

基于特征层融合的 EEG-NIRS 识别方法研究

周宇星¹, 樊丞成², 王震², 徐信毅¹, 林萍¹, 李晓欧^{1,2}

(1.上海理工大学健康科学与工程学院, 上海 200093;

2.上海健康医学院医疗器械学院, 上海 201318)

✉ zyuxing1999@outlook.com; fancc@sumhs.edu.cn; wangz_21@sumhs.edu.cn;

1254841987@qq.com; linley09@163.com; lixo@sumhs.edu.cn



摘要:针对目前卷积神经网络对时序序列识别率较低、单模态数据信息量不充足等问题,提出了一种基于特征层融合的卷积神经网络(CNN)和双向长短期记忆(BiLSTM)网络组合的方法。首先在毒品成瘾患者的脑电图和近红外光谱数据上,利用全新的CNN和BiLSTM组合网络分别对双模态数据进行特征提取,将最后一层BiLSTM的输出作为特征并进行特征串联,然后对串联特征进行分类识别。特征融合实验结果表明,文章提出的CNN-BiLSTM模型的分类效果最高准确率达到97.3%,并且双模融合方法进一步提高了分类准确率。

关键词:特征融合;卷积神经网络;双向长短期记忆网络;分类准确率

中图分类号:TP391 **文献标志码:**A

Research on EEG-NIRS Recognition Method Based on Feature Layer Fusion

ZHOU Yuxing¹, FAN Chengcheng², WANG Zhen², XU Xinyi¹, LIN Ping¹, LI Xiaou^{1,2}

(1.School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China;

2.School of Medical Instruments, Shanghai University of Medicine & Health Sciences, Shanghai 201318, China)

✉ zyuxing1999@outlook.com; fancc@sumhs.edu.cn; wangz_21@sumhs.edu.cn;

1254841987@qq.com; linley09@163.com; lixo@sumhs.edu.cn

Abstract: Aiming at the low recognition rate of temporal sequences by convolutional neural networks and insufficient information in single mode data, this paper proposes a method of combining Convolutional Neural Networks (CNN) and Bidirectional Long Short-Term Memory (BiLSTM) networks based on feature layer fusion. Firstly, on the EEG (Electroencephalogram) and near-infrared spectral data of drug addiction patients, the new CNN and BiLSTM combination network are used to extract features from the bimodal data. Then, the output of the last layer of BiLSTM is used as features for feature concatenation. Then, the concatenated features are classified and recognized. The experimental results of feature fusion show that the proposed CNN-BiLSTM model achieves the highest classification accuracy of 97.3%, and bimodal fusion method further improves the classification accuracy.

Key words: feature fusion; Convolutional Neural Networks; Bidirectional Long Short-Term Memory network; classification accuracy

0 引言 (Introduction)

随着全球吸毒人员数量逐年攀升,吸毒正成为严重的全球

性社会问题,而戒毒工作开展的前提是对吸毒人员的成瘾程度进行评估。近年来,研究人员主要基于磁共振成像(Magnetic

Resonance Imaging, MRI)、脑磁图 (Magnetoencephalography, MEG)、脑电图 (Electroencephalogram, EEG) 和近红外光谱 (Near-infrared Spectroscopy, NIRS) 等技术对毒品成瘾患者进行研究,随着多模态数据分析的不断进步,将脑电图和近红外光谱融合应用的研究也不断兴起。基于传统的机器学习方法进行分类,存在精度差和特征提取不全面等问题,所以深度学习网络逐渐被应用于分类问题,主要包括卷积神经网络 (Convolutional Neural Network, CNN)、长短期记忆网络 (Long Short-Term Memory, LSTM) 和双向长短期记忆网络 (Bi-directional Long Short-Term Memory, BiLSTM) 等,但单一的传统网络在分类任务中的识别效果不佳,故本文提出一种基于 CNN 和 BiLSTM 组合的神经网络,用于成瘾程度分类研究^[1]。

1 预备知识 (Preparatory knowledge)

1.1 国内外研究现状

近年来,脑电图和近红外光谱都被应用于成瘾治疗工作。高军峰等^[2]提取脑电信号的时域、频域和小波系数对冰毒成瘾者进行模式识别,分类结果显示,准确率达到 80% 以上。李嘉欣^[3]基于线性和非线性方法,利用 EEG 数据构造海洛因成瘾人员的脑网络研究成瘾机制,并将脑网络中的小世界属性作为特征,使用支持向量机 (Support Vector Machine, SVM) 进行分类。结果显示,最高的分类准确率达到 86% 以上,通过分析发现,成瘾患者的异常行为主要与大脑的前额叶、脑岛、顶叶的异常有关。殷悦^[4]利用功能近红外光谱对甲基苯丙胺成瘾患者进行实验,通过运动干预帮助成瘾患者尽快康复,有效缓解成瘾患者对毒品的渴求,并提高患者的认知功能。李玥^[5]通过功能性磁共振和功能近红外光谱对游戏成瘾患者的脑活动进行分析研究,结果显示,游戏成瘾患者在游戏过程中,不断地进行视觉运动交互,而此时患者的腹外侧前额叶脑区的激活程度,有可能与游戏本身的视觉运动交互任务强度有关,即游戏交互的强度增强,腹外侧前额叶脑区的激活程度随之增强,背外侧前额叶的脑活动则可能涉及如注意力转移和认知资源分配等更高级的功能。然而,EEG 虽然具有较高的时间分辨率,但是在采集过程中易受到生理噪声和周围环境的影响且空间分辨率较差。近红外光谱具有较高的空间分辨率,但只能观测到大脑皮层表面区域且时间分辨率较差,所以仅利用单模态数据所能够反映的信息仍有局限^[6]。

随着多模态学习的兴起及其在各种领域的广泛应用,越来越多的学者投入多模态学习这项工作。多模态信息相较于单模态信息,能够提供更丰富的信息内容,可以弥补单模态信息的不足^[7]。李日成^[8]利用脑电图和近红外光谱结合的双模态脑成像技术,对理解动作意图的神经机制进行了多信息、多层面的研究。通过提取两种模态数据的共同空间模式特征并基于线性判别分析方法进行特征层融合,结果显示双模态脑成像技术获得的平均正确率比脑电图单模态提高了 4.2%,比近红外光谱单模态提高了 19.8%。GU 等^[9]利用脑电信号和近红外光谱信号双模信号对毒品成瘾患者进行成瘾程度分析,通过将双模态数据分别输入卷积神经网络中获得双模特征,然后基

于决策层融合方法分别赋予特征 50% 的权重后进行分类,结果显示,分类准确率达到 63.15%。LI 等^[10]利用心电图和心音图双模态数据进行心血管疾病的早期诊断和预防,首先通过建立卷积神经网络和长短期记忆网络组合的神经网络,分别提取心电图和心音图的深度编码特征,其次利用遗传算法对双模特征进行融合得到最佳的特征子集,最后利用支持向量机进行分类,结果显示,利用双模态数据的曲线下面积 (Area Under the Curve, AUC) 为 0.936,高于单模态。

1.2 CNN-BiLSTM 网络

深度学习是一种机器学习方法,属于机器学习的一个分支,在互联网、人工智能等领域都有应用,如语音识别、视觉对象识别领域等。深度学习的基本架构就是简单模型的多层堆叠,通过将低层的特征组合形成更高层的特征。深度学习算法与其他机器学习算法相比,数据量越大,其表现的性能越优越。目前,比较常用的深度学习网络包括卷积神经网络、循环神经网络和长短期记忆网络等^[11]。

卷积神经网络可以处理多维数据,信号、序列、图像、视频等都属于不同的多维数组形式,基本运算包括卷积运算、池化运算、全连接运算和识别运算。卷积层通过不断改变卷积核大小,从而选择出能正确表征特征的卷积核,并对输入数据进行特征提取。池化层的主要目的是降维,保留卷积层输出特征中最有用的信息。全连接层相当于生成一个与所需分类数量相等的分类器。CNN 能够同时进行特征提取和分类,使神经网络结构变得更简单且适应性更强^[12]。

长短期记忆网络是具有记忆长短期信息能力的神经网络,常被用来处理时间序列和非线性特征,其结构示意图如图 1 所示。LSTM 通过设计“门”的结构解决循环神经网络梯度爆炸和梯度消失的问题,主要有遗忘门、输入门和输出门三种,这些门大大提高了 LSTM 处理时间信息的能力。三种门分别将前一个状态的输出 h_{t-1} 和当前状态的输入 x_t 作为输入。遗忘门决定了前一个状态有多少特征信息需要遗留,其中遗忘程度 f_t 都在 $[0, 1]$ 之间。输入门则是对当前状态进行新的补充,利用 Sigmoid 和 tanh 函数组合后得到当前的 i_t ,用于控制当前状态的信息更新程度。输出门则是确定需要输出的信息,控制滤除当前状态的特征信息^[13]。

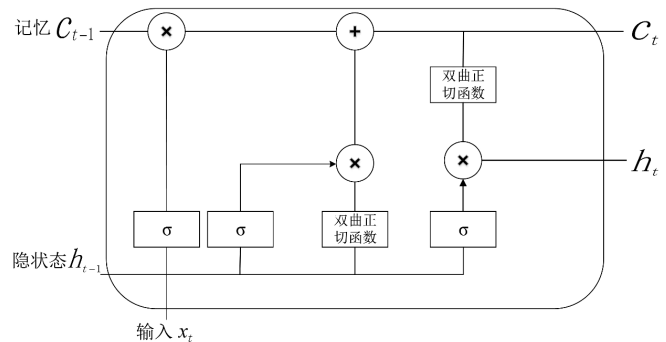


图 1 LSTM 结构示意图

Fig. 1 LSTM structure diagram

BiLSTM 是 LSTM 网络的改进模型,其结构示意图如图 2 所示。前向 LSTM 层和后向 LSTM 层共同构成隐藏层,其中前向 LSTM 沿正向处理时间序列,后向 LSTM 沿反向处理时间序列,然后将两个 LSTM 的输出拼接起来作为输出^[14],具体过程表示如下:

$$\vec{h}_t = \text{LSTM}(x_t, \vec{h}_{t-1}) \quad (1)$$

$$\overleftarrow{h}_t = \text{LSTM}(x_t, \overleftarrow{h}_{t+1}) \quad (2)$$

$$y_t = \vec{W}h_t + \overleftarrow{W}h_t + b_y \quad (3)$$

其中: $\vec{W}$ 为前向 LSTM 层到输出层的权重, $\overleftarrow{W}$ 为后向 LSTM 层到输出层的权重, b_y 表示输出层的偏置。

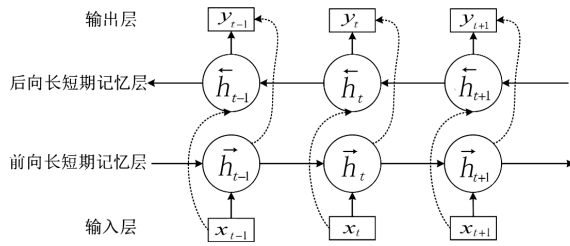


图 2 BiLSTM 结构示意图

Fig. 2 BiLSTM structure diagram

1.3 特征融合策略

相较于单模态,多模态学习通常需要考虑两个问题,即多模态学习性能和多模态数据融合。多模态相较于单模态能提供更多的信息,多模态还可以弥补单模态之间的不足,理论上多模态学习的效果优于单模态^[15]。目前,多模态数据融合主要有三种方式,包括数据层融合、特征层融合和决策层融合^[1]。数据层融合又称为信息层融合,是指对多源数据或多类数据进行多级别、多层面、多层次的处理。数据层的融合能够最大限度地保留原始数据,但也会导致数据处理量较大,信息冗余严重。特征层融合是指对不同的输入数据分别做特征提取,对提取后的特征数据做融合处理,例如进行简单的拼接,然后对融合后的特征进行分类。特征提取的过程中,原始数据的有效信息得以保留并去除了冗余和干扰成分,因此特征提取融合一方面减少了数据量,另一方面提高了分类精度。特征提取方法是特征层融合的关键,一些经改进的单模态特征提取方法目前已取得较高的准确率,所以理论上双模态的特征提取将会进一步提高融合后的分类准确率。决策层融合是指对不同模态的输入数据单独进行分类,再对每种分类结果通过投票或分配权重等方法,得到最终的分类型结果。决策层融合的容错率较高,即使某个模态的分类结果出现错误,也能够通过降低其权重的方式获得更准确的结果,并且决策层融合对原始数据没有限制,不同种类的数据都能够做决策层融合。但是,决策层融合的信息损失较大,分类识别性能会有所下降^[16]。

本文采用深度学习中的 CNN 和 BiLSTM 以及特征层融合的方式,设计了 CNN-BiLSTM 组合的深度学习模型,特征层融合模型结构图如图 3 所示^[17-18]。将两种模态的数据通过不

同的神经网络分别训练出特征后进行特征融合,然后利用分类器进行分类识别。

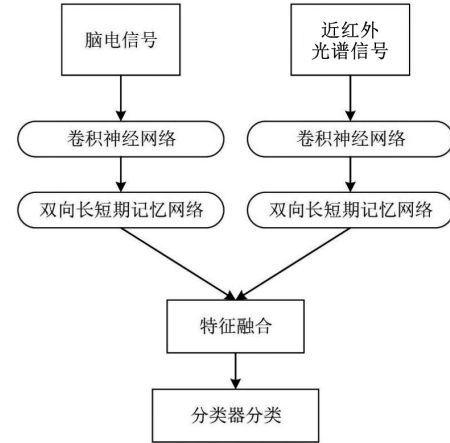


图 3 特征层融合模型结构图

Fig. 3 Structure diagram of feature layer fusion model

2 融合算法设计 (Fusion algorithm design)

2.1 数据采集

实验范式内容由三个阶段组成,第一阶段为采集静息态数据实验,即受试者先闭眼 5 min,再睁眼 5 min,历时 10 min。第二阶段为采集任务态数据实验,历时 6 min,受试者需要观看图片,图片分为毒品刺激图和中性图,共 16 张,当受试者看到毒品刺激图时,需要在纸上同步标记。第三阶段为采集任务态数据实验,历时 4.6 min,受试者所看图片均为中性图。

模型训练使用的数据为 EEG 和 NIRS 的双模态毒品成瘾数据,受试者来自上海市青东强制隔离戒毒所,均为 20~40 岁的男性在所戒毒人员。其中,EEG 数据由无线脑电图采集设备采集,脑电极数量为 64 个,采样率为 1 000 Hz,根据实际需求选择了 48 个电极。NIRS 数据由一个具有 204 个通道的功能性近红外光谱设备采集,有 24 个光点数量、32 个检测点数量。数据共分为三种成瘾程度分类,包括轻度、中度和重度,将轻度标记为 0,中度标记为 1,重度标记为 2。

2.2 数据网络构建

对于 EEG 数据设计了如图 4 所示的 EEG 神经网络结构,主要设计了四层卷积层和两层双向 LSTM 层,然后加入了丢弃层、全连接层、Softmax 层及输出分类层^[19]。其中,卷积层可以提取脑电特征;双向 LSTM 层能够提取时序特征;池化层能够减少数据量,起到特征提取和压缩的作用;丢弃层用来解决过拟合的问题;全连接层将输入乘以权重矩阵并添加偏置向量,可以将提取的特征综合。其中,第一个和第二个卷积层的卷积核大小为 32,第三个和第四个卷积层的卷积核大小为 16,双向 LSTM 层的指定输出大小均为 100。

针对 NIRS 数据,设计了如图 5 所示的 NIRS 神经网络结构^[20]。同样设计了四层卷积层和两层双向 LSTM 层,相较于 EEG 数据的网络,NIRS 数据的网络将最大池化层改为平均池

化层。其中,第一个卷积层的卷积核大小为 32,第二个卷积层卷积核大小为 16,双向 LSTM 层的指定输出大小均为 100。

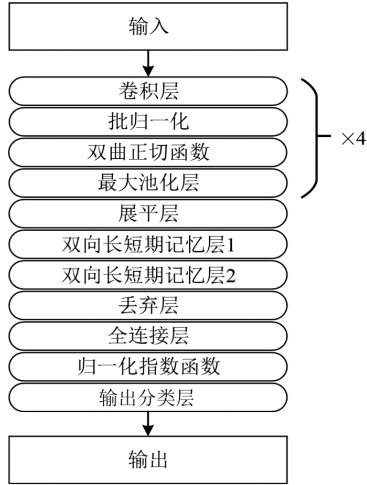


图 4 EEG 神经网络结构

Fig. 4 Structure of EEG neural network

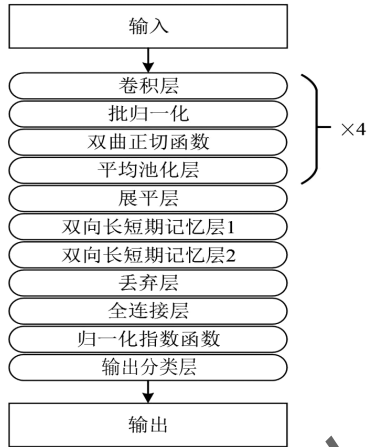


图 5 NIRS 神经网络结构

Fig. 5 Structure of NIRS neural network

2.3 特征融合方法

传统的特征融合方法包括特征的串联与并联。特征串联即对特征进行横向连接,假设双模特征向量为 e 和 p ,维度分别为 a 和 b ,使用串联方法融合后,特征向量组合为 $[e, p]$,维度为 $a+b$ 。特征并联即对特征进行纵向连接,融合后的特征向量为 $e+ip$,其中 i 为虚数单位,融合特征的维度为 a 和 b 中的较大者,对于较低维度的向量,融合后相应缺失地方则用 0 补位。

本文对双模特征进行横向的连接,特征融合方法如图 6 所示,将两个神经网络的第二个 BiLSTM 层的输出分别作为 EEG 和 NIRS 提取的特征,并将特征进行串联^[21]。

3 实验结果 (Experimental results)

在 EEG 数据实验中,使用的优化函数为自适应矩估计 (Adaptive Moment Estimation, Adam) 优化器,最大训练回合数设置为 100 次,初始学习率设置为 0.000 5,每次迭代使用的数据量设置为 10,数据打乱策略设置为每次迭代均打乱一次。

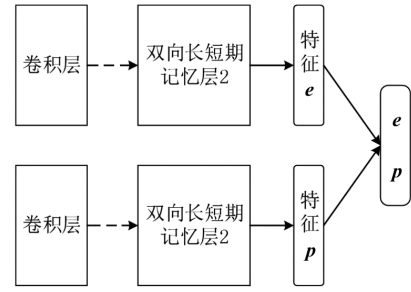


图 6 特征融合方法

Fig. 6 Feature fusion method

经全连接层输出的分类识别显示,EEG 的分类准确率为 74.2%,训练时长为 24 s。

在 NIRS 数据实验中,使用 Adam 优化器,最大训练回合数设置为 150,初始学习率设置为 0.000 5,每次迭代使用的数据量设置为 10,数据打乱策略设置为每次迭代均打乱一次。经全连接层输出的分类识别显示,NIRS 的分类准确率为 85.7%,训练时长为 131 s。

多模态实验中,分别将 EEG 和 NIRS 神经网络的第二层 BiLSTM 层输出的结果作为特征,然后进行特征融合,最后将融合后的特征利用决策树、KNN、SVM 进行分类,经过十折交叉验证,得到分类结果为 KNN 的分类准确率最高,达到 97.3%,单模态与双模态分类结果见表 1,成瘾性识别结果见表 2,混淆矩阵如图 7 所示。

表 1 单模态与双模态分类结果

Tab.1 Single-mode and bi-mode classification results

模态	分类准确率/%
脑电图	74.2
近红外光谱	85.7
脑电图+近红外光谱	97.3

表 2 成瘾性识别结果

Tab.2 Results of addiction recognition

分类器	分类准确率/%
决策树	96.4
K 最近邻(KNN)	97.3
支持向量机(SVM)	97.1

真实类	0	670	4	7
	1	10	655	12
	2	10	12	660
		0	1	2
		预测类		

图 7 混淆矩阵

Fig. 7 Confusion matrix

根据表 1 中的单模态与双模态分类结果显示,双模态融合后的分类准确率比单模态有较大的提高,比单模态 NIRS 提高

了 11.6%,比单模态 EEG 提高了 23.1%。不同的分类器也显示出不同的分类效果,其中 KNN 的分类准确率最高,达到 97.3%。

目前,对于成瘾性的相关研究中,GU 等^[9]利用决策层融合的方法和 CNN 网络对毒瘾患者所采集的双模态数据进行识别分类,准确率为 63.15%,本文所提方法的分类结果相较于 GU 等所提方法,分类准确率有较明显的提高。对于深度学习在双模态数据上的研究,吴礼祝等^[22]利用 CNN-BiLSTM 网络对睡眠数据中的 EEG 数据进行分期实验,结果显示分期的准确率为 92.33%。鞠晓慧^[23]利用 CNN-BiLSTM 网络对运动想象的 EEG 数据进行分类研究,结果显示 CNN-BiLSTM 能够有效提高分类精度。综上所述,本文设计的 CNN-BiLSTM 网络框架相较于目前对毒品成瘾研究的方法,在网络架构上,与传统的 CNN 或 LSTM 网络相比,使用了 CNN-BiLSTM 组合的方式。在融合方法上,使用了基于特征层融合的方法,避免了决策层融合所带来的信息损失的问题,从而能获得较高的分类准确率。

4 结论(Conclusion)

为了提高双模态数据的分类准确率,本文提出了一种基于特征层融合的 CNN-BiLSTM 混合神经网络模型,并利用 EEG 和 NIRS 数据进行实验验证。实验结果表明,在模型构造阶段,利用 CNN 提取空间特征和 BiLSTM 提取时间特征的特点,能够充分保留 EEG 和 NIRS 的时间维度特征,同时使用 CNN 和 BiLSTM 相结合的神经网络结构有效地提高了特征分类识别效果,双模融合的分类效果也远高于单模态分类结果。本文优化了以量表为依据的成瘾程度传统判断方式,利用客观生理数据进行成瘾程度的识别判断,可为后续临床上的成瘾患者治疗提供帮助。在未来的研究中,还可以针对不同的双模态数据进行有针对性的数据预处理,进一步提高神经网络输入数据的有效性。此外,还可以利用更高效的融合策略提高对有效特征的融合效果,以进一步提高模型的性能。

参考文献(References)

- [1] 祁铨颖,贺萍. 跨模态数据融合综述[J]. 软件工程,2022,25(10):1-7.
- [2] 高军峰,张家琦,韦思宏,等. 冰毒成瘾者脑电信号的分类研究[J]. 电子科技大学学报,2020,49(6):955-960.
- [3] 李嘉欣. 基于 EEG 的海洛因成瘾人群的脑网络研究[D]. 兰州:兰州大学,2020.
- [4] 殷悦. 运动干预对毒品成瘾者的脑机制影响的研究[D]. 大连:大连理工大学,2021.
- [5] 李玥. 电子游戏相关实时脑功能活动的功能磁共振成像(fMRI)/功能近红外光谱(fNIRS)研究[D]. 武汉:华中科技大学,2018.
- [6] CICALESE P A,LI R H,AHMADI M B,et al. An EEG-fNIRS hybridization technique in the four-class classification of Alzheimer's disease[J]. Journal of neuroscience methods,2020,336:108618.
- [7] 杨杨,詹德川,姜远,等. 可靠多模态学习综述[J]. 软件学报,2021,32(4):1067-1081.
- [8] 李日成. 基于 EEG-NIRS 双模态动作意图的分类识别[D]. 南京:东南大学,2017.
- [9] GU X L,YANG B H,GAO S W,et al. Application of bi-

modal signal in the classification and recognition of drug addiction degree based on machine learning[J]. Mathematical biosciences and engineering,2021,18(5):6926-6940.

- [10] LI P P,HU Y M,LIU Z P. Prediction of cardiovascular diseases by integrating multi-modal features with machine learning methods[J]. Biomedical signal processing and control,2021,66:102474.
- [11] RAVI V,CHAGANTI R,ALAZAB M. Recurrent deep learning-based feature fusion ensemble meta-classifier approach for intelligent network intrusion detection system[J]. Computers and electrical engineering,2022,102:108156.
- [12] 吴迪,王梓宇,赵伟超. ELMo-CNN-BiGRU 双通道文本情感分类模型[J]. 计算机工程,2022,48(8):105-112.
- [13] 王晨阳,段倩倩,周凯,等. 基于遗传算法优化卷积长短记忆混合神经网络模型的光伏发电功率预测[J]. 物理学报,2020,69(10):149-155.
- [14] 魏佳恒,郭惠勇. 基于贝叶斯优化 BiLSTM 模型的输电塔损伤识别[J]. 振动与冲击,2023,42(1):238-248.
- [15] LI H,WANG X P,LIU C C,et al. Integrating multi-domain deep features of electrocardiogram and phonocardiogram for coronary artery disease detection[J]. Computers in biology and medicine,2021,138:104914.
- [16] 史艳华. 基于 EEG-fNIRS 的少通道双模态脑机接口关键技术研究[D]. 南京:东南大学,2017.
- [17] 谷学静,周自朋,郭宇承,等. 基于 CNN-LSTM 混合模型的动态手势识别方法[J]. 计算机应用与软件,2021,38(11):205-209.
- [18] 黄伟建,李永涛,黄远. 基于混合神经网络和注意力机制的混沌时间序列预测[J]. 物理学报,2021,70(1):235-243.
- [19] MIDDYA A I,NAG B,ROY S. Deep learning based multimodal emotion recognition using model-level fusion of audio-visual modalities[J]. Knowledge-based systems,2022,244:108580.
- [20] 刘伟伟,桑胜波,张宏鹏. 基于 CNN+LSTM 的改进心音分类模型研究[J]. 电子设计工程,2022,30(2):38-42.
- [21] 陶鹏,邓绍江. 自适应权重特征融合的持续身份认证[J]. 重庆大学学报,2023,46(1):103-112.
- [22] 吴礼祝,卢伊虹,郑梓烨,等. 基于双通道脑电信号的在线实时睡眠分期系统[J]. 计算机系统应用,2023,32(1):87-98.
- [23] 鞠晓慧. 基于卷积神经网络的运动想象脑电信号分类研究[D]. 济南:山东师范大学,2020.

作者简介:

周宇星(1999-),女,硕士生。研究领域:生物医学信号处理。
樊丞成(1989-),男,博士,讲师。研究领域:生物工程力学。
王震(1986-),男,博士,讲师。研究领域:生物信息学。
徐信毅(1996-),男,硕士生。研究领域:生物医学信号处理。
林萍(1997-),女,硕士生。研究领域:生物医学信号处理。
李晓欧(1975-),男,博士,教授。研究领域:人体智能感知技术,穿戴式医疗器械。