

## 轻度认知障碍分类中全脑功能连接的特征压缩分析

马佳<sup>1</sup>, 吴海锋<sup>1,2</sup>, 李顺良<sup>1</sup>

(1. 云南民族大学电气信息工程学院, 昆明 650500; 2. 云南省无人自主系统重点实验室, 昆明 650500)

**摘要:** 利用静息态功能磁共振成像技术获取脑区的功能连接(Functional connection, FC)被广泛地应用于轻度认知功能障碍(Mild cognitive impairment, MCI)的分类研究中。然而,全脑FC用于分类通常存在信息冗余和特征维度灾难问题,为此,提出一种“G-Lasso + 特征压缩”的新方法来解决以上问题。首先,利用盲源分离技术获得全脑功能脑区的活跃信号时间序列,采用G-Lasso构建FC稀疏网络;其次,计算MCI患者、正常被试及所有被试在组平均上的稀疏FC,并结合欧氏距离进行簇Class 1~Class 3中心判决,获取簇间差异特征信息;最后,将每个被试的稀疏FC表达为簇中心的线性组合,获取压缩的FC作为关键特征完成分类。实验采用公开的数据库测试本文方法,结果表明,所提方法进行Class判决后获得簇间特征具有显著差异且提供了有效的标志信息,进一步压缩获取关键特征的分类准确率(89.8%)比仅使用稀疏方法提高了5%~10%。该结果表明,为了解决全脑FC存在的问题,需要考虑到特征选择和降维,但有诸多不确定因素信息,可以适当地将“稀疏+压缩”进行结合。

**关键词:** 功能连接;稀疏网络;轻度认知障碍;特征压缩;分类

**中图分类号:** TP301      **文献标志码:** A

## Feature Compression Analysis of Whole-Brain Functional Connection in Classification of Mild Cognitive Impairment

MA Jia<sup>1</sup>, WU Haifeng<sup>1,2</sup>, LI Shunliang<sup>1</sup>

(1. School of Electrical and Information Engineering, Yunnan Minzu University, Kunming 650500, China; 2. Yunnan Key Laboratory of Unmanned Autonomous System, Kunming 650500, China)

**Abstract:** The use of resting-state functional magnetic resonance imaging technology to obtain functional connection (FC) of brain regions is widely used in classification studies of mild cognitive impairment (MCI). However, the classification of whole-brain FC usually has the problems of information redundancy and feature dimension disaster. Therefore, a new method of “G-Lasso + feature compression” is proposed to solve the above problems. Firstly, the blind source separation technology is used to obtain the active signal time series of the whole brain functional brain region, and the FC sparse network is constructed by G-Lasso. Secondly, the sparse FC of MCI, normal subjects and all subjects on the group average is calculated, and the cluster Class 1—Class 3 center decision is performed in combination with the Euclidean distance to obtain the difference feature information between clusters. Finally, the sparse FC of each participant is expressed as a linear combination of the cluster center, and the compressed FC is obtained as the key feature to complete the classification. The results show that the proposed method obtains significant

differences in inter-cluster features after Class decision and provides effective sign information. The classification accuracy of the key features obtained by further compressing (89.8%) is 5%—10% higher than that of the sparse method alone. The results show that in order to solve the problems of whole-brain FC, feature selection and dimensionality reduction need to be considered, but there are many uncertain factors, and “sparse + compression” can be appropriately combined.

**Key words:** functional connection; sparse network; mild cognitive impairment (MCI); feature compression; classification

## 引 言

阿尔茨海默病(Alzheimer's disease, AD)是一种神经退行性疾病,不可逆转,目前尚无有效药物和治疗手段,只有及早干预才能延缓其发展。由于轻度认知功能障碍(Mild cognitive impairment, MCI)被认为是AD的必经阶段<sup>[1]</sup>,因此准确诊断MCI患者对早期干预和延缓AD具有重要意义。

近年来,静息状态功能磁共振成像(resting-state functional magnetic resonance imaging, rs-fMRI)技术在认知科学和神经疾病研究中起到了重要作用,由其所获得的血氧水平依赖(Blood oxygenation level dependent, BOLD)信号可反映结构脑区的自发性神经波动<sup>[2]</sup>。另外,对rs-fMRI的体素结构进行盲源分离还可获得活跃脑区的时间进程(Time course, TC)信号,其可反映功能性脑区的自发性神经波动<sup>[3]</sup>。无论BOLD信号还是TC信号,均可使用相关性算法<sup>[4]</sup>来获得脑区间功能性连接(Functional connection, FC)。FC表达了不同结构或功能脑区间的相互作用模式,而MCI已被认为是一种脑区FC丧失症<sup>[5]</sup>,因此研究FC进一步寻找MCI生物标记物将为临床脑疾病的诊断和评估提供新的视角。

通常,FC的计算是在全脑区域下进行,需计算多个脑区信号间的相关系数,因此所获得的FC也是多个结构脑区或功能脑区的结果。然而,全脑FC用于MCI与正常控制(Normal control, NC)组间差异分析或分类时,至少存在以下问题:(1)信息冗余,冗余的FC并不能展现疾病组与正常组的差异;(2)维度灾难,多个FC作为分类器的输入使得特征维度增加,致使分类性能下降。特征选择是解决以上问题的一种常用方法<sup>[6-7]</sup>,其目标是通过代价函数或计算模型得到相应特征对分类器的贡献程度,保留贡献较高的特征,剔除贡献较低的特征,但是特征选择需要解决的一个核心问题是选择的特征数目,保留过多的特征数会带来过拟合,而较少的特征又会带来欠拟合。另外一种解决全脑FC的方法是降维,它能有效解决维度灾难。常用的降维方法有主成分分析和独立成分分析<sup>[8]</sup>,但该方法依赖于各分量间的相关程度。还有一种用于全脑FC降维方法的是无监督聚类<sup>[9]</sup>,它将多个被试者的全脑FC聚类为有限的几类模式,而每个被试的全脑FC就是这几类模式的表达。聚类可以使得全脑FC的维度得到急剧下降,但由于采用无监督学习,结果会受聚类数目和判决准则等参数的影响导致不稳定<sup>[10]</sup>,且聚类结果所表征的生物特性也得不到合理的解释<sup>[11]</sup>。

除了以上解决全脑FC问题的方法外,还有一种重要方法是稀疏表达<sup>[12]</sup>,它引入正则化项,使得全脑FC的数目降低,从而起到去除冗余信息和降低维度的作用。现有研究中,采用稀疏线性回归模型,将特定功能的脑区信号表示为其他脑区的线性组合,求解稀疏解的同时过滤掉不重要或虚假的连接<sup>[12]</sup>。特别在疾病辅助诊断研究中,获取全脑有效的FC模式以及去除冗余信息后能提高最终的性能<sup>[13-14]</sup>。例如, Lee等<sup>[13]</sup>采用正则化的最小绝对收缩和选择算子(Least absolute shrinkage and selection operator, Lasso)来构建稀疏的FC网络,对自闭症的分类性能有所提升; Jie等<sup>[14]</sup>引入超图学习方法,结合稀疏线性回归模型构建功能连接超网络模型,以此得到了脑疾病的生物标记物;李瑶等<sup>[15]</sup>提出一种Sparse group Lasso方法来构建稀疏组间FC超网络,并对抑郁症患者进行分类研究也得到了较好

的分类表现;Liang等<sup>[16]</sup>采用图形Lasso(Graphical Lasso,G-Lasso)模型通过一对一场景进行训练,从而学习精神分裂症和抑郁症患者神经认知特征的条件独立结构。以上研究均基于Lasso方法来得到稀疏FC,能在一定程度上满足全脑FC分析的有效性,但也存在局限性。从目前的研究结果看,单纯地使用稀疏表达在解决全脑FC上并不一定达到最好的结果,其原因在于Lasso的稀疏表达更接近嵌入式的特征选择方法<sup>[6]</sup>,仅能去除冗余特征,但不能大规模降低特征维度。因此,稀疏表达并不能更好地解决特征维度灾难问题。

针对以上全脑FC问题,本文提出了一种“G-Lasso+特征压缩”的新方法来对MCI患者进行分类预测。首先,利用盲源分离技术中独立成分分析(Independent component analysis,ICA)获得功能脑区的TC序列信号,利用G-Lasso模型获得稀疏FC;然后,计算组平均上的稀疏FC,并根据分组集合定义了簇Class 1~Class 3,用来表征疾病态和健康态等标签信息,结合欧氏距离判决得到簇间差异特征;再次,采用最小二乘模型将被试的全脑稀疏FC表达为簇Class中心的线性相关,压缩全脑FC特征;最后,将压缩的FC关键特征输入至分类器中完成分类。本文的主要贡献如下:(1)提出了稀疏结合特征压缩的新方法,可有效地解决全脑FC的信息冗余和特征维度灾难问题,避免传统特征选择数不当带来的过拟合,而且在不考虑噪声影响下,理论上全脑原FC可被无损恢复;(2)提出一种特征压缩分析方法,既可获取簇间差异性特征来分析患者与健康组间的连接差异,又可将小样本多特征的数据急剧降至为二维或三维。实验中,采用公开数据阿尔茨海默病神经影像学(Alzheimer's disease neuroimaging initiative,ADNI)数据库测试了本文算法。实验结果表明,本文所提方法能获得具有显著差异的簇间特征,并提供了生物标志物信息;其分类准确率最佳能达到89.8%,比仅使用Lasso、G-Lasso及特征选择的分类准确率提高了5%~12%,比G-Lasso结合特征降维方法和其他相关研究都有明显的提高。

## 1 相关工作

分析大脑特定区域之间的FC,可以评估特定脑区的功能协同作用和连接模式,采用稀疏表示模型可过滤掉不重要或虚假的连接信息,且能反映多个脑区之间信息交互,是表征脑区生物标志物的有效方式<sup>[17]</sup>。

通常情况下,计算FC之前会将大脑分割为许多感兴趣解剖区域(Regions of interest, ROIs)<sup>[2]</sup>,一般使用先验模板(比如AAL、Brodmann等)<sup>[18-19]</sup>和数据驱动(如ICA)等方式<sup>[4]</sup>来获取ROIs。使用先验模板需要具有一定的相关经验,而基于盲源分离的ICA技术在捕获活跃的ROIs信号时可以避免先验知识不足所带来的影响。对脑图像进行ICA处理,可以得到相应ROIs的TCs序列信号和空间分布图<sup>[3]</sup>。通过计算TCs序列信号之间的相关性得到相关系数矩阵,用于反映在时间上各ROI之间关联程度的FC模式<sup>[9]</sup>。这种连接模式所建成的网络模型通常是全连接的,也就是网络中节点(ROIs)和连接边的关系(相关性)代表所有两两脑区间的直接交互作用<sup>[17]</sup>。然而,针对全脑FC网络分析,探索大脑在不同状态下的连接模式<sup>[20]</sup>、脑功能异常以及与脑疾病相关的特征时,将会面临信息冗余和维度灾难等相关问题。此外,传统网络结构及其复杂且可能存在虚假或无关紧要的连接,尤其在动态FC脑网络的构建中<sup>[21]</sup>。因此,在机器学习分类研究中常使用降维方法来解决全脑FC的多维问题,以及使用特征选择和特征提取等方法来获取具有代表意义的特征子集<sup>[22-23]</sup>,为临床脑疾病的诊断和评估提高准确率。例如,Guo等<sup>[22]</sup>在FC网络分析中,定义了不同聚类系数提取特征来描述功能脑网络拓扑,并选择具有显著差异的特征进行分类研究。然而在特征选择时需要相关的算法来确定特征数目和特征值。复杂的特征选择算法更有效,但会提高计算的复杂度,反之,简单的算法计算复杂度低,但会导致选择的特征有效性和可靠性降低。

将稀疏模型引入到全脑FC网络分析中,通过特定功能ROIs间的线性组合来产生稀疏解<sup>[24-27]</sup>,去

除无关的连接信息从而起到特征选择的作用,同时又能降低维度。Xue等<sup>[26]</sup>基于稀疏方法,以低阶和高阶的方式将空间信息组合到功能脑网络估计模型中,然而确定空间距离是一个困难的问题。Du等<sup>[27]</sup>提出基于稀疏和低秩约束来估计具有模块化结构的功能脑网络,但该方法只专注于模块内连接,而不强调模块间连接。因此,引入图模型的FC稀疏网络将克服传统脑网络的缺点,有助于提取生物标志物来检测神经系统疾病患者与健康对照组间的连接差异和脑疾病诊断。G-Lasso方法仍然基于线性回归模型来建立稀疏网络,为了识别大脑区域间存在的有效FC,使用了嵌套留一法交叉验证方法来估计 $\ell_1$ 范数<sup>[25]</sup>。G-Lasso方法最早由Friedman等<sup>[28]</sup>提出,该方法以块坐标下降算法为出发点,其目标是通过最小化一个正则化的负对数似然函数来估计稀疏的逆协方差矩阵,矩阵中的非零元素表示变量存在,零元素表示变量条件独立,且通过 $\ell_1$ 范数控制该矩阵的稀疏性。

然而,在rs-fMRI这样高维数据中,脑区间功能协同作用复杂多变<sup>[29]</sup>,对应的全脑FC特征也将会较多且存在冗余,虽然G-Lasso方法能去除部分冗余信息而得到有效连接模式,但其特征维度并不会大规模降低,这对小样本数据提取差异特征进行分类任务是比较困难的。目前,聚类算法被作为降维方式在脑疾病的FC网络分析中得到广泛应用<sup>[9,14,17]</sup>。聚类采用无监督学习,根据数据本身特性易于成簇,即同质FC划分为一簇,增强簇间差异表征能力使个体特征更易体现。但是,聚类方式与多个因数有关,如预设聚类数、初始聚类中心选择以及距离计算方法等,由于不同参数对应不同的结果,导致结果具有多样性。另外,如何客观地测试特征选择和降维方法的性能也是一个问题。如前所述,特征大小和维度也许无法提前预知,因此使用传统机器学习的分类算法所得到的分类准确率将存在一定偏差。

## 2 方法与内容

### 2.1 方法总体框架

针对全脑FC存在的问题,本文提出“G-Lasso+特征压缩”方法来构建FC稀疏网络并获取相应的特征对MCI进行分类研究,包括以下步骤:数据采集与预处理、基于G-Lasso构建FC稀疏网络模型、判决簇Class中心和特征表达来完成特征压缩以提取出关键特征、采用支持向量机(Support vector machine, SVM)进行分类预测。图1给出了本文的总体框架流程图。

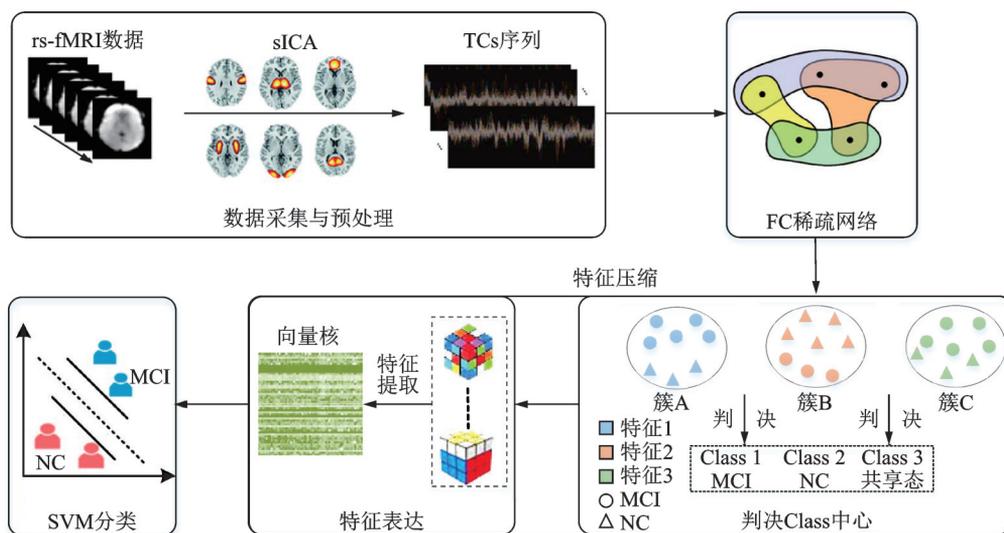


图1 本文方法的总体框架

Fig.1 Overall framework of the proposed method

## 2.2 数据采集和预处理

对医学图像数据进行处理、分析能降低信息不确定性,提高临床辅助诊断效率<sup>[8]</sup>,为此本文采用可公开下载的ADNI库(<http://adni.loni.usc.edu/>)中rs-fMRI数据对MCI患者和NC被试进行研究。该数据库提供的rs-fMRI是在3T功能磁共振扫描仪下让被试保持静息状态扫描得到,即被试全程闭眼、不做任务、不考虑任何事情以及保持清醒的状态<sup>[2]</sup>。

rs-fMRI数据是4D(三维空间体素和一维时间)模型,首先需要进行预处理操作得到低维的特征信息。典型的预处理技术主要有图像配准、空间位置校正、图像重采样和归一化等<sup>[30]</sup>。图像预处理后需要进一步进行ICA处理来获取大脑活跃的脑区信号<sup>[3]</sup>。ICA是一种盲源分离的方法,在脑疾病研究中对rs-fMRI的数据信号识别有良好的效果<sup>[4,9,31]</sup>,而因为rs-fMRI数据的空间维度(体素结构)要远大于时间维度(扫描的时间点),因此使用空间ICA(Space ICA, sICA)牺牲时间来获得空间独立<sup>[31]</sup>。此外, MCI作为退行性神经系统疾病,其临床表现是微妙的,大脑在一个大规模、高度连接的网络中受损,而不是在一个孤立的区域。鉴于此,应用sICA技术来获取脑网络相关的脑区TC序列信号。

## 2.3 G-Lasso求解稀疏线性回归模型

利用TC序列信号构建脑区FC稀疏网络模型,网络结构中的节点代表脑区,边代表特定脑区信号与多个脑区间连接关系。若一被试获取了 $M$ 个TCs序列信号,可以定义为 $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_m, \dots, \mathbf{X}_M]^T \in \mathbf{R}^{d \times M}$ ,  $\mathbf{X}_m$ 代表第 $m$ 个TC序列,  $d$ 表示时间长度。每个TC(如 $\mathbf{X}_m$ )序列信号可以作为一个脑区的响应向量,且能通过其他 $M-1$ 个TCs序列信号的线性组合表示为

$$\mathbf{X}_m = \mathbf{A}_m \boldsymbol{\alpha}_m + \boldsymbol{\tau}_m \quad m = 1, 2, \dots, M \quad (1)$$

式中:  $\mathbf{A}_m = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_{m-1}, \mathbf{0}, \mathbf{X}_{m+1}, \dots, \mathbf{X}_M]$ 表示除了第 $m$ 个TC序列外的其他所有序列的数据矩阵,同时该数据矩阵中第 $m$ 个TC序列对应的值设为0;  $\boldsymbol{\alpha}_m$ 表示权重向量,即其他脑区对第 $m$ 个脑区的影响程度;  $\boldsymbol{\tau}_m$ 表示噪声项。因此 $\boldsymbol{\alpha}_m$ 中非零元素代表了与所对应的脑区存在边,即有连接关系;零元素则表征脑区间无FC关系。

考虑到任意一被试的TC序列 $\mathbf{X}_{d \times M}$ ,它是服从高斯分布,且具有零均值( $M$ 个TCs序列信号均值为0)和正定协方差矩阵 $\boldsymbol{\Sigma}$ (脑区之间存在相关性)。G-Lasso方法的目标是通过最小化一个 $\ell_1$ 正则化的负对数似然函数,来估计具有稀疏性的逆协方差矩阵 $\boldsymbol{\Theta} = \boldsymbol{\Sigma}^{-1}$ ,这里 $\boldsymbol{\Theta}$ 被称为精度矩阵<sup>[28]</sup>,它表示了相关性信息。G-Lasso中最小化 $\ell_1$ 正则化的负对数似然函数为

$$\min_{\boldsymbol{\Theta} > 0} f(\boldsymbol{\Theta}): = -\ln \det(\boldsymbol{\Theta}) + \text{tr}(\mathbf{S}\boldsymbol{\Theta}) + \lambda \|\boldsymbol{\Theta}\|_1 \quad (2)$$

式中:  $\det$ 表示行列式,  $\text{tr}$ 表示迹,  $\mathbf{S} = \frac{\mathbf{X}^T \mathbf{X}}{M}$ 为样本协方差矩阵,  $\|\boldsymbol{\Theta}\|_1$ 表示 $\boldsymbol{\Theta}$ 绝对值的总和。  $\lambda > 0$ 为控制 $\ell_1$ 收缩量的调节参数,当 $\lambda$ 取值较大时,估计的 $\boldsymbol{\Theta}$ 会得到较多的零元素,即代表FC网络越稀疏;反之,FC网络越密集。

使用次梯度符号表示法,式(2)的最优条件(也称为“正态方程”)可表示为

$$-\boldsymbol{\Theta}^{-1} + \mathbf{S} + \lambda \boldsymbol{\Gamma} = \mathbf{0} \quad (3)$$

式中 $\boldsymbol{\Gamma}$ 为 $\boldsymbol{\Theta}$ 的分量符号矩阵,即

$$\begin{cases} \gamma_{jk} = \text{sign}(\theta_{jk}) & \theta_{jk} \neq 0 \\ \gamma_{jk} \in [-1, 1] & \theta_{jk} = 0 \end{cases} \quad (4)$$

这里使用 $\gamma_{jk} \in \text{sign}(\theta_{jk})$ 符号来表示 $\theta_{jk}$ 的符号函数,即相关性的正、负关系。由于式(3)的全局稳定条件要求 $\theta_{ij}$ 为正,这意味着

$$\omega_{ii} = s_{ii} + \lambda \quad i = 1, 2, \dots, m \quad (5)$$

式中估计的协方差矩阵  $\mathbf{W} = \boldsymbol{\Theta}^{-1}$ 。

G-Lasso 使用块坐标求解式(3)。考虑到  $\boldsymbol{\Theta}$  和  $\boldsymbol{\Gamma}$  的分区

$$\boldsymbol{\Theta} = \begin{pmatrix} \boldsymbol{\Theta}_{11} & \boldsymbol{\theta}_{12} \\ \boldsymbol{\theta}_{21} & \theta_{22} \end{pmatrix}, \boldsymbol{\Gamma} = \begin{pmatrix} \boldsymbol{\Gamma}_{11} & \boldsymbol{\gamma}_{12} \\ \boldsymbol{\gamma}_{21} & \gamma_{22} \end{pmatrix} \quad (6)$$

式中:  $\boldsymbol{\Theta}_{11}$  为  $(d-1) \times (d-1)$  的矩阵,  $\boldsymbol{\theta}_{12}$  为  $(d-1) \times 1$  的向量,  $\theta_{22}$  为一标量,  $\boldsymbol{\Gamma}_{11} \in \text{sign}(\boldsymbol{\Theta}_{11})$  为  $(m-1) \times (m-1)$  的矩阵,  $\boldsymbol{\gamma}_{12} \in \text{sign}(\boldsymbol{\theta}_{12})$  是  $(m-1) \times 1$  的列向量,  $\boldsymbol{\gamma}_{22} \in \text{sign}(\theta_{22})$  为一标量。使用块分区矩阵的逆性质,分区方式同式(6),  $\mathbf{W}$  和  $\mathbf{S}$  可分别写为

$$\mathbf{W} = \begin{pmatrix} \mathbf{W}_{11} & \boldsymbol{\omega}_{12} \\ \boldsymbol{\omega}_{21} & \omega_{22} \end{pmatrix}, \mathbf{S} = \begin{pmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & s_{22} \end{pmatrix} \quad (7)$$

估计的协方差矩阵  $\mathbf{W}$  的右上角块  $\boldsymbol{\omega}_{12}$  满足

$$\hat{\boldsymbol{\omega}}_{12} = \arg \min_{\boldsymbol{\omega}_{12}} \{ \boldsymbol{\omega}_{21} \mathbf{W}_{11}^T \boldsymbol{\omega}_{12} : \| \boldsymbol{\omega}_{12} - \mathbf{S}_{12} \|_{\infty} \leq \lambda \} \quad (8)$$

G-Lasso 一次求解式(3)中一行或一列,通过交换行和列,其余部分保持不变,每一个阶段后更新对  $\mathbf{W}$  的估计,一直重复,直到  $\mathbf{W}$  收敛。

利用凸对偶性,式(8)相当于解决以下对偶问题

$$\min_{\boldsymbol{\alpha}_m} \left\{ \frac{1}{2} \| \mathbf{W}_{11}^{1/2} \boldsymbol{\alpha}_m - \mathbf{W}_{11}^{-1/2} \mathbf{S}_{12} \|^2 + \lambda \| \boldsymbol{\alpha}_m \|_1 \right\} \quad (9)$$

式中:  $\boldsymbol{\alpha}_m = \boldsymbol{\theta}_{12} / \theta_{22}$ , 应满足  $\boldsymbol{\omega}_{12} = -\mathbf{W}_{11} \boldsymbol{\alpha}_m$ 。若  $\mathbf{W}_{11}$  为  $\mathbf{S}_{11}$ , 式(9)类似于一个 Lasso 回归问题,该问题可以使用元素坐标下降来有效求解出  $\boldsymbol{\alpha}_m$ 。一般  $\hat{\boldsymbol{\alpha}}_m$  是稀疏的,且通过  $\hat{\boldsymbol{\omega}}_{12} = -\mathbf{W}_{11} \hat{\boldsymbol{\alpha}}_m$  更新一次对  $\mathbf{W}$  的估计,每更新  $m$  次  $\hat{\boldsymbol{\omega}}_{12}$  则  $\mathbf{W}$  完全更新一次,一直重复直到收敛。

估计出协方差矩阵  $\mathbf{W}$  后,其精度矩阵(逆协方差矩阵)  $\hat{\boldsymbol{\Theta}} = \mathbf{W}^{-1}$  可得

$$\hat{\theta}_{22} = 1 / (\boldsymbol{\omega}_{12} - \hat{\boldsymbol{\alpha}}_m \hat{\boldsymbol{\omega}}_{12}), \hat{\boldsymbol{\theta}}_{12} = \hat{\boldsymbol{\alpha}}_m \hat{\theta}_{22} \quad (10)$$

最后,在 G-Lasso 算法更新过程中,一次  $\mathbf{W}$  的更新中共有  $m$  个问题,将每次更新时计算出来的  $m$  组系数  $\hat{\boldsymbol{\alpha}}_m$  存储在一个矩阵  $\hat{\mathbf{B}}$  中。G-Lasso 算法在收敛后即可利用  $\hat{\mathbf{B}}$  计算式(10),因此可以快速求出精度矩阵  $\hat{\boldsymbol{\Theta}}$ 。

## 2.4 特征压缩

通过 G-Lasso 稀疏表示求解后,每个被试者将会得到一个稀疏的 FC 精度矩阵,可将其转换为对应的稀疏网络模型,如图 2 所示。然而,G-Lasso 模型使用参数  $\lambda$  来控制模型的稀疏度,估计的  $\lambda$  值确定了有效的 FC 模式。通常情况下,如果  $\lambda$  取值较大则达到足够稀疏,但一些有用的信息将被丢失;取值较小就无法起到稀疏表达作用,会存在一些虚假连接。另外, $\lambda$  参数相当于选择具有强关联的脑区信号特征,也就是说,如果一个特定的脑区与其他一个或多个脑区的相关性很高时,就会选择这些脑区。而不同被试保留的这些脑区可能是不同的,因此提取的个体差异特征或组间差异特征难以寻找到疾病组和正常对照组间的生物标志物,所以还需要进一步处理来获取具有显著性差异的特征。

在脑区功能连接网络的传统分析方法中,有多项指标被定义来反映节点与整个网络的特性,特别在机器学习辅助脑疾病的诊断中,大多研究都采用聚类方式来定义有效的聚类系数提取差异特征<sup>[7,13,17]</sup>。典型的聚类算法由于其本身的局限性(如预设聚类数的主观性、初始点随机选择),会导致聚类结果具有模糊性和多样性<sup>[29]</sup>。因此,本文提出了一种特征压缩的降维方法,在不考虑噪声影响下,可对原稀疏 FC 急剧降维,而且理论上脑区 FC 可实现无损压缩。过程如下:首先根据标签信息划分被试集合,在组平均上计算稀疏 FC 作为判决的簇中心,采用欧氏距离准则进行簇 Class 判决,可急剧降维至二维或三维,获取簇间差异特征;然后将原稀疏 FC 表达为簇 Class 中心的线性组合,采用最小二乘求解过程可提取关键特征。

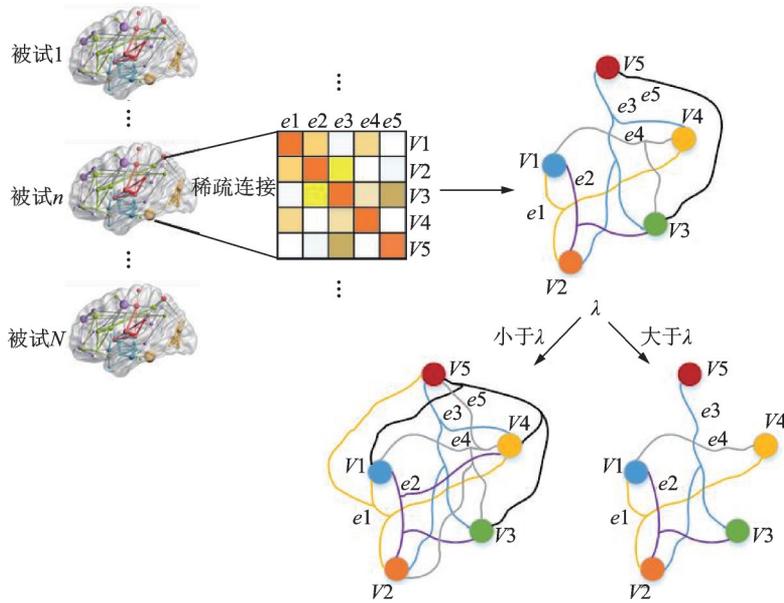


图2 脑区FC稀疏网络图

Fig.2 FC sparse network diagram of brain regions

由  $K$  折交叉验证划分集合时可预先知道被试标签信息,如“MCI”组、“NC”组。此外,考虑到所有样本数据中可能存在共享信息,因此定义“ALL”组代表所有组。所有被试组成的集合  $\mathcal{N}$ ,其训练集 ( $\mathcal{X}^S$ )和测试集 ( $\mathcal{X}^T$ )满足

$$\mathcal{X}^S \cup \mathcal{X}^T = \mathcal{N} \& \mathcal{X}^S \cap \mathcal{X}^T = \emptyset \& \frac{|\mathcal{X}^S|}{|\mathcal{X}^S + \mathcal{X}^T|} = 1 - 1/K \tag{11}$$

由于精度矩阵  $\Theta$  是对称的,所以取其上三角元素即可,其构成矢量  $R \in \mathbf{R}^{1 \times D}, D = M(M-1)/2$ , 则第  $n(1, 2, \dots, N)$  个被试可表示为

$$R^n = f_{ut}(\Theta^n) \tag{12}$$

式中  $f_{ut}(\cdot)$  为提取上三角元素构成矢量。若对所有  $R^N$  进行簇判决后将得到  $C$  个簇  $\mathfrak{R}_1, \mathfrak{R}_2, \dots, \mathfrak{R}_C$ , 即

$$\langle \mathfrak{R}_1, \mathfrak{R}_2, \dots, \mathfrak{R}_C \rangle = f_{clust}\{R^n\} \tag{13}$$

式中  $f_{clust}$  代表判决算法。相应的簇中心根据标签信息进行定义,令  $\bar{R}_c \in \mathbf{R}^{1 \times D}$  表示为  $\mathcal{X}^S$  中第  $c$  组被试的组平均,可表示为

$$\bar{R}_c = \frac{1}{|\mathfrak{R}_c^S|} \sum_{n \in \mathfrak{R}_c^S} R^n \tag{14}$$

式中  $c \in \{1, 2, 3\}$ , 分别代表 MCI 组、NC 组和 ALL 组,由此可确定簇中心。相比传统聚类算法判决过程中随机选择初始点,其能够避免聚类结果多样性。

Class 判决过程采用欧氏距离准则来对初始中心和原稀疏 FC 进行划分,其示意图如图 3 所示。首先,定义 3 个簇 Class 1、Class 2 和 Class 3, 分别代表

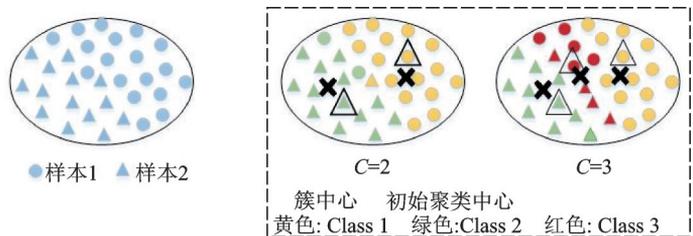


图3 判决Class过程示意图

Fig.3 Schematic diagram of the Class decision process

MCI态、NC态和共享态;然后,当 $C=2$ 时,若距离第 $c \in \{1, 2\}$ 组的初始中心( $\bar{R}_c$ )近就判为该组; $C=3$ 时也同样方法进行判决。此时,判决Class的数目就已经明确,即要么为2,要么为3。

此外,为进一步统计簇Class间差异特征,定义了性能指标占比 $\Delta$ 及最小聚合度 $O$ ,即

$$\Delta^c = \frac{|\mathfrak{K}_c^c|}{\sum |\mathfrak{K}_c|} \times 100\%, O = \frac{|\mathfrak{K}_c^*|}{\sum |\mathfrak{K}_c|} \times 100\% \quad (15)$$

式中: $\mathfrak{K}_c^c$ 代表判决为Class中第 $c$ 簇的元素集合,如 $\mathfrak{K}_c^1, \mathfrak{K}_c^2$ 分别代表簇中MCI组、NC组的元素集合。 $\mathfrak{K}_c^* = \min \{ \mathfrak{K}_c^1, \mathfrak{K}_c^2, \dots, \mathfrak{K}_c^c \}$ 代表簇Class中最小的簇,若最小聚合度过小,则表明判决不合理,需要重新划分集合并计算组平均。综上,Class的设置如下所示:(1) $C=2$ ,Class 1、Class 2分别表示MCI态、NC态;(2) $C=3$ ,Class 1、Class 2和Class 3分别表示MCI态、NC态和共享态。

即使是小样本的数据,而其脑区间连接关系复杂多样,具有较多的特征信息。由G-Lasso回归模型对被试构建出FC稀疏网络,对其进行MCI态、NC态和共享态判别,随后识别得到的各状态特征值。但是这些特征值不是全部都对分类起作用,这就需要提取具有贡献能力或重要性意义的特征。为此,本文通过将每个被试的稀疏FC表达为簇中心连接模式的线性组合,即采用最小二乘求解以优化特征的表达和区分能力,求解权重系数特征便可提取出关键特征,其过程如图4所示。需要注意的是,图4给出的是急剧降为三维( $C=3$ )时所进行的压缩过程,若 $C=2$ 时,也是同样的原理,不同的是没有Class 3。

图4中利用最小二乘对每个被试的稀疏FC进行压缩,也就是把每个被试的精度矩阵 $\Theta$ 的元素信息通过MCI态、NC态及共享态特征信息的线性组合进行表达,可表示为

$$\Theta^n = \mathbf{A} \epsilon^n \quad (16)$$

式中: $\mathbf{A} = [\mathfrak{K}_1, \mathfrak{K}_2]$  ( $C=2$ )或 $\mathbf{A} = [\mathfrak{K}_1, \mathfrak{K}_2, \mathfrak{K}_3]$  ( $C=3$ )表示各状态的特征矢量。求解式(16)中的权重系数 $\epsilon$ 为

$$\epsilon^n = \mathbf{A}^\dagger \Theta^n \quad (17)$$

式中 $(\cdot)^\dagger$ 表示伪逆。权重系数特征求解过程可视为特征表达并用来提取关键特征。最后,将所有被试进行特征压缩得到的关键特征矢量进行统计,得到

$$\mathbf{E} = [\epsilon^1, \epsilon^2, \dots, \epsilon^N] \quad (18)$$

由此可提取出具有统计意义的特征子集 $\mathbf{E}$ 作为分类器的输入特征。

## 2.5 分类

本文通过“稀疏+压缩”的新方法建立FC稀疏网络并提取出功能脑区的关键特征作为输入特征来构建分类模型。上节特征压缩中,在组平均上计算稀疏FC可反映出组间差异脑区信息,随后定义了稳定的Class判决模式,可获取簇间差异特征;由特征表达可提取出FC稀疏网络的关键特征;最后,获取以上相应的特征子集,分别对这些特征进行分类性能预测。分类任务中,采用基于高斯核的SVM作为分类器,因为对于小样本数据其具有良好的分类性能<sup>[32]</sup>。再由 $K$ 折交叉验证划分集合来评估分类模型的泛化性能,分类实验中使用了5折交叉验证,即将数据集随机分为5个互不重叠的子集,每个子集轮流作为测试集,其余4个子集用作训练集,共重复5次实验,最终计算这5次结果的平均作为模型的性能

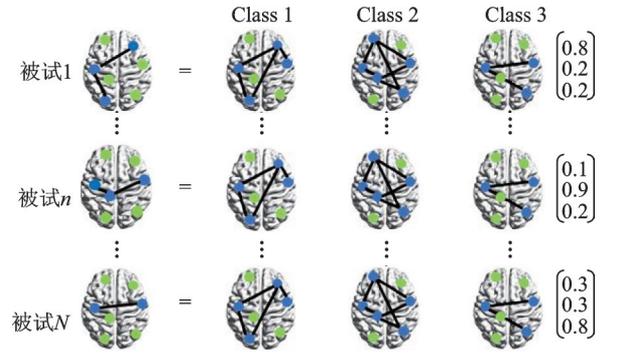


图4 特征表达过程

Fig.4 Feature expression process

表现。同时,为了确保分类结果的可信度和精确度,实验中需要重复多次5折交叉验证后计算其均值作为最终的分类准确率。

### 3 实验与分析

#### 3.1 被试信息和实验设置

实验采用的公开的 ADNI 数据库包括 ADNI-1、ADNI-GO、ADNI-2 和 ADNI-3 阶段,其中 MCI 被试的 rs-fMRI 数据大多在 ADNI-2~ADNI-4 阶段中。因此,本实验从该阶段下载了共 72 名被试数据,其中 MCI 共 35 名,NC 组共 37 名。被试的相关参数信息见表 1。两组被试的平均年龄在 70 岁左右,性别分布没有显著差异。

表 1 被试相关信息

Table 1 Information about subjects

参数	描述或数值	参数	描述或数值
MCI	疾病组被试	TR/ms	3 000
NC	正常组被试	TE/ms	30
数据库	ADNI-2-4	像素大小/ (mm×mm)	3.3×3.3
磁场强度/T	3.0	切片数量/个	48
采集设备	Philip	时间点	140 个 TR
翻转角/(°)	80		

运用数据处理和脑成像分析(Data processing & analysis of brain imaging, DPABI)工具箱(<http://rfmri.org/dpabi>)进行数据预处理,主要过程包括:(1)去除前 10 个不稳定的时间点,保留剩下的 130 个时间点;(2)时间层矫正,使每个切片上的数据具有相同时间点;(3)头动校正,将头动校正到同一位置,并为图像质量控制提供依据;(4)空间标准化,将图像配准到一个三维蒙特利尔 MNI(Montreal neurological institute)标准空间模版,并将图像进行分辨率重采样为 3 mm×3 mm×3 mm;(5)平滑,采用半高全宽高斯平滑核(8 mm×8 mm×8 mm)进行平滑,减少图像配准的误差,提高信噪比;(6)去除一些混淆因素,将不符合标准的头动参数(如头部的旋转、平移和原点定位等)作为无关变量去除。

最终,根据被试的头动标准、图像成像质量以及结构和功能像的配准情况对 rs-fMRI 图像进行筛选,保留了 32 个 MCI 和 32 个 NC 被试数据。然后,使用 GIFT 工具箱(<http://mialab.mrn.org/software/gift>)对保留的图像数据进行 sICA 处理来获取 TCs 序列信号,事先要估计 TCs 序列信号所对应的独立成分(Independent components, ICs)的数目。为了避免估计 ICs 数时先验知识不足和受主观因素的影响,使用特定模板来获取 ICs 和对应的脑网络信息。本文采用 neuromark\_fmri\_1.0 模板(<https://trendscenter.org/data/>),该模板可将脑区划分为 53 个 ICs,并划分到相应的 7 个子网络中,分别为下皮层(Subcortical, SC)网络、听觉(Auditory, AUD)网络、感觉运动(Sensorimotor, SM)网络、视觉(Visual, VIS)网络、认知控制(Cognitive control, CC)网络、默认模式(Default mode, DM)网络和小脑(Cerebellar, CB),具体如图 5 所示。实验中所涉及算法的具体步骤及参数设置如表 2 所示。

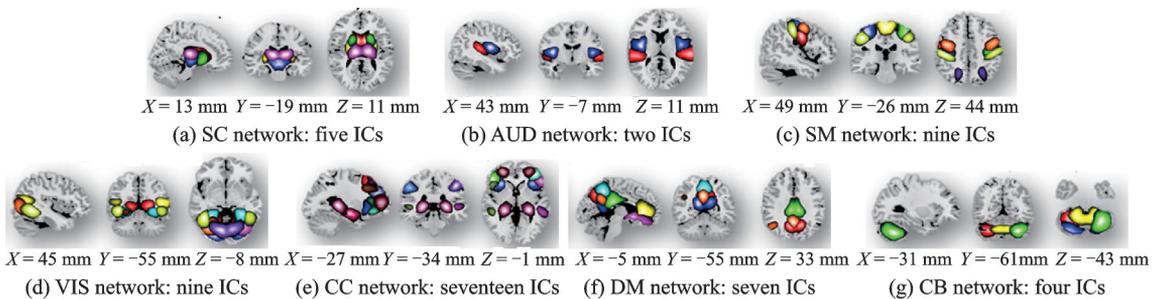


图 5 获取的独立成分

Fig.5 Obtained independent components

### 3.2 脑区功能连接

本实验首先对被试的 TCs 序列信号进行 G-Lasso 估算精度矩阵,图 6 给出了一个被试的 53 个 ICs 提取的 TCs 序列信号经过 G-Lasso 稀疏后的连接模式结果图。如图 6,选择任意的一个被试进行 G-Lasso 估计脑区间 FC 稀疏网络,可以观察到  $\lambda$  取值越大,得到的连接就越稀疏,且大部分脑区被强制为零关联; $\lambda$  取值越小,连接就越紧密且复杂。当  $\lambda < 0.2$  时,其连接模式与  $\lambda = 0.2$  时相似,若直接观察其连接情况是难以反映出 MCI 者和 NC 者间的显著性差异信息。因此,G-Lasso 处理后利用式(11~14)划分集合信息,并计算在组平均上以及簇 Class 判决来获取组间/簇间的生物标志物。

由上文分析,稀疏表示方法能去除一些虚假的或者不重要的连接信息,然而选择合适的  $\lambda$  值是一个关键的步骤,如果  $\lambda$  取值过小可能会导致模型过拟合,反之,则可能会导致模型过于稀疏而丢失部分有用的信息。由式(2)中目标函数,稀疏化 FC 则将丢弃那些使得剩余连接最大化可能性的连接<sup>[10]</sup>,因此在 G-Lasso 模型中使用元素坐标下降不断更新  $W$ ,直到收敛,即式(8,9)可估计得到最优的  $\lambda$ ,再由式(10)得到相应的精度矩阵。

图 7 分别给出了每组被试的最优  $\lambda$  取值。从图中可以看出  $\lambda$  取值在 0.02~0.1 之间,每个被试都确定一个  $\lambda$  值。这也说明了不同被试间其脑区 TCs 序列信号的 FC 模式不同,因为不同被试脑区间交互作用是不同的,从而避免了相同稀疏度带来的影响。此外,由于本实验中获取了 53 个 ICs 的 TCs 序列信号,精度矩阵的上三角特征数为  $53 \times (53 - 1) / 2$ 。若  $\lambda$  取值过大,FC 稀疏网络足够的稀疏,但可能丢失部

表 2 相关算法的参数设置

Table 2 Parameter settings for related algorithms

算法	参数	描述
G-Lasso	输入	TC 序列信号 $R^{130 \times 53}$
	初始化	$W + S + \lambda \Gamma = 0$
	块分区矩阵	$W, S$
	快速坐标下降法	求解式(9)
	更新 $\omega_{12}$	$\omega_{12} = -W_{11} \alpha_m$
	更新协方差矩阵	$W$
	重复 $m = 1, 2, \dots, M$	直到式(9)收敛
	权重向量矩阵 $\hat{B}$	$(M - 1)$ 维 $\hat{\alpha}_m$
	计算 $\hat{\theta}_{12}, \hat{\theta}_{22}$	式(10)
	输出	精度矩阵 $\Theta$
特征压缩	输入	$\Theta$
	判决 Class 数	2, 3
	判别准则	欧氏距离
	第 1 步: Class 判决	MCI 态、NC 态、共享态
	聚合度	$> 5\%$
	输出	Class 1~Class 3
第 2 步: 压缩	输入	$\Theta, S_c$
	线性模型	最小二乘表达
	权重特征	$\epsilon$
输出	$E$	

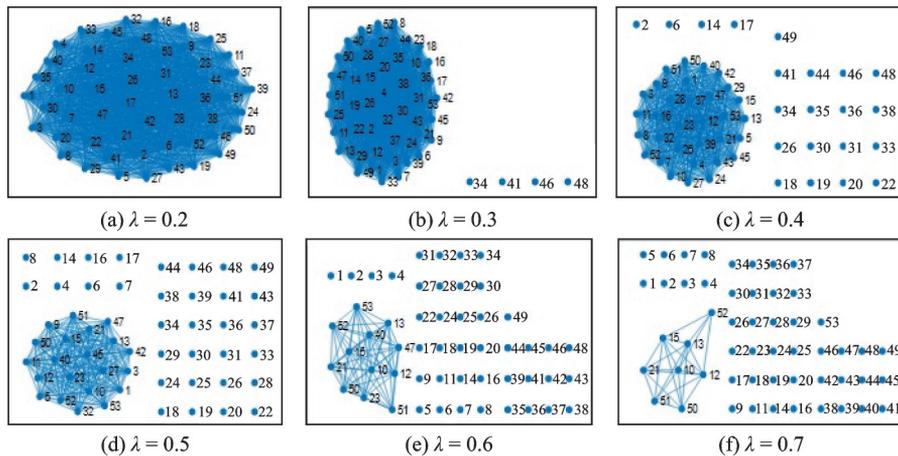


图 6 TCs 序列信号的稀疏连接示例图

Fig.6 Sparse connection example diagram of TCs sequence signal

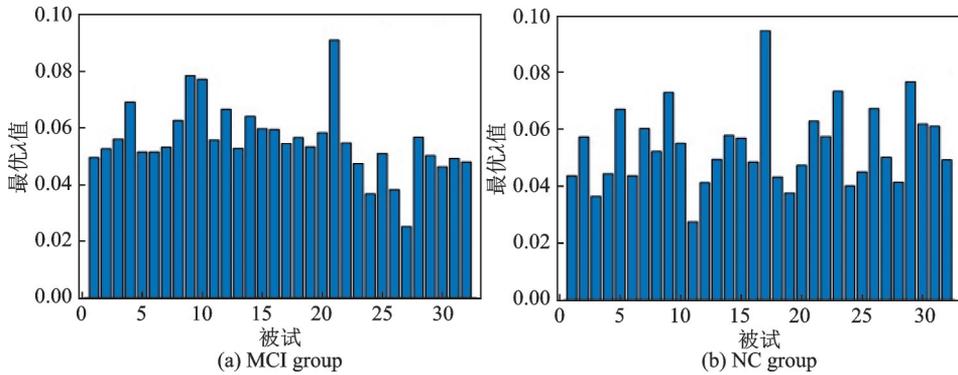


图7 被试最优λ的取值

Fig.7 The optimal λ value of subjects

分有用特征信息,因此该λ取值针对本实验数据来说是较为合理的。对于其他模板提取较多的ICs数据,相应的λ取值可能会大一些。

### 3.3 组间、簇间特征差异分析

在rs-fMRI数据预处理时,采用特定模板得到了53个ICs,即将全脑区域划分成53个特定功能区,及对应的7个脑网络中,分别为SC网络、AUD网络、SM网络、VIS网络、CC网络、DM网络和CB网络。图8给出了G-Lasso模型求解稀疏FC在组平均水平上7个特定功能脑网络间的连通性,其中图8(a)为λ=0时稀疏表示所估计的Pearson相关矩阵的组间特征差异;图8(b)为图7中被试取相应λ值估计的精度矩阵的组间特征差异。从图8可以看出,Pearson相关矩阵中组间特征差异较明显,MCI组的连接强度低于NC组,反映了NC者的FC强于MCI者,即MCI在脑功能活动中脑区间的连接可能减弱或者发生异变<sup>[33]</sup>。然而,经稀疏后得到的精度矩阵,其组间特征差异不明显,可能是因为稀疏变换后每个被试的脑区间FC模式不同,所以组平均上难以获取生物标志物来进行统计性差异分析。

图9给出了稀疏FC进行Class判决后的簇间特征差异结果图。C=2时,判决为Class 2(NC态)簇中心连接强度强于判决为Class 1(MCI态)簇中心;C=3时也同样展现出Class 2(NC态)簇中心连接强度更强。这表明与NC相比,MCI可能需要更多区域间交互来补偿网络效率的损失,这些相互作用模式的变化表明某些大脑区域已受到疾病的影响,与MCI患者早期功能异常的证据一致<sup>[4,5,33]</sup>。网络内部也存在明显的变化,特别是在感知网络(即SM,VIS和CB)中知觉与运动过程有关,受外部环境的信息传递影响,MCI患者可能在静息状态下认知加工有所丧失。从占

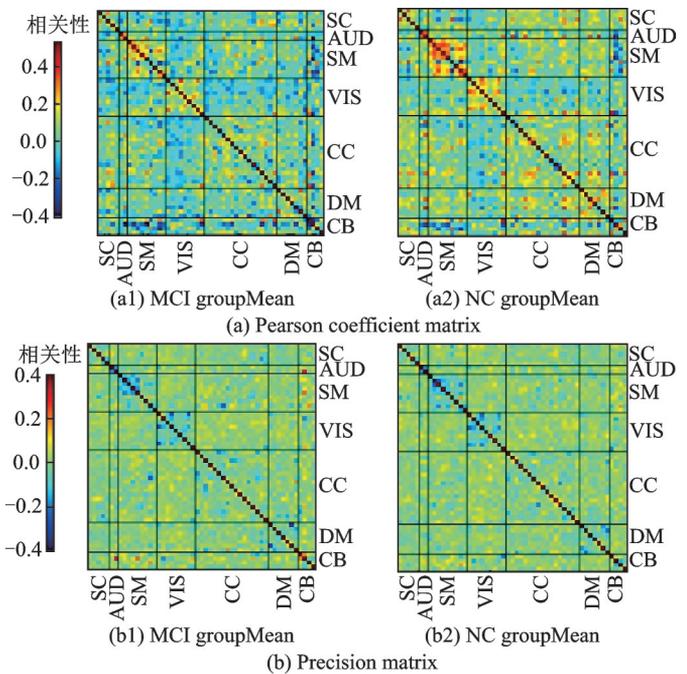


图8 组间差异结果图

Fig.8 Result plot of differences between groups

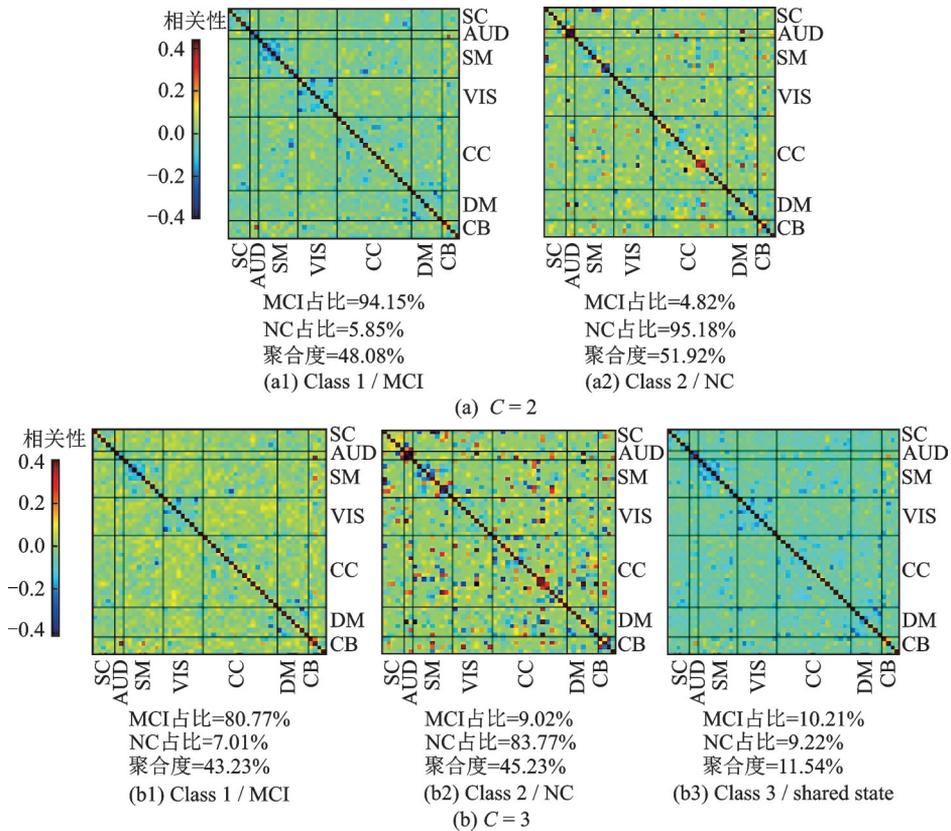


图9 簇间差异结果图

Fig.9 Result plot of differences between clusters

比和聚合度来看,不管是降至二维还是三维,簇 Class 1(MCI态)中MCI占比达到80%以上,显然高于NC占比;同样簇 Class 2(NC态)中NC占比显然高于MCI占比,且簇中两者占比差达到75%以上;簇 Class 3(共享态)的聚合度小,且占比差较小,表明降至三维时可能只有较少的共享信息存在。重要的是,以上簇间特征差异信息可以为脑疾病分类提供有效的标志物。本文将根据这种标志物来进一步获取脑异常特征,进一步研究其分类性能表现。

### 3.4 分类表现

为进一步评估所提“G-Lasso+特征压缩”方法的性能优劣,计算其在FC稀疏网络分析中的分类性能,即计算相应特征子集的分类准确率、灵敏度(可准确识别MCI患者的比例)、特异度(可准确识别NC被试的比例)和曲线下面积(Area under the curve, AUC)指标进行评估。同时,与传统的稀疏FC构建方法(Lasso及G-Lasso)和特征选择等方法进行比较。以往基于距离准则的特征选择方法在MCI患者分类研究中获得了较好的分类表现<sup>[7]</sup>,本文鉴于此方法进行特征选择来进行比较,其包括最小类内距离、最大类间距离和类内比类间距离准则。需要注意的是特征数目选择了每个被试者的总特征数到其1/3间的特征数。稀疏表示结合特征压缩的方法,即“Lasso+特征压缩和G-Lasso+特征压缩”进行对比。此外,通过Kmeans聚类且采用欧氏距离判决,可以获取聚类中心属性特征<sup>[26]</sup>,因此可以设置聚类数为2~5进行对比分析。表4给出了不同方法的分类性能,结果表明,G-Lasso+特征压缩具有最好的分类性能,分类准确率可达到89.8%,说明本文所提方法对信息冗余和特征降维具有一定的作用;Lasso

表4 分类性能比较

Table 4 Classification performance comparison

方法	分类特征	准确率/%	敏感度/%	特异度/%	AUC	
Lasso	所有特征	80.5	76.3	70.2	0.79	—
	组间特征	77.2	80.6	73.6	0.77	—
	簇间特征	79.6	82.0	77.3	0.79	—
G-Lasso	所有特征	82.4	80.4	87.6	0.85	—
	组间特征	80.8	80.6	81.0	0.81	—
	簇间特征	<b>84.8</b>	<b>85.5</b>	<b>82.3</b>	<b>0.87</b>	—
特征选择	最小类内	74~81	—	—	—	—
	最大类间	77~82	—	—	—	—
	类内/类间	75~81	—	—	—	—
Lasso+特征压缩	关键特征	82.7	84.3	78.3	0.84	—
G-Lasso+特征压缩	关键特征	<b>89.8</b>	<b>90.3</b>	<b>87.6</b>	<b>0.94</b>	—
G-Lasso +K-means	聚2簇	57	—	—	—	—
	聚3簇	67	—	—	—	—
	聚4簇	60	—	—	—	—
	聚5簇	65	—	—	—	—
稀疏表示 <sup>[5]</sup>	所有特征	84.0	84.0	84.0	0.75	—
稀疏表示 <sup>[26]</sup>	所有特征	76.9	82.2	71.1	—	—
稀疏和低秩 <sup>[27]</sup>	所有特征	80.1	78.3	81.8	0.87	—
组Lasso <sup>[22]</sup> /MDD	类间特征	80.3	84.2	75.0	—	—
	弹性网络	86.3	92.1	81.5	—	—
稀疏组Lasso <sup>[17]</sup> /MDD	单一节点	84.8	92.1	82.1	87.1	—
	成对节点	77.5	84.2	71.4	77.8	—
	类间特征	87.8	89.5	85.7	87.6	—

方法构建的稀疏FC,其分类性能较差,因为只能选择存在组结构中的任意一个TC序列信号,这会导致稀疏表示求解中失去一些重要连接信息;G-Lasso方法仅只是对FC进行特征选择,相比Lasso方法其分类性能有所提升,但特征维度依然较高导致分类结果并不是最好的;特征选择方法的分类准确率受到特征数目多少的影响,即使由最优的特征数目得到的最优分类准确率也只有82%,而G-Lasso方法获取的簇间特征的分类准确率可达到84.8%。

稀疏表示结合特征压缩的方法不仅可获取簇间特征差异信息,其获取的关键特征还可以进一步提高分类性能。表4结果表明,“Lasso+特征压缩”方法相比Lasso方法可提高2%~5%,同样“G-Lasso+特征压缩”方法可提高5%~9%。然而,“G-Lasso+Kmeans方法”得到的分类准确率并不太理想,最好的结果在聚3簇时,也仅有67%。因为Kmeans算法根据数据本身特性自主形成聚类中心类别,这种特性是不明确的,而且随机选择初始中心会导致结果不稳定。因此,本文所提出的特征压缩方法对MCI患者分类预测方面具有明显的效果。

除此之外,表4还给出了关于稀疏表示和特征降维的其他相关文献的分类结果,比较后可以得出,本文所提“G-Lasso+特征压缩”方法的分类结果是最好的。文献[5]同样是在ADNI库中采集rs-fMRI数据,对12名MCI和25名NC进行实验,提出组约束稀疏方法,通过SVM预测分类结果,分类准确率可达84%;文献[26]对MCI和NC进行实验,通过稀疏表示来获取空间相关的FC,使用 $t$ 检验来进行特征降维并进行分类计算,得到76.9%的分类准确率;文献[27]通过稀疏和低秩约束方法来估计具有清晰模块化结构的稀疏FC矩阵,将MCI受试者与健康对照组进行分类研究,其分类准确率可达80.1%。

此外,Guo等<sup>[22]</sup>提出组Lasso方法来改善FC网络的创建,使得稀疏方法具有选择组效应的能力,虽然该方法用于抑郁症(Major depressive disorder, MDD)的分类研究,但在全脑FC研究中具有一定的贡献。Li等<sup>[17]</sup>在组Lasso方法基础上考虑了脑区间选择组之间的重叠关系,再进一步构建稀疏FC,得到了较理想的分类性能。说明稀疏表示结合类别间的差异特征可以提高对精神疾病患者的分类准确率,也再次暗示本文的新框架,即“G-Lasso+特征压缩”方法,既考虑了多个脑区之间有效连接,又考虑到表征组间差异的特征信息的选择能力,同时达到特征选择和降维的作用,能较好地弥补以往研究的不足。

图10给出了在10次实验中,“G-Lasso+特征压缩”过程中 $C=2$ 和 $C=3$ 的分类准确率曲线。从图10可以得到, $C=2$ 的分类准确率高于 $C=3$ 时,可能因为G-Lasso方法在求解精度矩阵时已经去除了部分虚假连接,保留了重要的信息,而 $C=3$ 时判决的Class 3(共享态)会使MCI和NC丢失部分重要的特征信息。

图11给出了簇间特征和关键特征的分类准确率统计结果。可以看出, $C=2$ 时两种特征的分类准确率分别达到最高,特别是进行了特征表达后,获取的关键特征准确率达到了89.8%。也可以看出, $C=3$ 时,簇间特征的分类准确率低于关键特征约10个百分点,也就意味着最小二乘表达能够优化特征的表达和区分能力,对稀疏FC进行了无损压缩,提取关键特征的同时保留了重要的信息。

## 4 方法论及参数影响

### 4.1 方法论

不同的FC计算方法,在MCI患者研究中会影响对其辅助诊断效率。传统计算的FC大多是全连接网络,会面临信息冗余和维度灾难等问题。在现有研究中使用稀疏表示进行FC稀疏网络构建,常基于Lasso模型构造线性回归模型,并使用 $\ell_1$ 范数来控制网络模型的稀疏度。Lasso求解过程中,由于只选择了单个变量,全脑区域中一些重要的FC难以选择出来,可能丢失部分有效的连接信息。从图论模型的角度思考,每个节点代表一个大脑区域,每条边可以代表多个脑区之间的连接关系,这就可以构建FC稀疏网络来表征脑区连通性,如G-Lasso模型将大脑区域的信号的线性组合来表示特定脑区,同时引入 $\ell_1$ 范数惩罚项来过滤掉无关紧要或虚假的连接,并以块坐标下降算法为出发点能够快速获取稀疏的逆协方差矩阵,即精度矩阵。但是G-Lasso模型估计的稀疏FC仅能去除部分冗余特征,不能大规模降低特征维度,并不能很好地解决维度灾难问题,所以提出了特征压缩的特征降维方法。将G-Lasso结合特征压缩的新方法,可以获取全脑脑区间有效的稀疏连接,可将稀疏FC特征急剧降维至二维或三维,既能够实现特征选择又起到特征降维的目的。

也有使用深度学习对MCI进行分类预测研究,如文献[6]使用迁移网络对MCI和NC进行分类,结果表明相关系数特征的分类准确率为63.67%,区域数特征的分类准确率为73.67%;文献[33]中使用深度学习算法对MCI进行分类实验,文中提出的JP系数特征最好的分类准确率能达到84.33%。深度学习虽然在退行性疾病领域的辅助诊断提供了新的方法<sup>[34]</sup>,但其模型搭建过程中需要大量数据训练样本

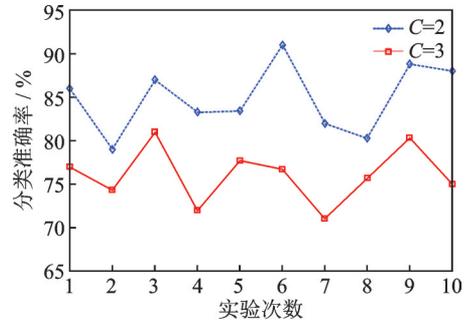


图10 簇间特征分类曲线

Fig.10 Classification curves of inter-cluster features

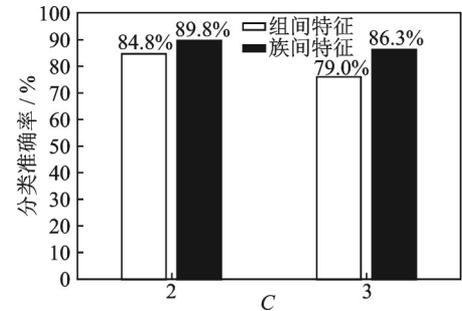


图11 特征压缩分类准确率

Fig.11 Feature compression classification accuracy

和众多的参数设计,对于神经影像数据库中这样的小样本多特征的数据在一定程度上可能会有限制。

除此之外,本文所提方法的参数调优过程也会影响其分类性能。本文在G-Lasso方法下构建的FC稀疏网络涉及到参数 $\lambda$ 值调优,为了探索此方法最优分类效果下的参数值,对其 $\lambda$ 进行了实验分析。

## 4.2 正则化参数 $\lambda$ 的影响

以往研究已表明 $\lambda$ 参数影响求解线性回归模型的拓扑结构,该参数指 $\ell_1$ 范数惩罚项,用于控制全脑FC的稀疏度。同一被试选用不同 $\lambda$ 参数值会得到不同的FC稀疏网络模型,而有研究为了使不同被试生成相同的网络拓扑<sup>[18-19,28]</sup>,便选取固定的 $\lambda$ 值。考虑到被试之间存在脑区的连接差异,因此本研究中,采用元素坐标下降方案对全脑FC稀疏网络构建过程中的正则化参数进行调优,得到各被试 $\lambda$ 取值如图7所示。为了探讨 $\lambda$ 取值对分类性能的影响,指定所有被试取相同的 $\lambda$ 值,通过在特定范围内的变化产生不同超边,最后生成全脑FC稀疏网络模型,再进行特征压缩提取关键特征来计算分类准确率。实验中,参考图7中各被试 $\lambda$ 取值,可进一步设所有被试的 $\lambda$ 取值范围为 $\{0.01, 0.03, \dots, 0.15\}$ ,在 $C=2$ 时得到的分类结果如图12所示。从图中可以看出,当 $\lambda$ 在0.03和0.05时表现出较高的分类准确率,且当 $\lambda=0.05$ ,关键特征的分类准确率达到最高,即86.7%;组间差异特征也在 $\lambda=0.05$ 达到较好的结果。簇间差异特征在 $\lambda=0.03$ 时有较高的分类准确率。

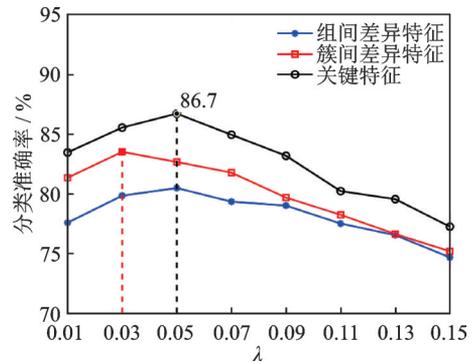


图12 参数 $\lambda$ 取值的分类准确率

Fig.12 Classification accuracy of parameter  $\lambda$  values

而在本文使用的G-Lasso模型中,所有被试取不同的 $\lambda$ 值,分类准确率可达到89.8%。从图7中可以看出大部分被试取值在0.03到0.05之间,这也说明本文G-Lasso方法使得FC稀疏网络更精确,从而获取更有意义的特征信息来提高分类准确率。

## 5 结束语

针对rs-fMRI数据中全脑FC存在的信息冗余和特征维度灾难等问题,本文提出了一种能更好分析全脑FC稀疏网络模型的框架,使得MCI和NC间有显著的差异性特征,因此能获得更高的分类准确率。先通过G-Lasso方法对获取的TC序列信号进行稀疏表示,加入 $\ell_1$ 范数惩罚项来控制脑区功能连接的稀疏度,求解得到精度矩阵;再次,进行特征压缩,将每个被试的稀疏FC判决为簇Class 1~Class 3;原稀疏FC通过簇中心的线性组合进行表达,采用最小二乘求解并提取出关键的特征子集;最后对不同特征进行分类性能预测。通过实验结果分析,G-Lasso方法能使得不同被试在不同 $\lambda$ 下获得有效且稀疏的FC;特征压缩过程中,簇间差异特征为分类研究提供了有效的标志信息。分类结果表明,“G-Lasso+特征压缩”方法的分类表现优于现有的关于全脑FC稀疏网络的构建方法。由此暗示着仅只进行稀疏表示或特征选择均不能达到意向的结果,若对特征降维进行适当的扩展,将会获得更为有效和重要的特征。这也表明“稀疏+压缩”方法能有效地解决全脑FC的多特征问题,即冗余信息和维度灾难。

本文的实验结果验证了所提方法的有效性,且对小样本多特征数据具有更好的分类性能,但也存在以下局限性:(1)根据特定模板仅获取了53个脑区时间序列,使得 $\lambda$ 参数会更敏感;(2)由于MCI患者数据信息不好采集,因此实验中被试数目较少,且还要划分集合,导致分类结果具有一定的局限性。在未来工作中,可以考虑对MCI患者数据提取更多的脑区时间序列,从静态和动态的方面进一步获取不同的时效特征;也可以从多种神经类疾病的数据获取不同的脑区信号,结合有监督聚类方法来挖掘网络的特定功能特征。

## 参考文献:

- [1] Alzheimer's Association. 2019 Alzheimer's disease facts and figures[J]. *Alzheimer's & Dementia*, 2019, 15(3): 321-387.
- [2] PANNUNZI M, HINDRIKS R, BETTINARDI R G, et al. Resting-state fMRI correlations: From link-wise unreliability to whole brain stability[J]. *Neuroimage*, 2017, 157: 250-262.
- [3] MILLER R L, YAESOUBI M, TURNER J A, et al. Higher dimensional meta-state analysis reveals reduced resting fMRI connectivity dynamism in schizophrenia patients[J]. *PLoS One*, 2016, 11(3): e0149849.
- [4] CUI W, MA Y, REN J, et al. Personalized functional connectivity based spatio-temporal aggregated attention network for MCI identification[J]. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2023, 31: 2257-2267.
- [5] WEE C Y, YAP P T, ZHANG D, et al. Group-constrained sparse fMRI connectivity modeling for mild cognitive impairment identification[J]. *Brain Structure and Function*, 2014, 219: 641-656.
- [6] 孔伶俐, 吴海锋, 曾玉, 等. 迁移学习特征提取的 rs-fMRI 早期轻度认知障碍分类[J]. *智能系统学报*, 2021, 16(4): 662-672.  
KONG Lingxu, WU Haifeng, ZENG Yu, et al. Early mild cognitive impairment classification based on transfer learning feature extraction from rs-fMRI[J]. *Journal of Intelligent Systems*, 2021, 16(4): 662-672.
- [7] 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, 14(2): 552-564.
- [8] 张辉, 王建国. 一种基于主分量分析的 SAR 图像变化检测算法[J]. *电子与信息学报*, 2008, 30(7): 1727-1730.  
ZHANG Hui, WANG Jianguo. A SAR image change detection algorithm based on principal component analysis[J]. *Journal of Electronics & Information Technology*, 2008, 30(7): 1727-1730.
- [9] ALLEN E A, DAMARAJU E, PLIS S M, et al. Tracking whole-brain connectivity dynamics in the resting state[J]. *Cerebral Cortex*, 2014, 24(3): 663-676.
- [10] ELLIS C A, MILLER R L, CALHOUN V D. A novel explainable fuzzy clustering approach for fMRI dynamic functional network connectivity analysis[J]. *bioRxiv*, 2023, 526110.
- [11] CHEN Y Y, CUI Q, XIE A L, et al. Abnormal dynamic functional connectivity density in patients with generalized anxiety disorder[J]. *Journal of Affective Disorders*, 2020, 261: 49-57.
- [12] HUANG S, LI J, SUN L, et al. Learning brain connectivity of Alzheimer's disease by sparse inverse covariance estimation[J]. *NeuroImage*, 2010, 50(3): 935-949.
- [13] LEE H, LEE D S, KANG H, et al. Sparse brain network recovery under compressed sensing[J]. *IEEE Transactions on Medical Imaging*, 2011, 30(5): 1154-1165.
- [14] JIE B, WEE C Y, SHEN D, et al. Hyper-connectivity of functional networks for brain disease diagnosis[J]. *Medical Image Analysis*, 2016, 32: 84-100.
- [15] 李瑶, 赵云芄, 李欣芸, 等. 基于 sparse group Lasso 方法的脑功能超网络构建与特征融合分析[J]. *计算机应用*, 2020, 40(1): 62-70.  
LI Yao, ZHAO Yunpeng, LI Xinyun, et al. Brain functional hypernetwork construction and feature fusion analysis based on sparse group Lasso method[J]. *Journal of Computer Applications*, 2020, 40(1): 62-70.
- [16] LIANG S, VEGA R, KONG X, et al. Neurocognitive graphs of first-episode schizophrenia and major depression based on cognitive features[J]. *Neuroscience Bulletin*, 2018, 34: 312-320.
- [17] LI Y, LI Q, LI T, et al. Construction and multiple feature classification based on a high-order functional hypernetwork on fMRI data[J]. *Frontiers in Neuroscience*, 2022, 16: 848363.
- [18] 汪瀚, 吴海锋, 王俨, 等. 轻度认知障碍的 BOLD 信号 MEMD 降噪处理[J]. *控制与决策*, 2023, 38(10): 2823-2831.  
WANG Han, WU Haifeng, WANG Yan, et al. MEMD denoising of BOLD signals in mild cognitive impairment[J]. *Control and Decision*, 2023, 38(10): 2823-2831.
- [19] 袁悦铭. 基于静息态功能磁共振的动态脑连接分析方法及其在主观认知下降研究中的应用[D]. 深圳: 深圳大学, 2021.  
YUAN Yueming. Dynamic brain connectivity analysis based on resting-state functional MRI and its application in subjective cognitive decline research[D]. Shenzhen: Shenzhen University, 2021.
- [20] QURESHI M N I, OH J, CHO D, et al. Multimodal discrimination of schizophrenia using hybrid weighted feature concatenation of brain functional connectivity and anatomical features with an extreme learning machine[J]. *Frontiers in Neuroinformatics*, 2017, 11: 59.
- [21] 周洲, 钟元. 动态功能连接方法及在神经精神疾病中的应用研究[J]. *磁共振成像*, 2021, 12(1): 73-76.

- ZHOU Zhou, ZHONG Yuan. Dynamic functional connectivity methods and their applications in neuropsychiatric disorders[J]. Magnetic Resonance Imaging, 2021, 12(1): 73-76.
- [22] GUO H, LI Y, XU Y, et al. Resting-state brain functional hyper-network construction based on elastic net and group Lasso methods[J]. Frontiers in Neuroinformatics, 2018, 12: 25.
- [23] 李鹏祖, 李瑶, IBEGBU N J, 等. 基于多特征融合的重叠组套索脑功能超网络构建及分类[J]. 计算机科学, 2022, 49(5): 206-211.
- LI Pengzu, LI Yao, IBEGBU N J, et al. Construction and classification of brain functional hypernetworks based on overlapping group Lasso with multi-feature fusion[J]. Journal of Computer Science, 2022, 49(5): 206-211.
- [24] SHI L, TONG H, DAIANU M, et al. Visual analysis of brain networks using sparse regression models[J]. ACM Transactions on Knowledge Discovery from Data (TKDD), 2018, 12(1): 1-5.
- [25] ZHANG J, ZHOU L, WANG L, et al. Functional brain network classification with compact representation of SICE matrices [J]. IEEE Transactions on Biomedical Engineering, 2015, 62(6): 1623-1634.
- [26] XUE Y, ZHANG L, QIAO L, et al. Estimating sparse functional brain networks with spatial constraints for MCI identification [J]. PLoS One, 2020, 15(7): e0235039.
- [27] DU Y, WANG G, WANG C, et al. Accurate module induced brain network construction for mild cognitive impairment identification with functional MRI[J]. Frontiers in Aging Neuroscience, 2023, 15: 1101879.
- [28] FRIEDMAN J, HASTIE T, TIBSHIRANI R. Sparse inverse covariance estimation with the graphical lasso[J]. Biostatistics, 2008, 9(3): 432-441.
- [29] 保涵, 吴海锋, 晏洁, 等. 轻度认知障碍功能连接的频率和聚类分析[J]. 计算机应用研究, 2023, 40(2): 526-532.
- BAO Han, WU Haifeng, YAN Jie, et al. Frequency and clustering analysis of functional connectivity in mild cognitive impairment[J]. Journal of Computer Application Research, 2023, 40(2): 526-532.
- [30] 李栋, 吴海锋, 保涵, 等. 轻度认知障碍磁共振信号中固有频率动态功能性连接的聚类研究[J]. 数据采集与处理, 2022, 37(4): 798-813.
- LI Dong, WU Haifeng, BAO Han, et al. Clustering study of dynamic functional connectivity based on intrinsic frequency in MRI signals of mild cognitive impairment[J]. Journal of Data Acquisition and Processing, 2022, 37(4): 798-813.
- [31] CHIU H Y S, JAMES C J. Introducing a deflationary approach to space-time ICA that uses temporal methods in brain signals processing[C]//Proceedings of 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). [S.l.]: IEEE, 2020: 213-216.
- [32] BURGESS C J C. A tutorial on support vector machines for pattern recognition[J]. Data Mining and Knowledge Discovery, 1998, 2(2): 121-167.
- [33] 孔伶俐, 吴海锋, 曾玉, 等. 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 Lingxu, WU Haifeng, ZENG Yu, et al. Diagnosis of mild cognitive impairment using deep learning and brain functional connectivity across different frequency domains[J]. Journal of Computer Applications, 2021, 41(2): 590-597.
- [34] 张柏雯, 林岚, 吴水才. 深度学习在轻度认知障碍转化与分类中的应用分析[J]. 医疗卫生装备, 2017, 38(9): 105-111.
- ZHANG Baiwen, LIN Lan, WU Shuicai. Application and analysis of deep learning in transformation and classification of mild cognitive impairment[J]. Medical Equipment and Information, 2017, 38(9): 105-111.

#### 作者简介:



马佳(1999-),女,硕士研究生,研究方向:生物医学信号处理、机器学习,E-mail: 1609186984@qq.com。



吴海锋(1977-),通信作者,男,博士,教授,硕士生导师,研究方向:机器学习、信号处理、FRID 标签识别等,E-mail: whf5469@gmail.com。



李顺良(1997-),男,硕士研究生,研究方向:生物医学信号处理、机器学习。

(编辑:王静)