

一种基于双模态的睡眠分期研究

王亚群^{1,2,3}, 杨青^{1,2,3}, 文斗^{1,2,3}, 王莹^{1,2,3}, 王翔宇^{1,2,3}

(1. 华中师范大学 人工智能与智慧学习湖北省重点实验室 湖北 武汉 430079;

2. 华中师范大学 计算机学院 湖北 武汉 430079;

3. 国家语言资源监测与研究网络媒体中心 湖北 武汉 430079)

摘要: 现有研究普遍专注于单个信号,忽略了多模态信号在特定睡眠阶段提供的睡眠信息,以及提取睡眠信号时,随着网络的加深导致重要信息丢失,从而降低模型分类的能力。针对上述问题,设计了一种基于脑电图(electroencephalogram, EEG)、眼电图(electrooculogram, EOG)的深度神经网络模型,以端到端的方式进行睡眠分期,称为 MCNN-LSTMs 模型。多尺度卷积神经网络(multiscale convolutional neural network, MCNN)用于提取脑电信号和眼电信号的特征,双层长短期记忆(long short-term memory, LSTM)神经网络对提取到的特征进行融合,然后输入分类器进行睡眠分期。在公共数据集 Sleep-EDF 上评估了所设计的方法在睡眠分期上的性能。实验表明,使用两种通道(EEG-EOG)时,在 Sleep-EDF-20 数据集上的分类准确率最高为 92.60%,在 Sleep-EDF-78 数据集上的分类准确率最高为 91.10%,优于单通道信号以及其他对比方法。所提方法验证了多种信号对睡眠分期的有效性,并为研究睡眠分期提供了重要思路。

关键词: 睡眠分期; 多模态; 卷积神经网络; 多层网络; LSTM

中图分类号: TP399

文献标志码: A

文章编号: 1671-6841(2025)03-0081-07

DOI: 10.13705/j.issn.1671-6841.2023257

A Study on Sleep Staging Based on Bimodal Analysis

WANG Yaqun^{1,2,3}, YANG Qing^{1,2,3}, WEN Dou^{1,2,3}, WANG Ying^{1,2,3}, WANG Xiangyu^{1,2,3}

(1. Hubei Provincial Key Laboratory of Artificial Intelligence and Smart Learning, Central China Normal

University, Wuhan 430079, China; 2. School of Computer, Central China Normal University,

Wuhan 430079, China; 3. National Language Resources Monitoring & Research Center for

Network Media, Wuhan 430079, China)

Abstract: The existing research generally focused on a single signal, ignoring the sleep information provided by other model signals in a specific sleep stage. The loss of important information, as the network deepened when extracting sleep signals, could reduce the classification ability of the model in view of the problems, a deep neural network model based on electroencephalogram (EEG) and electrooculogram (EOG) was designed to sleep stage in an end-to-end manner, which was called MCNN LSTMs model. The features extracted from EEG and EOG signals were fused by a two-layer long short term memory (LSTM) neural network after multi-scale convolutional neural network, and then input into the classifier for sleep staging. The performance of the proposed method on sleep staging was evaluated on a public sleep EDF dataset. Experiments showed that when two channels (EEG-EOG) were used, the classification accuracy reached 92.60% on Sleep-EDF-20 dataset and 91.10% on Sleep-EDF-78 dataset, which was better than single channel signal and comparison methods. The effectiveness of multiple signals for sleep staging was verified, and an important idea for the study of sleep staging was provided.

Key words: sleep stage; multimodal; convolutional neural network; multi-layer network; LSTM

收稿日期: 2023-11-15

基金项目: 湖北省重点研发计划项目(2020BAB017); 武汉市科技计划项目(2019010701011392); 国家语委科研中心项目(ZDI135-135)

第一作者: 王亚群(1998—), 女, 硕士研究生, 主要从事数据挖掘及计算机技术应用研究, E-mail: 2371975519@qq.com。

通信作者: 杨青(1965—), 女, 副教授, 主要从事数据挖掘及计算机技术应用研究, E-mail: yangqing@cenu.edu.cn。

0 引言

睡眠几乎占我们生命活动的三分之一,良好的睡眠对于维持一个人的身心健康至关重要。现在越来越多的人作息不规律,导致了各种睡眠障碍的发生。但大多数睡眠障碍并不容易被检测。多导睡眠图(polysomnography, PSG)是测量睡眠质量和睡眠障碍的“金标准”^[1]。PSG 包括脑电图、眼电图、肌电图(electromyogram, EMG)和心电图(electrocardiogram, ECG)。从这些信号中收集数据后,使用美国睡眠医学学会(American academy of sleep medicine, AASM)的标准将数据分割为 30 s 的时间段(称为 epoch),并划分不同的睡眠阶段:唤醒期(W)、快速眼动期(REM)、非快速眼动期(NREM)(包括 N1, N2, N3 期, N1 期和 N2 期为浅睡眠期, N3 期为深睡眠期)。

早期,睡眠分期这一任务由专家进行肉眼判定,然而手动划分睡眠阶段的过程对睡眠专家的专业知识要求较高,并耗费大量时间与人力,难以大规模应用^[2]。因此,一个自动睡眠分期系统为睡眠技术人员提供长期睡眠监测和分期的辅助具有重要意义。机器学习的出现为睡眠分期研究提供了新的希望。但传统的机器学习方法是基于特征工程的,需要专家依靠先验知识手动设计特征并提取最相关的特征,然后,基于提取的特征,使用支持向量机(support vector machine, SVM)^[3]、随机森林(random forest, RF)^[4-5]、决策树(decision tree, DT)^[6]等传统的机器学习算法对睡眠阶段进行分类。虽然这些方法能达到比较高的分类性能,但是如果数据集庞大,手动设计特征将是一个巨大的工程。因此,基于自动提取特征的深度学习方法逐渐成为睡眠分期领域的主要方法,其与传统机器学习相比,能够减少特征依赖^[7]。Supratak 等^[8]提出了 DeepSleepNet 睡眠分期网络,针对原始单通道 EEG 数据,开创性地使用两个 CNN 分支分别进行时、频特征的提取,然后使用双向长短时记忆网络从 EEG 历元中自动学习睡眠阶段之间的转换规则并进行训练,该模型预测总体准确率达到 82%。Mousavi 等^[9]设计了 SleepEEG-Net 神经网络模型,该模型以原始单通道 EEG 信号作为输入,使用两个 CNN 分支分别进行特征的提取,然后利用双向 RNN 进行序列学习,最后预测准确率达到 84.26%。Yang 等^[10]使用 1D-CNN 自动从原始脑电信号中提取特征,然后利用 HMM 纠正相邻脑电时期的睡眠阶段过渡先验信息,该模型总

体准确率达到 83.98%。Fan 等^[11]从原始 EOG 信号中提取特征,利于递归神经网络捕获长期顺序信息,该模型预测总体准确率达到 81.20%。Li 等^[12]在 CNN 中添加了一个注意力模块来学习 EEG 信号局部序列的权重,然后使用两层 Bi-LSTM 对连续时期的全局相关性进行编码,该模型总体准确率达到 80.80%。Toma 等^[13]采用 Fpz-Cz 的 EEG 和 EOG 信号作为输入,提出了一种支持多源数据融合的卷积递归神经网络 CRNN,该模型总体准确率达到 90.30%。Sharma 等^[14]使用 EEG、EMG 和 EOG 信号开发了一种新的机器学习模型,将 PSG 信号分割成多个 30 s 时间段,使用正交滤波器组对每个时间段进行五级一维小波分解,并计算每个子带基于 Tsallis 熵的特征,在 SHHS 数据集上取得了 90.70% 的准确率。然而近些年的研究大多数均从单个 EEG 或 EOG 信号对睡眠分期进行研究,忽略了其他生理信号对睡眠分期结果的影响,尽管已有一些研究对多个信号进行探索,但大多数依然依赖于人工设计的特征,而针对多个信号进行端到端的研究目前较少,同时在睡眠分期领域的准确率依旧存在着很大的上升空间。采用双分支的卷积神经网络作为特征提取确实能够有效地提取原始信号的时频特征,但同时也使模型的参数量巨大,训练时间长。

为了解决上述问题,本文设计了一种新的端到端的双模态深度学习模型。在该模型中,分别从原始 EEG 和 EOG 信号中自动学习不同阶段的判别性特征,而不是手动提取特征。为使模型更加轻量,本文使用了单分支的卷积神经网络对原始信号进行特征提取,特征提取模块由多个卷积层和池化层串联组成,为了捕捉表征不同睡眠阶段的 EEG 和 EOG 信号在不同频段的变化,设计了一个单分支的多尺度卷积神经网络融合浅层特征和深层特征。此外,为了融合 EEG 和 EOG 信号,并充分学习 EEG 和 EOG 特征之间的关系,利用双层 LSTM 网络融合学习到的特征表示并进行分类。

1 基于 MCNN-LSTMs 模型的设计

本文使用的睡眠分期框架如图 1 所示。模型由三部分组成:第一部分是特征提取模块,由单个分支卷积层组成,利用五个堆叠的卷积层提取睡眠 EEG 和 EOG 的信号特征。第二部分是双层 LSTM 特征融合模块,使用双层 LSTM 模型对 EEG 和 EOG 特征进行融合,获得这两种模态的联合特征表示,并提高分类精度。第三部分是分类器分类模块,将特征输

出到全连接层中,由 Softmax 输出最终睡眠分期结果。

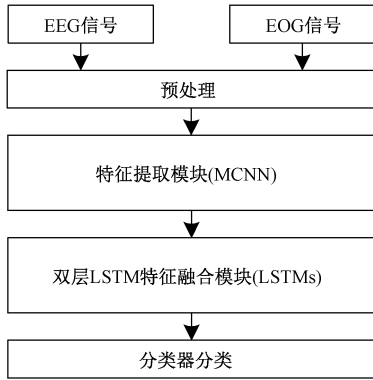


图 1 睡眠分期框架

Figure 1 Sleep staging framework

1.1 预处理

本文模型的输入是一个 30 s(epoch)的 EEG 和 EOG 历元。为了验证本文模型针对原始单通道脑电和眼电数据对睡眠阶段分类的性能,不对原始数据作复杂的处理,以保证实验的泛化性,预处理步骤如下。

1) 将连续的原始单通道 EEG 和 EOG 分割为以 30 s 为周期的一系列历元,并为每个历元(即睡

眠阶段)分配一个标签。

2) 对 30 s 的 EEG 和 EOG 历元进行 z-score 标准化处理,将原始数据集归一化为均值为 0、方差为 1 的数据集。

1.2 EEG 和 EOG 特征提取

本文使用卷积神经网络提取脑电信号和眼电信号的特征,特征提取部分采用的是单个分支的 CNN,而不是两个具有不同大小的卷积核的分支,其借鉴 VGGNet^[15]的模型设计思想,其中 VGG16 网络相比 AlexNet 网络的一个改进是采用连续的几个 3×3 的卷积核代替 AlexNet 中的较大卷积核(11×11, 7×7,5×5)。对于相同的感受野,采用堆积的小卷积核优于采用大的卷积核,因为多层非线性层可以增加网络深度来保证学习更复杂的模式,而且代价更小(参数更少),本文采用了五个堆叠的卷积层提取睡眠 EEG 和 EOG 的特征。图 2 中的 Conv1d(128, 8, 1)是指使用具有 128 个滤波器的一维卷积层,其内核大小为 8,步幅为 1。类似地,MaxPool1d(8, 8)指的是内核大小为 8,步幅为 8 的一维 Max 池化层。此外,为了防止过拟合和增强模型的泛化能力,在网络中加入了 Dropout 方法。

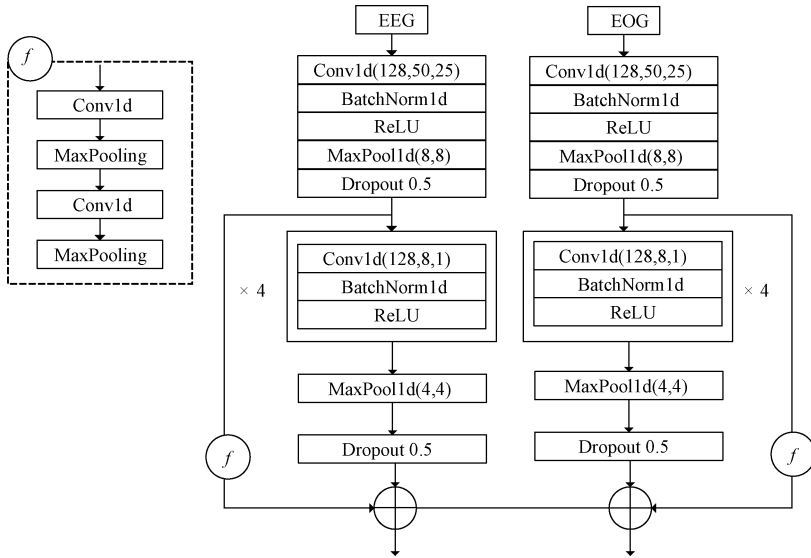


图 2 特征提取模块 MCNN

Figure 2 Feature extraction module MCNN

由于随着卷积网络的深入,提取的特征对细节的描述能力会随之变差,而浅层特征包含更多的细节信息,所以本文设计了一个映射块 f ,用于提取更精确的脑电和眼电特征。映射块 f 由几个卷积单元组成,最终的特征映射结果是由中间映射块的输出与最后一个卷积单元的输出叠加得到的,然后将用于提取 EEG 和 EOG 特征的两个分支的最终输出特

征输入双层 LSTM 层进行特征融合。通过使用多尺度卷积,本文设计的模型可以从全局中捕捉到突出的波形。例如,在 N1 周期内,可以在低振幅的混合频率中捕捉到缓慢的眼球运动等细节,特征提取模块 MCNN 如图 2 所示。

假设单通道 EEG 有 N 个 30 s 的 EEG 历元 $\{x_1, x_2, \dots, x_n\}$,单通道 EOG 有 N 个 30 s 的 EOG 历元

$\{y_1, y_1, \dots, y_n\}$, 从第 i 个 EEG 历元 x_i 中提取第 i 个特征 a_i 的表示如公式(1), 从第 j 个 EOG 历元 y_j 中提取第 j 个特征 b_j 的表示如公式(2),

$$a_i = CNN_\theta(x_i), \quad (1)$$

$$b_j = CNN_\theta(y_j), \quad (2)$$

式中: CNN_θ 表示将单通道脑电历元和眼电历元转化为特征向量的 CNN, θ 为 CNN 的可学习参数。 a_i 和 b_j 的大小取决于输入 EEG 和 EOG 的采样率。

设不经映射块 f 的单分支 CNN 的输出为 Z_2 , 中间映射块 f 的输出为 $f(Z_1, W)$ 。通过映射 f 得到的矢量的维数应该与 Z_2 一致。在包含了来自低层的细或粗的特征信息后, 两者求和, 最终的输出 Z_{out} 可以表示为

$$Z_{out} = Z_2 + f(Z_1, W), \quad (3)$$

其中: W 是要学习的权重矩阵。这样, 所得到的特征向量就被期望具有高语义相似度和高细粒度相似度的特征。

1.3 基于双层 LSTM 的特征融合

多模态融合是从多种来源接收的数据中筛选、提取和组合所需特征的过程, 融合的方法主要有三种: 特征级融合、决策级融合和模型级融合。特征级融合^[16](也称为早期融合)直接将每个模态的所有特征合并为一个特征向量, 再将其输入分类算法中, 其局限是不能对复杂的关系进行建模, 且很难学习不同模态之间的相互关系。决策级融合^[16](也称为后期融合)首先对各模态的特征进行独立处理和分类, 然后将分类结果进行融合, 形成最终的决策向量。然而, 决策级融合的局限是忽略了不同模态之间的相互作用。与特征级和决策级融合相比, 模型级融合学习了模型内部的多模态交互, 使神经网络具有更大的优势^[17], 是更深层次的融合方法, 它的实现主要取决于使用的融合模型。LSTM 是当前比较流行的睡眠分期网络模型, 因为它能够通过递归结构学习睡眠过程神经信号的动态时间性。受文献[18]的启发, 本文选用双层 LSTM 模型用于融合 EEG 和 EOG 的特征, 将多层网络与传统的 LSTM 模型相结合来充分学习 EEG 和 EOG 特征之间的关系。融合思路如下: 将 EEG 特征输入第一层 LSTM 得到的是每个神经元的隐藏层状态, 然后将 EOG 特征与第一层得到的隐藏层状态相拼接输入第二层 LSTM, 得到第二层每个神经元的隐藏层状态, 最后将融合后的特征输入全连接 (fully connected, FC) 层, 使用 Softmax 分类器得到最终的睡眠分期结果。双层 LSTM 融合模型 (LSTMs) 如图 3 所示。

第一层 LSTM 结构中的相关公式为

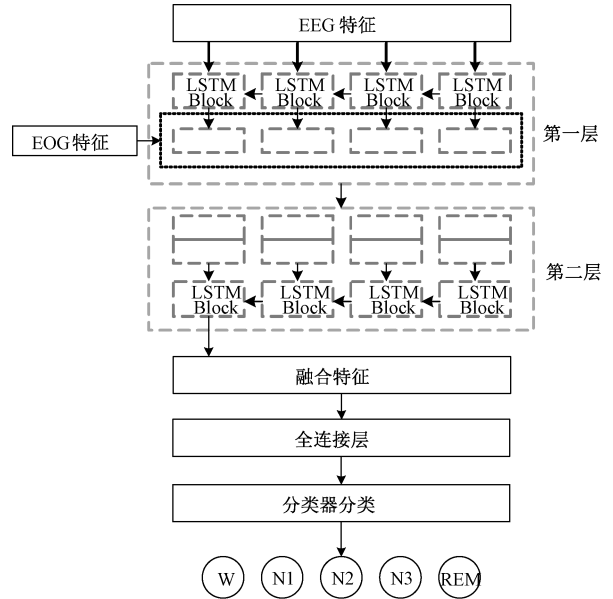


图3 双层 LSTM 融合模型 (LSTMs)

Figure 3 Double layer LSTM fusion model (LSTMs)

$$f_i^1 = \sigma(W_f^1[h_{i-1}^1, g_i^E] + b_f^1), \quad (4)$$

$$i_i^1 = \sigma(W_i^1[h_{i-1}^1, g_i^T] + b_i^1), \quad (5)$$

$$C_i^1 = \tanh(W_c^1[h_{i-1}^1, g_i^T] + b_c^1), \quad (6)$$

$$o_i^1 = \sigma(W_o^1[h_{i-1}^1, g_i^T] + b_o^1), \quad (7)$$

$$h_i^1 = o_i^1 \times \tanh(C_i^1), \quad (8)$$

式中: g_i^E 为输入的 EEG 特征; C_i^1 为细胞单元暂时性的状态; f_i^1 为 LSTM 遗忘门信号; i_i^1 为输入门信号; o_i^1 为输出门信号; h_i^1 为 LSTM 网络最后的输出信号; $W_f^1, W_i^1, W_c^1, W_o^1$ 为权重矩阵; $b_f^1, b_i^1, b_c^1, b_o^1$ 为偏置项。

第二层 LSTM 将 EOG 特征和第一层 LSTM 的输出融合, 作为第二层 LSTM 的输入。所用公式为

$$f_i^2 = \sigma(W_f^2[h_{i-1}^2, h_i^1, g_i^O] + b_f^2), \quad (9)$$

式中: g_i^O 为输入的 EOG 特征; f_i^2 为第二层 LSTM 的输出。最后, 将融合之后的特征输入到全连接层, 使用 Softmax 分类器输出分类结果。

2 实验与结果分析

2.1 实验数据和设置说明

本文使用的公共数据集是 Sleep-EDF-20 和 Sleep-EDF-78^[19], 各个睡眠阶段 30 s 时期的样本数量如表 1 所示。Sleep-EDF-20 是包含 20 名受试者的数据文件, Sleep-EDF-78 是 Sleep-EDF-20 的扩展版本, 包含了 78 名受试者两晚的 PSG 数据。每个 PSG 数据包含两个脑电图通道 (Fpz-Cz, Pz-Oz), 一个水平的 EOG 通道和一个 EMG 通道, 所有 EEG 和

EOG的采样率均为100 Hz。睡眠专家根据R&K标准^[20],将这些记录手动分为八类(W、N1、N2、N3、N4、REM、MOVEMENT、UNKNOWN)。实验中采用AASM标准将N3和N4阶段合并为一个睡眠阶段N3,同时舍弃MOVEMENT和UNKNOWN期,因为它们不属于睡眠阶段。实验使用EEG的Fpz-Cz通道和EOG通道在这两个数据集上评估本文模型,并验证EEG和EOG联合特征优于单一的EEG或EOG特征这一假设。数据集中每个睡眠阶段的30 s期的数量如表1所示。

表1 每个睡眠阶段的30秒期的数量

Table 1 Number of 30 second periods in each sleep stage

单位:个

数据集	W	N1	N2	N3	REM
Sleep-EDF20	8 285	2 804	17 799	5 703	7 717
Sleep-EDF78	65 951	21 522	69 132	13 039	25 835

本实验在Sleep-EDF20和Sleep-EDF78数据集上使用十折交叉验证方法来评估本文模型性能。另外,由于Sleep-EDF数据库中各个睡眠分期阶段样本数差距较大。比如W和N2阶段的样本数量远大于其他阶段的样本数。为了解决这个问题,本文采用了加权交叉熵损失函数,将N1阶段和N3阶段的权重大小设置为1.5,其他阶段的权重大小设置为1,在一定程度上降低类不平衡问题。除此之外,对数据集进行合成少数类过采样技术SMOTE^[21]处理,通过考虑现有少数类样本之间的相似性来生成合成数据点,让每个类的睡眠阶段样本数趋于平衡。实验使用Adam优化器训练100个历元,学习率设为 10^{-4} ,Adam优化器的三个参数 β_1 、 β_2 、 ε 分别为0.9、0.999和 10^{-8} ,训练使用的mini batch大小为10,序列长度为20。

2.2 评价方法

本文采用了三个指标来评估各种模型对睡眠分期的性能,即每类 $F1$ 值、宏观平均 $F1$ 值(MF1)、总体精度(ACC)。 $F1$ 、MF1、ACC的定义为

$$F1_c = 2 \times \frac{P_c \times R_c}{P_c + R_c}, \quad (10)$$

$$MF1 = \frac{\sum_{c=1}^C F1_c}{C}, \quad (11)$$

$$ACC = \frac{\sum_{c=1}^C TP_c}{N}, \quad (12)$$

式中: TP_c 为 C 类真阳性; $F1_c$ 为 C 类的每类 $F1$ 分

数; C 为睡眠阶段数; N 为训练历元总数。

2.3 实验结果及分析

为了评估本文设计的模型的性能,实验采用了十折交叉验证,将每个数据集中的受试者分为10组。例如,在含有10名被试者的Sleep-EDF-20数据集上,每轮选取1组被试者作为测试数据,剩下的9组作为训练数据。最后将10轮测试样本的预测睡眠阶段合并,计算各种性能指标。为了证明同时使用原始单通道EEG和EOG信号在睡眠分期中的优势,分别对单通道EEG和单通道EOG进行对比实验,不同的是,在对单个信号进行分析时,将双层LSTM模型改为传统的LSTM模型。图4和图5分别展示了本文模型在训练过程中使用单通道EEG(Fpz-Cz)、使用单通道EOG与使用EEG的Fpz-Cz通道+EOG通道共同作用的准确率变化,从中可以看出本文模型在训练过程中表现出稳定的性能,在两个数据集上均表现较好,并且EEG的Fpz-Cz通道和EOG通道在MCNN-LSTMs模型上共同作用的准确率明显高于单通道的EEG和单通道的EOG,所以验证了EEG和EOG同时作用的效果是大于单个EEG或EOG的。

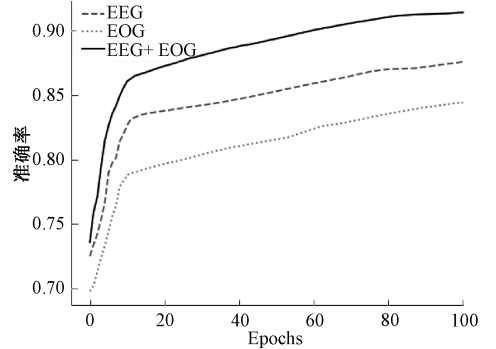


图4 在Sleep-EDF-20数据集进行训练的准确率比较
Figure 4 Comparison of accuracy of training on the Sleep EDF-20 dataset

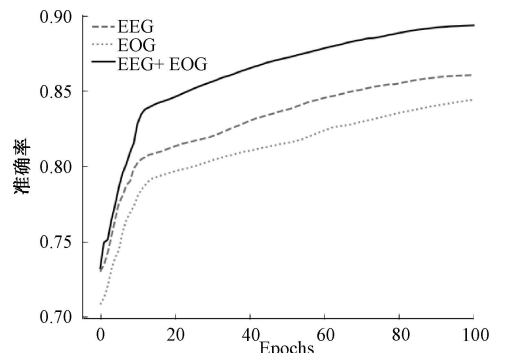


图5 在Sleep-EDF-78数据集进行训练的准确率比较
Figure 5 Comparison of accuracy of training on the Sleep EDF-78 dataset

此外,本文利用三种不同的融合策略来融合提取到的 EEG 特征和 EOG 特征,分别为特征级融合、决策级融合以及本文基于 LSTMs 的模型级融合。实验结果表明,通过特征级融合后,睡眠分期结果的 $MF1$ 和准确率分别为 0.731 和 0.816。通过决策级融合后,睡眠分期结果的 $MF1$ 和准确率分别为 0.747 和 0.843。而本文基于 LSTMs 的模型级融合

后,睡眠分期结果的 $MF1$ 和准确率分别为 0.773 和 0.911。由此可见,本文设计的 LSTMs 融合方法获得了更好的性能,证明本文设计的双层 LSTM 融合模型的可行性。

本文方法与另外一些利用深度学习模型从单通道 EEG(Fpz-cz)或 EOG 中提取特征的方法在 ACC 、 $MF1$ 、和每类 $F1$ 值,以及训练时间上的对比如表 2 所示。

表 2 MCNN-LSTMs (本文方法)与其他研究方法结果对比

Table 2 Comparison of results between MCNN-LSTMs (the method proposed in this article) and other research methods

数据集	现有方法(研究的信号)	每类 $F1$ 分数					总体评价指标		训练时间/h
		W	N1	N2	N3	REM	ACC	$MF1$	
Sleep-EDF-20	DeepSleepNet(EEG) ^[8]	0.867	0.455	0.851	0.833	0.826	0.819	0.766	2.5
	SleepEEGNet(EEG) ^[9]	0.894	0.444	0.847	0.846	0.796	0.815	0.766	1.5
	CNN+RNN(EOG) ^[11]	0.880	0.360	0.830	0.710	0.660	0.793	0.693	0.5
	CAttSleepNet(EEG) ^[12]	0.900	0.410	0.880	0.830	0.840	0.842	0.782	1.0
	MCNN-LSTMs(本文 EEG+EOG)	0.892	0.456	0.886	0.903	0.789	0.926	0.785	0.6
Sleep-EDF-78	DeepSleepNet(EEG) ^[8]	0.909	0.450	0.792	0.727	0.711	0.778	0.718	7.2
	SleepEEGNet(EEG) ^[9]	0.898	0.421	0.752	0.704	0.706	0.742	0.696	4.6
	1D-CNN-HMM(EEG) ^[10]	0.878	0.351	0.866	0.905	0.868	0.840	0.769	3.5
	CAttSleepNet(EEG) ^[12]	0.920	0.420	0.850	0.770	0.770	0.809	0.735	3.0
	CRNN(EEG+EOG) ^[13]	0.837	0.224	0.800	0.773	0.705	0.903	0.668	4.0
	MCNN-LSTMs(本文 EEG+EOG)	0.885	0.423	0.868	0.898	0.793	0.911	0.773	1.5

由于本文的研究是基于原始的 EEG 和 EOG 信号,所以在对比实验上仅针对近些年使用深度学习模型从原始单通道 EEG 中或 EOG 中提取特征进行评估的方法。与在相同 EEG 通道或 EOG 通道评估的其他研究方法相比,本文方法在两个数据集上的每类 $F1$ 分数、 ACC 以及 $MF1$ 上均表现较好,并且训练时间更短,取得了相似或更好的性能。最新的研究 Toma 等^[13]提出的 CRNN 模型在 Sleep-EDF-78 数据集上的 ACC 和 $MF1$ 分别达到了 90.30% 和 66.80%,而本文设计的 MCNN-LSTMs 模型在 Sleep-EDF-20 数据集上总体 ACC 和 $MF1$ 分别达到了 92.60% 和 78.50%,在 Sleep-EDF-78 数据集上总体 ACC 和 $MF1$ 分别达到了 91.10% 和 77.30%,效果更好,证明了本文设计的双层 LSTM 融合网络的可行性,再一次验证了 EEG 通道和 EOG 通道共同作用效果是大于单个生理信号,因此本文设计的模型在睡眠分期任务中具有较好的性能。

3 结语

准确的睡眠分期是对人类睡眠质量分析的重要步骤。为了设计无须人工干预、准确率高的端到端的自动睡眠分期系统,本文设计了一种利用原始单

通道 EEG 和 EOG 信号进行睡眠分期的双模态深度神经网络,使用设计的多尺度卷积神经网络分别从 EEG 和 EOG 信号中提取特征,将提取到的 EEG 特征和 EOG 特征由双层 LSTM 模型进行融合,然后输入分类器进行睡眠分期。在 Sleep-EDF 两个数据集上评估了一个五分类任务,即 W、N1、N2、N3、REM 睡眠阶段,最终在两个数据集上的总体准确率分别为 92.60%、91.10%。最终的实验结果证明了结合两种不同生理信号特征的睡眠分期系统优于仅使用单一生理信号的睡眠分期系统。但是在端到端的睡眠分期的研究上仍面临较大的挑战,虽然本文在实验过程中已经对数据集进行了平衡性处理,但是 N1 睡眠阶段的 $F1$ 分数仍旧较低。此外本文验证了结合两种生理信号的性能是高于单个生理信号的,所以未来的研究可以结合更多模态的数据对睡眠阶段进行分类。

参考文献:

- [1] AMINOFF, MICHAEL J. Handbook of clinical neurology [M]. Amsterdam: North-Holland Publishing Company, 1968.
- [2] 陶雨洁, 杨云. 基于原始单通道脑电图的高效睡眠自动分期方法[J]. 郑州大学学报(理学版), 2022, 54(3): 40-44.

- TAO Y J, YANG Y. Efficient automatic sleep staging method based on original single-channel electroencephalogram[J]. *Journal of Zhengzhou university (natural science edition)*, 2022, 54(3): 40-44.
- [3] KOLEY B, DEY D. An ensemble system for automatic sleep stage classification using single channel EEG signal[J]. *Computers in biology and medicine*, 2012, 42(12): 1186-1195.
- [4] FRAIWAN L, LWEESY K, KHASAWNEH N, et al. Automated sleep stage identification system based on time-frequency analysis of a single EEG channel and random forest classifier[J]. *Computer methods and programs in biomedicine*, 2012, 108(1): 10-19.
- [5] MOGHADDAM A P, MOUSAVI S. Learning decision tree using neural network for stability and flexibility[J]. *Frontiers in health informatics*, 2012, 1(3): 39-44.
- [6] GÜNEŞ S, POLAT K, YOSUNKAYA Ş. Efficient sleep stage recognition system based on EEG signal using k-means clustering based feature weighting[J]. *Expert systems with applications*, 2010, 37(12): 7922-7928.
- [7] 张恩铭, 袁玥, 滕飞, 等. 基于 RGCN 的阻塞性睡眠呼吸暂停预测研究[J]. *郑州大学学报(理学版)*, 2023, 55(6): 71-76.
- ZHANG E M, YUAN Y, TENG F, et al. Prediction of obstructive sleep apnea based on RGCN[J]. *Journal of Zhengzhou university (natural science edition)*, 2023, 55(6): 71-76.
- [8] SUPRATAK A, DONG H, WU C, 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: a publication of the IEEE engineering in medicine and biology society*, 2017, 25(11): 1998-2008.
- [9] MOUSAVI S, AFGHAH F, ACHARYA U R. SleepEEG-Net: automated sleep stage scoring with sequence to sequence deep learning approach[J]. *PLoS One*, 2019, 14(5): e0216456.
- [10] YANG B F, ZHU X L, LIU Y T, et al. A single-channel EEG based automatic sleep stage classification method leveraging deep one-dimensional convolutional neural network and hidden Markov model[J]. *Biomedical signal processing and control*, 2021, 68: 102581.
- [11] FAN J H, SUN C L, LONG M, et al. EOGNET: a novel deep learning model for sleep stage classification based on single-channel EOG signal[J]. *Frontiers in neuroscience*, 2021, 15: 573194.
- [12] LI T T, ZHANG B F, LV H H, et al. CAtSleepNet: automatic end-to-end sleep staging using attention-based deep neural networks on single-channel EEG[J]. *International journal of environmental research and public health*, 2022, 19(9): 5199.
- [13] TOMA T I, CHOI S. An end-to-end convolutional recurrent neural network with multi-source data fusion for sleep stage classification[C]//2023 International Conference on Artificial Intelligence in Information and Communication. Piscataway: IEEE Press, 2023: 564-569.
- [14] SHARMA M, YADAV A, TIWARI J, et al. An automated wavelet-based sleep scoring model using EEG, EMG, and EOG signals with more than 8000 subjects[J]. *International journal of environmental research and public health*, 2022, 19(12): 7176.
- [15] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [EB/OL]. (2014-09-04) [2023-10-15]. <https://arxiv.org/pdf/1409.1556v6.pdf>.
- [16] GANDHI A, ADHVARYU K, PORIA S, et al. Multimodal sentiment analysis: a systematic review of history, datasets, multimodal fusion methods, applications, challenges and future directions[J]. *Information fusion*, 2023, 91: 424-444.
- [17] HUANG J, TAO J H, LIU B, et al. Multimodal transformer fusion for continuous emotion recognition[C]//ICASSP 2020—2020 IEEE International Conference on Acoustics, Speech and Signal Processing. Barcelona, Spain. Piscataway: IEEE Press, 2020: 3507-3511.
- [18] NIE W Z, YAN Y, SONG D, et al. Multi-modal feature fusion based on multi-layers LSTM for video emotion recognition[J]. *Multimedia tools and applications*, 2021, 80(11): 16205-16214.
- [19] 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): E215-220.
- [20] 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.
- [21] CHAWLA N V, BOWYER K W, HALL L O, et al. SMOTE: Synthetic minority over-sampling technique[J]. *Journal of artificial intelligence research*, 2002, 16: 321-357.