

文章编号: 2095-2163(2024)03-0128-05

中图分类号: TP18, TP399

文献标志码: A

基于平滑非线性能量算子划分的尖峰相关特征癫痫发作自动检测算法

何雪兰, 吴江, 蒋路茸

(浙江理工大学信息科学与工程学院, 杭州 310018)

摘要: 针对癫痫发作自动检测算法多集中于时域、频域等传统特征, 无法全面表征癫痫脑电信号的信息等问题, 本文结合癫痫脑电图中异常波振幅和频率提高的现象, 提出一种基于平滑非线性能量算子划分的尖峰相关特征癫痫发作自动检测算法。该算法使用传统的时域、频域特征, 结合尖峰相关性特征对脑电信号进行刻画, 使用有监督的机器学习分类器, 测试癫痫发作自动检测的有效性和可靠性。本文将提出的方法在开源数据集 CHBMIT 上进行了评估, 获得了 96.52% 的准确率、95.65% 的敏感性和 97.09% 的特异性。实验结果表明, 基于平滑非线性能量算子划分的尖峰相关特征, 能够作为癫痫脑电信息的补充, 提高癫痫发作检测的性能。

关键词: 癫痫发作检测; 机器学习; 尖峰相关性; 平滑非线性能量算子

Automatic seizure detection algorithm based on spike-related features of smoothed nonlinear energy operator division

HE Xuelan, WU Jiang, JIANG Lurong

(School of Information Science and Engineering, Zhejiang Sci-Tech University, Hangzhou 310018, China)

Abstract: Most current seizure automatic detection algorithms focus on traditional features such as time domain and frequency domain, which cannot fully characterize the information of epileptic EEG signals. This paper proposes an automatic seizure detection algorithm based on spike correlation features divided by a smooth nonlinear energy operator, taking into account the phenomenon that the amplitude and frequency of abnormal waves in epileptic EEG will increase. The algorithm uses traditional time-domain and frequency-domain features, combined with spike correlation features to characterize the EEG signal, and uses supervised machine learning classifiers to test its effectiveness and reliability for automatic seizure detection. The research evaluates the proposed method on the open source dataset CHBMIT and obtains 96.52% on accuracy, 95.65% on sensitivity and 97.09% on specificity. The experimental results show that the proposed spike-related features based on the smoothed nonlinear energy operator segmentation can be used as a complement to the epileptic EEG information to improve the performance of seizure detection.

Key words: seizure detection; machine learning; spike correlation; smoothed nonlinear energy operator

0 引言

癫痫是一种神经系统疾病, 由大脑神经元异常放电引起^[1], 常常表现为突发性、反复性和复发性等特点。癫痫发作的临床症状复杂多样, 如阵发性痉挛、意识丧失、认知功能障碍等^[2]。这些发作事件对患者的认知水平及正常生活都产生了明显影响。因此, 癫痫的诊断和治疗对于预防癫痫发作和改善生活质量至关重要。

头皮脑电图是一种用于临床记录脑活动的无创

信号采集方法^[3], 用于记录大脑活动时的电波变化。头皮脑电图包含丰富的生理、心理和病理信息, 是评估癫痫和其他脑部疾病的有效工具^[4]。在脑电图的记录中, 癫痫发作和癫痫样放电(如棘波、尖波和棘慢波复合体)是癫痫的重要生物标志物^[5], 并被广泛应用于临床评价。目前, 临床上基于脑电图的识别与分析是医生进行癫痫检测的黄金标准, 但对海量的临床脑电数据进行人工筛查, 不仅给医生带来沉重的负担, 还存在较强的主观性、判断标准不统一等问题^[6-7], 影响分析的效率和准确性。因

基金项目: 浙江省基础公益项目(LGF19F010008); 北京邮电大学泛网无线通信教育部重点实验室(BUPT)(KFKT-2018101); 浙江省重点研发计划(2022C03136); 国家自然科学基金(61602417)。

作者简介: 何雪兰(1999-), 女, 硕士研究生, 主要研究方向: 癫痫检测; 吴江(1978-), 男, 博士, 高级工程师, 主要研究方向: 无线通信技术, 工业物联网。

通讯作者: 蒋路茸(1982-), 男, 博士, 教授, 主要研究方向: 生理电信号处理、复杂网络和无线传感器网络。Email: jianglurong@zstu.edu.cn

收稿日期: 2023-03-15

此,设计一种自动的癫痫发作检测方法是亟待解决的问题。

为了克服传统诊断方法的局限性、提高医疗效率,伴随着机器学习的快速发展,癫痫发作的自动检测已成为行业内关注的重点。研究者们根据头皮脑电图的时域、频域或非线性特征建立了特征工程方法^[8-10],并通过具有一个或多个特征的分类器检测癫痫发作。Mursalin等学者^[11]从时域、频域和基于熵的特征中选择突出特征,使用随机森林分类器学习选定特征集合的特性,获得了更好的分类结果。

杨舒涵等学者^[12]使用时域和非线性特征对脑电信号进行表征,结合XGBoost分类器,实现了癫痫的自动检测。Zarei等学者^[13]使用离散小波变换DWT和正交匹配追踪(Orthogonal Matching Pursuit, OMP)提取EEG中不同的系数,计算非线性特征和统计特征,使用SVM进行分类,获得了较好的检测性能。吴端波等学者^[14]使用aEEG尖峰和cEEG棘波提取的方法计算棘波率,使用阈值法对癫痫进行发作检测。

上述模型虽然都能取得较好的分类结果,但是也存在以下问题:

(1)多数研究在特征提取阶段仅从时域、频域或时频域中表征脑电信号信息,这些特征所涵盖的信息量并不足以全面描述一段EEG信号。

(2)在癫痫发作的自动检测中,强调周期性的信号转换对于有效、可靠地区分癫痫发作的重复特征至关重要,而互相关是时域上广泛用于表示信号周期性的方法。

针对上述问题,本文提出一种基于平滑非线性能量算子划分的尖峰相关(Spike Correlation, SC)特征的癫痫发作自动检测算法。SC是关于自适应提取的脑电图尖峰信号段之间时间延迟的最大互相关。使用平滑非线性能量算子衡量癫痫脑电信号中出现的异常波,将脑电信号在癫痫发作期和非发作期的尖峰相关特征作为度量患者大脑活动的一个重要补充。本文提出的算法主要使用巴特沃斯滤波器对脑电信号进行滤波,去除外部伪迹的干扰,然后从传统特征角度出发,提取时域、频域特征,再结合提出的尖峰相关特征,进一步表征癫痫发作时的异常信息。最后结合有监督的机器学习分类模型,实现癫痫发作的自动检测。

1 方法

癫痫发作自动检测整体流程设计如图1所示,

其中包含预处理、特征提取和分类等3个模块。

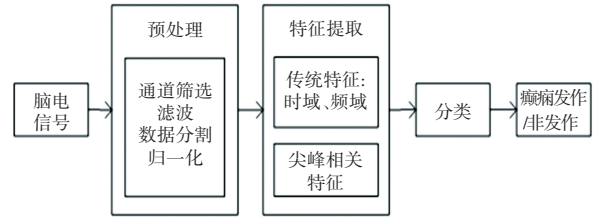


图1 癫痫发作自动检测流程图

Fig. 1 Flow chart of seizure detection

1.1 脑电信号预处理

头皮脑电数据通过放置在头皮固定位置的电极采集得到。由于外置电极,这种采集方式很容易受到外部干扰,导致采集到的数据被噪声污染。此外,由于受试者在采集过程中生理活动产生的内部伪迹(如:眨眼、心脏跳动等)^[15],也会对数据产生干扰,影响分类结果。因此,针对内部伪迹,本文首先对采集到的脑电信号进行通道筛选,剔除受眼部运动干扰严重的2个电极FT1和FT2;同时,由于左侧耳电磁极易受到心电伪迹的干扰,因此也剔除了靠近耳部的2个电极FT9和FT10。所以,在通道筛选阶段,共选择了脑电图中20个通道信号。

滤波是一种常见的去除脑电信号外部伪迹的方法,本文采用1~48 Hz的带通巴特沃斯滤波器进行滤波,抑制其他频率范围的信号^[16]。

根据数据集中标注的癫痫发作开始和结束时间,为了保证波形的完整性,设置重叠率为50%的滑动窗口,将脑电信号分割成4 s的数据片段,最后对所有片段进行归一化处理。由于通道筛选和滤波后的脑电信号幅值的浮动一般是在可接受范围内,最大最小标准化能够较大幅度地还原真实EEG信号波形。因此,本文采用最大最小标准化对原始EEG信号进行归一化操作,推得的公式为:

$$X_{\min-\max} = \frac{X - \bar{X}}{\max(X) - \min(X)} \quad (1)$$

1.2 特征提取

原始脑电信号数据量庞大,且不具有代表性,而特征提取方法可以提炼出能够表征癫痫发作特征的数据,用于模型的建立。因此,本文主要使用传统时域、频域特征和基于平滑非线性能量算子的尖峰相关性特征,对脑电数据进行特征提取。

1.2.1 传统特征提取

研究主要从时域和频域两个角度对脑电信号进行传统特征提取。本文主要提取时域上每个通道的最大值、最小值、平均值、峰度(Kurtosis)、偏斜度

(Skewness) 和线长 (Line Length); 频域上主要提取每个信号频域分量的振幅。其中, 峰度、偏斜度和线长的数学定义分别见式(2) ~ (4):

$$Kurtosis = \frac{E[(x - \text{mean}(x))^4]}{\{E[(x - \text{mean}(x))^2]\}^2} \quad (2)$$

$$Skewness = E\left[\left(\frac{x - \text{mean}(x)}{\text{std}(x)}\right)^3\right] \quad (3)$$

$$Line\ Length = \frac{1}{n} \sum_{i=1}^n \text{abs}(x_{i+1} - x_i) \quad (4)$$

其中, x 表示脑电信号片段; E 表示对括号中数值求期望; x_i 表示采样点 i 的值; n 表示片段 x 中采样点数。

1.2.2 尖峰相关特征提取

根据癫痫发作时脑电信号异常波的振幅和频率发生改变的特异性现象, 本文提出将尖峰相关特征作为表征癫痫发作时异常特征的补充。非线性能量算子 (NLEO) 是一种对信号进行能量度量的方法^[17], 能够跟踪信号的瞬时能量。对于离散信号 x_n , 其非线性能量算子表达如式(5)所示:

$$\varphi[x(n)] = x(n-l)x(n-p) - x(n-q)x(n-s) \quad (5)$$

通常, 当癫痫脑电信号中出现异常放电时, 脑电波的振幅和频率会有所提高, 可以更好地突出异常波在平稳状态下的放电波形, 但非线性能量对脑电信号中可能存在的噪音信号也具有很高的敏感度。为了进一步提高 NLEO 对非平稳信号的表征能力和抗干扰能力, 文献[18]提出了一种 NLEO 的改进方法, 即平滑非线性能量算子 (SNLEO), 将计算所得的能量与一个窗函数进行卷积运算, 在一定程度上减小低波幅噪音信号对输出结果的影响。SNLEO 计算见式(6):

$$\varphi[x(n)] = w(n) * \varphi[x(n)] \quad (6)$$

其中, w 是一个矩形的窗函数, “*” 表示卷积操作。在非线性能量算子的计算中, 本文使用的参数值为 $l=1$, $p=2$, $q=0$ 和 $s=3$, 并采用 7 个点的窗函数进行卷积计算。

获得 SNLEO 后, 需要设定一个合适的阈值, 尽可能多地筛选出可能是尖峰的样本, 同时最小化漏检率。本文使用自适应阈值, 对 SNLEO 进行尖峰筛选识别。本文采用影响检测尖峰数量没有大范围变化的阈值作为最优阈值。最优阈值的搜索范围为 SNLEO 的 10%~90%^[19], 相邻 2 个峰值的中间被确定为一个尖峰的起始点或结束点。由于数据在划分过程中导致波形的不连续问题, 本文将检测到的第

一个和最后一个尖峰丢弃, 以确保每个片段具有完整的尖峰形态。如果检测出尖峰, 则将每个划分好的尖峰与后续 5 个尖峰片段相关联。本文使用尖峰相关性 (Spike Correlation, SC) 来定义该矩阵, 并将 SC 的平均值和标准差作为癫痫发作检测的特征。SC 计算见式(7~8):

$$SC_{i,j} = \max_m R_{x_i x_j}(m) \quad (7)$$

$$R_{x_i x_j}(m; i, j) = \frac{E[x_i(n)x_j(n+m)]}{\sigma_{x_i}\sigma_{x_j}} \quad (8)$$

其中, x_i, x_j 是脑电 EEG 信号的片段, 这里 $i = [2, \dots, S-6]$, $j = [i+1, \dots, i+5]$; S 表示在一个片段中检测到的峰值数; σ 表示脑电图片段的标准差。

估计 SC 特征的处理过程如图 2 所示。将一个片段的第一个和最后一个丢弃, 而后根据得到的尖峰计算其与后面 5 个尖峰的相关性。根据图 2(a) 中样例计算出的尖峰相关矩阵如图 3 所示。

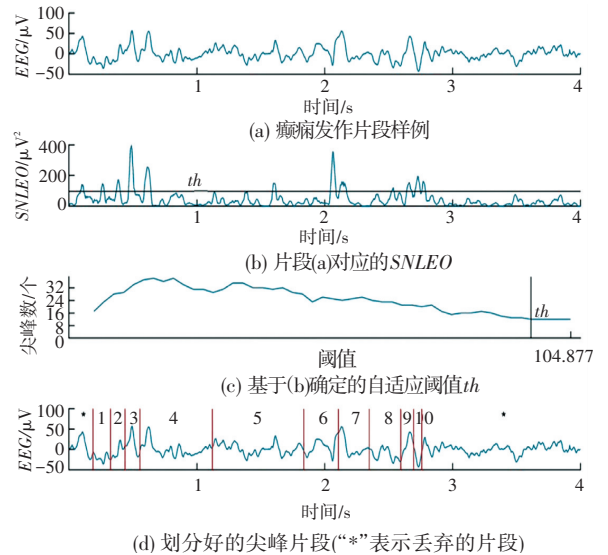


图 2 使用自适应阈值的 SNLEO 计算尖峰相关性示意图

Fig. 2 Schematic diagram of SNLEO calculation of spike correlations using adaptive threshold

	2	3	4	5	6	7	8	9	10
1	0.96	0.95	0.79	0.96	0.98	0.76	0.91	0.98	0.96
2		0.90	0.83	0.94	0.94	0.95	0.86	0.90	0.98
3			0.95	0.94	0.95	0.94	0.98	0.93	0.98
4				0.44	0.51	0.87	0.76	0.95	0.89
5					0.68	0.65	0.72	0.90	0.99
6						0.89	0.60	0.98	0.97
7							0.94	0.92	0.99
8								0.99	0.96
9									1.00

图 3 尖峰片段得到的最大相关矩阵

Fig. 3 Maximum correlation matrix obtained from spike fragments

此外,计算了 *SNLEO* 的平均值、标准差和平均最大 *SNLEO* 值 *spikiness*。其中, *spikiness* 被定义为 *SNLEO* 中峰值的最大值除以 *SNLEO* 的平均值^[20], 以及检测到的峰值数量 (*snum*)、平均持续时间 (*swidth*) 和平均峰值间间隔 (*sgap*)。基于 *SNLEO* 划分的尖峰相关特征的具体描述见表 1。

表 1 尖峰相关特征的描述

Table 1 Description of spike-related characteristics

特征	描述
<i>mean(SC)</i>	尖峰相关性矩阵的平均值
<i>std(SC)</i>	尖峰相关性矩阵的标准差
<i>mean(SNLEO)</i>	<i>SNLEO</i> 的平均值
<i>std(SNLEO)</i>	<i>SNLEO</i> 的标准差
<i>spikiness</i>	平均最大 <i>SNLEO</i> 值
<i>snum</i>	峰值数量
<i>swidth</i>	平均持续时间
<i>sgap</i>	平均峰值间间隔

1.3 分类模型

使用传统机器学习分类器 RF 和 SVM 来评估本文提出的方法,这些分类器经常被用于癫痫发作的自动检测。

2 实验

2.1 数据集

本研究采用公开的头皮脑电数据集 CHB-MIT。该数据集共记录了美国波士顿儿童医院的 23 名癫痫患者的头皮脑电数据,每个患者的数据都是由多个.edf文件组成,采样频率 256 Hz,共含有 157 次癫痫发作。大多数文件包含有 23 个 EEG 通道信号,并采用国际标准 10-20 系统使用的 EEG 电极位置命名这些通道记录。

由于癫痫发作时间远小于发作间期的时间,为了保证数据集正负样本的均衡性,本文采用欠采样的方式在发作间期随机采样和癫痫发作样本数量相当的负样本。

2.2 评价指标

为了验证本文方法的有效性,采用准确率 (*Acc*)、敏感性 (*Sen*)、特异性 (*Spe*)、*F1* 值和 *AUC* 等指标进行实验评估。计算方法见式(9)~式(11):

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \quad (9)$$

$$Sen = \frac{TP}{TP + FN} \quad (10)$$

$$Spe = \frac{TN}{TN + FP} \quad (11)$$

其中, *TP*、*FP*、*FN* 和 *TN* 分别为真阳性、假阳性、假阴性和真阴性。

本文产生的所有实验结果都是在配置为 Intel (R) Core (TM) i7-9700 CPU @ 3.00 GHz, 16 GB RAM 的计算机上实现的。实验模型使用 Python 3.7 和 Scikit-learn 构建。

2.3 结果分析

本文先对提取的传统时域、频域特征分别使用 RF 和 SVM 分类模型进行测试,所得实验结果见表 2。由表 2 可知,SVM 分类模型表现最佳。

表 2 基于传统特征的实验结果

Table 2 Experimental results based on traditional characteristics

特征	分类器	<i>Acc</i>	<i>Sen</i>	<i>Spe</i>
传统特征	RF	0.862 1	0.753 3	0.933 9
	SVM	0.959 0	0.938 5	0.973 6

在确定分类模型 SVM 的基础上,将传统特征和尖峰相关特征结合,探讨尖峰相关特征对癫痫脑电信号的表征能力。添加前后对比结果见表 3。

表 3 尖峰相关特征对比的分类结果

Table 3 Classification results of spike-related feature comparison

分类器	特征	<i>Acc</i>	<i>Sen</i>	<i>Spe</i>
SVM	传统特征	0.959 0	0.938 5	0.973 6
	传统特征+尖峰相关特征	0.965 2	0.956 5	0.970 9

由表 3 可知,尖峰相关特征能够对癫痫脑电信号信息进行表征。加入尖峰相关特征后,检测结果在 *Acc* 上提升了 0.62%,在 *Sen* 上提升了 1.8%,在 *Spe* 上有所降低。在实际的临床应用中,正确识别发作样本比正确识别非发作样本更重要,因此 *Sen* 指标更能准确衡量方法的优劣。本文提出的方法虽然在 *Spe* 上略有降低,但 *Sen* 指标上有一定程度的提升。

3 结束语

本文提出了一种基于平滑非线性能量算子划分的尖峰相关特征的癫痫发作自动检测算法。该算法使用传统的时域、频域特征,结合尖峰相关性特征对脑电信号进行刻画,使用 RF 和 SVM 分类器来测试癫痫发作自动检测的有效性和可靠性。将所提方法在开源数据集 CHB-MIT 上进行了评估,SVM 分类器获得了更好的结果,其准确率、敏感性和特异性分别为 96.52%、95.65% 和 97.09%。此外,研究开展的特征消融实验结果表明,提出的基于平滑非线性能

量算子划分的尖峰相关特征,能够作为癫痫脑电信息的补充,进一步提高癫痫发作检测的性能。

参考文献

- [1] PATEL D C, TEWARI B P, CHAUNSAI L, et al. Neuron - glia interactions in the pathophysiology of epilepsy [J]. *Nature Reviews Neuroscience*, 2019, 20(5): 282-297.
- [2] SPECCHIO N, WIRRELL E C, SCHEFFER I E, et al. International League Against Epilepsy classification and definition of epilepsy syndromes with onset in childhood; Position paper by the ILAE Task Force on Nosology and Definitions[J]. *Epilepsia*, 2022, 63(6): 1398-1442.
- [3] SCHAD A, SCHINDLER K, SCHELTER B, et al. Application of a multivariate seizure detection and prediction method to non-invasive and intracranial long-term EEG recordings[J]. *Clinical Neurophysiology*, 2008, 119(1): 197-211.
- [4] BENBADIS S R, BENICZYK S, BERTRAM E, et al. The role of EEG in patients with suspected epilepsy[J]. *Epileptic Disorders*, 2020, 22(2): 143-155.
- [5] 王学峰. 癫痫的脑电图:传统观点、新认识和新领域[J]. *中华神经科杂志*, 2004, 37(3): 7-9.
- [6] 刘晓燕, 黄珍妮, 秦炯. 不同类型小儿癫痫持续状态的临床及脑电图分析[J]. *中华神经科杂志*, 2000, 33(2): 73-73.
- [7] MATURANA M I, MEISEL C, DELL K, et al. Critical slowing down as a biomarker for seizure susceptibility [J]. *Nature Communications*, 2020, 11(1): 2172.
- [8] 彭睿旻, 江军, 匡光涛, 等. 基于 EEG 的癫痫自动检测:综述与展望[J]. *自动化学报*, 2022, 48(2): 335-350.
- [9] HOSSEINI M P, HOSSEINI A, AHI K. A review on machine learning for EEG signal processing in bioengineering [J]. *IEEE Reviews in Biomedical Engineering*, 2020, 14: 204-218.
- [10] ACHARYA U R, HAGIWARA Y, DESHPANDE S N, et al. Characterization of focal EEG signals: A review [J]. *Future Generation Computer Systems*, 2019, 91: 290-299.
- [11] MURSALIN M, ZHANG Y, CHEN Y, et al. Automated epileptic seizure detection using improved correlation-based feature selection with random forest classifier [J]. *Neurocomputing*, 2017, 241: 204-214.
- [12] 杨舒涵, 李博, 周丰丰. 基于机器学习的跨患者癫痫自动检测算法[J]. *吉林大学学报(理学版)*, 2021, 59(1): 101-106.
- [13] ZAREI A, ASL B M. Automatic seizure detection using orthogonal matching pursuit, discrete wavelet transform, and entropy based features of EEG signals[J]. *Computers in Biology and Medicine*, 2021, 131: 104250.
- [14] 吴端坡, 王紫萌, 董芳, 等. 基于 aEEG 尖峰和 cEEG 棘波提取的癫痫发作检测算法[J]. *实验技术与管理*, 2020, 37(12): 57-62.
- [15] 骆睿鹏, 冯铭科, 黄鑫, 等. 脑电信号预处理方法研究综述[J]. *电子科技*, 2023, 36(4): 36-43.
- [16] OCBAGABIR H T, ABOALAYON K A I, FAEZIPOUR M. Efficient EEG analysis for seizure monitoring in epileptic patients [C]//2013 IEEE Long Island Systems, Applications and Technology Conference (LISAT). Farmingdale, USA: IEEE, 2013: 1-6.
- [17] BOONYAKITANONT P, LEK-UTHAI A, CHOMTHO K, et al. A review of feature extraction and performance evaluation in epileptic seizure detection using EEG [J]. *Biomedical Signal Processing and Control*, 2020, 57: 101702.
- [18] MUKHOPADHYAY S, RAY G C. A new interpretation of nonlinear energy operator and its efficacy in spike detection [J]. *IEEE Transactions on Biomedical Engineering*, 1998, 45(2): 180-187.
- [19] TAPANI K T, VANHATALO S, STEVENSON N J. Incorporating spike correlations into an SVM-based neonatal seizure detector [C]//EMBECE & NBC 2017: Joint Conference of the European Medical and Biological Engineering Conference (EMBECE) and the Nordic-Baltic Conference on Biomedical Engineering and Medical Physics (NBC). Singapore: Springer, 2018: 322-325.
- [20] TAPANI K T, VANHATALO S, STEVENSON N J. Time-varying EEG correlations improve automated neonatal seizure detection [J]. *International Journal of Neural Systems*, 2019, 29(4): 1850030.