

• 生物电子学 •

基于小波变换与SVM的ADHD病人分类

谭颖^{1,2}, 张涛², 谭睿³, 沈小涛¹, 校景中¹

(1. 西南民族大学计算机科学与技术学院 成都 610041; 2. 电子科技大学生命科学与技术学院 成都 610054;

3. 西南交通大学生命科学与工程学院 成都 610031)

【摘要】提出基于小波变换的特征提取方法对ADHD病人进行分类研究。采用115名ADHD-200的竞赛静息态功能磁共振数据,首先提取了90个脑区的平均时间序列信号,然后利用小波变换多分辨率分析特性对信号进行3层分解;计算了各个尺度下小波系数的能量值,对能量值进行归一化处理后,将其作为分类特征向量;最后结合SVM分类器采用留一交叉验证法对ADHD病人进行分类。结果表明该方法有助于ADHD病人的分类与诊断。

关键词 注意缺陷与多动; 机器学习; 支持向量机; 小波变换

中图分类号 TP301

文献标志码 A

doi:10.3969/j.issn.1001-0548.2015.05.025

Classification Based Wavelet Translate and SVM in the ADHD

TAN Ying^{1,2}, ZHANG Tao², TAN Rui³, SHEN Xiao-tao¹, and XIAO Jing-zhong¹

(1. School of Computer Science and Technology, Southwest University for Nationalities Chengdu 610041;

2. School of Life Science and Technology, University of Electronic Science and Technology of China Chengdu 610054;

3. School of Life and Science Engineering, Southwest Jiaotong University Chengdu 610031)

Abstract In this study, we propose an approach to extract features based wavelet transform for the ADHD classification. One hundred and fifteen subjects' resting state fMRI data were adopted, which come from ADHD-200 competition. We first extracted the time series of ninety brain areas, and decomposed them into three levels using the wavelet transform for each subject. Secondly, the energy values of any scale were computed and normalized, which construct the classification feature vectors. Finally, we combined the SVM to classification in the ADHD based leave-one-out cross validation. The results demonstrate that the wavelet transform feature extract approach is useful in classification and diagnosis for ADHD.

Key words attention deficit/hyperactivity disorder; machine learning; support vector machine; wavelet-translate

儿童注意缺陷与多动症(attention deficit/hyperactivity disorder, ADHD)是一种常见的神经性发育性障碍,通常表现为注意力不集中,多动或者二者联合的症状^[1]。据统计学龄期的儿童发病率高达5%左右,ADHD的研究已成为世界的一个公共卫生问题^[1-2],引起了人们广泛的关注。目前,对ADHD的诊断和治疗都有较大的难度,缺乏客观的指标。模式识别的方法越来越广泛地应用于ADHD的研究中^[3-4],基于不同的功能磁共振数据特征参数被提取(如功能连接、低频振幅、不同模板下提取的脑区平均时间序列等)用于ADHD的分类。然而,基于这些特征的分类结果并不是很理想,2011年,在美国1000

Functional Connectomes Project项目组组织的全球范围的ADHD分类大赛('The ADHD200 Global Competition')中,基于这些特征参数,21个参赛团队中最高分类准确率只有61.3% (195个测试样本中119个分类正确)^[4-5]。

小波分析作为一种非常重要的信号时频分析手段,具有良好的局部化性质,可以聚焦到信号的任意细节水平,近年来在磁共振信号处理中的应用越来越广泛^[6]。多种基于小波变换的特征提取的方法被用于构建分类特征参数,如求取小波系数平均值、绝对平局值、能量值等。其中小波系数能量值作为分类参数具有重要的意义^[6-7]。

收稿日期: 2015-04-07; 修回日期: 2015-09-15

基金项目: 国家自然科学基金(81473337); 国家社会科学基金(13BZJ032); 四川省应用基础项目(2013JY0189)

作者简介: 谭颖(1974-),男,博士,副教授,主要从事功能磁共振数据处理方面的研究。

功能磁共振成像(functional magnetic resonance imaging, fMRI)是一种无创的被广泛应用于测量大脑活动的工具。功能磁共振数据的信噪比低、数据量大,但是文献[8]认为传统的分类算法仍然可以对这些数据进行分类,特别是SVM (support vector machine)具有小样本学习和较好的泛化能力,在fMRI数据分类中得到了青睐。本文选取的功能磁共振数据来源于ADHD分类大赛,该数据量大,而且是多中心数据。50名ADHD儿童和65名健康儿童的静息态磁共振信号被用于研究。本文利用小波变换实现对信号的多尺度分解和构建小波系数能量值特征参数,并结合SVM分类器进行分类。最后找到能够有效区分ADHD病人与正常儿童的相关生物标记,为实际临床诊断提供有价值的参考。

1 实验数据及预处理

1.1 数据来源

本文研究的静息态功能磁共振数据来自ADHD-200竞赛网北京中心数据集数据。总共有194名儿童,其中正常儿童114人,女孩46人,男孩68人。平均年龄11岁;ADHD儿童80人,女孩10人,男孩70人,平均年龄12岁,所有的儿童都没有精神分裂、情感障碍、精神发育迟滞、神经系统以及发育障碍疾病,且均无特殊家族遗传性疾病史。

1.2 数据预处理

对所有静息态fMRI数据采用SPM8软件包进行预处理,主要包括时间校正、头动校正、空间标准化(3mm × 3mm × 3mm)以及8 mm高斯核平滑。其中头部平动大于2 mm并且转动大于2°和空间标准化匹配不完整的被试被排除,最终有50名ADHD儿童和65名正常儿童的静息态磁共振数据用于进一步的处理。

根据自动解剖标定(automated anatomical labeling, AAL)模板,配准后的静息态功能MRI数据被分割为116个感兴趣区域(78个皮层区域、12个皮下区域和26个小脑区域)。对每一个被试考虑90脑区为感兴趣区域。首先提取90个脑区每个区域的平均时间序列。为了消除整理噪声的影响,本文对这些时间序列进行带通滤波(0.01~0.08 Hz),而且回归了6个头动参数、白质信号、脑脊液信号以及它们的一阶导数,这个过程主要是去除不太可能参与神经元特定区域相关性的变量^[7-8]。由于目前对于是否回归全脑平均信号存在较大的争议,本文没有进行全脑

信号回归处理。

2 相关基本理论

2.1 小波变换

2.1.1 基本原理

设函数 $\psi(t) \in L^2(R)$ 满足容许条件:

$$C_\psi = \int_{-\infty}^{+\infty} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega < \infty \quad (1)$$

式中, $L^2(R)$ 为平方可积的函数空间; $\hat{\psi}(\omega)$ 为 $\psi(t)$ 的傅里叶变换,则称 $\psi(t)$ 为一基本小波或母小波(mother wavelet)。将母小波函数进行平移和伸缩变换,可以获得一组小波基函数:

$$\psi_{a,b}(t) = |a|^{1/2} \psi\left(\frac{t-b}{a}\right) \quad b \in R, a \in R, a \neq 0 \quad (2)$$

假如时间序列信号为 $f(t)$,则 $f(t)$ 的小波变换定义为:

$$W_\psi f(a,b) = |a|^{1/2} \int_R f(t) \bar{\psi}\left(\frac{t-b}{a}\right) dt \quad (3)$$

小波变换的本质就是利用这一组基函数对信号进行分解,从而实现对信号的时频分析。在实际数据处理中可供选择的小波基一般有Daubechies(dbN)系列、Symlet(symN)系列、Mexican小波以及Morlet小波等。这些小波都具有一定紧支撑性、对称性和平滑性的正交小波^[11]。

2.1.2 BOLD一维信号的小波分解

一维信号进行离散小波分解,分别得到信号的近似成份和细节成份,即高频部分和低频部分,而低频部分通常包含了信号的主要信息。实际信号处理中要对信号进行多尺度小波分解,该特性可以更精确的获得信号的深层信息。本文选择对血氧依赖水平(blood oxygenation level dependent, BOLD)信号进行3层分解,主要是因为磁共振信号的频率较低(0~0.25 Hz),并且为了降低分类特征维度。图1给出了BOLD信号一维小波分解的具体流程^[7]。 A_1 、 A_2 表示为信号小波分解的低频近似成份; D_1 、 D_2 表示为信号小波分解的高频细节成份。

2.1.3 BOLD信号的小波基选择

小波函数具有多样性,在磁共振时间序列中,选择不同的小波基处理同一数据,可能会得到不同的结果,因此寻找某一BOLD时间序列处理的最优小波基十分必要。

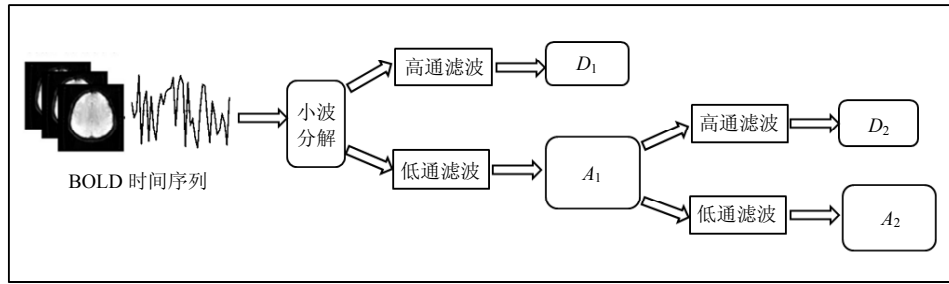


图1 小波分解示意图

本文介绍一种最优小波基选取方法, 其判断的准则是根据Mallat算法对原始信号 f_1 仅作单尺度小波分解与重构处理^[11-12], 然后计算重构信号 f_2 与原始信号 f_1 之间的误差。误差计算公式为:

$$\|f_1 - f_2\| = \sqrt{\sum_{i=1}^N (f_{1i} - f_{2i})^2}$$

式中, N 为信号的采样点的个数。

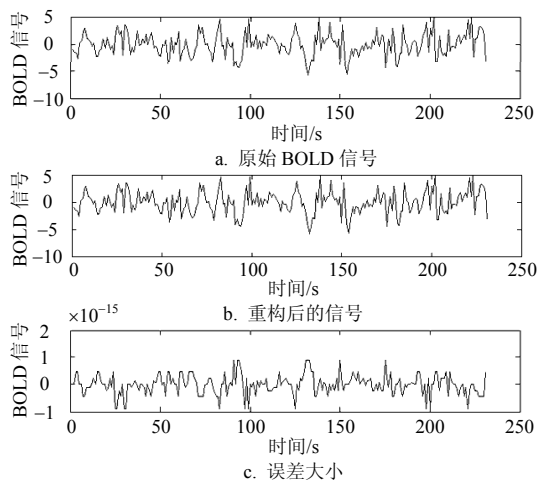
下面以一个BOLD信号在不同小波基下分解与重构的误差进行分析为例, 来说明最优小波基的选取。其中db和sym系列小波分解和重构之后的误差如表1和表2所示。

表1 db系列小波分解重构之后的误差 $\times 10^{-11}$

小波基	db ₁	db ₅	db ₆	db ₇	db ₈
$\ f_1 - f_2\ $	0.223	7.68	6.32	6.94	13.4

表2 sym系列小波分解重构之后的误差 $\times 10^{-11}$

小波基	sym ₄	sym ₅	sym ₆	sym ₇	sym ₈
$\ f_1 - f_2\ $	3.30	1.37	5.07	5.97	1.57

图2 db₁小波分解与重构效果图

通过表1与表2发现, db₁小波基分解重构的误差最低, 因此本文在处理BOLD信号中将选择db₁作为小波基。图2是选用db₁小波对一个BOLD信号进行小

波变换与重构的结果。由图2可以看出, 原始信号和重构信号只有微小的误差。

2.2 支持向量机

支持向量机(support vector machine, SVM)是一种最早的机器学习方法。该方法是建立在VC (vapnik-chervonenkis dimension)维理论和结构风险最小化原理的基础上, 具有很好的泛化能力, 目前在疾病的分类中得到了广泛的应用。SVM核心的原理是通过核函数把输入空间的样本映射到高维特征空间, 使得高维空间的点积运算在原低维空间进行, 从而在高维核空间中求取较低VC维的最优分类超平面^[2,10]。SVM的分类原理如下:

对于 n 个 m 维线性不可分样本有:

$$(x_1, y_1), \dots, (x_n, y_n) \in R^m \times \{-1, 1\} \quad (4)$$

在高维空间中构造的最优超平面满足:

$$y_i[\omega \cdot \phi(x_i) + b] \geq 1 \quad i = 1, 2, \dots, n \quad (5)$$

将该问题转化为二次规划问题:

$$\min J(\omega, \xi) = \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^n \xi_i \quad (6)$$

使 $y_i[\omega \cdot \phi(x_i) + b] \geq 1 - \xi_i, \xi_i \geq 0, i = 1, 2, \dots, n$ 。 C 为惩罚因子, ξ 为松弛因子。对于一些不能被超平面正确分类的噪音样本, 通过引入 ξ 来构造多个超平面,

如果样本被错分, 则其 $\xi > 0$, $\sum_{i=1}^n \xi_i$ 是错分样本误差的一个上界。惩罚因子 C 在超平面与最近的训练点之间的距离最大与分类误差最小之间寻求最佳折中。

根据拉格朗日乘子法, 最优分类超平面的求解可转化为以下约束问题:

$$Q(a) = \sum_{i=1}^N a_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N a_i a_j y_i y_j K(x_i, x_j) \quad (7)$$

其约束条件为:

$$\sum_{j=1}^N a_j y_j = 0 \quad a_i \geq 0 \quad i = 1, 2, \dots, N \quad (8)$$

式中, $\{a_i\}_{i=1}^N$ 为拉格朗日乘子, 其中大部分的 a_i 为0, 而不等于0的 a_i 所对应的样本即为支持向量机, 对应

的判别函数为:

$$D(x) = \text{sgn}(\sum_{i=1}^n a_i^* y_i (xx_i) + \omega_0^*) \quad (9)$$

而 $K(x_i, x_j)$ 为核函数满足 Mercer 定理, 常用的核函数主要有^[10]:

1) 线性核为 $K(x_i, x_j) = \langle x_i, x_j \rangle$;

2) 多项式核为 $K(x_i, x_j) = (\langle x_i, x_j \rangle + c)^p, c \geq 0$,

P 是自然数;

3) 高斯核 (径向基 RBF) 为 $K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right), \sigma > 0$;

4) Sigmoid 核为 $K(x_i, x_j) = \tanh(\gamma x_i^T x_j + r)$ 。

3 ADHD病人静息态磁共振BOLD信号特征提取与分类方法

3.1 能量表征及特征参数提取

信号经小波变换后可以得到很多的描述信号模式的特征, 获取各种候选特征子集为进一步的分类做好准备。文献[4]介绍了以小波变换的结果进行特征表示的各种常用方法。其中能量作为特征的方法被广泛的应用。

设 $E_{jk} = |D_j(k)|^2$ 为信号 $f(t)$ 在第 j 尺度 k 点的小波能量, 那么 $E_j = \sum_{k=1}^N E_{jk}$ 就为第 j 尺度下 $k=1, 2, \dots, N$ 个采样点的信号能量和, 其中 $D_j(k)$ 为 BOLD 信号经小波变换后第 j 尺度下 k 各数据^[13,15]。

大量磁共振信号研究表明, 人脑活动主要与固有的低频振荡(0.01~0.08 Hz)有关。本文中预处理的数据已经经过滤波处理, 因此着重考虑经小波分解后各频率段信号能量作为分类特征对 ADHD 病人进行分类。根据文献[13-14], 基于小波变换的静息态 BOLD 信号特征提取的算法如下:

1) 利用 AAL 模板提取经过预处理后的 90 个脑区的平均时间序列信号;

2) 选用 db₁ 最优小波基对提取的平均时间序列信号进行 3 层小波分解, 分别获得低频信号成份 CA_3 和 3 个高频信号成份 CD_3, CD_2, CD_1 ;

3) 提取各尺度小波系数, 计算小波系数能量值, 构造用于分类的特征向量为:

$$T = [E(CA_3), E(CD_3), E(CD_2), E(CD_1)]$$

4) 对特征向量进行归一化, 总的信号能量记做 E_0 , 则: $E_0 = E(CA_3) + E(CD_3) + E(CD_2) + E(CD_1)$,

因此归一化后的特征向量为:

$$\bar{T} = [E(CA_3)/E_0, E(CD_3)/E_0, E(CD_2)/E_0, E(CD_1)/E_0]$$

5) 用 SVM 分类器对 ADHD 病人和正常人进行分类, 并进行决策和获得分类权重显著高的脑区。

图 3 给出了基于小波变换特征提取与 SVM 进行分类的算法框图。

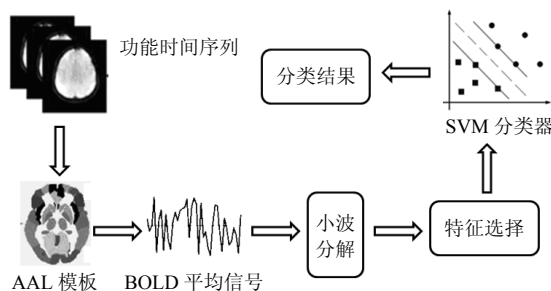


图3 BOLD时间序列经小波变换后进行特征提取和SVM分类的处理框图

3.2 分类器性能的评估

本文采用了较大的样本量, 用留一法来评估分类器的性能。利用准确率(accuracy, ACC)、敏感性(sensitivity, SEN)和特异性(specificity, SPE)这些指标来量化分类器的性能。各项指标计算方法如下^[4,10]:

$$ACC = \frac{TP+TN}{TP+TN+FP+FN}$$

$$SEN = \frac{TP}{TP+FN}$$

$$SPE = \frac{TN}{TN+FP}$$

$$P = \frac{\sum_{i=1}^m \left| \{ACC(\hat{S}) \geq ACC(S)\} + 1 \right|}{m+1}$$

式中, True Positive(TP)为真阳性的个数, 即病人正分的个数; True Negative(TN)为真阴性的个数, 即正常人正分的个数; False Positive(FP)为假阳性的个数, 即正常人错分的个数; False Negative(TN)为假阴性的个数, 即病人错分的个数; 为原始数据训练数据得到的准确率, $ACC(\hat{S})$ 为随机打乱训练数据集 m 次上得到的准确率, 本文取 $m=1\ 000$, 其中 P 值越小则准确率越难以随机得到, 若 $P < 0.05$, 则表示分类准确率在统计上具有显著性。另外, 在分类过程中可以获得每一个测试样本到决策面的距离, 将每一个距离作为一个阈值, 可以得到相应的受试者工作特征(receiver operating characteristic, ROC), ROC 曲线下面积(area under the curve, AUC)也可以作为一个量化分类器好坏的指标。

4 结果与分析

小波多尺度分解可以得到各个尺度频段上的信号特征, 因此将不同尺度上信号的能量值作为分类的特征向量是可行的^[12]。本文选择了SVM作为分类器, 使用了LIBSVM工具包, 惩罚参数 c 为默认值($c=1$)。在评估分类器的性能方面, 使用的是嵌入式留一法交叉验证, 这种方法可以对真实的泛化能力进行相对无偏的估计^[10]。最优SVM模型和最优特征子集由内部交叉验证获得, 然后用于对测试集样本进行测试^[10]。另外, 核函数对分类器的分类效果也有显著的影响, 本文分别讨论了线性核与高斯核对分类结果的影响, 表3给出了分类性能结果、AUC ($p=0.025$)值及方法比较, 图4给出了基于线性核对应的ROC曲线及曲线下面积。获得的平均准确率为62.7%, 这个结果好于ADHD-200 全球竞赛中冠军队基于图像特征的61.3%的分类结果。因此, 基于小波分解以信号的能量作为特征的方法是有效的。

表3 以信号能量作为特征的线性和非线性分类器性能比较

SVM分类器	p 值	ACC/%	SEN/%	SPE/%	AUC/%
线性	0.025	62.7	60.1	64.5	65.97
非线性(径向基)	0.030	58.2	56.3	65.6	62.22

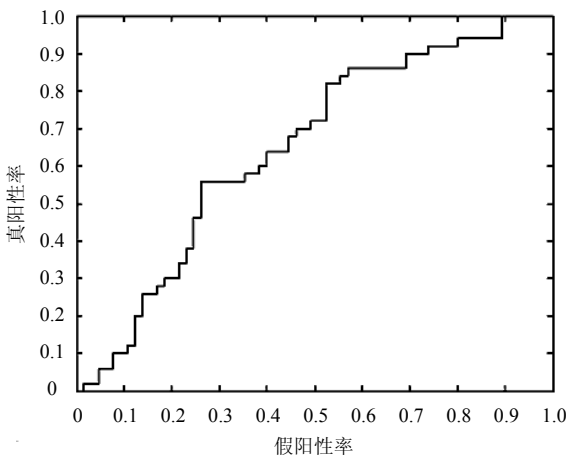


图4 基于信号能量作为特征的分类ROC曲线

不同的区域的信号特征在分类的过程中表现出不同的分类权重, 基于权重的大小可以帮助我们初步推断那些脑区可能在疾病中扮演重要的角色。通常, 如果一些脑区的权重大于所有区域权重的均值加上一倍的标准差, 就把这些区域看作是有显著高的权重^[16]。图5显示了基于这个判别标准下13个脑区拥有显著高的分类权重, 并给出了相应脑区的位置分布。左侧为脑区的标准化权重降序排列图, 黑线表示为权重为均值加一倍标准差的分界线; 右侧为权重显著高的脑区的位置分布, 节点的大小代表了

权重的大小。这些脑区主要是: 左侧背外侧上回、左侧眶部额上回、左侧额中回、右侧眶部额中回、左侧岛盖部额下回、右侧眶部额下回、右侧中央沟盖、左右眶内额上回、左侧后扣带回、左侧距状裂周围皮层、右侧豆状壳核、右侧丘脑。

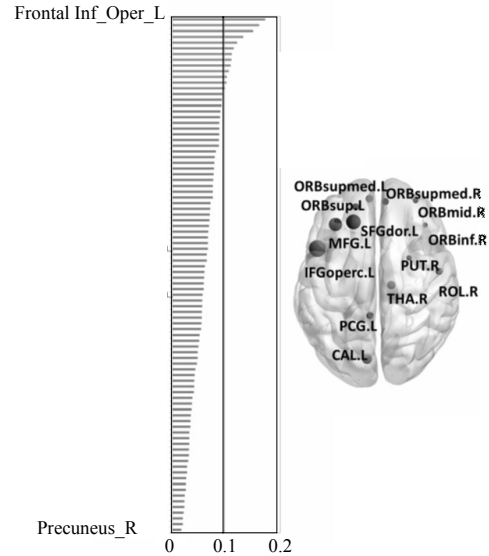


图5 基于SVM分类的90个脑区分类权重示意图

本文发现权重较高的脑区主要在额顶区域, 表明ADHD病人相对于正常人来说额顶区域的BOLD信号能量存在异常, 与之前的研究结论相一致^[17-18]。

5 结束语

本文提取了大尺度功能时间序列平均信号, 通过小波多尺度分解获取了大脑信号的深层信息特性。通过计算信号的内在小波系数能量值来构造了分类特征向量, 并利用SVM分类器来进行分类。该分类器的分类准确率虽然只有62.7%, 但是优于ADHD-200全球竞赛第一名的分类结果^[4], 且从信号能量分布的角度, 这有助于对ADHD病人的影像学诊断及其机制的理解。本文发现13个脑区有显著的分类权重, 这些脑区的BOLD信号能量的差异可能成为区分ADHD病人与健康人的重要的生物标记。另外也发现小波变换在对功能磁共振信号分解与特征提取上是一种有效的方法, 以后将进一步探讨如何提高小波变换在功能影像学分类中的准确率。

参 考 文 献

- [1] POLANCZK G, DE LIMA M S, HORTA B L, et al. The worldwide prevalence of ADHD: a systematic review and metaregression analysis[J]. Am J Psychiatry, 2007, 164(6): 942-948.
- [2] SELLERS R, MAUGHAN B, PICKLES A, et al. Trends in parent and teacher rated emotional, conduct and ADHD

- problems and their impact in prepubertal children in Great Britain: 1999-2008[J]. *Journal of Child Psychology and Psychiatry*, 2015, 56(1): 49-57.
- [3] ANURADHA J, RAMACHANDRAN V, ARULALAN K V, et al. Diagnosis of ADHD using SVM algorithm[C]// *Proceedings of the Third Annual ACM Bangalore Conference*. [S.l.]: ACM, 2010.
- [4] GUO X, AN X, KUANG D, et al. ADHD-200 Classification based on social network Method[C]// *Intelligent Computing in Bioinformatics*. Taiyuan, China: Springer International Publishing, 2014: 233-240.
- [5] 李长斌. 基于脑活动网络特征的ADHD分类研究[D]. 北京: 北京交通大学, 2014.
LI Chang-bin. Research on classification of ADHD based on the characteristic in brain activity-network[D]. Beijing: Beijing Jiaotong University, 2014.
- [6] AMIN H U, MALIK A S, AHMAD R F, et al. Feature extraction and classification for EEG signals using wavelet transform and machine learning techniques[J]. *Australasian Physical & Engineering Sciences in Medicine*, 2015, 38(1): 139-149.
- [7] 张晓文, 杨熠普, 许晓鸣. 基于小波变换的特征构造与选择[J]. *计算机工程与应用*, 2003, 19(19): 25-28.
ZHANG Xiao-wen, YANG Yu-pu, XU Xiao-ming. Wavelet based feature representation[J]. *Computer Engineering and Applications*, 2003, 19(19): 25-28.
- [8] MITCHELL T M. Learning to decode cognitive states from brain images[J]. *Machine Learning*, 2004, 57(1-2): 145-175.
- [9] WEI Liao. Altered functional connectivity and small-world in mesial temporal lobe epilepsy[J]. *PLoS One*, 2010, 5(1): e8525.
- [10] 刘凤. 基于磁共振成像的多变量模式分析方法学研究[D]. 成都: 电子科技大学, 2014.
LIU Feng. MRI-based multivariate pattern analysis methodologies and applications[D]. Chengdu: University of Electronic Science and Technology of China, 2014.
- [11] 孙延奎. 小波分析及其应用[M]. 北京: 机械工业出版社, 2005.
SUN Yan-kui. Wavelet analysis and its application[M]. Beijing: China Machine Press, 2005.
- [12] 成谢锋, 杨贺. 5种小波在心音信号处理中的分析与比较[J]. *南京邮电大学学报*, 2015, 35(2): 39-46.
CHENG Xie-feng, YANG He. Analysis and comparison of five kinds of wavelet in processing heart sound signal[J]. *Journal of Nanjing University of Posts and Telecommunications*, 2015, 35(2): 39-46.
- [13] JIAG Y E, LIN Y P. Error Estimation of INS ground alignment through observability analysis[J]. *IEEE Transactions on Aerospace and Electronic Sstems*, 1992, 28(1): 92-97.
- [14] 肖汉光, 蔡从中. 特征向量的归一化比较性研究[J]. *计算机工程与应用*, 2009, 45(22): 117-119.
XIAO Han-guang, CAI Cong-zhong. Comparison study of normalization of feature vector[J]. *Computer Engineering and Applications*, 2009, 45(22): 117-119.
- [15] 张丽琼, 王炳和. 基于小波变换的脉象信号特征提取方法[J]. *数据采集与处理*, 2004, 19(3): 322-325.
ZHANG Li-qiong, WANG Bin-he. Feature extraction methods for pulse signal based on wavelet transform[J]. *Journal of Data Acquisition & Processing*, 2004, 19(3): 322-325.
- [16] TIAN L, WANG J. Hemisphere and gender-related differences in small-world brain networks: a resting-state functional MRI study[J]. *NeuroImage*, 2011, 54(1): 191-202.
- [17] VAIDYA C J, BUNGE S A, DUDUKOVIC N M, et al. Altered neural substrates of cognitive control in childhood ADHD: Evidence from functional magnetic resonance imaging[J]. *Am J Psychiatry*, 2005, 162(9): 1605-1613.
- [18] SOLANTO M V, MARKS D J, WASSERSTEIN J, et al. Efficacy of meta-cognitive therapy for adult ADHD[J]. *The American Journal of Psychiatry*, 2010, 167(8): 958-968.

编辑 黄 莘