

梁袁泽, 张学军. 基于经验模态分解的精细复合多尺度排列熵癫痫脑电信号分类方法[J]. 智能计算机与应用, 2024, 14(5): 44-51. DOI:10.20169/j.issn.2095-2163.240506

基于经验模态分解的精细复合多尺度排列熵癫痫脑电信号分类方法

梁袁泽¹, 张学军^{1,2}

(1 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, 南京 210023;

2 南京邮电大学 射频集成与微组装技术国家地方联合工程实验室, 南京 210023)

摘要: 癫痫是一种常见的脑部疾病,通过脑电图能准确地定位人脑中的致病区域。文章提出一种基于经验模态分解的精细复合多尺度排列熵的癫痫脑电信号自动检测方法,用于解决区分致病区和非致病区的癫痫脑电信号难的问题。首先将原信号分割成多个子信号,并对各子信号进行经验模态分解,然后从分解后的不同经验模态函数中提取精细复合多尺度排列熵特征并利用支持向量机进行分类。通过对癫痫脑电的公共数据集测试,实验结果表明准确率、灵敏度和特异度三个性能指标分别达到 90.3%, 85.0% 和 96.0%, ROC 曲线下面积达 0.98。

关键词: 癫痫; 经验模态分解; 精细复合多尺度排列熵; 支持向量机

中图分类号: TN911.7

文献标志码: A

文章编号: 2095-2163(2024)05-0044-08

Classification of epileptic EEG signals based on

Empirical Mode Decomposition and refined composite multiscale permutation entropy

LIANG Yuanze¹, ZHANG Xuejun^{1,2}

(1 School of Electronic and Optical Engineering, School of Flexible Electronics (Future Technologies), Nanjing University of Posts and Telecommunications, Nanjing 210023, China; 2 National Joint Engineering Laboratory of RF Integration and Microassembly Technology, Nanjing University of Posts and Telecommunications, Nanjing 210023, China)

Abstract: Epilepsy is a common brain disease, which can accurately locate the epileptic regions in the human brain through EEG. In this paper, an automatic detection method of epileptic EEG signals based on Empirical Mode Decomposition and refined composite multiscale permutation entropy is proposed to solve the problem of distinguishing epileptic EEG signals between epileptic and non-epileptic areas. Firstly, the original signal is divided into multiple sub signals, and each sub signal is subject to empirical mode decomposition. Then, refined composite multiscale permutation entropy is extracted from different decomposed empirical mode functions and classified by Support Vector Machine. For the public data set of epileptic EEG, the final experimental results show that the accuracy, sensitivity and specificity of the three performance indicators reach 90.3%, 85.0% and 96.0% respectively, and the product under the ROC curve reaches 0.98.

Key words: epileptic; Empirical Mode Decomposition; refined composite multiscale permutation entropy; Support Vector Machine

0 引言

癫痫(Epilepsy)是一种大脑神经元突发性异常放电,会导致短暂大脑功能障碍的慢性神经系统疾病,其特点是短暂、突然的脑电紊乱^[1]。脑电图(Electroencephalogram, EEG)是一种通过记录生物电信号来描述大脑活动的方法,包含多种大脑生理状态的脑电信息。常用的视觉观察和临床试验等传统脑电图检测方法繁琐、耗时,诸多临床试验表明局

灶性癫痫患者大多产生耐药性,需要手术治疗^[2-3]。其中,局灶性(致病区)癫痫发作时,电信号表现在大脑有限的区域;而非局灶性(非致病区)癫痫的特征是影响整个大脑的全身性癫痫发作,脑电研究者们致力于患者在早期检测出癫痫并进行临床治疗。为尽早对患有耐药癫痫患者进行手术,需要快速有效地判断癫痫类型,因此获取检测方法则尤显重要。

Huang 等学者提出的经验模态分解(Empirical Mode Decomposition, EMD)等方法适用于处理非线性

基金项目: 国家自然科学基金(61977039); 江苏省研究生科研与实践创新计划项目(KYCX21_0712)。

作者简介: 梁袁泽(1996-),男,硕士研究生,主要研究方向:脑电信号处理,脑机接口技术。

通讯作者: 张学军(1969-),男,博士,副教授,主要研究方向:脑电信号处理,脑机接口技术,无线射频识别。Email: xjzhang@njupt.edu.cn

收稿日期: 2023-04-06

性和非平稳信号,这种时域分解法具有自适应和高效的特点,且 EMD 是基于数据的局部特征时间尺度,在非线性和非平稳系统中表现优秀^[4-6]。

非线性分析法主要有高阶矩^[7]、相关维度^[8]、概率密度相关函数^[9]、各类熵^[8, 10]等。近年来, Azami 等学者^[11-12]提出 2 种新的复杂性度量特征,即:精细复合多尺度散布熵 (Refined Composite Multiscale Dispersion Entropy, RCMDE) 和精细复合多尺度模糊熵 (Refined Composite Multiscale Fuzzy Entropy, RCMFE),这 2 种特征作为非线性分析法在处理非线性不平稳生物信号时效果较好,因此在生理医学信号分析中的应用前景十分广阔^[13-14]。

本文结合排列熵,提出一种基于经验模态分解的精细复合多尺度排列熵来区分致痫区和非致痫区癫痫患者的术前诊断算法,这种区域的定位有助于癫痫的术前分析。本文第 1 节介绍改进 EMD 的实现过程,第 2 节阐述了基于排列熵 (PE) 改进的 RCMPE 算法模型的理论推导过程,以及后续实验流程分析。第 3 节通过德国波恩大学癫痫公共数据集设计分类实验。第 4 节分析结果得到 RCMPE 的分类效果达到 88.3%,在此基础上进行主成分分析后的分类准确率达到 90.3%。

1 方法

1.1 经验模态分解

经验模态分解 (Empirical Mode Decomposition, EMD) 是 Huang 等学者^[5]于 1998 年针对非线性非平稳系统提出的一种自适应时频处理方法。其结果是对信号进行多次迭代,获得一系列具有不同特征尺度的时间序列,与傅里叶变换、小波变换的核心思路是一致的。相较于傅里叶变换和小波变换,EMD 不依赖基函数,仅依据数据自身的时间尺度特征来进行信号分解,具备自适应性。EMD 中的假设主要是基于下述条件:

- (1) 信号至少同时存在一个极大值和极小值;
- (2) 通过 2 个极值点之间的序列差定义尺度因子;
- (3) 极值点可由数据微分得到。

因此,脑电信号的 EMD 分解实现步骤如下:

(1) 统计整段信号 $x(t)$ 的极大值和极小值,并拟合上包络线 $e_{up}(t)$ 和下包络线 $e_{low}(t)$;

(2) 计算上下包络线的平均值 $m_1(t)$:

$$m_1(t) = \frac{e_{up}(t) + e_{low}(t)}{2} \quad (1)$$

(3) 用 $x(t)$ 减 $m_1(t)$ 得到第一个 IMF 分类

$h_1(t)$, 然后把 $h_1(t)$ 代入步骤 (1) 的 $x(t)$ 中重复 (1)、(2) 操作,得到的 IMF 分量 $h_c(t)$ 包含迭代前时间序列的最短周期分量。

(4) 原始脑电时间序列 $x(t)$ 可被分解为:

$$x(t) = h_c(t) + \sum_{i=1}^c m_i(t) \quad (2)$$

1.2 精细复合多尺度排列熵

精细复合多尺度排列熵算法 (Refined Composite Multiscale Permutation Entropy, RCMPE) 基于排列熵算法并在原算法基础上加以改进,核心操作是:

(1) 采用 Costa 提出的多尺度粗粒化方法对时间序列进行粗粒化处理,生成连续无重叠的粗粒化序列^[15];

(2) 对粗粒化序列,调整尺度因子 τ 创建对应粗粒化过程不同起始点的不同序列,并用 RCMPE 特征定义这些移位序列的色散模式的平均值的排列熵值。

1.2.1 排列熵

排列熵 (Permutation Entropy, PE) 基于信号值之间的顺序关系、即排列模式。作为一种分析系统非线性的新方法,主要衡量一维时间序列复杂度,反映一维时间序列复杂度的性能方面和 LyaPunov 指数相似^[16]。与样本熵 (Sample Entropy, SampEn) 和近似熵 (Approximate Entropy, ApEn) 等相比,PE 理论简单,参数少,对伪迹和噪声具有较强的鲁棒性,计算速度快。此外,当信号长度大于 2 000 时,PE、SampEn 和 ApEn 的熵值趋于稳定,信号的幅值和相位对三者都没有影响,其中 PE 能较好地描述频率特性的变化。因此,PE 被广泛应用于生理信号和图像处理中。Li 等学者^[17]利用脑电图信号预测大鼠的失神发作,结果表明,PE 可以跟踪脑电图记录的动态变化,PE 比样本熵更能预测脑电失神发作信号。Wu 等学者利用 PE 揭示典型缺失患者的大脑活动异常,可以检测出传统方法所不能显示的脑电活动异常的结论^[18-19]。排列熵的原理在于不考虑数据具体值,设长度为 N 的时间序列 $\{x(i), i = 1, 2, \dots, N\}$, 对其进行空间重构,得到如下的时间序列:

$$Y = \begin{bmatrix} x(1) & x(1+\lambda) & \cdots & x[1+(m-1)\lambda] \\ x(2) & x(2+\lambda) & \cdots & x[2+(m-1)\lambda] \\ \vdots & \vdots & \cdots & \vdots \\ x(j) & x(j+\lambda) & \cdots & x[j+(m-1)\lambda] \\ \vdots & \vdots & \cdots & \vdots \\ x(K) & x(K+\lambda) & \cdots & x[K+(m-1)\lambda] \end{bmatrix} \begin{matrix} \ddot{y} \\ \ddot{y} \\ \ddot{y} \\ \ddot{y} \\ \ddot{y} \\ \ddot{y} \end{matrix} \quad (3)$$

$j = 1, 2, \dots, K, K = N - (m-1)\lambda$

其中, K 表示重构分量的个数; m 表示嵌入维数; λ 表示时延; $x(j)$ 表示重构矩阵的第 j 行分量。

矩阵 Y 中的行数作为重构向量的计数, 共有 $K = N - (m - 1)\lambda$ 个重构分量。然后将重构矩阵 Y 中的 $\{x(j), x(j + \lambda), \dots, x(j + (m - 1)\lambda)\}$ 升序排列, j_1, j_2, \dots, j_d 作为重构矩阵的映射, 即:

$$x(i + (j_1 - 1)\lambda) \leq x(i + (j_2 - 1)\lambda) \leq \dots \leq x(i + (j_m - 1)\lambda) \quad (4)$$

当重构分量中存在相等的值, 即 $x(i + (j_p - 1)\lambda) = x(i + (j_q - 1)\lambda)$, $p \neq q$, 此时就按照 j_p 和 j_q 值的大小来排序, 也就是当 $j_p < j_q$, 有:

$$x(i + (j_p - 1)\lambda) < x(i + (j_q - 1)\lambda) \quad (5)$$

因此, 对于任意一个时间序列重构所得的矩阵 Y 中每一行都可以得到一组符号序列:

$$S(l) = \{j_1, j_2, \dots, j_d\} \quad (6)$$

其中, $l = 1, 2, \dots, k$ 且 $k \leq m!$, m 维相空间映射不同的符号序列 $\{j_1, j_2, \dots, j_d\}$ 总共有 $m!$ 种, 符号序列 $S(l)$ 是其中的一种排列。计算每一种符号序列出现的概率 P_1, P_2, \dots, P_k , 此时, 时间序列 $X(i)$ 的 k 种不同符号序列的排列熵 H_p 可以按照信息熵的形式定义为:

$$H_p(m) = - \sum_{j=1}^k p_j \ln(p_j) \quad (7)$$

当 $P_j = \frac{1}{d!}$ 时, $H_p(m)$ 达到最大值 $\ln(m!)$ 。简化计算, 通常用 $\ln(m!)$ 将 $H_p(m)$ 进行归一化处理, 即:

$$0 \leq H_p = \frac{H_p}{\ln(m!)} \leq 1 \quad (8)$$

H_p 值的大小表示时间序列 $\{x(i), i = 1, 2, \dots, N\}$ 的随机程度: H_p 的值越小, 说明时间序列越趋近, 反之, 则时间序列越混乱。

1.2.2 精细复合多尺度排列熵

本文提出精细复合多尺度排列熵 (Refined Composite Multiscale Permutation Entropy, RCMPE) 旨在解决 MPE 在处理较长时间序列特征提取缺少差异性的问题, 本节将详细说明 RCMPE 算法。

首先, 多尺度粗粒化方法对时间序列 $\{x(i), i = 1, 2, \dots, N\}$ 进行粗粒化处理, 生成多个随机的粗粒化时间序列, 计算每个粗粒度时间序列的排列熵作为比例因子的函数, 然后根据香农熵得到时间序列 $\{x(i), i = 1, 2, \dots, N\}$ 的精细复合多尺度排列熵^[20]。后续步骤分述如下。

(1) 对给定的离散时间序列 $\{x(i), i = 1, 2, \dots, N\}$, 通过对长度递增的 t 非重叠窗口内的数据点进

行平均, 构建多个粗粒度时间序列。粗粒度时间序列由下列方程构造^[18, 21]:

$$y_j^{(\tau)} = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} x_i, \quad 1 \leq j \leq \frac{N}{\tau} \quad (9)$$

每个粗粒度时间序列的长度等于原始时间序列的长度除以尺度因子 τ 。

(2) 研究对 $y_j^{(\tau)}$ 相空间重构 $Y_j^{(\tau)} = \{y_{(j)}^{(\tau)}, y_{(j+t)}^{(\tau)}, y_{(j+2t)}^{(\tau)}, \dots, y_{(j+(m-1)t)}^{(\tau)}\}$, 这里 t 是延迟因子, m 是嵌入维数。

(3) 可以按递增顺序排列:

$$[y(j + (j_1 - 1)t) \leq y(j + (j_2 - 1)t) \leq \dots \leq y(j + (j_m - 1)t)], \quad 1 \leq j_m \leq \frac{N - (m - 1)\tau t}{\tau} \quad (10)$$

(4) 位置序号向量 $\pi_k = \{j_1, j_2, \dots, j_m\}$, $m! \leq N$, $f(\pi)$ 是每种模式出现的次数, 共有 $m!$ 种模式, 记录第 k 个初始化起始点下某种模式出现的概率 $p_k^{(\tau)}$:

$$p_k^{(\tau)} = \frac{f(\pi_k)}{\frac{N}{\tau} - m + 1}, \quad 1 \leq k \leq m! \quad (11)$$

$$RCMPE = - \sum_{i=1}^{m!} \overline{\varphi(\pi_k)} \ln \overline{\varphi(\pi_k)} \quad (12)$$

$$\varphi(\pi_k) = \frac{1}{\tau} \sum_{i=1}^{\tau} p_k^{(i)} \quad (13)$$

其中, $\overline{\varphi(\pi_k)}$ 是时间序列 $Y_j^{(\tau)}$ 中 j_m 出现的平均概率。

2 基于 EMD 的精细复合多尺度排列熵脑电识别方法

基于 EMD 的精细复合多尺度排列熵脑电识别方法整体流程如图 1 所示。

假设输入信号为 $X = \{x(i), i = 1, 2, \dots, N\}$, 其中 N 是信号的长度, 基于 EMD 的精细复合多尺度排列熵特征提取方法的具体实现步骤为:

(1) 将原始脑电信号经过 0.53~40 Hz 带通滤波器去除伪迹和噪声, 经 12 位模数转换后, 以 173.61 Hz 的采样率将数据连续输出到数据采集磁盘上。由于离散时间序列结束和开始之间的不连续会产生寄生频率分量, 根据平稳性标准, 将每个预处理后的时间序列分割出 q 段不同的子数据集, 达到最后一个和第一个数据点的幅度差在连续数据点幅度差的范围内且时间序列结束和开始处的斜率积为正的效果;

(2) 对子数据集进行 EMD 分解, 针对癫痫脑电信号长度较长的特点, 利用 EMD 自适应特性获取极大值极小值数量差在 5~10 之间的迭代函数;

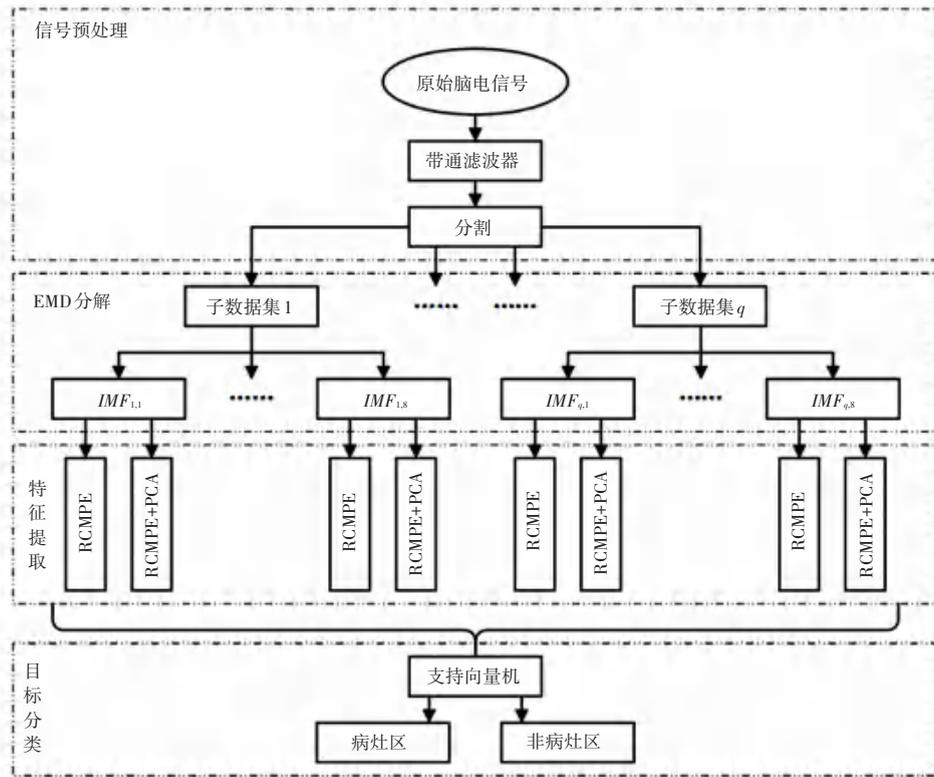


图 1 本文研究方法流程图

Fig. 1 Flow chart of research method in this paper

(3) 计算由 EMD 分解得到的迭代函数 $h_c(t)$ 的精细复合多尺度排列熵 $RCMPE_q$, 将每个子数据集分解得到的迭代函数的熵值特征进行组合, 构成该输入信号在位置序号为 $\pi_k = \{j_1, j_2, \dots, j_m\}$ 下的特征融合向量 $s = \{RCMPE_1, RCMPE_2, \dots, RCMPE_{\pi_k}\}$;

(4) 经 PCA 降维后将特征融合向量 s 输入 SVM 算法中进行分类, 从而完成对原始脑电信号致痫区和非致痫区的区分。

3 实验

3.1 数据集

本文使用的癫痫脑电信号数据是由波恩大学癫痫学系提供的公共数据集^[22], 该数据集共有五组集

合, 采样频率为 173.61 Hz, 每个子集包含 100 个持续 23.6 s 的单通道 EEG 片段, 有 4 097 个数据点, 每个片段从连续的多通道 EEG 中经过 0.53~40 Hz 滤波和视觉去除伪影处理后切割出来。其中, 集合 A 和 B 表示 5 名健康志愿者处在清醒状态下睁眼和闭眼由标准电极放置方案采集到的表面 EEG。集合 C、D 和 E 来自术前诊断后的 EEG, 集合 C 采集于海马区, 集合 D 和集合 E 采集于大脑致痫区, 集合 C 和 D 采集的是处在非癫痫发作期间的脑电片段, 集合 E 采集的是处在癫痫发作期间的脑电片段, 见表 1。本文将 5 组数据全部用于实验数据, 并在后续实验中设计分类案例测试算法可靠性。

表 1 波恩大学 EEG 癫痫数据集说明

Table 1 Description of EEG epilepsy data set of Bonn University

组别	对象	采集状态	记录时间	电极位置	电极类型
A	健康志愿者	睁眼	清醒状态	国际标准 10 - 20 系统	非侵入式头皮表面电极
B	健康志愿者	闭眼	清醒状态	国际标准 10 - 20 系统	非侵入式头皮表面电极
C	癫痫患者	发作间歇期	无癫痫发作的间歇期	海马区	颅内电极
D	癫痫患者	发作间歇期	无癫痫发作的间歇期	致痫区	颅内电极
E	癫痫患者	发作活跃期	癫痫活跃期	致痫区	颅内电极

针对波恩大学癫痫脑电数据集,本文提出 7 种在癫痫诊断中具有重要临床意义和研究价值的实验

分类案例,见表 2。经过预处理后的癫痫脑电波形如图 2 所示。

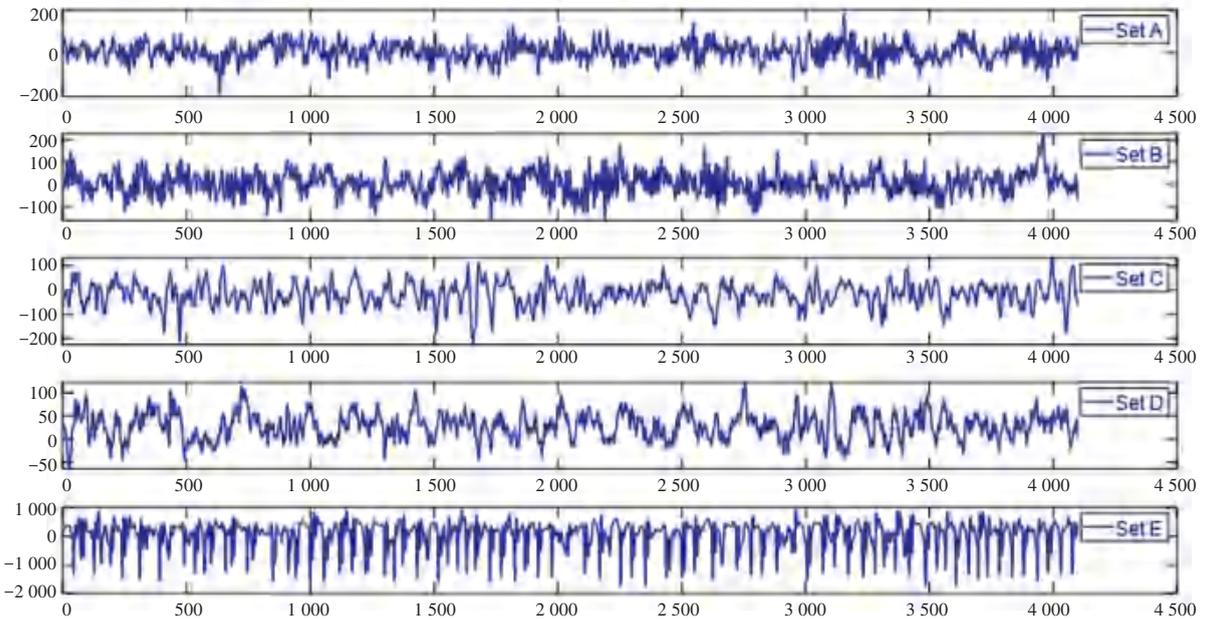


图 2 经过预处理后的癫痫脑电波

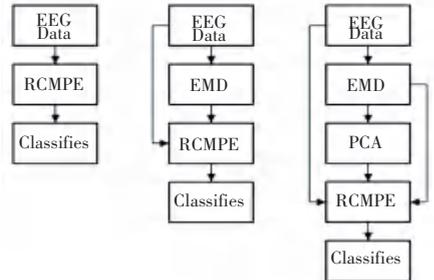
Fig. 2 Epileptic EEG after pretreatment

表 2 实验案例

Table 2 Experimental cases

实验案例	临床说明	集合
1	正常睁眼-癫痫发作期	A - E
2	正常闭眼 - 癫痫发作期	B - E
3	无癫痫发作间歇期(非致病区) - 癫痫发作期	C - E
4	无癫痫发作间歇期(致病区) - 癫痫发作期	D - E
5	正常 - 癫痫患者	AB - CE
6	正常 - 无癫痫发作间歇期	AB - CD
7	正常 - 致病区	AB - DE

本文设计了 3 组算法实验。第 1 组实验如图 3 (a) 所示,使用 RCMPE 算法提取原始脑电信号的特征,并将特征作为支持向量机分类器的输入。第 2 组设计实验如图 3(b) 所示,首先使用 EMD 对脑电信号进行局部分解,然后进行 RCMPE 特征提取,并将其与原始时间序列和迭代后的局部时间序列相结合输入分类器。第 3 组算法设计实验如图 3(c) 所示,在 RCMPE 特征提取后加入 PCA,最后与原始时间序列和迭代后的局部时间序列相结合作为分类器的输入。



(a) 第 1 组实验 (b) 第 2 组实验 (c) 第 3 组实验

图 3 算法设计实验流程图

Fig. 3 Flow chart of algorithm design experiment

3.2 特征提取

3.2.1 EMD 分解

脑电信号各频带段中因为测试时间不同以及被试的不同,其极值点的分布随着神经活动会发生轻微的变化,导致小波变化中需要选取小波基与确定分解尺度的问题难以解决。而 EMD 作为一种自适应的信号分解方法,将癫痫脑电时间序列中的不同尺度或趋势分量逐级分解出来,通过上下包络线的逐步迭代提高预测精度。

主要注意 EMD 算法在对信号进行分解前需要设定分解模态的个数。若个数过大会造成过分解,产生无用分量,并且会增加计算复杂度;若个数过小会造成欠分解,使部分带限信号分解不出来,造成原信号信息的丢失。本文先随机选择一组致病区脑电数据段,并进行 1~8 重 EMD 分解,如图 4 所示。

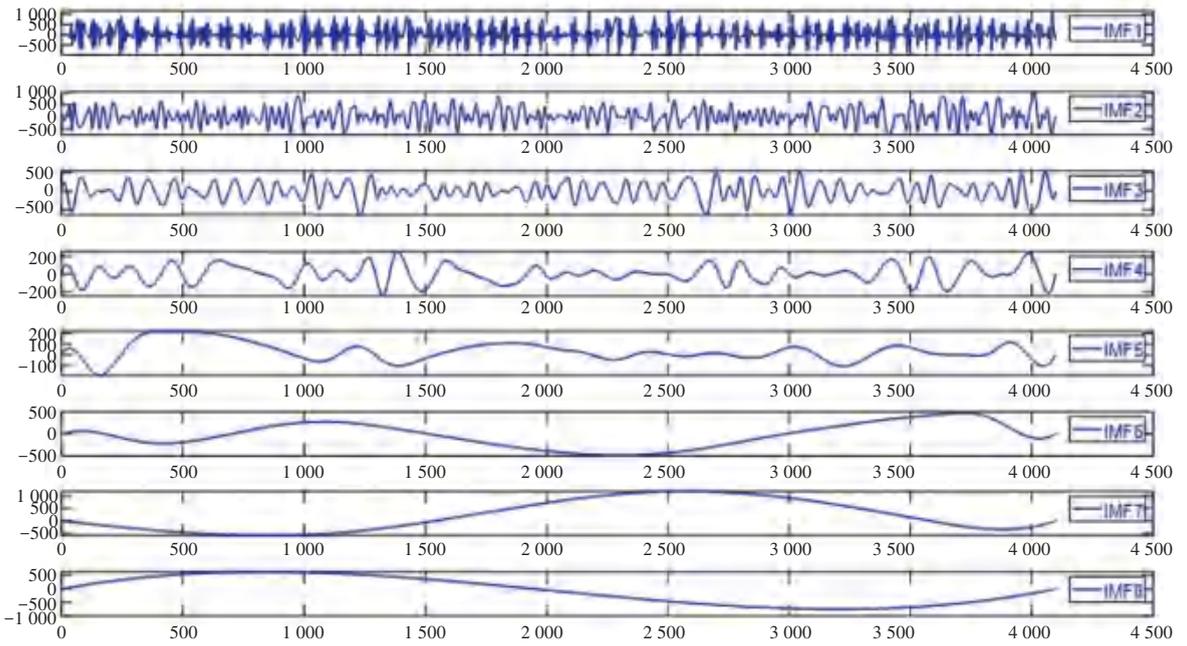


图 4 EMD 迭代分量

Fig. 4 EMD iteration components

3.2.2 RCMPE 特征提取

由式(9)~(12)可以看出,计算 RCMPE 需要分别设置类别个数 k , 嵌入维度 m , 时间延迟 t 以及尺度因子 τ 这四个参数。类别个数 k 过大会导致较小差异的 2 个量被归为同一类^[21]。

相空间重构后,若嵌入维度 m 过小,重构后的时间序列可用的状态缺少差异性,导致算法无意义;若 m 过大则会导致重构后的时间序列均匀过度,徒增计算量,但无法有效地计算时间序列的尖峰变化,因此 m 一般取 2 ~ 6。时间延迟 t 通常取 1。为分析尺度因子 τ 对数据的影响,将相同长度下的癫痫 EEG 信号分别用于 RCMPE 算法分析,并与 RCMFE 算法进行精细对比。RCMPE 的平均复杂度如图 5 所示,RCMFE 的平均复杂度如图 6 所示。由图 5、图 6 可知:

(1) $A - D$ 数据的 RCMPE 变化波动小,且随尺度变化有重叠, E 数据的 RCMPE 平均值明显小于其它 4 种。而 RCMFE 中 A 与 B 重叠, C, D 与 E 重叠,表明 RCMPE 比 RCMFE 具有更好的分离能力,能够有效区分单通道信号的复杂性;

(2) 随着尺度因子的增加, $A - D$ 数据的 RCMPE 平均值缓慢降低, E 数据的 RCMPE 平均值呈对数上升,当处在较大尺度因子时 5 组数据的 RCMPE 平均值开始逐步重叠,表明 τ 过大会导致数据均匀化;

(3) 当 τ 大于 5 时,RCMPE 曲线各个状态的平均值明显小于 RCMFE,表明经过 RCMPE 提取的特

征具有更高的稳定性。故尺度因子 τ 取值 5 ~ 30。

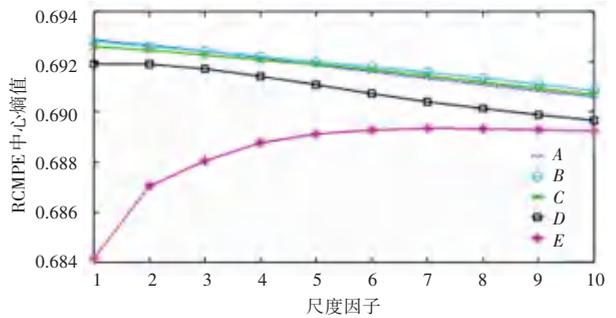


图 5 RCMPE 的平均复杂度

Fig. 5 Average complexity of RCMPE

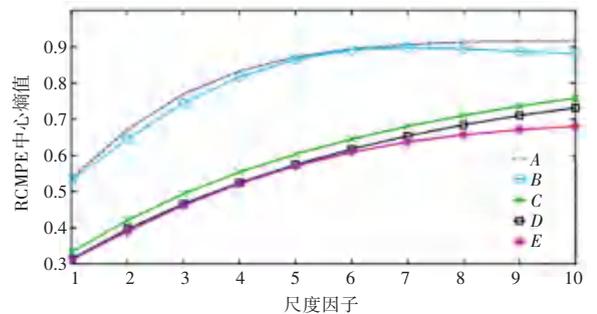


图 6 RCMFE 的平均复杂度

Fig. 6 Average complexity of RCMFE

非致病区和致病区癫痫的迭代函数的 RCMPE 熵值具有差异性,其中 A 表示无致病区脑电信号, E 表示致病区癫痫脑电信号, $D1, D2, D3$ 表示 RCMDE 在尺度因子为 3 的情况下的 3 种特征。RCMPE-AE-IMF1 三重对比结果如图 7 所示。从箱型图中,

可以看出在不同的 IMF 和不同尺度因子下无致痫区和致痫区癫痫脑电信号的特征 RCMPE 的中位数以及分布区间都是不同的。因此,RCMPE 可以作为区分病灶性和非病灶性癫痫脑电数据的分类特征^[23-24]。

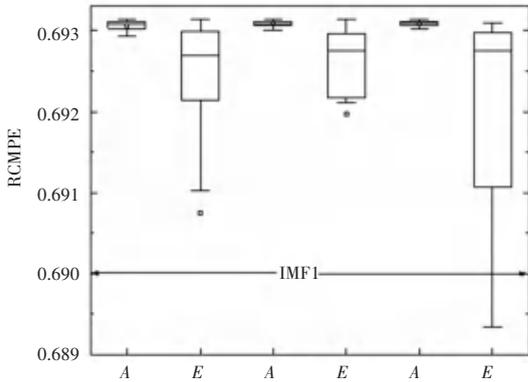


图7 RCMPE-AE-IMF1 三重对比

Fig. 7 Triple comparison of RCMPE-AE-IMF1

4 结果和分析

将 A、B、C、D、E 五组脑电数据段,按上述方法进行处理,然后将特征融合向量作为输入提供给 SVM。SVM 选用线性核函数,重复进行 10 次 5 折交叉验证实验,采用特异度、灵敏度和准确度三个指标对最终分类结果进行度量^[25],结果见表 3。

表 3 不同特征下的分类结果

特征	特异性	灵敏度	准确度
RCMPE	88.0	75.0	80.1
EMD+RCMPE	93.0	83.0	88.3
EMD+RCMPE+PCA	96.0	85.0	90.3

经过 EMD 处理后,RCMPE 的分类效果达到 88.3%。在此基础上再进行主成分分析后的分类准确率达到 90.3%,有着更好的分类效果,详见表 3。ROC 曲线则如图 8 所示。

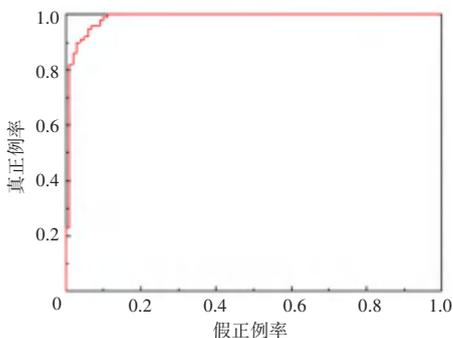


图 8 ROC 曲线

Fig. 8 ROC curve

5 结束语

本文针对癫痫患者的正常区域、致痫区和非致痫区诊断识别问题,提出一种基于 EMD 和精细复合多尺度排列熵融合的特征提取算法,并将特征融合向量输入 SVM 算法中进行分类。应用此方法对波恩大学 EEG 癫痫公共数据集进行处理,得到平均准确率为 90.3%,特异性为 96.0%,灵敏度为 85%,ROC 曲线下面积为 0.98。

本算法先通过改变迭代中心频率减少迭代次数,避免过度分解而导致数据失真,解决 EMD 处理非平稳离散信号的局限性。此外,传统的 PE 算法只考虑单一尺度,在评价复杂生理系统时性能表现有缺失。引入 MPE 后随着粗粒度时间序列长度的减少,而无法满足 N 远大于 $m!$,会导致分类效果差。为解决这 2 个问题,本文提出 EMD 和 RCMPE 特征融合后再进行 SVM 分类,实验结果表明 RCMPE 特征融合分类算法对于癫痫脑电信号的分类效果有显著的提升,为临床医学诊断提供了一种新思路。

参考文献

- [1] WEINSTEIN R L, SCHWARTZ S L, BRAZG R L, et al. Accuracy of the 5-day freestyle navigator continuous glucose monitoring system [J]. *Diabetes Care*, 2007, 30(5): 1125-1130.
- [2] LIN Yuhang, DU Peishan, SUN Hongze, et al. Identifying refractory epilepsy without structural abnormalities by fusing the common spatial patterns of functional and effective EEG networks [J]. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2021, 29(4): 708-717.
- [3] JEONG J W, LEE M H, KURODA N, et al. Multi-scale deep learning of clinically acquired multi-modal MRI improves the localization of seizure onset zone in children with drug-resistant epilepsy [J]. *IEEE Journal of Biomedical and Health Informatics*, 2022, 26(11): 5529-5539.
- [4] RIAZ F, HASSAN A, REHMAN S, et al. EMD-based temporal and spectral features for the classification of EEG signals using supervised learning [J]. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2015, 24(1): 28-35.
- [5] HUANG N E, SHEN Z, LONG S R, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis [J]. *Mathematical, Physical and Engineering Sciences*, 1998, 454(1971): 903-995.
- [6] FARAHMAND S, SOBAYO T, MOGUL D J. Noise-assisted multivariate EMD-based mean-phase coherence analysis to evaluate phase-synchrony dynamics in epilepsy patients [J]. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2018, 26(12): 2270-2279.
- [7] NGO C Q, CHAI R, NGUYEN T V, et al. Electroencephalogram spectral moments for the detection of nocturnal hypoglycemia [J]. *IEEE Journal of Biomedical and Health Informatics*, 2019, 24

- (5); 1237-1245.
- [8] AKAR M, HEKIM M, ORHAN U. Mechanical fault detection in permanent magnet synchronous motors using equal width discretization-based probability distribution and a neural network model [J]. *Turkish Journal of Electrical Engineering and Computer Sciences*, 2015, 23(3): 813-823.
- [9] LU Jun, XIE Kan, MCFARLAND D J. Adaptive spatio-temporal filtering for movement related potentials in EEG-based brain-computer interfaces [J]. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2014, 22(4): 847-857.
- [10] BHATTACHARJEE A, SAHA S, FATTAH S A, et al. Sleep apnea detection based on rician modeling of feature variation in multiband EEG signal [J]. *IEEE Journal of Biomedical and Health Informatics*, 2018, 23(3): 1066-1074.
- [11] AZAMI H, FERNÁNDEZ A, ESCUDERO J. Refined multiscale fuzzy entropy based on standard deviation for biomedical signal analysis [J]. *Medical & Biological Engineering & Computing*, 2017, 55(11): 2037-2052.
- [12] AZAMI H, ROSTAGHI M, ABÁSOLO D, et al. Refined composite multiscale dispersion entropy and its application to biomedical signals [J]. *IEEE Transactions on Biomedical Engineering*, 2017, 64(12): 2872-2879.
- [13] 景鹏, 张学军, 孙知信. 基于弹性变分模态分解的癫痫脑电信号分类方法 [J]. *物理学报*, 2021, 70(1): 369-76.
- [14] 姚文坡, 刘铁兵, 戴加飞, 等. 脑电信号的多尺度排列熵分析 [J]. *物理学报*, 2014, 63(7): 427-433.
- [15] COSTA M, GOLDBERGER A, PENG C-K. Multiscale entropy to distinguish physiologic and synthetic RR time series [C]// *Proceedings of the Computers in Cardiology*. Madrid, Spain: IEEE, 2022, 29: 137-140.
- [16] BHATTACHARYYA A, TRIPATHY R K, GARG L, et al. A novel multivariate-multiscale approach for computing EEG spectral and temporal complexity for human emotion recognition [J]. *IEEE Sensors Journal*, 2020, 21(3): 3579-3591.
- [17] LI Huihui, DU Wenjing, FAN Kai, et al. The effectiveness assessment of massage therapy using entropy-based EEG features among lumbar disc herniation patients comparing with healthy controls [J]. *IEEE Access*, 2020, 8(1): 7758-7775.
- [18] XI Xugang, DING Jinsuo, WANG Junhong, et al. Analysis of functional corticomuscular coupling based on multiscale transfer spectral entropy [J]. *IEEE Journal of Biomedical and Health Informatics*, 2022, 26(10): 5085-5096.
- [19] WU E Q, ZHU Limin, ZHANG Weiming, et al. Novel nonlinear approach for real-time fatigue EEG data: An infinitely warped model of weighted permutation entropy [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2019, 21(6): 2437-2448.
- [20] HILAL M, BERTHIN C, MARTIN L, et al. Bidimensional multiscale fuzzy entropy and its application to pseudoxanthoma elasticum [J]. *IEEE Transactions on Biomedical Engineering*, 2019, 67(7): 2015-2022.
- [21] LABATE D, LA FORESTA F, MORABITO G, et al. Entropic measures of EEG complexity in Alzheimer's disease through a multivariate multiscale approach [J]. *IEEE Sensors Journal*, 2013, 13(9): 3284-3292.
- [22] ANDRZEJAK R G, LEHNERTZ K, MORMANN F, et al. Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state [J]. *Physical Review E*, 2001, 64(6): 061907.
- [23] ROSTAGHI M, KHATIBI M M, ASHORY M R, et al. Fuzzy dispersion entropy: A nonlinear measure for signal analysis [J]. *IEEE Transactions on Fuzzy Systems*, 2021, 30(9): 3785-3796.
- [24] WANG Xianzhi, SI Shubin, LI Yongbo. Variational embedding multiscale diversity entropy for fault diagnosis of large-scale machinery [J]. *IEEE Transactions on Industrial Electronics*, 2021, 69(3): 3109-3119.
- [25] REULECKE S, CHARLESTON-VILLALOBOS S, VOSS A, et al. Temporal analysis of cardiovascular and respiratory complexity by multiscale entropy based on symbolic dynamics [J]. *IEEE Journal of Biomedical and Health Informatics*, 2017, 22(4): 1046-1058.