

DOI:10.19651/j.cnki.emt.2417098

基于时-频-空域特征表征方法的自闭症儿童诊断^{*}

王亚民 潘礼正 闵云霄 石旻弘
(常州大学机械与轨道交通学院 常州 213164)

摘要: 自闭症谱系障碍是一组复杂的神经系统疾病,通常出现在幼儿期,目前,自闭症儿童的诊断主要依赖于行为观察和诊断量表,然而由于儿童的某些行为症状可能不明显,诊断结果往往存在一定的主观性。为了提高自闭症儿童早期诊断辨识的准确性,本文提出一种基于时-频-空三域特征以及改进快速相关滤波算法的诊断方法。首先利用脑电信号中时-频-空特征之间的互补性全面分析脑功能网络。其次,采用改进快速相关滤波算法进行特征优化,筛选出相关但非冗余的特征,最后采用BP-Adaboost分类器辨识诊断。通过实验对比分析发现,该模型效果较为优异,BP-Adaboost分类器具有较高的辨识精度,平均诊断准确率达到98.72%,该模型可用作辅助神经科医生诊断自闭症的辅助工具。

关键词: 自闭症;脑连接性;特征优化;模式识别

中图分类号: TN911.7 **文献标识码:** A **国家标准学科分类代码:** 510.40

Diagnosis of autistic children based on temporal-spectral-spatial feature representation

Wang Yamin Pan Lizheng Min Yunxiao Shi Minhong

(College of Mechanical Engineering and Rail Transit, Changzhou University, Changzhou 213164, China)

Abstract: Autism spectrum disorder is a group of complex neurological disorders that usually appear in early childhood. At present, the diagnosis of autistic children mainly relies on behavioral observation and diagnostic scales. However, some behavioral symptoms of children may not be obvious, the diagnosis results are general subjective. In order to improve the accuracy of early diagnosis and identification of autistic children, the paper proposes the diagnosis method based on temporal-spectral-spatial three-domain features and improved fast correlation based filter. Firstly, the complementarity between temporal-spectral-spatial features of EEG signals is used to analyze the brain functional network. Secondly, the improved fast correlation based filter algorithm is used to optimize the features and screen out the relevant but non-redundant features. Finally, BP-Adaboost classifier is used for identification and diagnosis. Through comparative analysis of experiments, it is found that the model has excellent effect, and the BP-Adaboost classifier has a higher identification accuracy, with an average diagnostic accuracy of 98.72%. The model can be used as an auxiliary tool to assist neurologists in diagnosing autism.

Keywords: autism spectrum disorder; brain connectivity; feature optimization; pattern recognition

0 引言

孤独症谱系障碍(autism spectrum disorder, ASD)是一组复杂的神经系统疾病,包括自闭症、儿童分裂性障碍、阿斯伯格综合症和一种未定义类型的广泛性发育障碍^[1],其临床上主要表现为3个核心特征:社会交往障碍、语言交流障碍以及刻板重复行为^[2],虽然目前没有能够治愈ASD的方法,但早期干预可以改善学习、沟通以及大脑发育,因

此需要一种高效且准确的方法来诊断ASD^[3]。

人类大脑包含大约860亿个神经元,是人类已知的最复杂的生物系统,能够调节思想,感知,记忆,感觉和行动^[4]。研究人员通常选择磁共振成像、电子计算机断层扫描和脑电信号(electroencephalography, EEG)作为研究ASD的客观工具。除此以外,研究人员还采用了几种不同的医学成像方法以进行不同疾病的自动诊断,包括正电子发射断层扫描,功能磁共振成像和扩散张量磁共振成像^[5]。

收稿日期:2024-10-13

^{*} 基金项目:江苏省研究生科研创新计划项目(SJCX22_1408)、江苏省研究生科研创新计划项目(KYCX23_3095)资助

然而,由于成像技术的昂贵成本等因素,这些方法并不被认为具有脑电图的优势。EEG 由于其出色的时间分辨率、方便性、非侵入性和低成本而被广泛使用^[6]。

特征工程已经逐步成为诊断 ASD 的一种重要手段,可以帮助临床医生提高分析质量并减少分析时间^[7]。Heunis 等^[8]通过递归归一化分析提取特征,用随机分析确定最优特征集,并将其输入支持向量机(support vector machine, SVM)分类器,取得了较好的实验结果。Abdolzadegan 等^[9]从 EEG 信号中提取了线性和非线性特征,并结合各种特征选择技术(信息增益,互信息,最小冗余最大相关性)进行特征优化,结合 SVM 和 K-近邻(K-nearest neighbor, KNN)分类器进行辨识性能分析。Jahmunah 等^[10]从 EEG 信号中提取了 18 个非线性特征,使用学生 t 检验进行特征选择,并采用 Marginal Fisher 分析进行数据缩减,最后分别选择具有 1、2 和 3 阶多项式内核的 SVM 分类器对最优特征集进行分类,其中 2 阶多项式内核的 SVM 辨识性能最优。目前已有较多特征表征方法并且取得了一定的成果,但现存的脑电训练模型大多仅单独使用了脑电信号的时域、频域、空域信息或是以上两种特征的组合。这些模型忽略了脑电信号时-频-空特征之间的互补性,会在一定程

度上限制脑电分类模型的性能表现。此外,在特征选择方面,现存的方法则是只考虑了特征与类别的相关性和特征间的冗余性,没有考虑特征组合发挥正向协同作用。因此,采用合适的特征表征方法对提高计算机辅助诊断神经系统疾病具有重要意义。

综上所述,本文提出了一种有效的特征表征方法以协助 ASD 的诊断。在对脑电信号进行预处理后,利用脑电信号中时-频-空(temporal-spectral-spatial, TSS)特征之间的互补性对脑电信号进行特征提取,从而有效地提取信号的关键信息。其次采用了改进快速相关滤波算法(improved fast correlation based filter, IFCBF)进行特征优化,去除冗余且负相关的特征。最后采用不同的分类方法验证本文所提出模型的普适性,包括 KNN、SVM、长短时记忆网络(long short-term memory, LSTM)、BP 神经网络和 BP-Adaboost。

1 数据和方法

为提高自闭症诊断精度,如图 1 所示,本文提出了一种 TSS-IFCBF 的状态表征方法,最后结合 BP-Adaboost 强分类器进行辨识诊断。

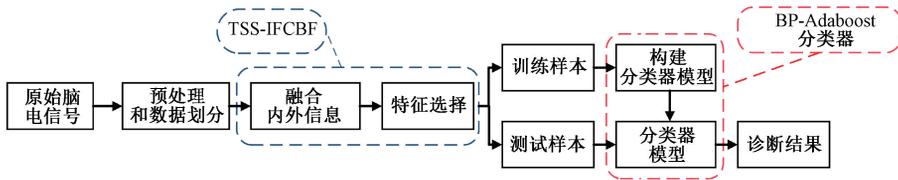


图 1 整体方法的框图

Fig. 1 Block diagram of the overall approach

1.1 数据采集

本文实验使用的数据集由沙特阿拉伯吉达的阿卜杜勒阿齐兹国王大学医院发布^[11],该数据集的受试者包括 12 名患有自闭症的儿童(9 男 3 女,年龄 6~20 岁)以及 4 名从未患有任何神经疾病的正常儿童(均为男孩,年龄 9~13 岁)。使用带有 Ag/AgCl 电极的 g. tec EEG 帽、G. tech USB 放大器和 BCI2000 软件在放松条件下从参与者收集信号,以生成无伪影的 EEG 数据。使用国际 10~20 导联系统(AFz 为 GND,右耳垂为参考^[12])从 16 个电极通道(FP1、FP2、F7、F3、Fz、F4、F8、T3、C4、Cz、C3、T5、Pz、O1、Oz 和 O2)收集数据,电位拓扑图如图 2 所示。

1.2 基于 TSS-IFCBF 的状态表征方法

为了更有效地反映自闭症患者和正常人的脑电信号差异,本文通过融合时-频-空三域的方式来描述原始脑电信号。时域包括样本熵、近似熵、排列熵、小波熵;频域包括重心频率、均方根频率、频率标准差、功率谱密度;人脑是一个有机整体的空间,信息传递任务需要大脑皮层各区域协同工作,在脑功能分析中,空域的信息也起着重要的作用。先前的研究表明,分析两个信号通道之间的同步可

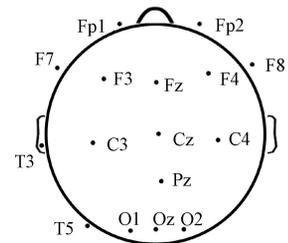


图 2 16 个电极通道的拓扑图

Fig. 2 Topology diagram of 16 electrode channels

能有助于揭示大脑各个空间之间的相互作用^[13]。因此,为了积极探索大脑空域的信息特征,本文通过式(1)计算脑电信号通道对的互信息值表征大脑空域信息。图 3 是正常组和孤独症组受试者的不同类型特征的脑地形图:(1)分别为样本熵、近似熵、排列熵、小波熵;(2)为功率谱密度;(3)分别为重心频率,均方根频率,频率标准差。从图 3 中看出孤独症组儿童的 16 个通道信号的功率谱密度高于正常组,而其他特征则是正常组高于孤独症组。

$$I(X;Y) = \sum_{x \in X} \sum_{y \in Y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)} \quad (1)$$

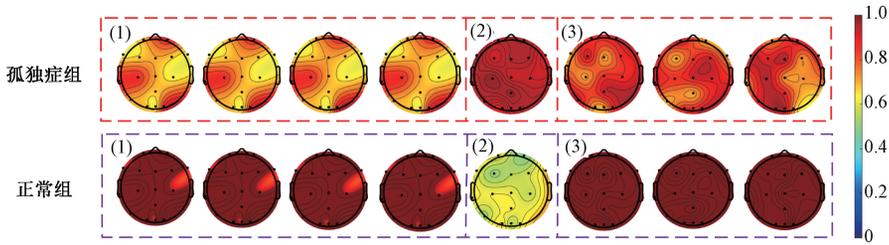


图 3 正常组和孤独症组受试者的不同特征脑地形图

Fig. 3 Brain topography maps with different characteristics of normal and autistic subjects

式中： $p(x, y)$ 是 X 和 Y 的联合概率分布函数， $p(x)$ 和 $p(y)$ 分别是 X 和 Y 的边缘概率分布函数。

本文提取脑电信号的级联特征来表征原始脑电信号，包括时-频-空三域特征。对自闭症 EEG 信号进行特征表征后，为去除不相关、冗余或嘈杂的特征，本文选择相关性快速过滤特征选择算法 (fast correlation-based filter, FCBF) 对特征进行优化。然而在直接使用 FCBF 删除冗余特征时，存在可能导致信息丢失的情况：当两个特征对于目标变量的相关性都较大时，被认定为近似马尔科夫链的特征仍属于关键特征，直接删除该特征可能导致有用信息被误删，因此本文通过 IFCBF 进一步优化特征表征。该算法主要步骤如下：

步骤 1) 计算每个特征 f_i 的 $SU_{i,c}$ 值，其中 i 是特征标签， c 是类向量；

步骤 2) 删除 $SU_{i,c} < \delta$ 不相关特征，其中 δ 是预定义的阈值；

步骤 3) 根据 SU 值降序列出剩余特征，并将结果标记为 S_{list} ；

步骤 4) 将 S_{list} 中的第一个元素定义为主要特征 f_m ，将 f_m 附加到 S_{select} 并将其从 S_{list} 中删除。对于剩余的特征 f_n (从 f_m 右边的特征到 S_{list} 中的最后一个特征)，计算 $SU_{m,n}$ 。如果 $SU_{m,n} > SU_{m,c}$ ，将 f_n 附加到 $S_{remnant}$ 并将其从

S_{list} 中删除；

步骤 5) 将 S_{list} 中 f_m 旁边的剩余特征视为新的主要特征，并重复步骤 4)，直到 S_{list} 为空集。最后，获得次特优特征子集 S_{select} 以及候补特征子集 $S_{remnant}$ ；

步骤 6) 将 $S_{remnant}$ 中的一个特征 X_i 附加到 S_{select} ，根据式(2)计算加入 X_i 前后 $J(S_{select})$ ，若 S_{select} 的整体价值下降，则删除该特征，反之保留该特征。以此类推，重复步骤 5) 直至分析完候补特征子集 $S_{remnant}$ 中所有特征，最后得到最优特征子集 S_{best} 。

$$J(Z) = \frac{n\overline{SU}(X, Y)}{\sqrt{n + n(n-1)\overline{SU}(X_i, X_j)}} \quad (2)$$

$$\overline{SU}(X, Y) = \sum_{X_i \in F} \frac{SU(X_i, Y)}{n} \quad (3)$$

$$\overline{SU}(X_i, X_j) = \sum_{X_i \in F} \sum_{X_j \in F, i \neq j} \frac{SU(X_i, X_j)}{n(n-1)} \quad (4)$$

式中： $J(Z)$ 代表一个由 n 个特征组成的特征集 Z 的价值。 $\overline{SU}(X, Y)$ 是特征-标签之间的平均相关性， $\overline{SU}(X_i, X_j)$ 是特征-特征之间的平均相关性。

其中，IFCBF 特征筛选过程分为 3 个阶段。首先用 SU 值区分特征，其次通过比较特征与特征之间的相关性，区分特征是否冗余，最后计算特征集合价值，区分特征是否具有互补性，该方法的流程图如图 4 所示。

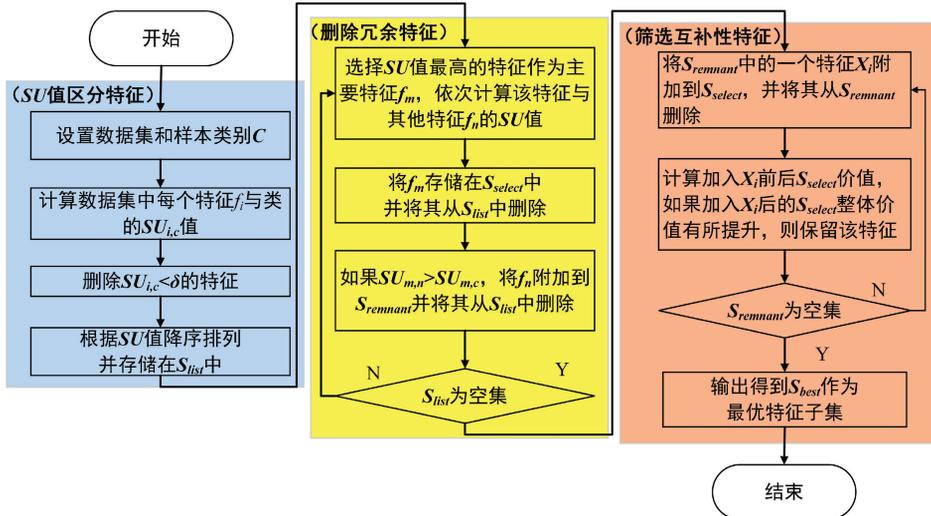


图 4 IFCBF 特征选择的流程图

Fig. 4 Flow chart of IFCBF feature selection

由上述步骤所示,想要获得最优特征子集 S_{best} , 阈值的选择方式是重要的影响因素, 考虑到阈值 δ 的选择和每个特征 f_i 与类标签 c 之间的 SU 值有关, 本文根据式(5)选取最佳阈值。

$$\delta = (SU_{i,c})_{\min} + \frac{k \times ((SU_{i,c})_{\max} - (SU_{i,c})_{\min})}{100} \quad (5)$$

式中: $k=10\sim 90$, δ 为阈值。

1.3 BP-Adaboost

BP神经网络是由输入层、隐藏层和输出层组成的前馈网络^[14], 如图5所示。BP神经网络是最典型的人工神经网络, 网络通过动态调整连接权值和阈值优化输入与输出矢量间的映射, 将实际输出与期望输出的偏差最小化, 从而实现复杂分类识别和非线性建模^[15]。

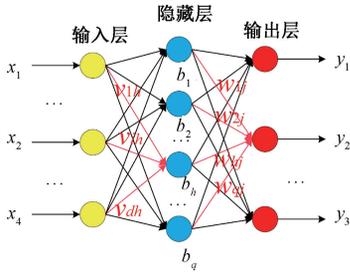


图5 BP神经网络结构

Fig. 5 BP neural network structure

自适应增强(Adaboost)算法的思想是合并多个弱分类器的输出以产生有效分类^[16]。它是通过改变数据分布和加性模型实现的。该算法定义了样本权重的概念, 用该权重来确定每个样本的重要程度, 并对多个基分类器的结论进行加权和, 从而得到一个强分类器。

BP-Adaboost模型^[17]即把BP神经网络作为弱分类器, 反复训练BP神经网络预测样本输出, 通过Adaboost算法得到多个BP神经网络弱分类器组成的强分类器。算法步骤如下:

步骤1) 初始化训练数据的分布权值:

$$D_t(i) = 1/m \quad (6)$$

式中: t 为当前迭代次数, m 为样本总数。

步骤2) 训练第 t 个弱分类器时, 用训练数据训练BP神经网络并且预测训练数据输出, 得到预测序列 $g(t)$ 的预测误差和 e_t , 根据误差和 e_t 计算预测序列权重 a_t , 计算公式为:

$$e_t = \sum_i D_t(i) \quad (7)$$

$$a_t = \frac{1}{2} \ln\left(\frac{1-e_t}{e_t}\right), (g(t) \neq y) \quad (8)$$

式中: $g(t)$ 为预测分类结果; y 为期望分类结果。

步骤3) 根据预测序列权重 a_t 调整下一轮训练样本的权重, 调整公式为:

$$D_{t+1}(i) = \frac{D_t(i)}{B_t} \times \exp[-a_t y_i g_t(x_i)] \quad (9)$$

式中: B_t 是归一化因子, 其使得 D_{t+1} 的所有元素和为1。

步骤4) 训练 T 轮后得到 T 组弱分类函数 $f(g_t, a_t)$, 将其组合得到了强分类器 $h(x)$ 。

$$h(x) = \text{sign}\left[\sum_{t=1}^T a_t \cdot f(g_t, a_t)\right] \quad (10)$$

2 实验结果及分析

为了确保实验的稳定性和合理性, 本文将原始 EEG 数据集以 8:2 的比例划分为训练集和测试集, 此外, 在量化分析时每类实验重复 5 次以减小偶然性造成的影响, 以 5 次平均值作为实验结果。

2.1 特征表征

本文结合时-频-空三域特征对 EEG 进行状态表征。为了进一步验证时-频-空三域特征的有效性, 本文比较了 7 种特征模型: 时域特征、频域特征、空域特征、时-频特征、时-空特征、频-空和时-频-空融合特征, 并选择 BP-Adaboost 作为分类器来演示性能, 为了确保实验的可靠性, BP-Adaboost 分类器参数均保持一致。从图6中可以看出, 时-频-空融合特征模型的平均辨识指标高于其他 6 种模型, 说明融合时-频-空三域信息更能有效表征大脑信号状态。

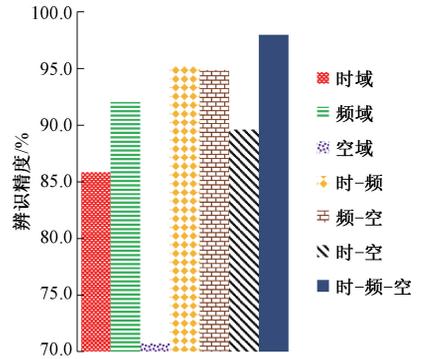


图6 不同特征模型的辨识精度

Fig. 6 Identification accuracy of different feature models

2.2 特征选择方法比较

为了有效表征 EEG 信号, 本文提出了融合时-频-空三域特征的表征方法, 相比于单独使用了脑电信号的时域、频域、空域信息或是以上两种特征的组合的方式, 融合时-频-空三域特征可能会导致模型维度偏高, 而训练模型中存在一些与分类目标无关或与其他特征具有高相关性的特征, 导致模型构建复杂、训练时间增加、性能下降, 本文采用 IFCBF 方法解决了这一问题。与采用所有类型的特征共同作为特征向量相比, 采用 IFCBF 方法进行特征选择时, 提高了辨识精度, 降低了模型的训练复杂度。为了验证本文阈值选择公式的有效性, 研究了不同 k 值的分类性能, 相应的辨识结果图7所示。从图7中可以看出, 随着 k 值增大, 两种特征选择方法的辨识准确率先上升后下降, 且在 k 为 20 时达到最大值, 说明存在一个理论最优 k 值使得性能达到最优状态, 充分证明了本文选择的阈值公式

的有效性。从图 7 中可以看出当 k 值大于 50 时,FCBF 算法的特征维度显著降低,导致辨识精度下降;IFCBF 算法由于并未直接删去冗余特征子集,而是将其作为候补特征

子集计算其与次优特征子集之间的相关性进行最优特征子集的优化,随着 k 值的变化,IFCBF 算法的辨识精度并没有显著变化。

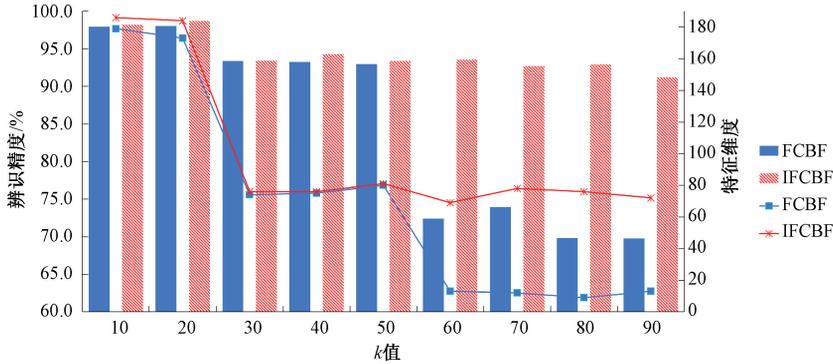


图 7 FCBF 和 IFCBF 方法的特征维度和辨识精度

Fig. 7 Feature dimension and identification accuracy of FCBF and IFCBF methods

为了观察经过 IFCBF 进行特征优化后通道内各个特征的贡献,本文统计了 TSS-IFCBF 模型中单个特征占总特征总数的平均百分比,如图 8 所示。从图中可以看出样本熵的特征贡献率最高,达到了 15%,而频率标准差和均方根频率的特征贡献率最低,这说明孤独症组和正常组儿童的样本熵差异显著,这可能是由于自闭症儿童的脑电信号复杂度较小,自相似性较高;同时为了分析脑电信号通道的贡献率,本文统计了所有受试者每个通道所选择的特征的平均百分比,从图 9 可以看出中央区 and 额叶区域通道内的时域特征具有很高的重要性,而频域特征方面则是额叶区域和顶叶区域的通道相对而言更加重要,空域方面则是前额叶和中央空间区域的通道与其他脑区的通道的连接性起着重要作用。此外,从图 9 中可以看出融合三域特征后,额叶区域的通道对辨识结果影响最大。

为了进一步证明 IFCBF 方法的有效性,IFCBF 方法和其他典型的特征选择方法进行了比较,选择互信息 (mutual information, MI)、最大相关性最小冗余性 (max-relevance and min-redundancy, mRMR)、ReliefF 和 FCBF 方法代替 IFCBF 方法进行特征选择,将所选特征进行归一化,并将 BP-Adaboost 用作分类器。从图 10 中可以看出,FCBF 和 IFCBF 特征选择方法的效果明显高于其他典型的特征选择方法,其主要原因是对称不确定性是一种有效的相关性度量,它比传统的距离度量和互信息度量更能描述变量之间的非线性关系。此外,与 FCBF 方法相比,采用 IFCBF 方法进行特征选择时,辨识精度普遍提高。这是因为原 FCBF 方法只计算候选特征与主导特征之间的冗余度,没有考虑到候选特征与次优特征子集中的特征组合后发挥正向协同作用,而 IFCBF 没有直接删去候补特征,而是评估加入候补特征后给整个已选特征集合带来的影响,进一步优化特征子集,得到最优特征子集。从以上分析可以看出,IFCBF 方法比 FCBF 方法更有效。

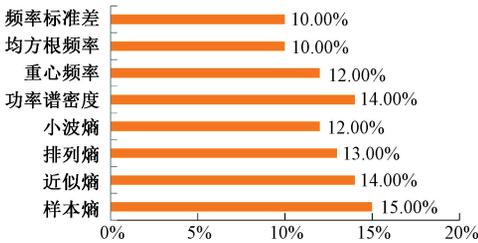


图 8 特征贡献率

Fig. 8 Feature contribution rate

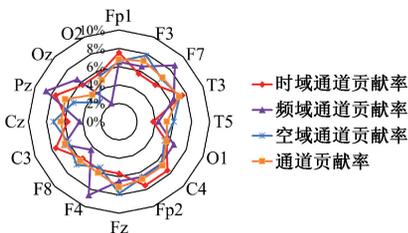


图 9 通道贡献率

Fig. 9 Channel contribution rate

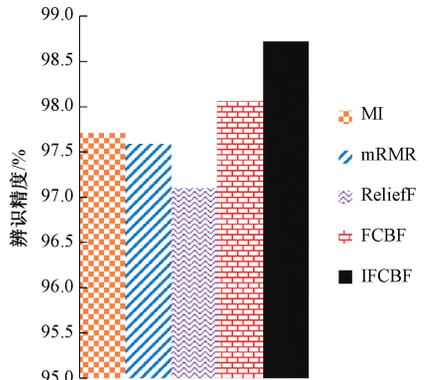


图 10 特征选择方法的对比

Fig. 10 Comparison of feature selection methods

2.3 辨识方法分析

在本实验中,对脑通道内外信息进行提取和融合,然后利用 IFCBF 方法进行特征优化,最后选择分类器实现识别。为了充分验证 BP-Adaboost 分类器的有效性,进行了 KNN、SVM、LSTM、BP 和 BP-Adaboost 之间的比较。KNN 方法是基于欧氏距离作为点距离,通过交叉验证确定 k 的最优值。SVM 采用径向基函数作为核函数,采用网格搜索法确定惩罚参数 c 和核函数参数 g 。五个分类器的识别结果如表 1 所示,表 1 中 Acc 表示准确率 (accuracy), Pe 表示精度 (precision), Re 表示召回率 (recall)。从表中可以看出,5 个分类器的平均性能指标均达到了 95% 以上,说明本文提出的融合时-频-空三域特征的特征方法以及选择 IFCBF 作为特征优化的策略具有普适性。此外,与 KNN、SVM 等典型模式识别方法相比, BP-Adaboost 具有更高的辨识精度。

表 1 不同方法的性能指标

Table 1 Performance indicators of different methods

性能指标/%	Acc	Pre	Re	F1
KNN	97.75	95.67	98.95	97.28
SVM	98.02	96.81	98.38	97.59
LSTM	97.17	96.39	96.66	96.53
BP	97.10	95.77	97.14	96.45
BP-Adaboost	98.72	98.29	98.57	98.43

为了进一步分析本模型的性能,本文绘制了 TSS-IFCBF-BP-Adaboost 模型所获得的混淆矩阵图和受试者工作特征(receiver operating characteristics, ROC)曲线图。图 11 的行代表真实类,而图 11 的列代表预测类。如图 11 所示,来自正常组和自闭症组的 1 034 和 1 515 个样本被正确分类,来自自闭症组和正常组的 18 和 15 个样本被错误预测。TSS-IFCBF-BP-Adaboost 模型的 ROC 曲线图如图 12 所示。从 ROC 曲线中可以注意到,它立即沿着真正速率轴上升,表明本模型具有优异的性能。从以上分析可以看出,本文所提出的 TSS-IFCBF-BP-Adaboost 模型是一种有效的自闭症诊断方法。

为了客观地验证本文方法的有效性,表 2 显示了现有最先进方法与本文方法的综合对比总结,同时为了保证对比实验的鲁棒性,表 2 中所有方法均使用相同的 ASD 脑电数据库。文献[18]使用快速傅立叶变换(FFT)进行特征提取,然后选择 Fisher 线性判别分析(fisher linear discriminant, FLD)进行分类。文献[19]使用多层感知器神经网络(multilayer perceptron network, MLPN)来实现自闭症的诊断。文献[12]采用短时傅里叶变换(short-time Fourier transform, STFT)将 EEG 转换为频谱图像,结合局部三元模式和 Census 变换提取纹理特征并以 SVM 作

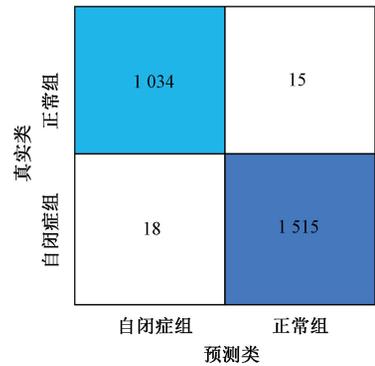


图 11 针对 TSS-IFCBF-BP-Adaboost 模型的混淆矩阵

Fig. 11 Confusion matrix for TSS-IFCBF-BP-Adaboost model

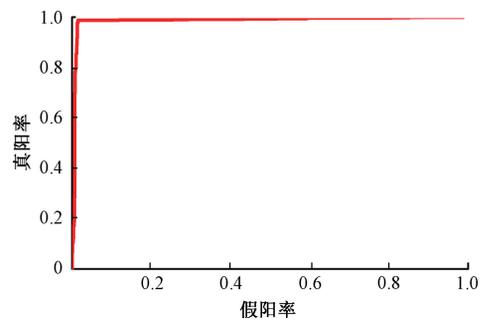


图 12 TSS-IFCBF-BP-Adaboost 模型效果的 ROC 曲线

Fig. 12 ROC curve of the effect of TSS-IFCBF-BP-Adaboost model

为分类器进行分类辨识。文献[20]同样采用 STFT 将 EEG 转换为频谱图像,然而为了避免特征工程的广泛使用,选择利用卷积神经网络训练图像模型,最后采用基于 3 个卷积神经网络的加权集成算法对 ASD 患者进行分类。文献[21-22]利用离散小波变换(DWT)将脑电信号分解为若干子带,从若干子带中提取香农熵(shannon entropy, SE)作为特征,分别选择 KNN 和 ANN 作为辨识方法。文献[23]使用连续小波变换提取了各种统计特征(statistical features, ST),即平均值,标准差,偏度和峰度,并使用 SVM 作为分类器。此外,文献[24]同样采用小波变换提取了统计特征,分类器则选择了梯度提升算法(gradient boosting, GB)。文献[25]提取时域特征(temporal features, TF),并使用具有时-频包装层的双向堆叠门控递归单元(bidirectional stacked gated recurrent unit, BSGRU)的概念实现了较高的准确度。在表 2 中,本文所提出的技术实现了 98.72% 的诊断准确率,优于上述文献的研究方案,这主要归功于以下几个因素:1)充分考虑了脑通道信号的时-频-空域特征的互补性,更加全面地对信号进行表征;2)利用所提出的 IFCBF 对特征进行高效优化,降低特征维度的同时提高了辨识精度;3)在弱分类器 BP 的基础上采用 Adaboost 算法组合成强分类器,提高了模型的泛化能力和学习能力。

表 2 与现有方法的性能比较

Table 2 Performance comparison with existing methods

文献	特征表征	分类器	准确率/%
[18]	FFT	FLDA	80.27
[19]	MLPN	MLPN	80.00
[12]	STFT	SVM	95.25
[20]	STFT	CNN	94.71
[21]	DWT-SE	KNN	94.21
[22]	DWT-SE	ANN	98.20
[23]	ST	SVM	95.00
[24]	ST	GB	97.15
[25]	TF	BSGRU	97.00
本文	TSS-IFCBF	BP-Adaboost	98.72

3 结 论

为了提高自闭症诊断精度,本文提出了一种基于时-频-空特征之间的互补性,改进 FCBF 和 BP-Adaboost 的自闭症辨识方法。首先通过融合时-频-空特征表征 EEG 信号,更加全面地揭示大脑区域之间的通信异常。其次,针对传统 FCBF 的相关性控制问题,提出了一种改进的 FCBF 方法,用于选择相关但非冗余的特征。通过大量实验比较分析,充分表明了本文所提信号表征方法的有效性。最后,采用 Adaboost 算法构建多 BP 弱分类器模型,并将 BP-Adaboost 分类器用于自闭症诊断。最终,该方法自动检测 ASD 的准确率为 98.72%,精确率为 98.29%,召回率为 98.57%,F1 评分为 98.43%。

参考文献

- [1] 李翔,马昕,李贻斌. 基于视觉行为分析的自闭症谱系障碍辅助诊断研究综述[J]. 生物医学工程学杂志, 2023,40(4):812-819.
LI X, MA X, LI Y B. A review of studies on visual behavior analysis aided diagnosis of autism spectrum disorders [J]. Journal of Biomedical Engineering, 2023, 40(4): 812-819.
- [2] MEMISEVIC H, DJIPA A. Content analysis of abstracts published in autism journals in 2021: The year in review [J]. Journal of Autism and Developmental Disorders, 2023, 53(12): 4849-4855.
- [3] OH S L, JAHMUNAH V, ARUNKUMAR N, et al. A novel automated autism spectrum disorder detection system [J]. Complex & Intelligent System, 2021, 7(5): 2399-2413.
- [4] SIULY S, ALCIN F O, BAJAJ V, et al. Exploring hermite transformation in brain signal analysis for the detection of epileptic seizure [J]. IET Science, Measurement & Technology, 2019, 13(1): 35-41.

- [5] 谢士遥,汤佳贝,蔡雨,等. 脑电 BCI 系统的软硬件开发平台发展现状[J]. 电子测量与仪器学报, 2022, 36(6):1-12.
XIE SH Y, TANG J B, CAI Y, et al. Review on software and hardware platforms for EEG-based BCI system [J]. Journal of Electronic Measurement and Instrumentation, 2022, 36(6): 1-12.
- [6] 朱康,吴晓红,郭远哲,等. 基于 MTACNet 网络的运动想象脑电分类[J]. 电子测量技术, 2023,46(16):24-31.
ZHU K, WU X H, GUO Y ZH, et al. Motor imagination EEG classification based on MTACNet [J]. Electronic Measurement Technology, 2023, 46(16): 24-31.
- [7] TOLSGAARD M G, BOSCARDIN C K, PARK Y S, et al. The role of data science and machine learning in health professions education: Practical applications, theoretical contributions, and epistemic beliefs [J]. Advances in Health Science Education, 2020, 25(5): 1-30.
- [8] HEUNIS T, ALDRICH C, PETERS M J, et al. Recurrence quantification analysis of resting state EEG signals in autism spectrum disorder systematic methodological exploration of technical and demographic confounders in the search for biomarkers [J]. BMC Medicine, 2018, 16(1): 1-17.
- [9] ABDOLZADEGAN D, MOATTAR M H, GHOSHUNI M. A robust method for early diagnosis of autism spectrum disorder from EEG signals based on feature selection and DBSCAN method [J]. Biocybernetics and Biomedical Engineering, 2020, 40(1): 482-493.
- [10] JAHMUNAH V, ARUNKUMAR N, ABDULHAY E W, et al. A novel automated autism spectrum disorder detection system [J]. Complex & Intelligent Systems, 2021, 7(5): 2399-2413.
- [11] LI J, KONG X L, SUN L L, et al. Identification of autism spectrum disorder based on electroencephalography: A systematic review [J]. Computers in Biology and Medicine, 2024, 170:108075.
- [12] TAWHID M N A, SIULY S, WANG H. Diagnosis of autism spectrum disorder from EEG using a time-frequency spectrogram image-based approach [J]. Electronics Letters, 2020, 56(25): 1372-1375.
- [13] WANG M, HU J K, ABBASS H A. BrainPrint: EEG biometric identification based on analyzing brain connectivity graphs [J]. Pattern Recognition, 2020, 105: 107381.
- [14] SHEN W Z, LI G, WEI X L, et al. Assessment of

- dairy cow feed intake based on BP neural network with polynomial decay learning rate [J]. *Information Processing in Agriculture*, 2022, 9(2): 266-275.
- [15] 何兴霖,赵丽,边琰,等.基于BP神经网络的运动想象脑电信号分类研究[J].*电子测量技术*,2022,45(21):123-129.
- HE X L, ZHAO L, BIAN Y, et al. Research on EEG signal classification of motor imagery based on BP neural network [J]. *Electronic Measurement Technology*, 2022, 45(21): 123-129.
- [16] 潘礼正,王顺超,丁忆,等.基于全视域特征表征与ELM-Adaboost方法的情绪辨识研究[J].*仪器仪表学报*,2022,43(11):162-171.
- PAN L ZH, WANG SH CH, DING Y, et al. Emotion recognition research based on the full-view feature representation and ELM-Adaboost[J]. *Chinese Journal of Scientific Instrument*, 2022, 43 (11): 162-171.
- [17] ZHAO C C, LIANG H Y, SHAN Y, et al. Bearing fault diagnosis method based on comprehensive information divergence and improved BP-AdaBoost algorithm[J]. *Structural Health Monitoring*, 2023, 22(5): 3047-3064.
- [18] ALSAGGAF E A, KAMEL M I, et al. Using EEGs to diagnose autism disorder by classification algorithm[J]. *Life Science Journal*, 2014, 11(6): 305-308.
- [19] ALISA N A. Autism spectrum disorder classification on electroencephalogram signal using deep learning algorithm[J]. *IAES International Journal of Artificial Intelligence(IJ-AI)*, 2020, 9(1):91-99.
- [20] ULLAH Z M, YU D. Grid-tuned ensemble models for 2D spectrogram-based autism classification [J]. *Biomedical Signal Processing and Control*, 2024, 93: 106151.
- [21] IBRAHIM S, DJEMAL R, ALSUWAILEM A. Electroencephalography (EEG) signal processing for epilepsy and autism spectrum disorder diagnosis[J]. *Biocybernetics and Biomedical Engineering*, 2018, 38(1):16-26.
- [22] RIDHA D, KHALIL A, SUTRISNO I, et al. EEG-based computer aided diagnosis of autism spectrum disorder using wavelet, entropy and ANN[J]. *Biomed Research International*, 2017, 2017:9816591.
- [23] MELINDA M, JUWONO F H, ENRIKO I K, et al. Application of continuous wavelet transform and support vector machine for autism spectrum disorder electroencephalography signal classification [J]. *Radioelectronic and Computer Systems*, 2023, 29(3): 73-90.
- [24] THIRUMAL S, THANGAKUMAR J. Investigation of hybrid feature selection techniques for autism classification using EEG signals [J]. *International Journal of Advanced Computer Science and Applications(IJACSA)*, 2022, 13(4):0130475.
- [25] WADHERA T, BEDI J, SHARMA S. Autism spectrum disorder prediction using bidirectional stacked gated recurrent unit with time-distributor wrapper: An EEG study[J]. *Neural Computing and Applications*, 2023, 35(13): 9803-9818.

作者简介

潘礼正(通信作者),博士研究生,副教授,主要研究方向为人机共融康复机器人、智能诊疗技术等。

E-mail:plz517@sina.com

王亚民,硕士研究生,主要研究方向为脑电信号处理与模式识别。

E-mail:wym2815486551@163.com