

# 轻度认知障碍的 BOLD 信号 MEMD 降噪处理

汪瀚, 吴海锋, 王俨, 王勇, 王霞<sup>†</sup>

(云南民族大学 电气信息工程学院, 云南 昆明 650504)

**摘要:** 早期诊断轻度认知障碍是干预阿尔茨海默症的有效途径. 目前常使用静息态功能磁共振成像和机器学习方法进行轻度认知障碍的辅助诊断, 关键是使用血氧水平依赖 (Blood Oxygenation Level Dependent, BOLD) 信号构建大脑的功能性连接. 针对大脑静息态 BOLD 信号中存在各种外界噪音干扰的问题, 提出了结合多元经验模态分解与皮尔逊相关的重构方法与极正极负重构准则, 将大脑默认模式网络的中心节点后扣带回皮层作为模板, 重构 BOLD 信号以降低外界噪音干扰. 实验结果表明, 基于极正极负重构准则降噪后的 BOLD 信号构建功能性连接, 相较降噪前的数据, 在分类性能方面可以提高数据的差异性, 在特征选择性能方面可以对数据集降维的同时进一步提升分类性能. 此外, 以上性能均优于传统重构准则. 最后, 对降噪后的最优特征子集进行统计性分析, 发现脑岛可能是默认模式网络的相关脑区, 小脑蚓体与后扣带回皮层可能构成了一种认知功能补偿网络, 这是以往研究中少有提出的结论.

**关键词:** 静息态功能磁共振成像; 轻度认知障碍; 血氧水平依赖; 功能性连接; 多元经验模态分解; 皮尔逊相关; 机器学习

中图分类号: TP391.4

文献标志码: A

DOI: 10.13195/j.kzyjc.2022.0143

开放科学(资源服务)标识码(OSID):



## Blood Oxygenation Level Dependent signals noise reduction for Mild Cognitive Impairment using Multivariate Empirical Mode Decomposition

Wang Han, Wu Hai-feng, Wang Yan, Wang Yong, Wang Xia<sup>†</sup>

(College of Electrical Information Engineering, Yunnan Minzu University, Kunming 650504, China)

**Abstract:** Early diagnosis of mild cognitive impairment is an effective way to intervene Alzheimer's disease. At present, resting fMRI and machine learning methods are often used in the auxiliary diagnosis of mild cognitive impairment. The key is to use Blood Oxygenation Level Dependent (BOLD) signals to construct the functional connection of the brain. Aiming at the problem of various external noise interference in brain resting BOLD signals, a reconstruction method combined with Multiple Empirical Mode Decomposition and Pearson correlation and polar positive negative reconstruction criterion are proposed. The posterior cingulate cortex, the central node of brain default mode network, is used as a template to reconstruct BOLD signals to reduce external noise interference. The experimental results show that the functional connection based on pole positive negative reconstruction criterion can improve the difference of data in classification performance compared with the data before noise reduction, and further improve the classification performance while reducing the dimension of data set in feature selection performance. In addition, the above performance is better than the traditional reconstruction criterion. Finally, the optimal feature subset after noise reduction is statistically analyzed. It is found that the insula may be the relevant brain area of the default mode network, and the cerebellar vermis and posterior cingulate cortex may constitute a cognitive function compensation network, which is a rare conclusion in previous studies.

**Keywords:** rest-state functional Magnetic Resonance Imaging; mild cognitive impairment; Blood Oxygenation Level Dependent; functional connectivity; Multivariate Empirical Mode Decomposition; Pearson correlation; machine learning

## 0 引言

阿尔茨海默病 (Alzheimer's Disease, AD) 是一种进行性神经退行性疾病, 又名老年痴呆症, 临床表现

为记忆与认知功能的减退<sup>[1]</sup>, 目前只有一些治疗手段可以干预 AD 的恶化, 尚未出现有效治疗 AD 的方法. 轻度认知障碍 (Mild Cognitive Impairment, MCI)

收稿日期: 2022-01-20; 录用日期: 2022-04-27.

基金项目: 云南省基础研究专项-面上项目 (202101AT070072). 云南省教育厅科学研究基金项目 (2022J0439). 国家自然科学基金项目 (62161052).

<sup>†</sup>通讯作者. E-mail: wangxiacsu@163.com.

是指由正常老化导致的认知衰退与 AD 之间的过渡阶段, 虽然未达到 AD 的认知衰退程度<sup>[2]</sup>, 但研究表明 MCI 有转化为 AD 的可能, 且转化率逐年递增, 有超过一半的 MCI 患者在 5 年内转化为 AD<sup>[3]</sup>. 因此 MCI 的早期诊断对 AD 的干预非常重要. 对于 MCI 辅助诊断方面的研究, 多使用基于静息态磁共振成像 (rest-state functional Magnetic Resonance Imaging, rs-fMRI) 提取的血氧水平依赖 (Blood Oxygenation Level Dependent, BOLD) 信号, 重点在于使用 BOLD 信号建立脑功能性连接 (Functional Connectivity, FC), 对 MCI 与正常认知 (Normal Cognition, NC) 间的差异性进行研究<sup>[4]</sup>.

然而研究表明 BOLD 信号会受到各种外界因素的干扰, 主要分为头部运动与外界噪声两个方面. 在静息态下, 即使是细微的头部运动也会使 BOLD 信号产生波动<sup>[5]</sup>. 而对于外界噪声, 既有电路热噪声、仪器漂移与硬件不稳定造成的波动<sup>[6]</sup>, 也有心跳、呼吸、血管活动与新陈代谢等生理噪声<sup>[7-9]</sup>. 这些复杂的外界干扰可能会恶化 BOLD 信号的质量从而影响 FC 的有效性, 甚至得出错误的研究结论.

针对上述问题, 数据驱动方法是一种主流的有效工具, 通过对数据底层结构进行分析, 来从 BOLD 信号中提取有价值的信息. 过去的研究使用了如独立成分分析 (Independent Component Correlation Algorithm, ICA)、主成分分析 (Principal Component Analysis, PCA)、模糊聚类分析 (Fuzzy Cluster Analysis, FCA) 等传统的数据驱动方法<sup>[10-12]</sup>, 但这些方法在分析中均存在着一定的局限性<sup>[13]</sup>. 多元经验模态分解 (Multivariate Empirical Mode Decomposition, MEMD) 作为一种多元数据驱动方法近几年被应用于 fMRI 相关研究中, 它使同一子频带内的多通道信号特征可以在一个固有模态函数 (Intrinsic Model Function, IMF) 分量中进行匹配和分类<sup>[14]</sup>. 文献 [15] 采用 MEMD 对任务态 BOLD 信号进行分解, 通过寻找与嗅觉任务相关的 IMF 分量对 BOLD 信号进行重构, 以提取嗅觉诱发的血液动力学反应<sup>[15]</sup>. 然而对于静息态 BOLD 信号, 由于无法寻找与特定任务相关的 IMF 分量, 目前尚无文献基于 MEMD 对静息态 BOLD 信号进行重构.

针对大脑静息态 BOLD 信号中掺杂的外界噪声干扰, 本文提出了结合多元经验模态分解与皮尔逊相关的重构方法 (Multivariate Empirical Mode Decomposition - Pearson correlation Restructure, MEMD-PR) 与极正极负重构准则. 首先, 使用 MEMD

将 BOLD 信号分解得到不同频率维度的 IMF 分量. 其次, 由于静息态时大脑功能网络主要为默认模式网络 (Default Mode Network, DMN)<sup>[16]</sup>, 而后扣带回皮层 (Posterior Cingulate Cortex, PCC) 是 DMN 的主要节点<sup>[17]</sup>, 所以本文以 PCC 节点为模板, 根据极正极负重构准则以筛选出与静息态相关的 IMF 分量. 最后, 根据筛选后的 IMF 分量对 BOLD 信号进行重构, 从而将对应于外界噪声的 IMF 分量滤除.

为了验证 MEMD-PR 方法的有效性, 本文设计了以下验证实验: 第一, 计算被试 BOLD 信号重构前后 PCC 节点与各脑区间的 FC, 以此作为分类特征输入, 再使用支持向量机 (Support Vector Machine, SVM) 进行分类准确度的测试, 实验结果表明降噪后 MCI 与 NC 间的差异性得到提升, 而且极正极负重构准则优于传统重构准则. 第二, 对降噪前后的分类特征输入使用近邻成分分析 (Neighborhood Component Analysis, NCA)<sup>[18]</sup> 特征选择算法提取最优特征子集, 再进行分类准确率的测试, 实验结果表明在分类性能提升程度与数据集降维程度两方面, 基于极正极负重构准则降噪后的性能是原始的 6 倍与 12 倍, 相较传统重构准则也有所提升, 表明 MEMD-PR 方法降噪后的数据, 其特征更为明显, 更有利于特征的提取. 上述结果验证了本文提出的 MEMD-PR 方法可以去掉部分外界噪声对大脑静息态 BOLD 信号的干扰. 最后, 对降噪前后最优特征子集进行统计性分析, 发现降噪前与以往研究结果具有较高一致性, 而降噪后的结果表明脑岛可能也是 DMN 的相关脑区, 此外小脑蚓体可能与 PCC 节点构成了某种认知功能补偿网络, 这是以往研究少有提到的.

## 1 相关工作

随着神经影像技术的发展, fMRI 技术由于其具有无创性的特点, 近年来被广泛应用于脑科学的相关研究. 其中, 无需受试者完成特定任务的 rs-fMRI 技术受到广泛关注, 其有效避免了受试者完成任务过程中对数据采集的影响. 使用 rs-fMRI 数据提取的 BOLD 信号构建 FC 是研究 MCI 与 NC 间差异性的常用方法, 而基于感兴趣区域 (Region of Interest, ROI) 的皮尔逊相关系数法由于其计算简单、高效易操作等特点, 可以有效构建 FC, 因此本文使用皮尔逊相关系数法构建 FC.

传统的皮尔逊相关系数法主要关注的是整个大脑的 FC, 根据已知的模板将大脑划分为大量 ROI, 再计算所有 ROI 间的皮尔逊相关系数. 但全脑 FC 的数据维度较大, 对于特定疾病的研究, 可能存在大量

的冗余信息. 已有研究表明大脑在静息态时并非无任何反应<sup>[19]</sup>, 而是在某些区域持续存在内在的自发反应, 并且受试者在任务态时这些反应便会减少<sup>[20]</sup>, 这表明大脑的功能网络存在一种默认模式, 被称为 DMN. 也有一些研究重点关注 DMN 的相关脑区, 其主要位于内侧前额叶皮层 (Medial Prefrontal Cortex, MPFC)、后扣带回皮质 (Posterior Cingulate Cortex, PCC)、顶叶下叶 (Inferior Parietal Lobe, IPL)、外侧颞叶皮层 (Lateral Temporal Cortex, LTC) 的背侧和腹侧, 以及海马结构 (HF)<sup>[21-23]</sup>. 在这些区域中, PCC 是 DMN 的主要节点, 在一项基于任务态的与 DMN 区域共激活的 meta 分析中, 该节点对应于一个峰值坐标, 并与其他 DMN 区域间表现出高重测信度的静息态连接<sup>[24]</sup>, 因此本文只重点关注 PCC 节点与各脑区间的 FC, 以避免全脑 FC 可能存在的维度诅咒.

针对 BOLD 信号中存在的各种干扰, 数据驱动方法是一种主流的有效工具, 传统的研究使用了 ICA、PCA、FCA 等数据驱动方法来提取有意义的信息, 然而它们的使用受到一定的限制. 近几年数据驱动方法主要围绕 ICA 的统计独立性与字典学习 (Dictionary Learning, DL) 的稀疏性进行发展. 熵界最小化和熵率界最小化是最近引入的两种 ICA 算法, 在 fMRI 数据上具有良好的性能<sup>[25-27]</sup>. DL 利用稀疏性作为分解的起点, 对 fMRI 数据的分析也很有效<sup>[28]</sup>. 由于各种 ICA 算法与 DL 方法中假设的建模不同, 难以进行比较, 文献 [29] 提出使用客观的全局指标 (如图论度量) 来评价多样性在 fMRI 分析中的作用, 发现最优方法将根据分析目标的不同而变化<sup>[29]</sup>.

经验模态分解 (Empirical Mode Decomposition, EMD) 是一种分析非线性和非平稳数据的数据驱动方法, 与 ICA 和 DL 方法不同, 它不需要事先的统计假设或预先定义的聚类数, 通过迭代的筛选过程将信号分解成有限数量的 IMF<sup>[30]</sup>, 近几年开始运用于 fMRI 研究<sup>[31]</sup>. 然而若对每个 BOLD 信号单独进行 EMD 分解, 每个 BOLD 信号对应的 IMF 分量数无法保证一致, 并且不同信号的各层 IMF 分量对应的固有频率维度也难以匹配, 针对多通道信号的 MEMD 则有效解决了上述问题. 文献 [32] 基于 MEMD 提出了一种不同频率维度的 FC 系数作为分类的特征输入, 以提高 MCI 与 NC 的分类准确率<sup>[32-33]</sup>, 体现了 MEMD 在 rs-fMRI 研究中的潜力. 通过 MEMD 对信号进行分解并重构是近几年新兴的一种信号处理方法. 文献 [34] 依据峭度准则和相关系数对机械故障主要信息对应的 IMF 分量进行筛选, 再对机械故障

信号进行重构以提取故障特征<sup>[34]</sup>. 在 fMRI 相关研究中, 文献 [15] 针对任务态 BOLD 信号, 通过互相关系数与 KMeans 算法寻找与嗅觉任务相关的 IMF 分量对 BOLD 信号进行重构, 以此提取有意义的嗅觉诱发血液动力学反应<sup>[15]</sup>. 然而对于静息态 BOLD 信号, 并没有与特定任务相关的 IMF 分量, 如何寻找与静息态相关的 IMF 分量并设计相应的 BOLD 信号重构方法, 目前尚无文献报道.

## 2 问题提出

对于静息态 BOLD 信号, 无法通过寻找与特定任务相关的 IMF 分量来对 BOLD 信号进行重构. 针对上述问题, 对静息态 BOLD 信号降噪需要解决的首要问题是如何寻找与静息态相关的 IMF 分量, 以及如何对 IMF 分量进行合理筛选的问题.

针对第一个问题, 考虑到大脑静息态时主要功能网络为默认模式网络, PCC 节点为其中心节点, 故本文以 PCC 节点为模板寻找与静息态相关的 IMF 分量. 针对第二个问题, 考虑到噪声对 BOLD 信号的影响可能导致 IMF 分量与 PCC 节点的相关性发生极性的改变, 本文提出极正极负重构准则以筛选用于重构 BOLD 信号的 IMF 分量.

## 3 MEMD-PR 方法

### 3.1 MEMD-PR 基本过程

本文中, 我们提出了 MEMD-PR 方法来对大脑静息态的 BOLD 信号进行重构, 重构过程如附录图 A1 所示. 首先, 从被试的 rs-fMRI 数据中提取 ROI 与 PCC 节点的 BOLD 信号, 经过 MEMD 分解得到每个 ROI 的若干个 IMF 分量; 然后, 将每个 IMF 分量与 PCC 节点的 BOLD 信号计算皮尔逊相关系数; 最后, 根据相关系数矩阵的结果, 我们提出了极正极负重构准则来筛选出与 PCC 节点相关的 IMF 分量, 并将选出的 IMF 分量进行叠加完成对 BOLD 信号的重构. 下面将对各步骤进行详细介绍.

### 3.2 MEMD 分解

为了在不同频率维度上对 BOLD 信号进行分析, 需要使用 EMD 分解技术, 但对每个 ROI 单独进行 EMD 分解会使各 ROI 对应的 IMF 分量的个数不同, 且各层 IMF 分量的频率维度也无法匹配. 为了解决这个问题, 需要使用多通道的 MEMD 技术, 以确保各 ROI 对应的 IMF 分量在频率维度以及数量上保持一致.

首先, 从被试的 rs-fMRI 数据中提取  $N$  个 ROI 以及 PCC 节点的 BOLD 信号, 令  $\mathbf{X}_n^{(q)} = [x_{i,n}^{(q)}] \in \mathbb{R}^{T \times 1}, n = 1, 2, \dots, N$  为从第  $q$  个被试的 rs-fMRI 数据

中提取的  $N$  个 ROI 的 BOLD 信号,  $T$  为信号的时间点个数. 然后, 对这些序列使用  $N$  通道的 MEMD 分解, 每个 ROI 均得到  $M$  层 IMF 分量与残差. 可表示为

$$\mathbf{X}_n^{(q)} = \sum_{m=1}^M \mathbf{C}_n^{(q,m)} + \mathbf{E}_n, n=1, 2, \dots, N, m=1, 2, \dots, M \quad (1)$$

其中,  $\mathbf{C}_n^{(q,m)}$  为第  $m$  层 IMF 分量,  $\mathbf{E}_n$  为残差.

### 3.3 计算皮尔逊相关系数矩阵

经过 MEMD 分解后, 得到被试的  $N$  个 ROI 的各层 IMF 分量与残差, 再将残差去除, 由于存在的外界噪声可能分布于不同的频率维度上, 那么剩下的各层 IMF 分量中可能存在对应于外界噪声的 IMF 分量.

为了去除外界噪声产生的干扰, 可以寻找与 PCC 节点相关的 IMF 分量对 BOLD 信号进行重构. 要衡量 IMF 分量与 PCC 节点的相关程度, 可以通过计算各层 IMF 分量与该被试 PCC 节点 BOLD 信号的皮尔逊相关系数矩阵, 来筛选出与 PCC 节点相关的 IMF 分量, 从而去除可能对应于外界噪声的 IMF 分量所产生的影响. 由于每个 ROI 时间序列分解得到的 IMF 个数相同且频率维度一致, 所以各 ROI 不同层 IMF 分量计算得到的皮尔逊相关系数可以彼此匹配. 计算相关系数矩阵是对一个被试进行计算, 为了简化符号表示, 省去  $\mathbf{X}_p^{(q)}$  与  $\mathbf{C}_n^{(q,m)}$  中符号  $q$ . 令  $\mathbf{C}_n^{(m)} = [c_{t,n}^m] \in \mathbb{R}^{T \times 1}$  为一被试第  $n$  个 ROI 的第  $m$  层 IMF 分量,  $\mathbf{X}_p = [x_{t,p}] \in \mathbb{R}^{T \times 1}$  为该被试 PCC 节点 BOLD 信号, 那么它们之间的皮尔逊相关系数可表示为:

$$r_n^{(m)} = \frac{(\mathbf{C}_n^{(m)} - \mu_n^{(m)})(\mathbf{X}_p - \mu_p)^H}{\sqrt{(\mathbf{C}_n^{(m)} - \mu_n^{(m)})(\mathbf{C}_n^{(m)} - \mu_n^{(m)})^H (\mathbf{X}_p - \mu_p)(\mathbf{X}_p - \mu_p)^H}} \quad (2)$$

其中  $\mu_n^{(m)}$  与  $\mu_p$  分别表示  $\mathbf{C}_n^{(m)}$  与  $\mathbf{X}_p$  的均值.

若提取的 ROI 个数为  $N$ , MEMD 分解得到的 IMF 分量个数为  $M$ , 则可得到任意一个被试的相关系数矩阵:

$$\mathbf{R} = [r_n^{(m)}] \in \mathbb{R}^{M \times N}, n=1, 2, \dots, N, m=1, 2, \dots, M \quad (3)$$

### 3.4 极正极负重构准则与 BOLD 信号的重构

获得每个 ROI 的各层 IMF 分量与 PCC 节点 BOLD 信号的相关系数矩阵后, 便可以此筛选出与 PCC 节点相关的 IMF 分量. 由于相关系数矩阵中正相关与负相关同时存在, 且已有研究表明 PCC 节点与负相关区域的相关系数明显弱于其与 DMN 正相

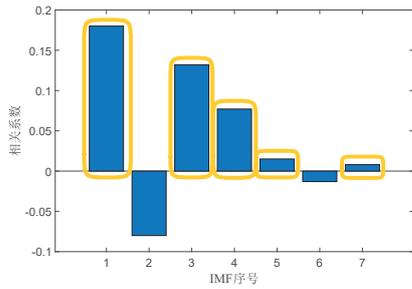
关区域的相关系数<sup>[35]</sup>, 这对于 MCI 显著性差异的相关研究是有意义的, 所以在筛选时正相关与负相关需要同时考虑. 在外界噪声干扰下, 可能存在某些 ROI 与 PCC 节点间的较强正相关变为较弱正相关, 或者某些 IMF 分量与 PCC 节点由较弱正相关变成了较弱负相关, 反之同理. 这些情况都可能会使 rs-fMRI 的相关研究结果出现误差, 甚至隐藏了本该出现的结果.

针对这些复杂的情况, 本文提出了极正极负重构准则, 目的是基于皮尔逊正相关系数与负相关系数, 筛选出与 PCC 节点相关的 IMF 分量, 去除外界噪声对大脑静息态 BOLD 信号的影响. 下面对极正极负重构准则进行具体说明.

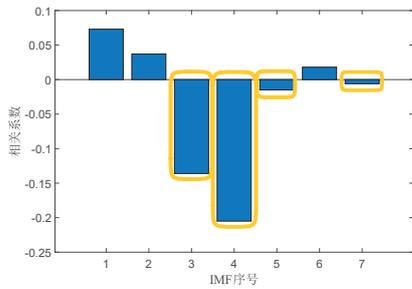
考虑到对于一个被试的各 ROI 经过 MEMD 分解得到的 IMF 分量数是一致的, 但无法确保不同被试得到的 IMF 分量数是相同的. 为了使重构准则对所有被试均适用, 令所有被试用于重构的 IMF 分量数为  $S = \min \{M_1, M_2, \dots, M_q\} - I$ , 其中  $I$  表示重构中至少去除的 IMF 分量数, 当  $I = 1$  时表明至少去除 1 个对应于噪声的 IMF 分量. 由于是对一个被试的 BOLD 信号进行重构, 后续说明为了简化符号表示, 省去符号  $q$ .

根据得到的皮尔逊相关系数矩阵  $\mathbf{R}$ , 观察每个 ROI 各层 IMF 分量的正相关系数与负相关系数的个数, 令第  $n$  个 ROI 的正相关系数个数为  $U^{(n)}$ , 负相关系数个数为  $V^{(n)}$ ,  $n = 1, 2, \dots, N$ , 对所有系数按正负相关性分别进行绝对值由大到小排序, 正相关系数按绝对值排序对应的 IMF 分量可记为  $\mathbf{C}_n^{(\text{pos}, u)} = [c_{t,n}^{(\text{pos}, u)}] \in \mathbb{R}^{T \times 1}$ ,  $u = 1, 2, \dots, U^{(n)}$ , pos 表示相关系数为正; 负相关系数按绝对值排序对应的 IMF 分量可记为  $\mathbf{C}_n^{(\text{neg}, v)} = [c_{t,n}^{(\text{neg}, v)}] \in \mathbb{R}^{T \times 1}$ ,  $v = 1, 2, \dots, V^{(n)}$ , neg 表示相关系数为负.

在外界噪声干扰下, 针对 ROI 与 PCC 节点间的较强正相关可能变为较弱正相关这种情况, 图 1 给出了 2 种各层 IMF 分量与 PCC 节点之间的皮尔逊相关系数结果, 可以发现 (a) 图中正相关系数比负相关系数多, 以正相关为主, 那么可以认为与 PCC 节点表现为负相关的 IMF 分量削弱了该 ROI 与 PCC 节点间的正相关程度, 所以应优先选择表现为正相关的 IMF 分量 (黄色框标出) 对 BOLD 信号进行重构. (b) 图中负相关系数比正相关系数多, 以负相关为主, 所以应优先选择表现为负相关的 IMF 分量进行重构.

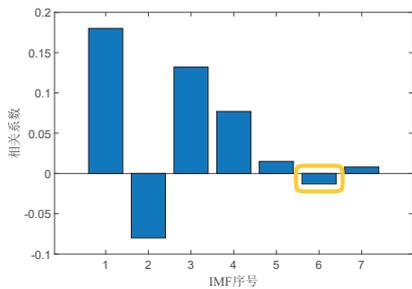


(a) 正相关为主情况图 1

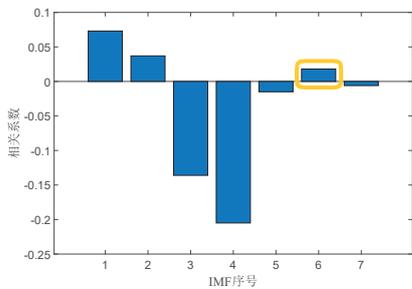


(b) 负相关为主情况图 1

图 1 皮尔逊相关系数结果示意图 1



(a) 正相关为主情况图 2



(b) 负相关为主情况图 2

图 2 皮尔逊相关系数结果示意图 2

同时还需考虑 IMF 分量与 PCC 节点由较强正相关变成了较弱负相关的情况,由图 2 可以发现,(a) 图中存在较弱的负相关(黄色框标出),这可能是由于噪声影响由较弱的正相关变为了较弱的负相关,所以该负相关对应的 IMF 分量应作为备选重构 IMF 分量.(b) 图中存在的较弱正相关同样可能是受到噪声的影响,应作为备选重构 IMF 分量.

根据上述思想,考虑以下 2 种情况对被试  $N$  个 ROI 的 BOLD 信号进行重构:

1) 正相关系数的个数大于等于负相关系数. 若正相关的 IMF 分量超过  $S$  个,则选择具有相关系数最大的  $S$  个正相关对应的 IMF 分量;若正相关的分

量不足  $S$  个,则选择所有具有正相关系数的 IMF 分量,并选择绝对值最小的负相关对应的 IMF 分量进行补齐不足  $S$  个的部分.可表示为:

$$\mathbf{X}_n = \begin{cases} \sum_{u=1}^S C_n^{(\text{pos},u)}, U^{(n)} \geq V^{(n)} \geq S \\ \sum_{u=1}^{U^{(n)}} C_n^{(\text{pos},u)} + \sum_{v=V^{(n)}+U^{(n)}-S}^{V^{(n)}} C_n^{(\text{neg},v)}, S > U^{(n)} \geq V^{(n)} \end{cases} \quad (4)$$

2) 负相关系数的个数大于正相关系数. 若负相关的 IMF 分量超过  $S$  个,则选择具有相关系数绝对值最大的  $S$  个负相关对应的 IMF 分量;若负相关的分量不足  $S$  个,则选择所有具有负相关系数的 IMF 分量,并选择绝对值最小的正相关对应的 IMF 分量进行补齐不足  $S$  个的部分.可表示为:

$$\mathbf{X}_n = \begin{cases} \sum_{v=1}^S C_n^{(\text{neg},v)}, V^{(n)} > U^{(n)} \geq S \\ \sum_{v=1}^{V^{(n)}} C_n^{(\text{neg},v)} + \sum_{u=V^{(n)}+U^{(n)}-S}^{U^{(n)}} C_n^{(\text{pos},u)}, S > V^{(n)} > U^{(n)} \end{cases} \quad (5)$$

## 4 实验设置

### 4.1 数据来源与预处理

为实现实验的可重复性,本文使用可公开下载的阿尔茨海默病神经影像库 (Alzheimer's Disease Neuroimaging Initiative, ADNI) 数据,网址为 <http://adni.loni.usc.edu/>,获取了 MCI 组与 NC 组分别 45 个被试的数据,详细数据参数见附录表 A1.

数据预处理使用数据处理和脑成像分析 (Data Processing & Analysis of Brain Imaging, DPABI) 工具箱<sup>[36]</sup>,使用自动解剖标记图谱 (Automated Anatomical Labeling, AAL)<sup>[37]</sup>提取 ROI,具体过程见附录表 A2.

### 4.2 BOLD 信号的重构与 FC 的构建

对数据预处理完成后,使用本文提出的 MEMD-PR 方法与极正极负重构准则对每个被试 116 个 ROI 的 BOLD 信号进行重构,并计算重构前后每个被试的 PCC 节点与各 ROI 间 BOLD 信号的皮尔逊相关系数,以此建立 PCC 节点与各脑区间的 FC,然后将该 FC 系数作为分类的特征输入.

由于 MEMD 是基于数据驱动的信号处理方法,无法事先确定每个被试得到的 IMF 数目,为了实现所有被试使用相同数目的 IMF 进行 BOLD 信号重构,需要分析 IMF 数目的选择策略.考虑到选择较少的 IMF 进行 BOLD 信号重构虽然有可能去除更多外界噪声,但也可能造成 BOLD 信号中关键信息的缺

失,故本文基于所有被试 IMF 的最小数目,选择比最小数目少一个的若干 IMF 进行重构.这样既能够去除部分噪声干扰,同时可以保留更多的有效信息.由于本文所有被试 IMF 的最小数目为 5,故选择 4 个 IMF 分量对所有 BOLD 信号进行重构.同时本文分别选择 3、2、1 个 IMF 进行重构设计对比,以验证 IMF 数目选择策略的合理性.

#### 4.3 基于 Kmeans 算法的传统重构准则

文献 [15] 提出了结合互相关系数与 Kmeans 算法的重构准则,对任务态 BOLD 信号进行重构.首先,计算各层 IMF 与模板间的互相关系数并取绝对值中的最大值,再对所有 ROI 的各层互相关系数计算均值.然后,基于 Kmeans 算法将所有互相关系数聚为高、中、低相关三组.最后,根据高相关组对应的 IMF 序号,对该被试所有 ROI 的 BOLD 信号进行重构.

#### 4.4 基于 SVM 的机器学习分类方法

为了验证方法的有效性,通常需要使用量化的指标来衡量,这里使用 MCI 组与 NC 组的分类准确率作为衡量指标.由于 fMRI 的相关研究通常基于少量的研究对象,且实验数据通常呈现非线性,所以本文使用在处理小样本、非线性等数据表现出许多优势的 SVM 来进行分类准确率的测试.SVM 在 MATLAB2021a 中实现.

本文使用 SVM 分别对基于极正极负重构准则与传统重构准则的重构组以及原始组进行分类,分类对象为 PCC 节点与各脑区间的 FC 系数.使用五折交叉验证策略分类 100 次,结果取均值.

#### 4.5 NCA 特征选择算法的应用

NCA 特征选择算法是一种非参数技术,属于监督学习,其通过优化正则化参数来使分类准确率达到最大化,并且在这一过程中实现数据降维.具体实验过程如下.

在分类实验五折交叉验证策略基础上,首先使用训练集寻找 NCA 特征选择的最优正则化参数.然后基于该参数在测试集中拟合 NCA 模型,得到所有特征输入的权重排名.再基于排名对整个数据集进行降维.最后对降维后的数据集进行 SVM 分类实验.对于 2 种重构准则组和原始组,以上整个流程重复 100 次,以降低测试的偶然性.

为了研究最优特征子集具有的生物学意义,本文分别对基于极正极负重构准则降噪前后的最优特征子集进行统计性分析.由于最优特征子集对应的 ROI 与 PCC 节点间可能具有特殊关联性,本文通过 AAL 模板将大脑划分成额叶、脑岛、边缘、枕叶、顶

叶、皮层、颞叶、小脑半球和小脑蚓体九个部分,再统计最优特征子集对应的 ROI 在九个脑区中的占比,以找出与 DMN 相关的大脑区域,将以上统计流程重复 100 次以避免实验的偶然性.

## 5 实验结果

### 5.1 MEMD-PR 方法对提升数据差异性的有效性

图 3 给出了每 10 次五折交叉验证取均值的 SVM 分类结果,可以发现经过 BOLD 信号重构的两组分类准确率曲线总体在原始组上方.将 100 次实验的所有分类结果计算均值,结果如表 1 中降维前分类准确率所示,相较于原始组,极正极负重构组的分类准确率由 55.1% 提升至 58.4%,提升了 3.3%,而 Kmeans 重构组提升了 2.1%.

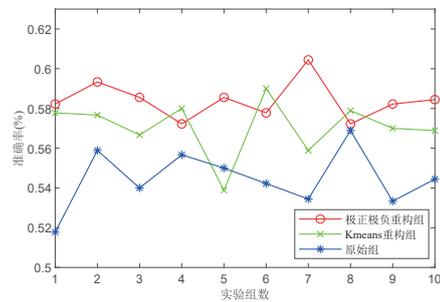


图 3 重构前后的 SVM 分类结果 2

结果表明经过 MEMD-PR 方法重构后的 BOLD 信号降低了外界噪音对分类性能造成的干扰,提升了 MCI 与 NC 数据集的差异性.此外,在分类性能提升的程度方面,本文提出的极正极负重构准则略高于传统基于 Kmeans 算法的重构准则,其原因可能是后者仅使用了高相关性的一至两个 IMF 对所有 ROI 进行统一重构,同时仅考虑了相关性的绝对值大小,这导致了 BOLD 信号中某些关键信息的缺失.

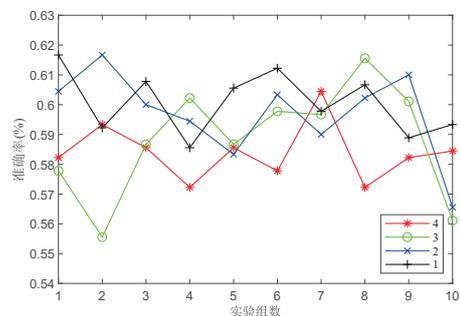


图 4 使用不同数目 IMF 进行重构的分类结果

为了验证本文中 IMF 数目选择策略的合理性,图 4 给出了分别使用 4、3、2、1 个 IMF 进行重构的分类结果,可以发现使用不同的 IMF 数目进行重构,分类准确率较原始组均有不同程度的提升,故 IMF 数目选择策略并不影响 MEMD-PR 方法提升数据差异性的有效性.为了降低外界噪声干扰的同时保留

更多的有效信息, 所以本文的 IMF 数目选择策略具有合理性.

## 5.2 MEMD-PR 方法对特征选择的有效性

由于 SVM 分类的特征输入为每个被试 PCC 节点与 116 个 ROI 间的 FC 系数, 本文使用 NCA 特征选择算法计算每个 FC 系数的特征权重并由大到小排序, 再选择部分权重最高的 FC 系数作为最优特征子集. 图 5 给出了原始组、极正极负重构组与 Kmeans 重构组对于不同特征子集数时的 SVM 分类结果, 每个值均为 100 次分类实验结果取均值. 特征子集选择数目为 1 时, 代表分类的特征输入为特征权重最大的 1 个 FC 系数, 特征子集选择数目为 116 时, 代表分类的特征输入为所有 FC 系数.

实验结果表明, 对于分类特征输入的 116 个 FC 系数, 其中既有重要特征也有冗余特征, 且过多或者过少的特征输入都会影响分类性能. 三组的最优特征子集分类性能相关结果如表 1 所示.

以上结果表明 NCA 特征选择算法运用于三组

的性能存在巨大差异, 极正极负重构组的分类准确率提升程度与数据集降维程度分别达到了原始组的 6 倍与 12 倍, 也优于 Kmeans 重构组. 一方面, 静息态 BOLD 信号中的外界噪音会对 NCA 特征选择算法性能造成影响, 采用本文提出的 MEMD-PR 方法对 BOLD 信号进行降噪后, 有利于 NCA 特征选择算法更好地提取特征, 同时对分类的特征输入进行降维. 另一方面, Kmeans 重构组的特征选择性能弱于极正极负重构组, 其原因可能是基于 Kmeans 算法的重构准则导致 BOLD 信号关键信息缺失引起的.

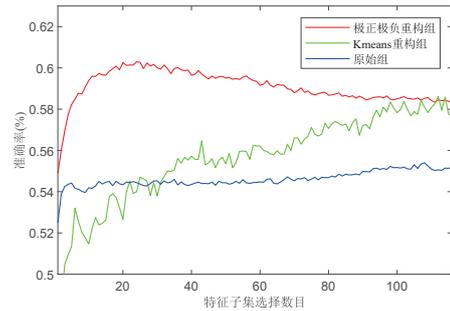


图 5 最优特征子集 SVM 分类性能

表 1 最优特征子集分类性能相关结果

组名	降维前特征子集数目	降维后特征子集数目	降维程度/%	降维前分类准确率/%	降维后最优分类准确率/%	准确率提升/%
原始组	116	108	6.9	55.1	55.4	0.3
Kmeans 重构组	116	112	3.4	57.2	58.6	1.4
极正极负重构组	116	20	82.8	58.4	60.2	1.8

## 5.3 最优特征子集的统计特性

由 5.2 节的结果可知极正极负重构组降维后最优分类准确率对应的特征子集数目为 20, 表明最优特征子集对应的 20 个 ROI 表现出了最强的差异性, 若选择这些 ROI 进行统计分析更利于寻找降噪前后数据统计特性的差异. 考虑到九个脑区分别对应的 ROI 个数不同, 若进行占比统计则需对指标进行标准化处理, 这里指标  $Index$  的表达式为

$$Index = \frac{K_g K_H}{K_G K_h} \quad (6)$$

其中  $K_g$  为该脑区中 FC 系数对应的 ROI 出现次数,  $K_G$  为最优分类准确率对应的特征子集数目,  $K_h$  为该脑区对应的 ROI 数目,  $K_H$  为 ROI 总数  $N$ .

由于统计分析实验中只基于本文提出的极正极负重构准则进行重构, 为了简化描述, 这部分中的重构组均指极正极负重构组. 图 6 给出了原始组与重构组的统计结果, 图中小脑代表小脑半球, 蚓体代表小脑蚓体.

根据原始组的统计结果可以发现, 额叶、枕叶、边缘、颞叶与顶叶的占比较高, 表明了这些区域可能与 DMN 较为相关, 这个结论与以往的研究结果也较

为相似<sup>[21-23]</sup>, 这体现本文的统计结果是可信的.

重构组的统计结果与原始组相比体现出一定的差异性, 这里我们将重构组的统计指标与原始组作差, 各脑区统计结果差异如附录中图 A2-A3 所示, 可以直观地发现, 相较于原始组, 重构组中额叶、边缘、枕叶、顶叶、皮层、颞叶与小脑半球占比差异不明显, 而脑岛、小脑蚓体占比有较为显著的提升, 这表明脑岛与小脑蚓体可能也与 PCC 节点联系较为紧密, 可能也是 DMN 的相关脑区, 这是以往 DMN 研究中少有提出的结论. 研究表明脑岛的激活与复杂的言语记忆任务相关<sup>[38]</sup>, 由于大脑静息态时 DMN 会出现许多自发性活动, 记忆任务正是其中一部分, 这与本文得到的脑岛可能为 DMN 相关脑区的结论相吻合. 虽然小脑蚓体通常与运动调节相关, 但研究表明老年被试的小脑与脑岛间的 FC 加强, 其构成的感知运动网络可以弥补老年被试的脑功能认知衰退<sup>[39]</sup>, 而小脑蚓体占比的提升, 这可能意味着小脑蚓体与 PCC 节点间构成了某种具有认知功能补偿作用的网络.

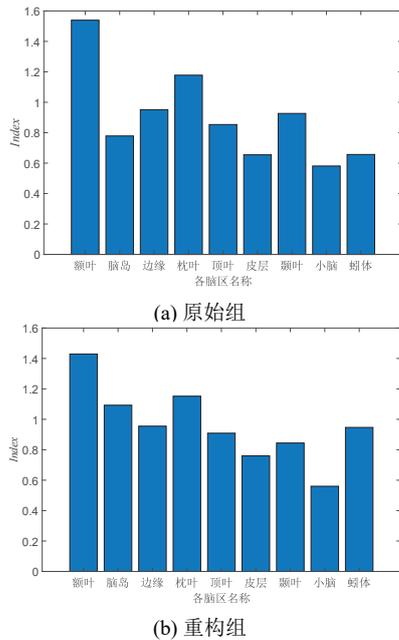


图 6 原始组与重构组的统计结果

由于通过 MEMD-PR 方法得到的重构组经过统计性分析可以获得一些原始组中未能发现的结论, 这表明外界噪声造成的干扰可能会影响具有生物学意义的发现过程. 这再一次说明本文提出的 MEMD-PR 方法降低了外界噪声的干扰, 有助于发现新的生物学结论.

## 6 结论与展望

已被证实, 大脑静息态 BOLD 信号作为 rs-fMRI 相关研究的核心存在各种外界噪声, 可能会影响信号的质量导致错误的研究结果. 针对上述问题, 本文提出了 MEMD-PR 方法与极正极负重构准则, 以默认模式网络中心节点 PCC 为模板, 通过对 BOLD 信号进行重构以实现降噪, 并设计实验验证其有效性. 下一步工作我们将尝试对应用于 rs-fMRI 研究的重构准则进行优化, 进一步研究 IMF 数目的选择策略, 如降低外界噪声干扰与保留有效信息的平衡性, 针对不同被试选择不同数目的 IMF 进行重构, 而对静息态 BOLD 信号进行重构的依据也不仅仅局限于 DMN 的中心节点 PCC, 同时还应考虑 DMN 的次要节点.

### 参考文献 (References)

- [1] Albert M S, DeKosky S T, Dickson D, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease[J]. *Alzheimer's & dementia*, 2011, 7(3): 270-279.
- [2] Petersen R C. Mild cognitive impairment: transition between aging and Alzheimer's disease[J]. *Neurología* (Barcelona, Spain), 2000, 15(3): 93-101.
- [3] Gauthier S, Reisberg B, Zaudig M, et al. Mild cognitive impairment[J]. *The lancet*, 2006, 367(9518): 1262-1270.
- [4] Binnewijzend M A A, Schoonheim M M, Sanz-Arigita E, et al. Resting-state fMRI changes in Alzheimer's disease and mild cognitive impairment[J]. *Neurobiology of aging*, 2012, 33(9): 2018-2028.
- [5] Caballero-Gaudes C, Reynolds R C. Methods for cleaning the BOLD fMRI signal[J]. *Neuroimage*, 2017, 154: 128-149.
- [6] Murphy K, Birn R M, Bandettini P A. Resting-state fMRI confounds and cleanup[J]. *Neuroimage*, 2013, 80: 349-359.
- [7] Chang C, Cunningham J P, Glover G H. Influence of heart rate on the BOLD signal: the cardiac response function[J]. *Neuroimage*, 2009, 44(3): 857-869.
- [8] Birn R M, Diamond J B, Smith M A, et al. Separating respiratory-variation-related fluctuations from neuronal-activity-related fluctuations in fMRI[J]. *Neuroimage*, 2006, 31(4): 1536-1548.
- [9] Wise R G, Ide K, Poulin M J, et al. Resting fluctuations in arterial carbon dioxide induce significant low frequency variations in BOLD signal[J]. *Neuroimage*, 2004, 21(4): 1652-1664.
- [10] McKeown M J, Makeig S, Brown G G, et al. Analysis of fMRI data by blind separation into independent spatial components[J]. *Human brain mapping*, 1998, 6(3): 160-188.
- [11] Erhardt E B, Rachakonda S, Bedrick E J, et al. Comparison of multi-subject ICA methods for analysis of fMRI data[J]. *Human brain mapping*, 2011, 32(12): 2075-2095.
- [12] Windischberger C, Barth M, Lamm C, et al. Fuzzy cluster analysis of high-field functional MRI data[J]. *Artificial Intelligence in Medicine*, 2003, 29(3): 203-223.
- [13] Li K, Guo L, Nie J, et al. Review of methods for functional brain connectivity detection using fMRI[J]. *Computerized medical imaging and graphics*, 2009, 33(2): 131-139.
- [14] Ur Rehman N, Mandic D P. Filter bank property of multivariate empirical mode decomposition[J]. *IEEE transactions on signal processing*, 2011, 59(5): 2421-2426.
- [15] Wang K W, Chen C Y, Chang H H, et al. A multivariate empirical mode decomposition-based data-driven approach for extracting task-dependent hemodynamic responses in olfactory-induced fMRI[J]. *IEEE Access*, 2019, 7: 15375-15388.
- [16] Raichle M E, MacLeod A M, Snyder A Z, et al. A default mode of brain function[J]. *Proceedings of the National Academy of Sciences*, 2001, 98(2): 676-682.
- [17] Toro R, Fox P T, Paus T. Functional coactivation map of the human brain[J]. *Cerebral cortex*, 2008, 18(11): 2553-2559.
- [18] Goldberger J, Hinton G E, Roweis S, et al.

- Neighbourhood components analysis[J]. *Advances in neural information processing systems*, 2004, 17.
- [19] Beckmann C F, DeLuca M, Devlin J T, et al. Investigations into resting-state connectivity using independent component analysis[J]. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 2005, 360(1457): 1001-1013.
- [20] Mayer J S, Roebroeck A, Maurer K, et al. Specialization in the default mode: Task-induced brain deactivations dissociate between visual working memory and attention[J]. *Human brain mapping*, 2010, 31(1): 126-139.
- [21] Buckner R L, Andrews-Hanna J R, Schacter D L. The brain's default network: anatomy, function, and relevance to disease[J]. *Ann NY Acad Sci*, 2008, 1124(1): 1-38.
- [22] Greicius M D, Krasnow B, Reiss A L, et al. Functional connectivity in the resting brain: a network analysis of the default mode hypothesis[J]. *Proc Nat Acad Sci*, 2003, 100(1): 253-258.
- [23] Raichle M E, Snyder A Z. A default mode of brain function: A brief history of an evolving idea[J]. *Neuro Image*, 2007, 37: 1083-1090.
- [24] Shehzad Z, Kelly A M C, Reiss P T, et al. The resting brain: unconstrained yet reliable[J]. *Cerebral cortex*, 2009, 19(10): 2209-2229.
- [25] Li X L, Adali T. Blind spatiotemporal separation of second and/or higher-order correlated sources by entropy rate minimization[C]//2010 IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, 2010: 1934-1937.
- [26] Li X L, Adali T. Independent component analysis by entropy bound minimization[J]. *IEEE Transactions on Signal Processing*, 2010, 58(10): 5151-5164.
- [27] Long Q, Bhinge S, Levin-Schwartz Y, et al. A graph theoretical approach for performance comparison of ICA for fMRI analysis[C]//2017 51st Annual Conference on Information Sciences and Systems (CISS). IEEE, 2017: 1-6.
- [28] Abraham A, Dohmatob E, Thirion B, et al. Extracting brain regions from rest fMRI with total-variation constrained dictionary learning[C]//International conference on medical image computing and computer-assisted intervention. Springer, Berlin, Heidelberg, 2013: 607-615.
- [29] Long Q, Bhinge S, Levin-Schwartz Y, et al. The role of diversity in data-driven analysis of multi-subject fMRI data: Comparison of approaches based on independence and sparsity using global performance metrics[J]. *Human brain mapping*, 2019, 40(2): 489-504.
- [30] 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]. *Proceedings of the Royal Society of London. Series A: mathematical, physical and engineering sciences*, 1998, 454(1971): 903-995.
- [31] McGonigle J, Mirmehdi M, Malizia A L. Empirical Mode Decomposition in data-driven fMRI analysis[C]//2010 First Workshop on Brain Decoding: Pattern Recognition Challenges in Neuroimaging. IEEE, 2010: 25-28.
- [32] 孔伶俐, 吴海锋, 曾玉, 陆小玲. 使用深度学习和不同频率维度的脑功能性连接对轻微认知障碍的诊断 [J]. *计算机应用*, 2021, 41(2): 590-597.  
(Kong L X, Wu H F, Zeng Y, Lu X L. Diagnosis of mild cognitive impairment using deep learning and brain functional connectivity with different frequency dimensions[J]. *Computer applications*, 2021, 41 (2): 590-597.)
- [33] Wu H, Kong L, Zeng Y, et al. Resting-state Brain Connectivity via Multivariate EMD in Mild Cognitive Impairment[J]. *IEEE Transactions on Cognitive and Developmental Systems*, 2021.
- [34] 武哲, 杨绍普, 刘永强. 基于多元经验模态分解的旋转机械早期故障诊断方法 [J]. *仪器仪表学报*, 2016, 37(2): 241-248.  
(Wu Z, Yang S P, Liu Y Q. Early fault diagnosis method of rotating machinery based on multiple empirical mode decomposition[J]. *Journal of instrumentation*, 2016, 37 (2): 241-248.)
- [35] Chang C, Glover G H. Time-frequency dynamics of resting-state brain connectivity measured with fMRI[J]. *Neuroimage*, 2010, 50(1): 81-98.
- [36] Yan C G, Wang X D, Zuo X N, et al. DPABI: data processing & analysis for (resting-state) brain imaging[J]. *Neuroinformatics*, 2016, 14(3): 339-351.
- [37] Tzourio-Mazoyer N, Landeau B, Papathanassiou D, et al. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain[J]. *Neuroimage*, 2002, 15(1): 273-289.
- [38] Stowe L A, Broere C A J, Paans A M J, et al. Localizing components of a complex task: sentence processing and working memory[J]. *Neuroreport*, 1998, 9(13): 2995-2999.
- [39] 曹卫芳. 基于功能磁共振成像技术的认知老化脑机制研究 [D]. 电子科技大学, 2018.  
(Cao W F. Brain mechanism of cognitive aging based on functional magnetic resonance imaging[D]. University of Electronic Science and technology, 2018.)

## 作者简介

汪瀚(1996—), 男, 硕士生, 从事生物医学信号处理、机器学习的研究, E-mail: wh961017@163.com;

吴海锋(1977—), 男, 教授, 硕士生导师, 从事深度学习、生物医学信号处理的研究, E-mail: whf5469@gmail.com;

王俨(1996—), 男, 硕士生, 从事信号处理、智能优化算法的研究, E-mail: 793407577@qq.com;

王勇(1999—), 男, 硕士生, 从事生物医学信号处理、机器学习的研究, E-mail: 1403788428@qq.com;

王霞(1985—), 女, 讲师, 硕士生导师, 从事智能优化算法、信号分析与处理的研究, E-mail: wangxiacsu@163.com.