

基于改进局部图结构的再认记忆脑电特征提取^{*}王凯^{1,2} 顾翔^{2,3} 李文杰^{1,2} 王苏弘⁴ 邹凌^{1,2}(1.常州大学微电子与控制工程学院 常州 213164; 2.常州市生物医学信息技术重点实验室 常州 213164;
3.常州大学阿里云大数据学院 常州 213164; 4.苏州大学附属第三医院临床心理科 常州 213003)

摘要: 为了研究再认记忆脑电的纹理特征,以及解决垂直对称局部图结构和对称局部图结构在提取脑电纹理特征时结构不稳定的问题。基于新旧范式设计了再认记忆实验,采集医学生和非医学生(均为35名)相关脑电,并且将这些脑电分为学习医学图片阶段、学习非医学图片阶段、再认旧医学图片阶段、再认旧非医学图片阶段、再认新医学图片阶段和再认新非医学图片阶段。首先,利用二维小波变换得到每位被试脑电的三个子频带,并提出改进集成局部图结构方法对原数据和3个子频带进行特征提取,改进算法纳入了具有稳定结构的扩展对称局部图结构和复合局部图结构;然后对特征进行归一化,避免结果过拟合,使用皮尔逊相关系数筛选出相关系数在0.8~1之间的特征矩阵列。在支持向量机等分类器上验证改进前后的算法,并使用正确率、精确率、召回率和F1评分这四个指标对模型进行评估。与改进前算法相比,改进后算法在支持向量机上的分类正确率分别提升3.8%,0.4%,0.3%,1.6%,5.1%和4.2%。分类结果说明医学生和非医学生在医学图片学习再认阶段存在明显差异,新加入扩展对称局部图结构和复合局部图结构比原算法中垂直对称局部图结构和对称局部图结构具有更好的分类性能。

关键词: 脑电;再认记忆;集成局部图结构;特征提取

中图分类号: TP331 **文献标识码:** A **国家标准学科分类代码:** 510.4

EEG feature extraction of recognition memory based on improved local graph structure

Wang Kai^{1,2} Gu Xiang^{2,3} Li Wenjie^{1,2} Wang Suhong⁴ Zou Ling^{1,2}(1. School of Microelectronics and Control Engineering, Changzhou University, Changzhou 213164, China;
2. Changzhou Key Laboratory of Biomedical Information Technology, Changzhou 213164, China;
3. Aliyun School of Big Data, School of Software, Changzhou University, Changzhou 213164, China;
4. Clinical Psychology, The Third Affiliated Hospital of Soochow University, Changzhou 213003, China)

Abstract: To investigate the texture features of recognition memory EEG and address the issue of structural instability in extracting EEG texture features using vertical symmetric local graph structure (VSLGS) and symmetric local graph structure (SLGS). A recognition memory experiment was designed based on new and old paradigms, and relevant EEG data were collected from 35 medical students and 35 non-medical students. The EEG data were categorized into six different stages: learning medical images, learning non-medical images, recognizing old medical images, recognizing old non-medical images, recognizing new medical images, and recognizing new non-medical images. Firstly, a 2D wavelet transform was applied to obtain three subbands for each participant's EEG data. Then, an improved integrated local graph structure method was proposed to extract features from the original data and the three subbands. This improved algorithm incorporated extended symmetric local graph structures (ESLGS) and composite local graph structures (CLGS) with stable structures. The features were then normalized to avoid overfitting, and feature matrix columns with correlation coefficients between 0.8 and 1 were selected using Pearson correlation coefficients. The improved algorithm was validated on classifiers such as support vector machines (SVM), and the model was evaluated using four metrics: accuracy, precision, recall, and F1 score. Compared to the original algorithm, the improved algorithm achieved an increase of 3.8%, 0.4%, 0.3%, 1.6%, 5.1%, and 4.2% in classification accuracy on support vector machines for each condition. The classification results indicate significant differences between medical students and non-medical students in the recognition stage of learning medical images. The addition of ESLGS and CLGS demonstrates better classification performance compared to the original algorithm, which utilized VSLGS and SLGS.

Keywords: EEG; recognition memory; integrated local graph structure; feature extraction

0 引言

再认记忆作为记忆功能的重要组成部分,与人类的日

常生活息息相关。双过程模型(dual-process models)认为,再认记忆包含熟悉和回忆两个过程^[1-2],得益于脑电图(electroencephalography, EEG)的发展,再认记忆的生理过

收稿日期:2023-12-21

^{*} 基金项目:江苏省重点研发计划项目(BE2021012-2, BE2021012-5)、常州市科技计划(CE20225034)、浙江省脑机协同智能基金重点实验室项目(2020E10010-04)资助

程被详细地展示出来。后续基于 EEG 的事件相关电位(event-related potential, ERP)研究也证实再记忆脑电包含在刺激呈现后 300~400 ms 期间额区皮层产生新旧效应的熟悉过程,以及刺激呈现后 400~800 ms 期间顶区皮层产生新旧效应的回忆过程^[3-5]。

为了更好地研究脑电的特征,许多研究引入机器学习的方法对脑电进行分析,郭玉雪等^[6]基于运动想象的脑电数据,结合了小波变换(wavelet transform, WT)和自编码网络,利用离散小波变换(Discraete WT, DWT)将 EEG 信号转化为时频矩阵,通过卷积自编码网络进行特征提取和分类,实现了 97.36% 的分类正确率。Albaqami 等^[7]使用小波包分解(wavelet packet decomposition, WPD)算法结合梯度提升决策树来对认知障碍等异常脑电进行特征提取以及分类,实现了 86% 的分类正确率。然而,利用小波变换对脑电信号处理,不能很好地分解和表示脑电这种包含大量细节纹理的信号。Tasci 等^[8]引入图像处理领域的算法,在局部二值模式(local binary patterns, LBP)算子的基础上提出量子局部二值模式(quantum LBP, QLBP)算子。通过结合 WPD 和 QLBP 算子的方法提取具有认知障碍的被试的脑电特征,实现了 97% 的分类正确率。使用图像处理的方法能够更好地提取出脑电图像的纹理特征。Tuncer 等^[9]提出一种基于集成局部图结构算法的癫痫脑电检测框架,结合 DWT 和集成局部图结构算法对癫痫脑电进行特征提取,最后将筛选完的特征送入分类器进行分类,达到了 98% 的分类正确率。但是,集成局部图结构算法中的垂直对称局部图结构和对称局部图结构稳定性不如其他局部图结构。

因此,本文在现有集成局部图算法^[9]基础上提出改进的集成局部图结构算法。该方法纳入复合局部图结构和扩展对称局部图结构^[10]来替代对称局部图结构和垂直对称

局部图结构。利用二维离散小波变换(2D-discraete wavelet transform, 2D-DWT)提取出子带,接着使用改进后的集成局部图结构进行特征提取,并用皮尔逊相关系数筛选出提取出相关系数在 0.8~1 之间的特征矩阵列。分别利用支持向量机、K 最邻近、双层神经网络进行分类。最后,将改进前和改进后的集成局部图结构算法在医学生和非医学生的再记忆脑电数据上进行验证。结果表明,改进后的算法相比于改进前的算法具有一定优越性。

1 方 法

1.1 被 试

在苏州大学医学院和非医学院招募大四临床医学专业学生 35 名(19 名男性,16 名女性)与非医学生 35 名(19 名男性,16 名女性),两组被试年龄均在 21~23 岁之间,且年龄方面无统计学差异($P > 0.05$)。所有被试均为右利手,视力或矫正视力正常,无精神疾病,未滥用药物,被试在实验前均签署实验知情同意书。

1.2 实验范式

范式选择经典新旧范式^[11],从医学图片库以及非医学图片库中各随机抽取 120 张图片组成两个相同规模的模块,每个模块都包含医学图片和非医学图片各 60 张,两部分的图片没有重复且中间有两分钟的休息时间。具体范式过程如图 1 再记忆范式所示,学习阶段:给予刺激图片位于屏幕中央,共呈现 60 张图片(医学图片和非医学图片各 30 张)每张图片显示时长 2 000 ms,间隔 1 500~2 000 ms;再认阶段:出现过的图片和未出现过的图片随机呈现,每张图片显示时间以及间隔与学习阶段相同,共 120 张(出现过和未出现过图片各 60 张),被试需要判别图片是否为出现过的图片。两个阶段间隔 1 min。

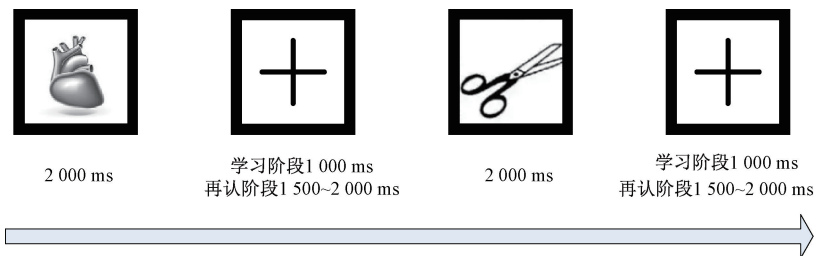


图 1 再记忆范式

1.3 脑电数据预处理

采集后的脑电经过初步预处理才能做进一步的分析,预处理全部流程均使用 MATLAB 的 EEGLAB 工具箱完成,其具体流程包括滤波、伪迹检测和坏导替换等。某个非医学生的预处理前后 EEG 波形图如图 2 某个非医学生 EEG 预处理前后的波形图对所示。

1.4 改进的集成局部图结构算法

改进前的集成局部图结构包含逻辑扩展局部图结构、垂直局部图结构、对称局部图结构、垂直对称局部图结构、

Z 字形垂直局部图结构、Z 字形水平局部图结构、Z 字形水平中间局部图结构和 Z 字形垂直中间局部图结构这 8 种不同的局部图结构^[9],目的在于提取脑电原数据以及经过二维小波变换处理后 3 个子频带的特征,然而,改进前算法中的对称局部图结构和垂直对称局部图结构存在稳定性问题,这可能会对分类结果造成影响。

本文在此基础上提出改进后的集成据图结构算法,新算法纳入复合局部图结构和扩展对称局部图结构,这两种算法在结构上比对称局部图结构和垂直对称局部图结构更平衡。

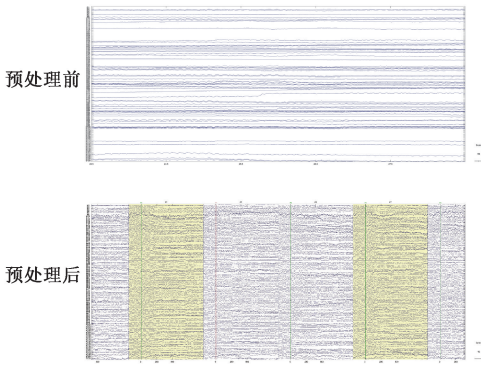


图 2 某个非医学生 EEG 预处理前后的波形图对比

1) 二维小波变换

在脑电矩阵上应用三级 2D-DWT 得到 3 个不同频段的子带,具体参照式(1)~(2):

$$[LL_1, LH_1, HL_1, HH_1] = 2D-DWT(\mathbf{X}) \quad (1)$$

$$[LL_k, LH_k, HL_k, HH_k] = 2D-DWT(LL_{k-1}) \quad (2)$$

其中, $k \geq 2$, $\mathbf{X}$ 表示原矩阵, LL_k 表示第 k 级 low-low 子带, LH_k 表示第 k 级 low-high 子带, 以此类推。图 3 为再记忆脑电的二维表示, 其中图 3(a)~(d) 分别为原矩阵 $\mathbf{X}$ 、 LL_1 、 LL_2 和 LL_3 的图形表示。

2) 改进的集成局部图结构

本文纳入复合局部图结构和扩展对称局部图结构来

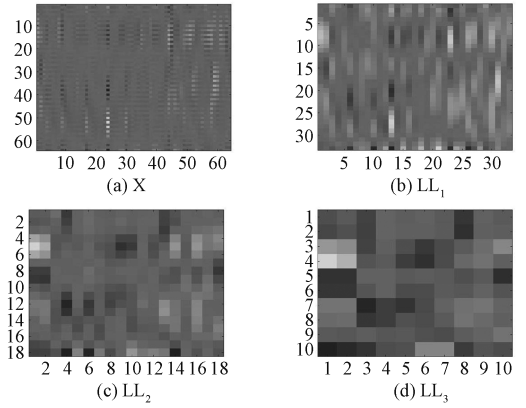


图 3 再记忆脑电的二维表示

替代原算法^[9]中对称局部图结构和垂直对称局部图结构。以复合局部图结构为例, 本文提出的复合对称局部图结构, 使用了 5×5 的重叠块进行特征的计算, 相比较上述两种单一的对称局部图结构有更平衡的布局且精度更高, 复合对称局部图结构示例图如图 4 复合局部图结构流程图所示, 具体方法如式(3)^[12]。

$$\text{sig}(m-n) = \begin{cases} 1, & m-n \geq 0 \\ 0, & m-n < 0 \end{cases} \quad (3)$$

其中, m 和 n 表示不同像素点的值。

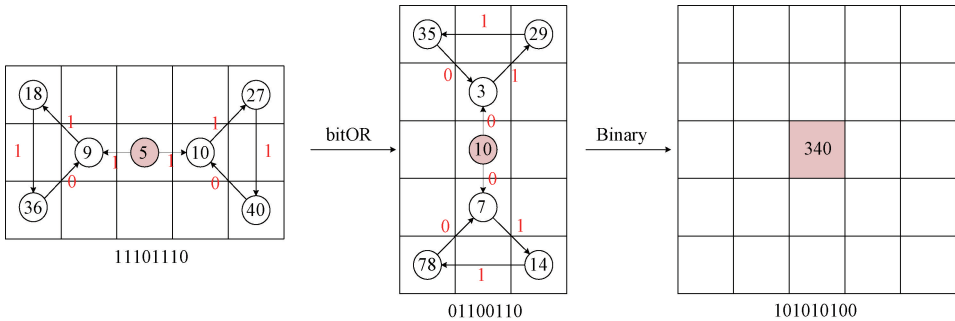


图 4 复合局部图结构流程图

3) 特征提取

通过集成局部图结构算法可以算出四层二维矩阵的不同区域的像素值, 并由此组成脑电特征向量, 然后将提取出的四层特征向量串联成需要进一步处理的特征, 具体表达式如式(4)~(7), 其中, $\mathbf{f}$ 表示特征向量, $\mathbf{X}$ 为原矩阵, LL_k 表示第 k 级 low-low 子带。

$$\mathbf{f}_1 = \text{enLGS}(\mathbf{X}) \quad (4)$$

$$\mathbf{f}_2 = \text{enLGS}(LL_1) \quad (5)$$

$$\mathbf{f}_3 = \text{enLGS}(LL_2) \quad (6)$$

$$\mathbf{f}_4 = \text{enLGS}(LL_3) \quad (7)$$

1.5 归一化

提取后的特征按行进行归一化, 得到每个被试的归一化特征, 方便后续进行分类。具体算法如式(8)所示^[13], 其中一位的被试归一化后特征如图 5 归一化后的特征所示。

$$X = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (8)$$

其中, x 为特征的值, X 为归一化后的特征值。

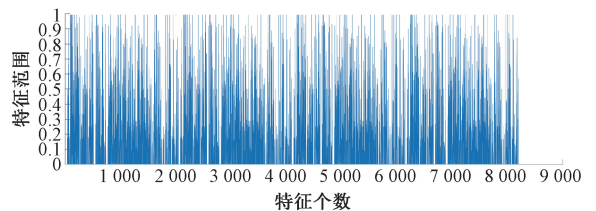


图 5 归一化后的特征

1.6 特征筛选

特征筛选部分采用皮尔逊相关系数, 皮尔逊相关系数计算如式(9)所示^[14]。

$$\rho = \frac{cov(X,Y)}{\sigma_X\sigma_Y} \tag{9}$$

其中, $\mathbf{X}, \mathbf{Y}$ 分别为特征矩阵不同两列, cov 为协方差, σ 为标准差。通过对平均相关性的排序筛选出相关系数处于 0.8~1 之间的特征矩阵列, 从而组成筛选后的特征矩阵, 以便后续分析。

1.7 分 类

在分类阶段常用的分类器有支持向量机, 人工神经网络, K 最邻近等^[15]。本文使用支持向量机、K 最邻近和双层神经网络这 3 种不同的分类器, 并用五倍交叉验证。在实际的实验中, 支持向量机的核函数选择线性函数, K 最邻近的 K 值经过多次试验, 确定为 3, 距离度量使用欧氏距离。双层神经网络内含两个全连接层, 每层 10 个节点, 激活函数选用 ReLU 函数。

1.8 分类性能指标评估方法

本文选择准确性、精确性、召回率以及 F1-score(F1) 作为评估结果的标准, 四项性能指标评估方法如式(10)~(13)。

$$acc = \frac{t_p + t_n}{t_p + t_n + f_p + f_n} \tag{10}$$

$$call = \frac{t_p}{t_p + f_n} \tag{11}$$

$$precision = \frac{t_p}{t_p + f_p} \tag{12}$$

$$F1 = \frac{2 \times precision \times call}{call + precision} \tag{13}$$

其中, t_p, t_n, f_p, f_n 分别为真医学生、真非医学生、假医学生、假非医学生。

2 结 果

按照再记忆任务阶段, 将脑电分为六个不同部分, 分别为学习医学图片阶段、学习非医学图片阶段、再认旧医学图片阶段、再认旧非医学图片阶段、再认新医学图片阶段和再认新非医学图片阶段。这 6 个阶段的脑电样本

分别用 I、II、III、IV、V 和 VI 命名, 每组医学生和非医学生均为 35 位。

改进前后算法的分类正确率如表 1 和 2 所示, 从表 1 和 2 都可以看出 I 组和 V 的分类结果相比较于其他 4 组的结果要好, 说明在进行再记忆范式时, 学习医学图片阶段以及再认新医学图片阶段, 医学生和非医学生的脑电特征存在的差异更加明显, 这也说明先验记忆会对再记忆的判断产生一定的影响。

表 1 不同分类算法对改进前算法(二维小波+旧集成局部图结构方法)的分类结果 %

组别	支持向量机	K 最邻近	双层神经网络
I	89.0	79.8	75.0
II	87.8	78.0	76.9
III	86.7	75.4	77.0
IV	85.4	76.4	83.4
V	79.0	76.3	90.0
VI	82.5	79.5	85.1

表 2 不同分类算法对改进后算法(二维小波+新集成局部图结构方法)的分类结果 %

组别	支持向量机	K 最邻近	双层神经网络
I	92.8	81.7	80.1
II	88.2	79.5	80.2
III	87.0	75.9	78.6
IV	87.0	79.3	84.4
V	84.1	80.8	92.3
VI	86.7	80.5	87.0

而从两个表的整体分类结果来看, 改进后算法的分类结果要优于改进前算法, 这说明改进后算法性能优于原算法。另外, 从分类算法的角度看, 支持向量机分类器的分类结果要比其他两个分类器的结果要好。

针对改进后算法结果绘制混淆矩阵如图 6 改进后算

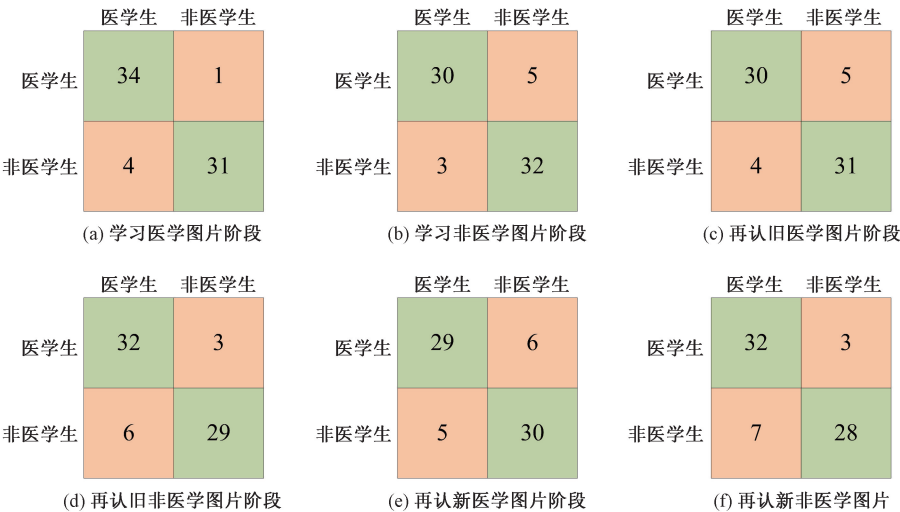


图 6 改进后算法各组支持向量机的混淆矩阵

法各组支持向量机的混淆矩阵所示,从中可以清楚看出被分类正确的医学生和非医学生以及分类错误的医学生和非医学生数量关系。

改进前后算法四项评估指标对比图如图 7 改进前后算法

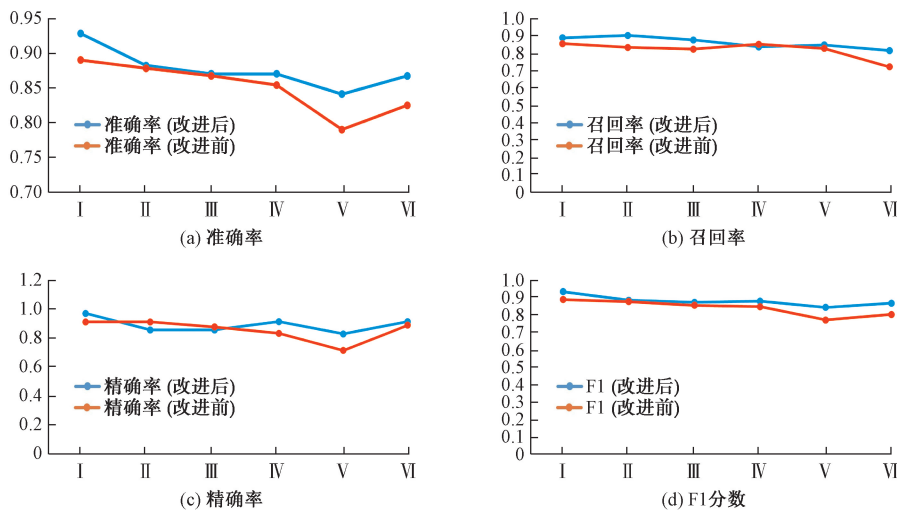


图 7 改进前后算法四项评估指标对比图

3 讨 论

从被试的角度来看,本文将两组被试的再记忆脑电按范式的不同任务分为学习医学图片阶段、学习非医学图片阶段、再认旧医学图片阶段、再认旧非医学图片阶段、再认新医学图片阶段和再认新非医学图片阶段这六个不同阶段,从最后特征分类的结果来看,不同阶段的分类是有必要的,每个阶段的分类正确率以及精确度存在一定的差别;从分类正确率来看,改进后的算法在支持向量机、K 最邻近和双层神经网络这三个分类器上得到的正确率都比改进前的算法高,这表明新纳入的复合局部图结构和扩展对称局部图结构方法增加了算法整体的性能。另外,两种集成局部图结构算法在支持向量机上的正确率最高,且两个不同专业的学生在支持向量机里最高分类正确率集中在 I 组,即学习医学图片的阶段。这一结果表明,在看到医学图片作为刺激的阶段,因为受到相关专业知识的培训,医学生的脑电与非医学生存在一定的差异,即先验记忆会影响再记忆的判断^[11]。

此外,从提取特征的角度看,改进后的算法具有更高的 F1 分数。这意味着改进后的算法能够更有效地从脑电信号中提取出与学习医学图片和再认新医学图片相关的特征,并能够更准确地将其分类^[11]。这一结果进一步证实了特征提取网络在识别和区分不同组别之间的差异方面的重要性。弥补了再记忆相关领域研究的缺乏,而且改进后的特征提取网络同样可以作为一种特征提取的方法提取其他脑电信号。

四项评估指标对比图所示,从性能指标柱状图的结果来看可以很明显看出,在支持向量机分类器的分类结果里,改进后的算法在性能方面要优于改进前的算法,I 组的结果是最好的,医学生在学习医学图片方面与非医学生有一定的差异。

4 结 论

本文提出改进集成局部图结构算法结合二维小波的方法对医学生和非医学生的再记忆脑电进行特征提取研究,分类的结果说明本文提出的改进后的集成局部图结构相比较于改进前的算法分类正确率更高,对特征分类的精确率也更高,更有力地证明了先验记忆对再记忆存在一定的影响,为再记忆领域的研究以及其他类型脑电的研究提供一种新的思路。

参考文献

- [1] 刘贵雄,贾永萍,买合甫来提·坎吉,等.熟悉性对被试操作效应的作用:来自无线索回忆再认的证据[J]. 心理学探新,2019,39(3).
- [2] 刘楚麒,陆爱桃,田海平,等.再记忆中的消极重复效应:基于中国被试的新证据[J]. 心理研究,2016,9(3): 36-42.
- [3] CAO J, SI F, LI X, et al. The effects of social anxiety on recognition memory for social threat words: An ERP study[J]. Biological Psychology, 2022, 174: 108420.
- [4] LEYNES P A, KALELKAR A U, SHAIK H T, et al. Event-Related Potential (ERP) evidence for fluency and disfluency effects on recognition memory [J]. Brain and Cognition, 2023, 167: 105961.
- [5] MIAO J, WEIGL M, KONG N, et al. Electrophysiological evidence for context reinstatement effects on object recognition memory[J]. Neurobiology

- of Learning and Memory, 2023: 107861.
- [6] 郭玉雪, 于洪丽, 么航, 等. 基于离散小波变换的卷积自编码运动想象脑电信号的分类[J]. 电子测量技术, 2023, 46(19):188-196.
- [7] ALBAQAMI H, HASSAN G M, SUBASI A, et al. Automatic detection of abnormal EEG signals using wavelet feature extraction and gradient boosting decision tree[J]. Biomedical Signal Processing and Control, 2021, 70: 102957.
- [8] TASCI G, GUN M V, KELES T, et al. QLBP: Dynamic patterns-based feature extraction functions for automatic detection of mental health and cognitive conditions using EEG signals[J]. Chaos, Solitons & Fractals, 2023, 172: 113472.
- [9] TUNCER T, DOGAN S, ERTAM F, et al. A novel ensemble local graph structure based feature extraction network for EEG signal analysis[J]. Biomedical Signal Processing and Control, 2020, 61: 102006.
- [10] YUNANTO A A, HERUMURTI D. Face recognition based on extended symmetric local graph structure[C]. IEEE, 2016: 80-84.
- [11] 齐云鹏, 王苏弘, 陈芋圻, 等. 相位转移熵在再记忆脑网络中的应用研究[J]. 计算机工程与应用, 2022, 58(15):87-94.
- [12] OLIVA J T, ROSA J L G. Classification for EEG report generation and epilepsy detection [J]. Neurocomputing, 2019, 335: 81-95.
- [13] JATOI M A, KAMEL N, MUSAVI S H A, et al. Bayesian Algorithm Based Localization of EEG Recorded Electromagnetic Brain Activity[J]. Current Medical Imaging, 2019, 15(2): 184-193.
- [14] SAHIB A K, ERB M, MARQUETAND J, et al. Evaluating the impact of fast-fMRI on dynamic functional connectivity in an event-based paradigm[J]. PLoS One, 2018, 13(1): e0190480.
- [15] KANG J, HAN X, SONG J, et al. The identification of children with autism spectrum disorder by SVM approach on EEG and eye-tracking data[J]. Computers in Biology and Medicine, 2020, 120: 103722.

作者简介

王凯, 硕士研究生, 主要研究方向为信号处理。

E-mail: wangkai9811@163.com

顾翔, 硕士研究生, 主要研究方向为信号处理。

E-mail: guxiang0914@163.com

李文杰, 硕士, 主要研究方向为信号与信息处理。

E-mail: lwj@cczu.edu.cn

王苏弘, 副主任医师, 主要研究方向为儿童精神病学、认知心理学。

E-mail: yarmine@foxmail.com

邹凌(通信作者), 博士, 硕士生导师, 主要研究方向为人工智能, 生物医学信息技术、信号处理。

E-mail: zouling@cczu.edu.cn