

DOI: 10.16652/j.issn.1004-373x.2026.05.003

引用格式: 吕卓言, 黄丽亚. 基于 Fisher-ISSA-BiLSTM 的酗酒脑电信号分类研究[J]. 现代电子技术, 2026, 49(5): 16-24.

基于 Fisher-ISSA-BiLSTM 的酗酒脑电信号分类研究

吕卓言, 黄丽亚

(南京邮电大学 电子与光学工程学院、柔性电子(未来技术)学院, 江苏 南京 210023)

摘要: 传统的脑电时频信号分类方法仅关注脑电信号的局部特性, 没有充分探究不同脑区之间的关联特征, 无法全面捕捉大脑活动的前后时间关联, 存在对大脑理解不够全面、分类准确率不高等问题。文中基于加州大学尔湾分校提供的酗酒脑电数据集, 运用相位锁定值(PLV)构建功能脑网络, 研究了 α 、 β 、 γ 、 θ 四个子频段和全频段的 EEG 脑网络拓扑特征, 探究了酗酒者与健康对照者 EEG 脑网络的拓扑属性差异。同时, 提出一种基于 Fisher 特征筛选和改良麻雀算法优化的 BiLSTM 分类算法(Fisher-ISSA-BiLSTM), 运用 Fisher 准则进一步筛选了脑网络特征, 通过双向长短期记忆网络充分分析了脑电信号前后时序的关联性, 运用改良的麻雀搜索算法(ISSA)优化 BiLSTM 的超参数。在原麻雀算法的基础上, 引入 Sobol 映射的初始化方式, 提升了麻雀种群的分布质量; 加入搜索因子, 避免算法过早地陷入局部最优; 引入自适应方向因子 $d_{i,j}^t$, 优化了麻雀跟随者位置的更新方向。相比其他超参数优化算法, 文中算法分类所需时间减少了约 4%~5%, 分类准确率达 92.6%, 相比传统的 LSTM 分类算法提升了约 20%, 对于运用脑电信号识别酗酒患者具有一定的实际意义。

关键词: EEG; 功能脑网络; Fisher 特征筛选; 麻雀搜索算法; 超参数优化; BiLSTM

中图分类号: TN911.7-34; TP391

文献标识码: A

文章编号: 1004-373X(2026)05-0016-09

Research on alcoholism EEG signal classification based on Fisher-ISSA-BiLSTM

LÜ Zhuoyan, HUANG Liya

(College of Electronic and Optical Engineering & College of Flexible Electronics (Future Technology), Nanjing University of Posts and Telecommunications, Nanjing 210023, China)

Abstract: The traditional time-frequency EEG (electroencephalogram) classification methods predominantly focus on local characteristics of EEG signals. They neglect inter-regional brain connectivity, and fail to fully capture the temporal correlation of brain activities, which result in incomplete understanding of the brain and low classification accuracy rate. On the basis of the UCI alcoholism EEG dataset, this study constructed functional brain networks by phase-locking value (PLV), investigated the EEG brain network topological characteristics across sub-bands (α , β , γ , and θ) and full-band, and explored the topological attribute differences of EEG brain networks between alcoholics and healthy controls. A classification algorithm, which integrates Fisher feature selection with bidirectional long short-term memory (BiLSTM) network which is optimized by improved sparrow search algorithm (ISSA), is proposed, and it is named Fisher-ISSA-BiLSTM. The Fisher criterion is used to further screen the characteristics of brain networks, the BiLSTM is used to fully analyze the correlation of the timing sequence of the EEG signals, and the ISSA is used to optimize the superparameter of BiLSTM. On the basis of the SSA, the initialization method of Sobol mapping is introduced to enhance the distribution quality of the sparrow population. Additionally, a search factor is incorporated to prevent the algorithm from falling into local optima prematurely. An adaptive direction factor $d_{i,j}^t$ is also introduced to optimize the update direction of the location of the sparrow followers. In comparison with the other hyperparameter optimization algorithms, the classification duration of the proposed algorithm is reduced by approximately 4% to 5%, and its classification accuracy rate reaches 92.6%, which is about 20% higher than that of the traditional LSTM classification algorithm. To sum up, the proposed algorithm has a certain practical significance for identifying alcoholics by using EEG signals.

Keywords: EEG; functional brain network; Fisher feature selection; SSA; hyperparameter optimization; BiLSTM

收稿日期: 2025-03-26

修回日期: 2025-05-28

基金项目: 国家自然科学基金项目(61977039)

0 引言

长期酗酒会导致大脑额叶、胼胝体等部位产生特异性损伤^[1],有效识别酗酒行为有助于防范风险。通过脑电信号分析识别酗酒问题已成为近年来的研究热点,文献[2]研究发现,酗酒者 α 、 θ 频段的能量较低,而 β 与 γ 频段能量较高。文献[3]研究发现,长期酗酒与小于8 Hz的低频段脑电活动增强相关。文献[4]通过研究酗酒患者和正常人EEG信号,发现两类受试者的EEG信号存在显著的功率谱差异。从以上研究得出,当前研究酗酒问题主要关注EEG信号的时频信号特征,但是EEG信号包含了非线性成分,仅依靠时频特征分析不能充分挖掘EEG信号蕴含的信息。大脑是一个典型的复杂网络,其认知行为涉及多个功能脑区之间的信息交换,当前大多数研究主要基于单一通道来探究局部特性,没有将大脑作为一个整体系统探究其不同导联间的关联特性,没有充分挖掘脑电信号特征时间的前后关联性。

本文以加州大学尔湾分校的酗酒数据集为基础,运用陷波滤波、经验模式分解、独立成分分析等预处理方法去除伪迹干扰,通过带通滤波、小波包分解,获得 θ 、 α 、 β 、 γ 频段的EEG数据。由于长期酗酒会对大脑的相位同步产生影响,相位锁定值(Phase-Locking Value, PLV)更能够反映相位同步性特征^[5],选取相位锁定值作为EEG脑网络的功能连接指标,分析各频段脑网络特征,研究酗酒患者与正常对照者的差异。通过统计学分析脑网络特征,得出具有显著性差异的导联,研究长期酗酒对大脑脑区的影响。最后,从EEG信号的时间关联度角度出发,结合脑网络与BiLSTM,提出一种Fisher-ISSA-BiLSTM的分类算法,研究EEG在时间跨度上的关联性。该算法通过Fisher准则筛选脑网络特征,将筛选后的特征序列输入到BiLSTM算法进行分类。超参数的选取对分类效果有较大影响,本文对比了不同超参数优化算法的优化效果,选取并改良麻雀算法优化超参数。该算法优化了麻雀种群初始化方式,加入了搜索因子和自适应方向因子,扩大了麻雀的搜索范围,提升了算法的分类准确率和运行效率。

本文算法的主要工作流程如图1所示。

1 数据集和实验范式

本文使用的脑电数据集来源于加州大学尔湾分校。该数据集涵盖了酗酒者和正常者的EEG数据。原始数据集共有122例样本,包含77名酗酒患者和45名正常对照者。实验数据采集了这两类受试者的64通道脑电数

据,电极放置位置遵循国际标准10-20系统,采样率设定为256 Hz。实验过程对两类受试者进行视觉相关实验,采集事件相关电位数据。受试者被安排在隔音室内的座椅上,接受显示屏展示的视觉刺激。展示的图片是从Snodgrass和Vanderwart的数据集^[6]中随机选择的。每位受试者接受120次随机顺序的视觉刺激,刺激分为三种类型:单刺激(S1),即仅呈现一次图片;双刺激匹配(S2 Match),即连续呈现两次图片,两次呈现的图片相匹配;双刺激不匹配(S2 No Match),即短时间内连续出现两次图片,两次呈现的图片不匹配。三类图像均以3.2 s的时间间隔呈现,每张图像显示300 ms。从每张图像出现开始,电极采集1 s的脑电信号。对于单刺激,受试者无需作出反应;对于双刺激,受试者需判断两张图像是否一致,进行按键操作响应。实验范式流程如图2所示。

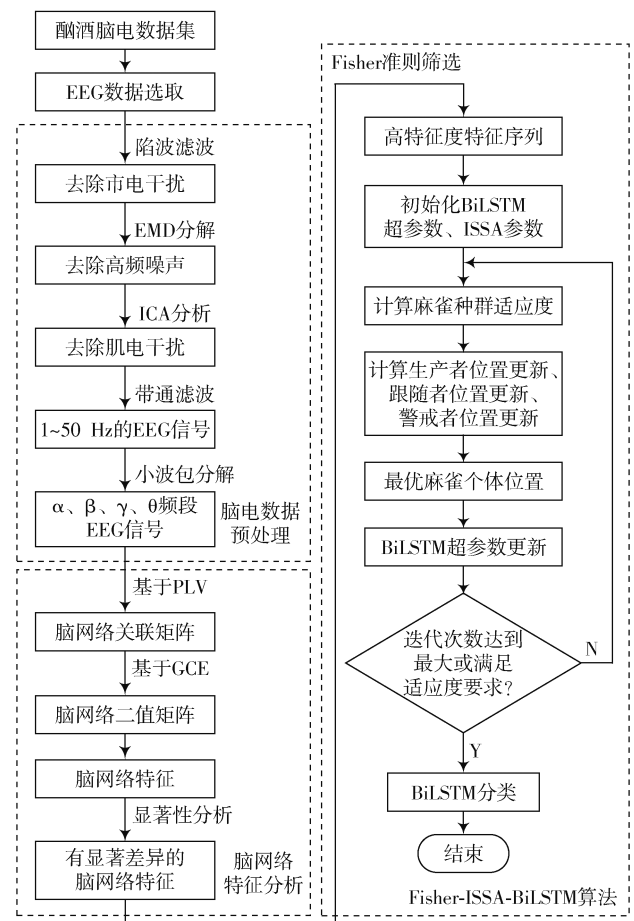


图1 本文算法的主要工作流程

因受试者的年龄、职业、心理素质等方面不同,判断不同图片产生的脑电信号存在差异,该差异并不完全取决于酗酒这一单一变量因素。为控制单一变量,降低不同刺激类型造成信号差别的影响,本文仅选取单刺激事件(S1)开展EEG分析。

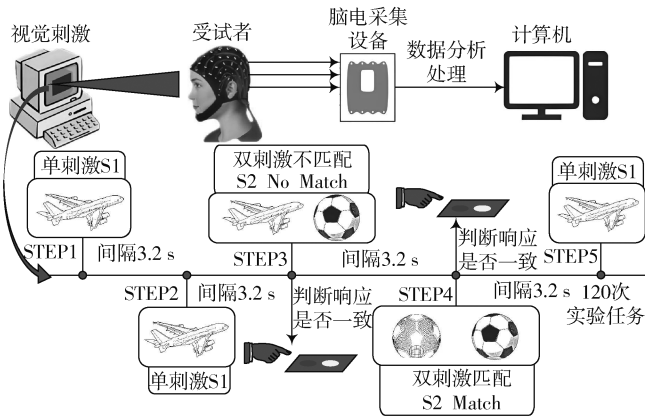


图2 实验范式流程

2 数据预处理

由于原始的 EEG 数据中包含市电、肌电等干扰,需要对 EEG 数据进行预处理,去除伪迹干扰。不同频段的 EEG 信号有助于识别更丰富的信息,由于小波包分解更适合非平稳随机信号的分析与检测,因此采用小波包分解进行频段划分^[7]。

预处理流程如图 3 所示。预处理后得到 θ 、 α 、 β 、 γ 和全频段的 EEG 信号数据。

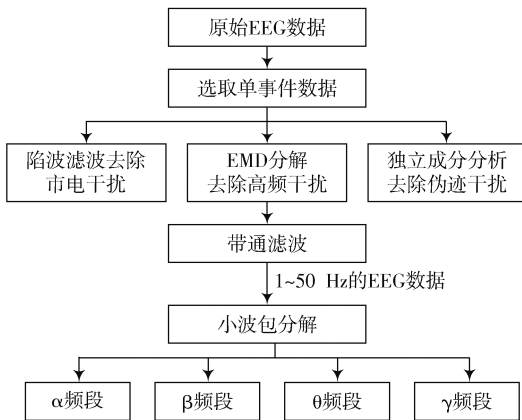


图3 EEG数据预处理流程

3 脑网络特征

3.1 构建脑网络矩阵

在 EEG 数据预处理的基础上,以 1 s 为时间窗,以 64 导电极为节点构建关联矩阵。酗酒通过影响脑电相位同步性,干扰大脑的认知功能^[8]。传统功能性连接指标,如皮尔逊相关系数和互信息,主要评估信号间的线性相关性,这些方法对相位同步性的直接反映能力有限。相位锁定值是衡量两个信号相位锁定程度的统计量,被广泛用于衡量信号间的相位同步性。选用 PLV 构建 α 、 β 、 γ 、 θ 和全频段的脑网络连接,获得关联矩阵,并对关联矩阵二值化。矩阵二值化的关键是阈值的选取,

当两个导联间的 PLV 大于阈值时,二值矩阵相应元素设为 1,反之;相应元素设为 0。本文采用全局代价效率 (Global Cost Efficiency, GCE) 的方法,通过设定强权值 (Proportion of the Strong Weight, PSW) 的比例对矩阵进行阈值化处理。

PSW 是指在阈值化过程中保留的强连接权值与所有连接权值总数的比率。文献[9]运用图论分析复杂网络,结果表明当全局代价效率 (GCE) 达到最大值时,对应的 PSW 所确定的阈值是最优的。

$$GCE = E - PSW \quad (1)$$

$$E = \frac{\sum_{a \neq b} L_{ab}}{N(N-1)} \quad (2)$$

式中:GCE 为全局效率与 PSW 值之间的差; L_{ab} 表示节点 a 和 b 间的最短距离; N 为导联数量。通过该阈值确定方法获取脑网络的阈值。

构建两个受试者的全频段二值矩阵,如图 4 所示。图 4a) 为“co2a0000364”数据的二值矩阵,图 4b) 为“co2c0000337”数据的二值矩阵。

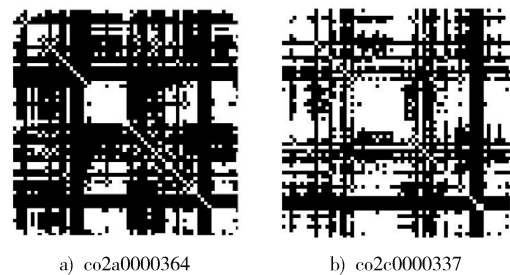


图4 两个受试者的全频段二值矩阵

3.2 脑网络特征分析

选取酗酒和正常受试者各频段脑网络特征的中值比较分析,通过 p 值分析,得出特征差异较大的导联和特征。在显著性水平为 0.05 的情况下,如果 p 值比显著性水平 0.05 小,样本集之间存在显著性差异,反之,样本集之间不存在显著性差异。存在显著性差异的导联如图 5 所示。

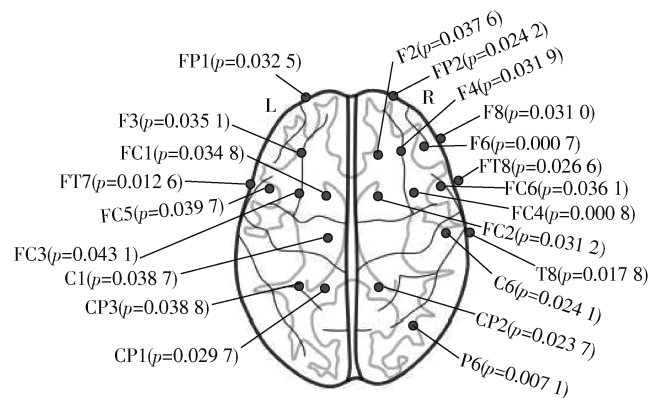


图5 存在显著性差异的导联

4 算法模型方法

基于脑网络特征分析,本文提出一种 Fisher-ISSA-BiLSTM 的分类算法,进行 EEG 信号的分类。算法分为以下三部分:一是运用 Fisher 准则有效识别和筛选脑网络特征,减少不相关特征的干扰;二是运用 BiLSTM 网络对脑网络特征参数进行分类;三是提出改良的麻雀搜索算法优化 BiLSTM 算法的超参数,提升算法的分类准确率,减少算法运行时间。Fisher-ISSA-BiLSTM 算法流程如图 6 所示。

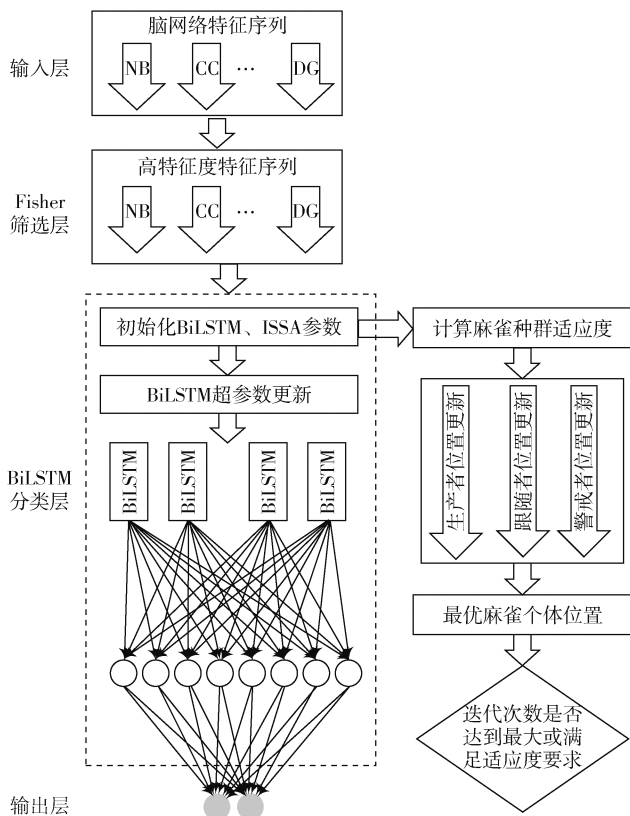


图 6 Fisher-ISSA-BiLSTM 算法流程

4.1 基于 Fisher 准则筛选特征

Fisher 准则能够处理高维数据,将高维数据投影到低维空间中,在处理噪声和异常数据时,具有较好的鲁棒性,算法具体如下。

定义数据集中共有 m 个类别,每个类别的样本数量为 n , μ_i 表示同一个类别中第 i 个样本的均值, $\bar{\mu}$ 表示所有样本的均值, μ_i^j 表示同一个类别中第 i 个样本中的第 j 个值,Fisher 是特征的 Fisher 判别比。在训练样本集中,特征的 Fisher 比值越大,表明该特征的区分能力越强,越有利于特征分类。

$$Fisher = \frac{\sum_{i=1}^m (\mu_i - \bar{\mu})^2}{\sum_{i=1}^m \sum_{j=1}^n (\mu_i^j - \mu_i)^2} \quad (3)$$

$$\mu_i = \frac{1}{n} \sum_{j=1}^n \mu_i^j \quad (4)$$

$$\bar{\mu} = \frac{1}{m} \sum_{i=1}^m \mu_i \quad (5)$$

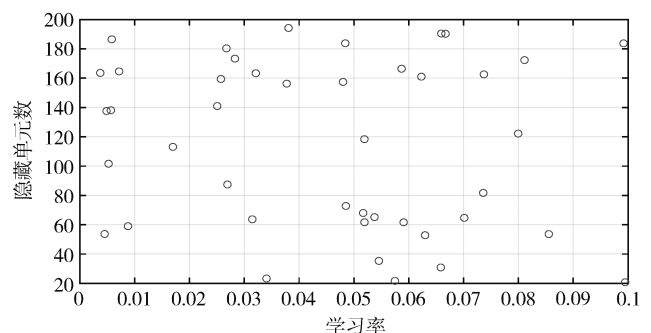
本文将 Fisher 准则应用于脑网络特征参数的筛选,对每维特征参数的 Fisher 比按从高到低排序,选取 Fisher 比最高的特征参数分量作为特征序列。

4.2 改良麻雀搜索算法

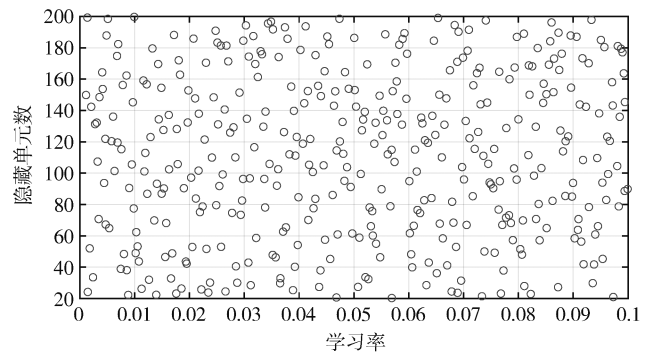
因传统的麻雀搜索算法迭代后期生产者快速收敛于 0 且向全局最优值的位置集中,这种收敛行为可能导致种群多样性的急剧下降,容易陷入局部最优解等问题^[10]。本文提出一种改良的麻雀搜索算法(ISSA),改进麻雀种群的初始化方式、位置更新机制、引入搜索因子,扩大了麻雀搜索范围,避免过早陷入局部最优。

4.2.1 改进麻雀种群的初始化方式

传统麻雀算法的初始种群是随机的,会造成种群的非均匀分布,削弱了算法的全局搜索能力。在初始化种群过程中引入混沌策略能更好地覆盖解空间。Sobol 映射是一种低差异混沌序列,它生成的点集在高维空间分布更均匀,提升了算法的全局覆盖率^[11]。本文算法采用 Sobol 映射初始种群,图 7a)为改进前麻雀搜索算法随机初始种群,图 7b)为 Sobol 映射生成初始种群。对比可知,Sobol 映射初始种群比随机初始种群产生的种群分布更加均匀。



a) 改进前麻雀搜索算法



b) Sobol 映射

图 7 不同方法生成的初始种群

4.2.2 改进麻雀种群的初始化方式

算法初期,生产者进入广泛搜索,有利于算法全面探索解空间。然而,生产者位置中的 $\exp\left(-\frac{i}{k \cdot \text{iter}_{\max}}\right)$ 具有随机性,当该值较大时,具有较强的全局搜索能力;当该值较小时,主要表现为局部搜索。若生产者位置变化过于剧烈,会导致生产者频繁聚集与扩散。因此,合理调控该参数值对于麻雀搜索算法至关重要。文献[12]提出将惯性权重这一概念引入群智能算法中,能够平衡粒子群算法全局搜索和局部搜索的能力。为约束 $\exp\left(-\frac{i}{k \cdot \text{iter}_{\max}}\right)$ 值在迭代中的变化,在生产者位置更新公式中引入搜索因子 ω 。搜索因子 ω 在算法迭代中能够线性变化,通过影响生产者位置,提升算法的收敛速度和精度。当 ω 值较大时,算法进入一种广泛的搜索模式,随着 ω 的减小,算法将聚焦于局部搜索,从而提升算法的收敛精度。搜索因子 ω 具体表达式如式(6)所示:

$$\omega = \left| \sin\left(\frac{\pi}{2} + \frac{\pi t}{2 \times \text{iter}_{\max}}\right) \right| \quad (6)$$

式中: iter_{\max} 为总迭代次数,取 500; t 为当前迭代次数。

搜索因子 ω 变化曲线如图 8 所示。

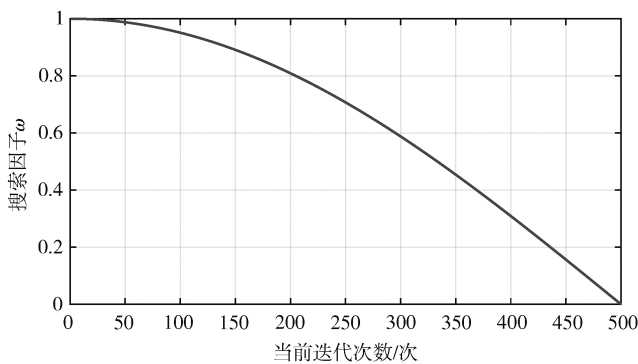


图 8 搜索因子 ω 变化曲线

4.2.3 改进生产者位置更新方式

如果生产者在前几轮迭代就接近全局最优解,会导致麻雀搜索区域受限,使算法陷入局部最优^[13]。为解决该问题,本文以生产者的个体位置更新方式为切入点,构建全局最优解与生产者位置之间的联系。改进后生产者位置如式(7)所示:

$$x_{ij}^{t+1} = \begin{cases} \omega x_{ij}^{t+1} \cdot \exp\left(-\frac{i}{k \cdot \text{iter}_{\max}}\right) + \omega(x_p^t - x_{ij}^t), & R_2 < ST \\ \omega x_{ij}^t + Q \cdot L, & R_2 \geq ST \end{cases} \quad (7)$$

式中: x_p^t 为当前迭代的全局最优位置; x_{ij}^t 为第 i 个麻雀在第 t 次迭代时在第 j 维中的位置信息; $k \in (0,1]$, 是一个随

机值; Q 是遵循标准正态分布的随机数; L 是一个元素均为 1 的 $1 \times d$ 矩阵; R_2 和 ST 分别表示预警值和安全阈值。若 $R_2 \geq ST$, 则表明环境中出现了捕食者的迹象, 生产者的主要任务为迅速识别潜在的威胁并逃离到安全区域; 若 $R_2 < ST$, 则表示当前种群处于相对安全的状态。生产者可以有效地利用其搜索能力, 为群体提供丰富的食物资源。

4.2.4 改进跟随者位置更新方式

由于跟随者位置更新公式 $Q \cdot \exp\left(\frac{x_{\text{worst}}^t - x_{ij}^t}{i^2}\right)$ 具有

随机性, 跟随者可能会探索到适应度更差的空间区域, 增加了获取最优解的难度^[14]。为了尽量避免这一情况的发生, 在保证跟随者位置更新具有随机性的基础上, 增加对跟随者飞行方向的考量。若跟随者本次迭代位置 x_{ij}^t 的适应度比前一次迭代位置 x_{ij}^{t-1} 的适应度更优, 说明麻雀可能往适应度更优的空间探索, 需要引导麻雀继续向这一空间方向探索; 相反, 麻雀可能往适应度更差的方向飞行, 需要引导麻雀避免相同的飞行方向。基于这一思想, 将麻雀先后迭代位置的差值 ($x_{ij}^t - x_{ij}^{t-1}$) 加入跟随者位置更新公式, 同时考虑这一差值对位置更新影响的大小。因此, 麻雀跟随者位置更新公式中引入自适应方向因子 d_{ij}^t , 表示前次粒子更新趋势对本次粒子迭代影响的大小。自适应方向因子 d_{ij}^t 由麻雀当前迭代位置 x_{ij}^t 的适应度值 $f(x_{ij}^t)$ 和前一次迭代位置 x_{ij}^{t-1} 的适应度值 $f(x_{ij}^{t-1})$ 决定, 它的区间范围是 $[-1,1]$, 计算公式如式(8)所示:

$$d_{ij}^t = f(x_{ij}^{t-1}) - f(x_{ij}^t) \quad (8)$$

基于这一思想, 改进后的跟随者位置更新公式如式(9)所示:

$$x_{ij}^{t+1} = \begin{cases} Q \cdot \exp\left(\frac{x_{\text{worst}}^t - x_{ij}^t}{i^2}\right) + d_{ij}^t(x_{ij}^t - x_{ij}^{t-1}), & i > \frac{n}{2} \\ x_p^{t+1} + |x_{ij}^t - x_p^{t+1}| \cdot A^+ \cdot L, & i \leq \frac{n}{2} \end{cases} \quad (9)$$

式中: x_p 为目前生产者所占据的最好位置; x_{worst} 为全局最差位置; n 是种群大小; A 为一个 $1 \times d$ 的矩阵, 其中每个元素随机赋值为 1 或 -1, 且 $A^+ = A^T(AA^T)^{-1}$ 。

5 实验结果

5.1 改良麻雀搜索算法的性能测试

本文使用单峰测试函数和多峰测试函数测试 ISSA 算法的性能, 并对 SSA、PSO、GWO、WOA 算法开展对比测试。测试函数维数为 30, 最优值为 0。为确保对比实验的公平性, 每个算法的种群规模统一设为 30, 迭代次

数均为 100,让算法独立运行 30 次,对得到的结果选取最优值。测试函数、取值范围、算法种类及实验结果如表 1、表 2 所示。

表 1 单峰测试函数及实验结果

函数	取值范围	算法	最优值
$F_1(x) = \sum_{i=1}^n x_i^2$	[-100,100]	PSO	2.640 8e-56
		GWO	7.023 3e-87
		WOA	3.012 5e-94
		SSA	0
		ISSA	0
$F_2(x) = \sum_{i=1}^n x_i + \prod_{i=1}^n x_i $	[-10,10]	PSO	2.864 9e-52
		GWO	2.587 6e-67
		WOA	1.022 6e-71
		SSA	1.022 6e-95
		ISSA	0
$F_3(x) = \sum_{i=1}^n \left(\sum_{j=1}^i x_j \right)^2$	[-100,100]	PSO	6.264 4e-55
		GWO	1.459e-84
		WOA	2.783 3e-92
		SSA	0
		ISSA	0
$F_4(x) = \max\{ x_i , 1 \leq i \leq n\}$	[-100,100]	PSO	2.605 2e-53
		GWO	4.026 7e-68
		WOA	1.231e-72
		SSA	1.231e-96
		ISSA	0
$F_5(x) = \sum_{i=1}^{n-1} [100(x_{i+1} - x_i)^2 + (x_i - 1)^2]$	[-30,30]	PSO	2.380 7e-52
		GWO	1.212 9e-68
		WOA	3.477 7e-72
		SSA	3.477 7e-97
		ISSA	0
$F_6(x) = \sum_{i=1}^n (x_i + 0.5)^2$	[-100,100]	PSO	6.549 1e-52
		GWO	9.418 6e-44
		WOA	4.574 7e-43
		SSA	4.406 9e-47
		ISSA	6.163e-93
$F_7(x) = \sum_{i=1}^n ix_i^4 + \text{random}[0,1)$	[-1.28,1.28]	PSO	1.297 2e-26
		GWO	3.299 8e-39
		WOA	1.715 1e-41
		SSA	3.451 2e-77
		ISSA	0

算法在运行单峰测试函数、多峰测试函数时的迭代

收敛情况如图 9 和图 10 所示。

表 2 多峰测试函数及实验结果

函数	取值范围	算法	最优值
$F_8(x) = \sum_{i=1}^n \left[-x_i \sin(\sqrt{ x_i }) \right]$	[-500,500]	PSO	4.355 2e-15
		GWO	3.441 3e-70
		WOA	6.959 7e-71
		SSA	0
		ISSA	0
$F_9(x) = \sum_{i=1}^n [x_i^2 - 10 \cos(2\pi) + 10]$	[-5.12,5.12]	PSO	2.364 3e-18
		GWO	1.065 8e-35
		WOA	4.371 7e-36
		SSA	7.105 4e-54
		ISSA	2.225e-92
$F_{10}(x) = -20 \exp\left(-0.2 \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}\right) - \exp\left(\frac{1}{n} \sum_{i=1}^n \cos(2\pi x_i)\right) + 20 + e$	[-32,32]	PSO	4.684 6e-16
		GWO	4.945 7e-44
		WOA	1.826 6e-41
		SSA	1.826 6e-89
		ISSA	0
$F_{11}(x) = \frac{1}{4000} \sum_{i=1}^n x_i^2 - \prod_{i=1}^n \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$	[-600,600]	PSO	4.402 7e-31
		GWO	6.125 7e-49
		WOA	2.359 8e-51
		SSA	2.359 8e-85
		ISSA	0

由以上实验运行结果和测试函数收敛曲线可知,经过 100 次迭代,ISSA、SSA、PSO、GWO 以及 WOA 算法基本能收敛到一个比较小的适应度值。无论是单峰测试函数,还是多峰测试函数,ISSA 算法收敛值均优于 SSA、PSO、GWO 以及 WOA 算法,且在单峰测试函数 $F_1 \sim F_5$ 、 F_7 和多峰测试函数 F_8 、 F_{10} 、 F_{11} 算法求解的最优值均达到理论最优值 0。

从算法的收敛速度看,在单峰测试函数和多峰测试函数中,ISSA 算法的收敛速度最快,基本在迭代 50 次以内能够收敛到最优解。综合收敛精度和收敛速度,ISSA 算法均明显优于 SSA、PSO、GWO 以及 WOA 算法,验证了 ISSA 算法的能力。

5.2 Fisher 特征筛选结果

由于脑网络特征属性较多,算法运算时间较长,因此运用 4.1 节的 Fisher 准则筛选得到 Fisher 比, Fisher 比能够一定程度上反映出特征的贡献度,通过 Fisher 比排序得到贡献度较高的特征序列。

算法运算结果如图 11 所示。

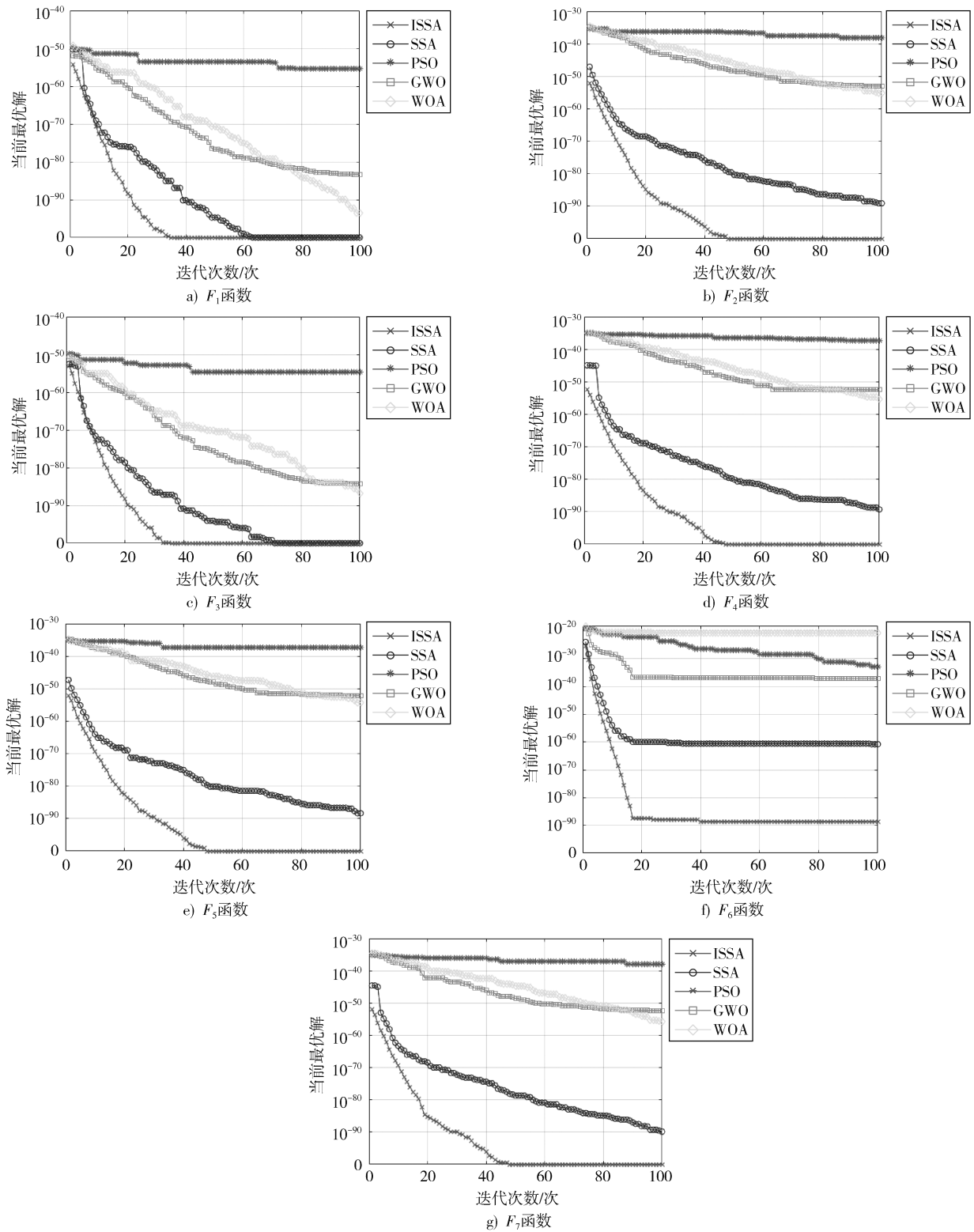


图9 算法在单峰测试函数上的收敛曲线

根据 Fisher 计算结果可知, α 、 β 频段的特征 Fisher 比值相对较高, 特征的辨别能力更强。将 Fisher 比值大于 25 的特征集合 {NB-all, CC- β , CPL- θ , CPL- γ , LE- α } 整合为第一梯队特征子集 G_1 ; 将 Fisher 比值在 20~25 的特征集合 {DG- α , NB- β , CC- θ , LE-all, LE- β } 整合为第二梯队

特征子集 G_2 ; 将 Fisher 比值在 10~20 的特征集合 {DG-all, DG- β , NB- α , CC-all, CC- α , CC- γ , LE- θ } 整合为第三梯队特征子集 G_3 ; 将 Fisher 比值在 0~10 的特征集合 {DG- θ , DG- γ , NB- θ , NB- γ , CPL-all, CPL- α , CPL- β , LE- γ } 整合为第四梯队特征子集 G_4 。

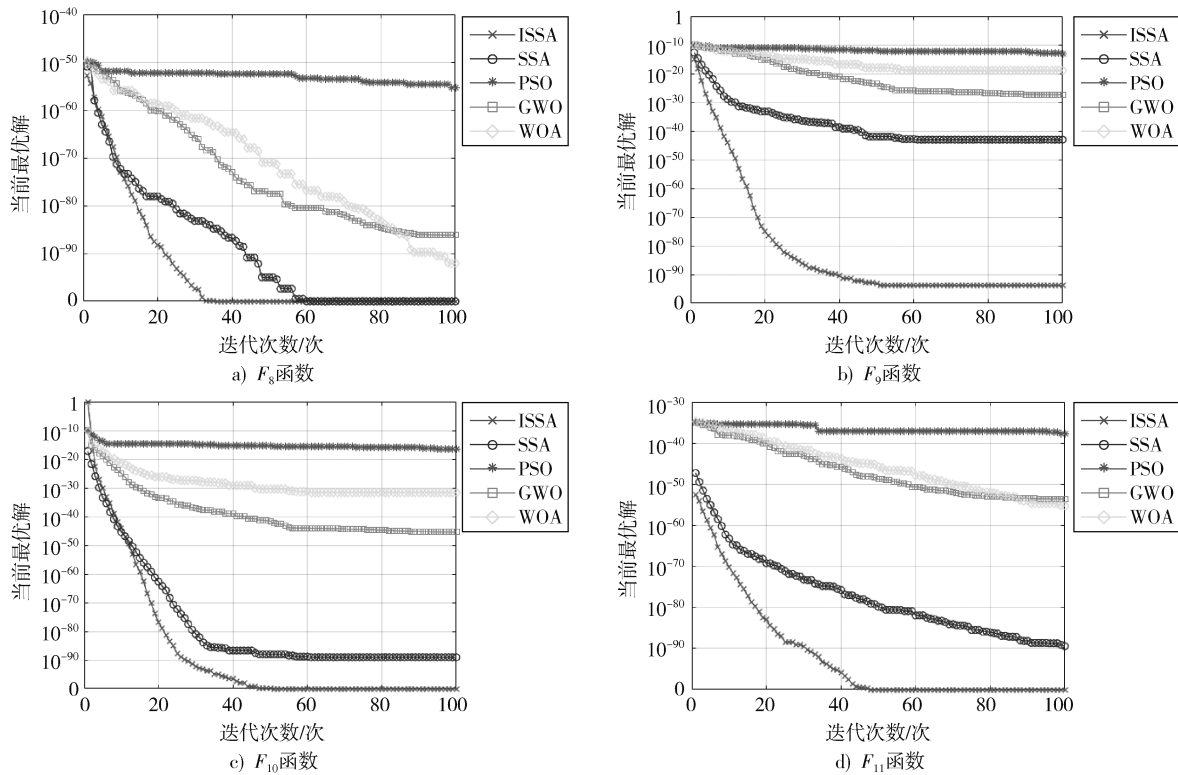


图 10 算法在多峰测试函数上的收敛曲线

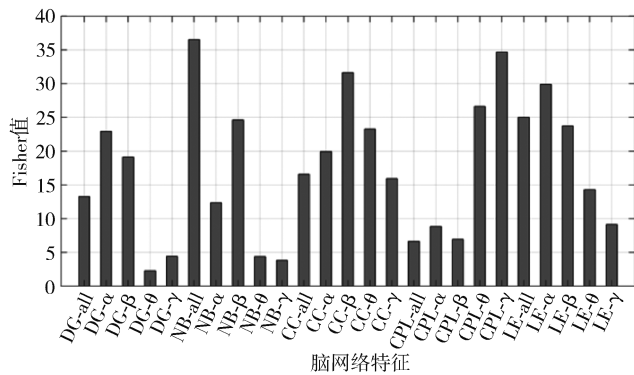


图 11 各频段脑网络特征的 Fisher 值

为了验证不同梯队特征子集对分类效果的贡献差异,将以上特征子集 G_1, G_2, G_3, G_4 输入 BiLSTM 算法进行分类比较,迭代次数为 100 次,分类准确率分别为 76.6%、71.2%、65.9%、57.2%,由此可知以上特征子集对分类贡献效果为 $G_1 > G_2 > G_3 > G_4$ 。取特征子集 G_1 和 G_2 的集合 G_{12} ,特征子集 G_1, G_2 和 G_3 的集合 G_{123} ,特征子集 G_1, G_2, G_3 和 G_4 的集合 G_{1234} ,运用 BiLSTM 算法进行分类,分类准确率分别为 75.3%、74.6% 和 71.7%,算法准确率并没有得到提升,而算法运行的时间大大增加,因此选取第一梯队特征子集 G_1 作为 Fisher-ISSA-BiLSTM 算法的特征输入进行分类。

5.3 超参数优化结果

本文以交叉熵损失函数为适应度函数,输入特征均

为经过 Fisher 筛选后的特征子集 G_1 。运用 ISSA 算法对 BiLSTM 的超参数进行优化,为验证 ISSA 的算法优化效果,分别用 SSA、PSO、WOA、GWO 算法对超参数寻优的效果进行对比。BiLSTM 算法的初始超参数取值范围如下:学习率 $L_r \in [0.000, 1, 0.5]$,网络隐含单元数 $N_h \in [50, 200]$,批量大小 $B_s \in [10, 200]$ 。将算法的初始超参数学习率 L_r 设为 0.01,网络隐含单元数 N_h 设为 100,批量大小 B_s 设为 20。算法迭代 100 次,不同算法得到的最优超参数组合不同,最优超参数组合见表 3。各算法的寻优迭代曲线如图 12 所示。

表 3 最优超参数组合结果

算法	L_r	N_h	B_s
Fisher-SSA-BiLSTM	0.099	118	117
Fisher-PSO-BiLSTM	0.147	135	89
Fisher-WOA-BiLSTM	0.127	97	68
Fisher-GWO-BiLSTM	0.082	132	142
Fisher-ISSA-BiLSTM	0.185	158	124

由图 12 可知, Fisher-ISSA-BiLSTM 算法相比其他算法拥有较好的寻优能力,且算法的收敛速度比其他算法更快,这是由于 Fisher-ISSA-BiLSTM 算法引入了搜索因子 ω 和自适应方向因子 d'_i ,使算法早期寻优范围更广,不易陷入局部最优,迭代过程中麻雀向适应度更优的方

向更新位置,提升了算法运行的效率。

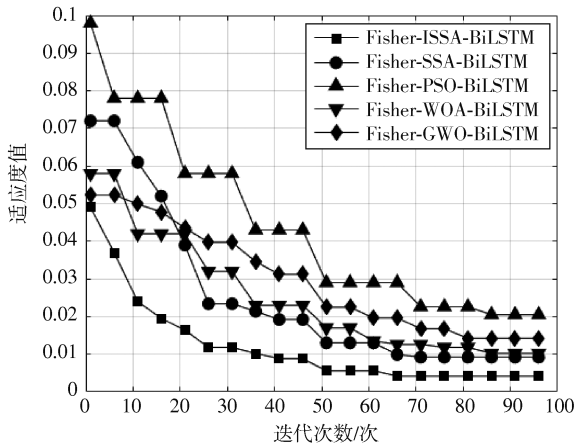


图 12 各算法的寻优迭代曲线

5.4 算法分类准确率

将算法优化后的超参数组合输入 BiLSTM 模型,实现数据的分类。将以上数据采用十折交叉检验的方法,得到平均的算法分类准确率。将本文算法与 Fisher-LSTM、Fisher-BiLSTM、Fisher-SSA-BiLSTM、Fisher-PSO-BiLSTM、Fisher-WOA-BiLSTM、Fisher-GWO-BiLSTM 进行比较,算法分类结果如表 4 所示。

表 4 算法分类结果 %

算法	平均分类准确率
Fisher-LSTM	72.2
Fisher-BiLSTM	76.6
Fisher-PSO-BiLSTM	86.9
Fisher-WOA-BiLSTM	89.7
Fisher-GWO-BiLSTM	88.4
Fisher-SSA-BiLSTM	90.9
Fisher-ISSA-BiLSTM	92.6

Fisher-BiLSTM 的平均分类准确率高于 Fisher-LSTM,这是由于 BiLSTM 能捕捉到 EEG 序列中的过去和未来的信息,并对其进行递归和反馈。同时,运用粒子优化超参数的算法平均分类准确率均显著高于 Fisher-BiLSTM,这是由于相比依靠经验确定超参数的取值,粒子优化超参数的策略能够从客观角度寻找最优超参数组合。在超参数粒子寻优算法中,Fisher-ISSA-BiLSTM 算法的平均分类准确率最好,这是由于算法加入了 Sobol 映射,使得算法的初始粒子分布更均匀,寻优范围更大,在生产者位置更新算法中加入了搜索因子 ω ,在算法前期扩大了搜索范围,避免了算法过早陷入局部最

优的情况。

5.5 算法运行时间

将 Fisher-ISSA-BiLSTM 算法的运行时间与其他超参数寻优算法对比,结果如表 5 所示。

表 5 算法运行时间对比 s

算法	运行时间
Fisher-PSO-BiLSTM	8 374
Fisher-WOA-BiLSTM	8 267
Fisher-GWO-BiLSTM	8 198
Fisher-SSA-BiLSTM	8 126
Fisher-ISSA-BiLSTM	7 892

由表 5 可知,Fisher-ISSA-BiLSTM 的运行时间较短,这是由于在初始化过程中引入了 Sobol 映射,使得麻雀种群的初始化分布相对均匀,减少了算法迭代的时间;同时,在麻雀跟随者位置更新算法中引入自适应方向因子 d_{ij} ,引导麻雀向适应度更优的方向飞行,提升了算法迭代的效率。

6 结 语

本文从脑电信号具有连续性的角度出发,提出一种 Fisher-ISSA-BiLSTM 的分类算法,相比 SVM 等机器学习分类算法,该算法充分挖掘了脑网络特征。为进一步探究脑电特征的前后关联性,对脑网络特征运用 Fisher 准则进一步筛选,采用双向长短时记忆 (BiLSTM) 网络进行特征识别和分类,并运用改良的麻雀搜索算法优化 BiLSTM 的参数,提高了酗酒者的分类准确率,减少了分类算法运行的时间,对于运用脑电信号识别酗酒患者具有一定的实际意义。

注:本文通讯作者为吕卓言。

参 考 文 献

- [1] UHLMANN A, BANDELOW B, STEIN D J, et al. Grey matter structural differences in alcohol-dependent individuals with and without comorbid depression/anxiety: an MRI study [J]. European archives of psychiatry and clinical neuroscience, 2019, 269(3): 285-294.
- [2] 刘桂青,曹锐,相洁.酗酒者皮层脑电同步性分析[J].中国生物医学工程学报,2016,35(6):671-676.
- [3] CAO J, BALL I K, SUMMERELL E, et al. Effect of ethanol on brain electrical tissue conductivity in social drinkers [J]. Journal of magnetic resonance imaging, 2024, 61(3): 1181-1187.
- [4] 丁尚文,王纯贤.基于支持向量机的酗酒脑电信号分类研究 [J].自动化仪表,2019,40(11):95-98.

(下转第 29 页)

