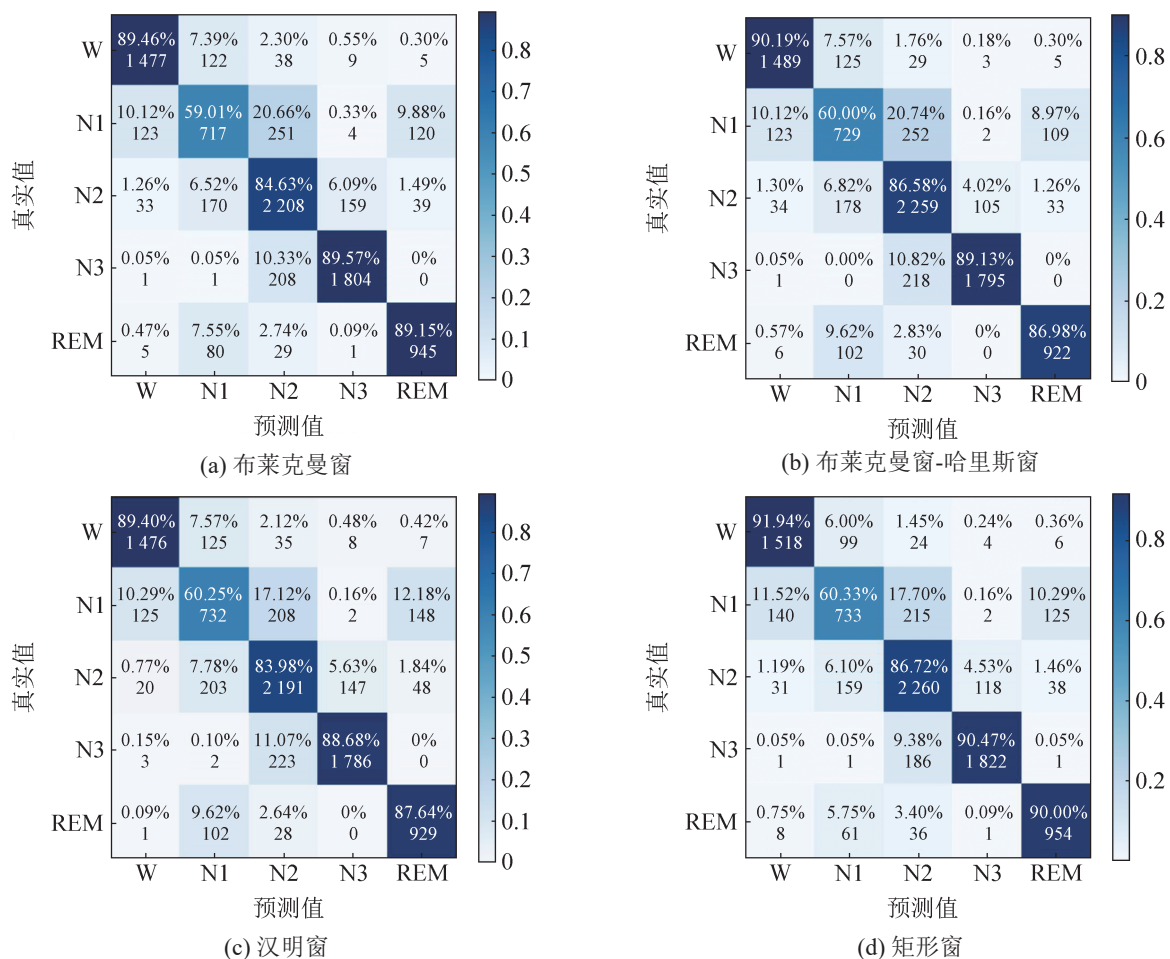


图7 混淆矩阵

Fig. 7 Confusion matrix

图9为使用不同结构的10折交叉验证的准确率折线图。GRU\_ATT通过权重操作来调整特征重要性, DC负责扩大感受野, 捕捉序列间转换关系, 二者结合有利于获取更加丰富的信息。虽然BiGRU+GRU\_ATT+DC也取得了不错的效果, 但考虑到精度和计算复杂度之间的平衡, 最终选择GRU\_ATT+DC作为TFF模块的基本组成来提取时间信息。

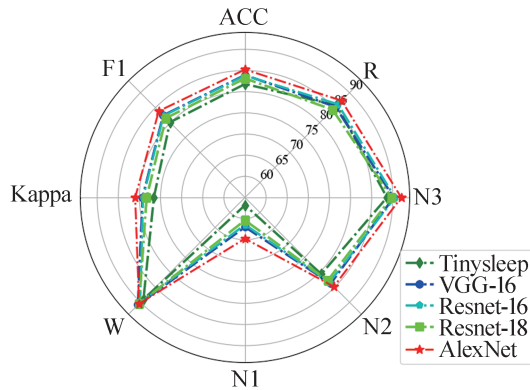


图8 时频特征提取

Fig. 8 Time-frequency feature extraction

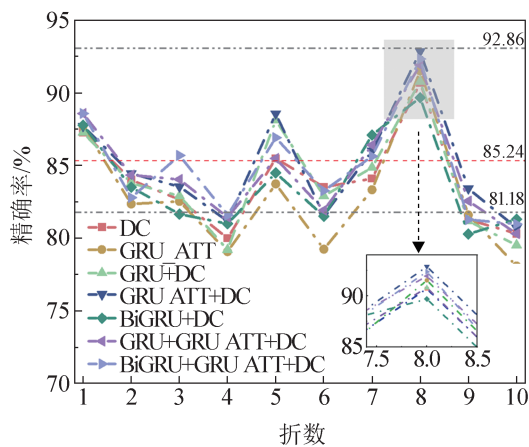


图9 时空特征提取

Fig. 9 Spatiotemporal Feature Extraction

### 3 结论

本文提出了一种基于自动睡眠分期的多模态残差时空融合模型MJ-Sleep, 用于多模态生理信号的睡眠分期任务。将残差图卷积和卷积神经网络相结合, 提取信号的时空信息和时频信息用于睡眠分期任务, 增加了模型的表征能力。借助残

差融合结构增加相邻时期生物信号通道的联系。该模型通过结合传统的循环单元和DC来获得更丰富的上下文信息。采用DC、DSC等操作, 减少模型在处理数据集所需的运行的时间。同时, 模型在ISRUC-S3数据集上的实验结果表明, 该模型优于其他先进模型, 且在N1阶段分类性能表现良好。在ISRUC-S1子集上的实验结果表明, 本文提出的模型具有一定的泛化能力。大量的消融实验, 验证了多模型和特征提取对模型性能的影响。与欧氏输入的深度学习方法和现有的图卷积方法相比, 该模型取得了更好的分类效果。尽管MJ-Sleep在睡眠阶段分类方面表现出很高的性能, 但仍有一些改进的空间。在该模型中, 残差图卷积操作发生在每个epoch中, 相邻epoch之间没有信号的交流。未来可考虑改进跳跃知识模块, 利用相邻历元交叉跳跃操作聚合不同epoch之间的信号特征, 学习更加丰富的过渡规则。

### 参考文献:

- [1] Kohn T P, Kohn J R, Haney N M, et al. The Effect of Sleep on Men's Health[J]. *Translational Andrology and Urology*, 2020, 9(S2): S178-S185.
- [2] Antoine Guillot, Valentin Thorey. RobustSleepNet: Transfer Learning for Automated Sleep Staging at Scale [J]. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2021, 29: 1441-1451.
- [3] Wolpert E A. A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects[J]. *Archives of General Psychiatry*, 1969, 20(2): 246-247.
- [4] Xiao Meng, Yan Hong, Song Jinzhong, et al. Sleep Stages Classification Based on Heart Rate Variability and Random Forest[J]. *Biomedical Signal Processing and Control*, 2013, 8(6): 624-633.
- [5] Aayushi Bansal, Rewa Sharma, Mamta Kathuria. A Systematic Review on Data Scarcity Problem in Deep Learning: Solution and Applications[J]. *ACM Computing Surveys*, 2022, 54(10S): 208.
- [6] Md Moshayur Rahman, Mohammed Imamul Hassan Bhuiyan, Ahnaf Rashik Hassan. Sleep Stage Classification Using Single-channel EOG[J]. *Computers in Biology and Medicine*, 2018, 102: 211-220.
- [7] A Jameer Basha, B Saravana Balaji, Poornima S, et al.

- Support Vector Machine and Simple Recurrent Network Based Automatic Sleep Stage Classification of Fuzzy Kernel[J]. *Journal of Ambient Intelligence and Humanized Computing*, 2021, 12(6): 6189-6197.
- [8] Pei Wei, Li Yan, Siuly S, et al. A Hybrid Deep Learning Scheme for Multi-channel Sleep Stage Classification[J]. *Computers, Materials & Continua*, 2022, 71(1): 889-905.
- [9] Phan H, Andreotti F, Cooray N, et al. SeqSleepNet: End-to-end Hierarchical Recurrent Neural Network for Sequence-to-sequence Automatic Sleep Staging[J]. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2019, 27(3): 400-410.
- [10] Neng Wenpeng, Lu Jun, Xu Lei. CRRSleepNet: A Hybrid Relational Inductive Biases Network for Automatic Sleep Stage Classification on Raw Single-channel EEG[J]. *Brain Sciences*, 2021, 11(4): 456.
- [11] Supratak A, Dong Hao, Wu Chao, et al. DeepSleepNet: A Model for Automatic Sleep Stage Scoring Based on Raw Single-channel EEG[J]. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2017, 25(11): 1998-2008.
- [12] Mathias Perslev, Sune Darkner, Lykke Kempfner, et al. U-sleep: Resilient High-frequency Sleep Staging[J]. *npj Digital Medicine*, 2021, 4(1): 72.
- [13] Jia Z, Lin Y, Wang J, et al. Salientsleepnet: Multimodal Salient Wave Detection Network for Sleep Staging[C]//2021 30th International Joint Conference on Artificial Intelligence (IJCAI). Montreal: Morgan Kaufmann, 2021: 1-10.
- [14] Kong G, Li C, Peng H, et al. EEG-based Sleep Stage Classification via Neural Architecture Search[J]. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2023, 31: 1075-1085.
- [15] Zhou Jie, Cui Ganqu, Hu Shengding, et al. Graph Neural Networks: A Review of Methods and Applications[J]. *AI Open*, 2020, 1: 57-81.
- [16] Zhang Si, Tong Hanghang, Xu Jiejun, et al. Graph Convolutional Networks: A Comprehensive Review[J]. *Computational Social Networks*, 2019, 6(1): 11.
- [17] Jia Ziyu, Lin Youfang, Wang Jing, et al. GraphSleepNet: Adaptive Spatial-temporal Graph Convolutional Networks for Sleep Stage Classification[C]//Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence. New York, NY, USA: ACM, 2021: 1324-1330.
- [18] Jia Ziyu, Lin Youfang, Wang Jing, et al. Multi-view Spatial-temporal Graph Convolutional Networks with Domain Generalization for Sleep Stage Classification[J]. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2021, 29: 1977-1986.
- [19] Ji Xiaopeng, Li Yan, Wen Peng. Jumping Knowledge Based Spatial-temporal Graph Convolutional Networks for Automatic Sleep Stage Classification[J]. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2022, 30: 1464-1472.
- [20] Li Miaofen, Liu Youmin, Zhi Shaodan, et al. Short-time Fourier Transform Using Odd Symmetric Window Function[J]. *Journal of Dynamics, Monitoring and Diagnostics*, 2022, 1(1): 37-45.
- [21] Zhou Haoran, Feng Yidan, Fang Mingsheng, et al. Adaptive Graph Convolution for Point Cloud Analysis [C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway, NJ, USA: IEEE, 2021: 4945-4954.
- [22] Akara Supratak, Guo Yike. TinySleepNet: An Efficient Deep Learning Model for Sleep Stage Scoring Based on Raw Single-channel EEG[C]//2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). Piscataway, NJ, USA: IEEE, 2020: 641-644.
- [23] Cai Xiyang, Jia Ziyu, Jiao Zehui. Two-stream Squeeze-and-excitation Network for Multi-modal Sleep Staging [C]//2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). Piscataway, NJ, USA: IEEE, 2021: 1262-1265.
- [24] Xu Ziliang, Yang Xuejuan, Sun Jinbo, et al. Sleep Stage Classification Using Time-frequency Spectra from Consecutive Multi-time Points[J]. *Frontiers in Neuroscience*, 2020, 14: 14.
- [25] Cai Qing, Gao Zhongke, An Jianpeng, et al. A Graph-temporal Fused Dual-input Convolutional Neural Network for Detecting Sleep Stages from EEG Signals [J]. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 2021, 68(2): 777-781.
- [26] Sirvan Khalighi, Teresa Sousa, José Moutinho Santos, et al. ISRUC-sleep: A Comprehensive Public Dataset for Sleep Researchers[J]. *Computer Methods and Programs in Biomedicine*, 2016, 124: 180-192.