

# 基于组 MCP 和复合 MCP 的人脑功能超网络分析及抑郁症分类研究

薛晓倩<sup>1</sup>, 李 瑶<sup>1</sup>, 梁家瑞<sup>2</sup>, Ibegbu Nnamdi Julian<sup>1</sup>, 孙 超<sup>1</sup>, 郭 浩<sup>1</sup>

<sup>1</sup>(太原理工大学 信息与计算机学院, 山西 晋中 030600)

<sup>2</sup>(太原理工大学 软件学院, 山西 晋中 030600)

E-mail: feiyu\_guo@sina.com

**摘 要:** 近年来, 脑功能超网络模型在脑疾病诊断中多有应用. 传统的脑功能超网络大多通过 LASSO 方法进行构建, 然而由于脑区间存在组效应问题, 在过去的几年里, 对 LASSO 方法进行延伸以进一步改善超网络成为主要研究内容, 由此出现各种分组模型方法. 但这些方法均存在同样的问题, 即惩罚函数对系数的过强压缩, 导致模型中目标变量回归系数的有偏估计, 使得噪声变量在压缩的同时, 目标变量也进行了一定程度的压缩. 因此, 本文考虑到该问题, 并在组效应的基础上, 提出两种基于 Minimax Concave Penalty (MCP) 的无偏稀疏模型用以改进原有方法: 组 MCP 方法和复合 MCP 方法. 实验结果表明, 两种方法均优于传统方法, 而两种方法由于对变量是否进入模型采取了不同解决方式, 因而构建的超网络结构差异较大, 复合 MCP 方法构建的超网络的超边分布范围较为集中, 而组 MCP 方法则较为分散; 此外, 复合 MCP 方法得到较好的分类表现和较高的分类权重. 本文提出的方法所构建的脑功能超网络可以更好地表达抑郁症患者与正常对照的结构差异, 具有重要的理论意义和临床价值.

**关键词:** 无偏稀疏模型; 组效应; 组 MCP; 复合 MCP; 分类

中图分类号: TP181

文献标识码: A

文章编号: 1000-1220(2022)01-0210-08

## Research on Analysis of Human Brain Functional Hyper-network and Classification of Depression Based on Group MCP and Composite MCP

XUE Xiao-qian<sup>1</sup>, LI Yao<sup>1</sup>, LIANG Jia-rui<sup>2</sup>, Ibegbu Nnamdi Julian<sup>1</sup>, SUN Chao<sup>1</sup>, GUO Hao<sup>1</sup>

<sup>1</sup>(College of Information and Computer, Taiyuan University of Technology, Jinzhong 030600, China)

<sup>2</sup>(College of Software, Taiyuan University of Technology, Jinzhong 030600, China)

**Abstract:** In recent years, the brain functional hyper-network model has been widely used in the diagnosis of brain diseases. Traditional brain functional hyper-networks are mostly constructed by the LASSO method. However, due to the problem of group effects in brain areas, in the past few years, the extension of the LASSO method to further improve the hyper-network has become the main research content, and various grouping models have appeared. But these methods all have the same problem, that is, the penalty function compresses the coefficients too strongly, which leads to the biased estimation of the regression coefficients of the target variables in the model, so that while the noise variables are compressed, the target variables are also compressed to a certain extent. Therefore, this paper takes this problem into account and proposes two unbiased sparse models based on Minimax Concave Penalty (MCP) and also based on the group effect to improve the original method: group MCP method and composite MCP method. The experimental results show that the two methods are better than the traditional methods, and because the two methods adopt different solutions to whether the variables enter the model, the structure of the hyper-network constructed is quite different. The hyper-edge distribution of the hyper-network constructed by the composite MCP method is more concentrated, while the group MCP method is more scattered; in addition, The composite MCP method has better classification performance and higher classification weight. The brain functional hyper-network constructed by the method proposed in this paper can better express the structural difference between depression patients and normal controls, which has important theoretical and clinical value.

**Key words:** unbiased sparse model; grouping effect; group MCP; composite MCP; classification

收稿日期: 2021-03-22 收修改稿日期: 2021-05-17 基金项目: 国家自然科学基金项目(61472270, 61672374, 61741212, 61876124, 61976150)资助; 山西省重点研发计划项目(201803D31043)资助; 山西省科技厅应用基础研究项目青年面上项目(201801D121135, 201803D31043)资助; 山西省教育厅高等学校科技创新研究项目(2016139)资助; 教育部赛尔网络下一代互联网技术创新项目(NGII20170712)资助. 作者简介: 薛晓倩, 女, 1996年生, 硕士研究生, 研究方向为人工智能和脑科学计算; 李 瑶, 女, 1996年生, 博士研究生, 研究方向为人工智能和医学影像分析; 梁家瑞, 男, 1998年生, 研究方向为人工智能; Ibegbu Nnamdi Julian, 男, 1993年生, 硕士研究生, 研究方向为人工智能和脑信息处理; 孙 超, 女, 1996年生, 硕士研究生, 研究方向为人工智能和脑信息处理; 郭 浩, 男, 1981年生, 博士研究生, 教授, CCF会员, 研究方向为脑科学与智能计算.

## 1 引言

研究发现,大脑的自发神经元活动和低频血氧水平依赖 (Blood Oxygen Level-Dependent, BOLD) 信号有密切的联系<sup>[1]</sup>,这表明在静息状态下,BOLD 信号构建的功能连接网络能够反映大脑区域之间的交互活动.在过去的几年中,静息态功能磁共振成像技术不断发展,脑功能网络作为大脑交互的简化表示已被广泛地应用在抑郁症<sup>[2]</sup>、阿尔茨海默症<sup>[3]</sup>等精神疾病的研究中.

在传统的功能连接方法中,往往是图形化模型<sup>[4]</sup>和基于相关性分析方法<sup>[5]</sup>等占据主要位置,然而它们都存在不足之处:图形化模型依赖于图的先验知识<sup>[6]</sup>;基于相关性分析方法则存在两个主要问题:1)其构建的连接不保证绝对可靠,可根据自定义阈值选择连接,因此存在虚假连接<sup>[7]</sup>;2)由于其连接网络是两两关系的映射,因此不能解决大脑信息传递的高阶互动过程<sup>[6]</sup>.

相关的研究证明,大脑的信息传递在局部场电位、神经元同位素示踪以及皮层活动中均存在明显的高阶交互<sup>[8]</sup>.因此,超网络被提出<sup>[9]</sup>.脑区对应网络中的节点,如果一个脑区与其他脑区之间有交互,则认为这些脑区中存在超边.

传统的超网络通过 LASSO 方法进行构建,然而该方法存在两个问题:1)LASSO 中的惩罚函数对系数的过强压缩,导致模型中目标变量回归系数并不是无偏估计或近似无偏估计;2)LASSO 是单一变量选择模型,缺乏组效应的解释能力,这使得所构建的脑功能超网络将有可能失去一些重要的具有明显神经学意义解释的连接.目前对后者的解决方法主要是通过引入分组的方法模型,如组 LASSO<sup>[10]</sup>、稀疏组 LASSO<sup>[11]</sup>以及弹性网<sup>[12,13]</sup>方法.然而组 LASSO 和弹性网均是基于 LASSO 方法的延伸,而稀疏组 LASSO 则是组 LASSO 方法的延伸,因此上述方法同样存在惩罚函数对系数的过强压缩的问题,导致构建的超网络也许过于严格,从而失去一些重要的连接.

因此,本文提出两种基于极小极大凹惩罚 (The Minimax Concave Penalty, MCP) 的无偏稀疏模型:组 MCP<sup>[14]</sup>和复合 MCP<sup>[15,16]</sup>,通过解决稀疏线性回归模型进行超网络构建. MCP 具有变量选择连续性和无偏性,不会出现惩罚函数对系数的过强压缩的问题.两种方法均能解决组效应问题,不同的是,组 MCP 方法将 MCP 惩罚仅应用到组级,解决了组效应问题;而复合 MCP 方法则把 MCP 惩罚既用作组级惩罚函数,又作为组内惩罚函数,分别在组级和组间对变量进行选择.两种方法对系数的压缩程度不同,构建出来的超网络也不同.

研究结果表明,两种方法均优于传统方法,而两种方法所构建的脑功能超网络结构有较大差异,考虑到是两种惩罚函数对回归系数的压缩方式导致的,复合 MCP 方法构建的超网络的超边数相较于组 MCP 方法分布集中,而组 MCP 方法的超边分布较为分散;选取超网络组间差异拓扑指标为分类特征,利用支持向量机 (Support Vector Machine, SVM) 方法构建分类模型.复合 MCP 具有最好的分类效果和较高的分类权重;最后,进一步分析了方法模型中的各个参数,探究方法参

数和分类器参数以及显著性脑区对方法是否产生影响.本文提出的方法一定程度上解决了无偏稀疏脑功能超网络构建过程中的组效应问题,所构建的脑功能超网络可以更好地表达抑郁症患者与正常对照的结构差异,具有重要的理论意义和临床价值.

## 2 材料和方法

### 2.1 数据采集和预处理

本文严格按照山西医学伦理委员会的要求 (会议号: 2012013),在该项研究实施之前,与每位参与者均达成了书面协议.根据赫尔辛基宣言,书面知情同意书由实验中的每一个被试签署.总共招募了 66 名被试,其中包括 38 名首发,无用药重度抑郁症患者 (Major Depression Disorder, MDD) (15 名男性) 和 28 名健康右利手志愿者 (13 名男性).静息状态下,应用 3T 磁共振扫描仪 (Siemens Trio 3-Tesla scanner, Siemens, Erlangen, Germany) 对他们进行功能磁共振成像 (fMRI) 扫描.被试的具体情况可参照表 1.

表 1 被试的具体信息

Table 1 Specific information of participants

类别	正常人	抑郁症患者	P 值
年龄	17-51	17-49	0.44 <sup>a</sup>
性别	13/15	15/23	0.57 <sup>b</sup>
利手 (右/左)	28/0	38/0	-
汉密尔顿量表	N/A	15-42	-

山西医科大学第一医院完成数据采集任务,并且由精通磁共振技术的放射科医师完成扫描任务.被试在进行扫描时也有相应的要求,即被试需要在放松的状态下闭上眼睛,但不能入睡,需要保持清醒,也不进行特定的思考.每个扫描的结果是 248 个连续的 EPI 功能图像,其中的扫描参数有如下设置:33 axial slices, echo time (TE) = 30ms, repetition time (TR) = 2000ms, thickness/skip = 4/0 mm, flip angle = 90°, matrix = 64 × 64 mm, field of view (FOV) = 192 × 192 mm. 前 10 个功能图像的时间序列由于被试对环境的自适应性以及初始磁共振信号的不稳定性而被丢弃.

表 1 中的 a 值表示通过双值双尾 t 检验得到 P 值; b 值表示通过双尾 Pearson 卡方检验得到 P 值.用 SPM8<sup>1</sup> 来完成数据的预处理过程.首先,头动校正和时间片校正必不可少,而在校正过程中,抑郁组和对照组中分别出现 2 例转动大于 3 度或者头动大于 3 毫米的被试,因此弃除这些被试的扫描数据.需要注意的是,最后的 66 例样本数据中不包含那些被丢弃的数据.接着,图像经过优化仿射变换会被标准化到 MNI (Montreal Neurological Institute, MNI) 的标准空间中.最后,为降低生物高频噪音和低频漂移的影响,对数据进行带通滤波 (0.01-0.10Hz) 和线性降维.

脑区分割时使用 AAL<sup>[17]</sup> (Anatomical Automatic Labeling, AAL) 模板,得到 90 个感兴趣解剖区域 (Region Of Interest, ROI),它们均等分布在两个脑半球中,这一模板已广泛应用在同类研究中<sup>[18]</sup>.脑网络中的一个节点可以由一个感兴趣解剖区域表示,因此 90 个感兴趣解剖区域可以看成 90 个节

<sup>1</sup> <http://www.fil.ion.ucl.ac.uk/spm>.

点分布在脑网络当中. 首先, 按照超图的基本定义<sup>[19]</sup>, 提取各脑区的时间序列信号, 脑区之间的关联关系可通过计算获得; 然后, 由于头动校正和白质信号以及平均脑脊髓液会对信号产生影响, 因此需要排除它们的影响, 进行脑区时间序列的回归分析; 最后, 依据脑区的时间序列, 在两种方法模型下构建脑功能超网络.

## 2.2 超网络构建

变量选择一直是构建模型的热门方法, 自 LASSO 方法被提出以来, 各种变量选择方法层出不穷. 在对 LASSO 模型分析研究后发现, 最初人们只是着眼于其仅作为一种单变量选择方法, 无法解决组效应问题, 因而提出各种组级选择方法. 事实上, LASSO 模型不仅存在以上问题, 其惩罚函数对系数还具有过强压缩的特性, 从而导致模型中目标变量回归系数的有偏估计. 而随后的诸多方法则是在对 LASSO 方法缺乏组效应解释能力情况下的改进, 同样存在系数的过强压缩问题.

因此, 本文提出使用 MCP 惩罚函数, 该函数具有 LASSO 不具有的良好性质, 将其应用到组级, 既不会对系数过度压缩, 又能解决组效应问题. 基于 MCP 惩罚, 提出两种变量选择方法构建超网络, 分别是组 MCP 方法和复合 MCP 方法.

### 2.2.1 组 MCP 方法构建超网络

组 MCP 是将 MCP 惩罚应用到组级上形成的. MCP 是一种单一变量选择方法, 具有同步选择一致性和渐近无偏性<sup>[20]</sup>. 这种性质意味着该模型渐近等效于最大似然模型的拟合, 在该模型中, 事先已知真正非零系数的恒等式<sup>[21]</sup>. 将 MCP 扩展到组级, 就得到了组 MCP, 能在组级上对变量进行选择.

组 MCP 定义如下<sup>[14]</sup>:

$$\min_{\alpha_k} \|x_k - A_k \alpha_k\|_2^2 + \sum_{p=1}^P \varphi_{\lambda, \gamma}(\alpha_k^p) \quad (1)$$

其中响应向量  $x_k \in \mathbb{R}^N$  表示第  $k$  个 ROI 的平均时间序列.  $A_k = [x_1, \dots, x_{k-1}, 0, x_{k+1}, \dots, x_K] \in \mathbb{R}^{N \times P}$  表示其他 ROIs 的平均时间序列的矩阵 (不包括第  $k$  个 ROI), 将第  $k$  个 ROI 的响应向量置零.  $\alpha_k \in \mathbb{R}^P$  是一个权重向量, 体现了其余 ROIs 对第  $k$  个 ROI 的影响程度.  $\alpha_k$  中有零元素和非零元素两种数据, 零元素对应的 ROIs 表明它们与选定的第  $k$  个 ROI 不存在交互, 被认为这些脑区与选定脑区之间是相互独立的, 通过这种将相互独立的连接置零的方法, 有效表示出存在相互作用的脑区.  $\alpha_k$  通过聚类被分成了  $P$  个非重叠的组,  $\alpha_k^p$  表示第  $p$  个组的回归系数, 也就是第  $p$  个组的权重向量,  $\gamma > 1$  是用来修正  $\varphi_{\lambda, \gamma}(\cdot)$  的调整参数;  $\lambda \geq 0$  是惩罚参数.  $\varphi(\cdot)$  为 MCP 惩罚:

$$\varphi_{\lambda, \gamma}(\theta) = \begin{cases} \lambda |\theta| - \frac{|\theta|^2}{2\gamma}, & |\theta| \leq \gamma\lambda \\ \frac{1}{2}\gamma\lambda^2, & |\theta| > \gamma\lambda \end{cases} \quad (2)$$

$\theta = \alpha_k^p$ ,  $\gamma > 1$  是用来修正  $\varphi_{\lambda, \gamma}(\cdot)$  的调整参数;  $\lambda \geq 0$  是惩罚参数. 由于 MCP 模型对绝对值小于等于  $\gamma\lambda$  的回归系数进行压缩, 而对绝对值大于  $\gamma\lambda$  的回归系数不再压缩, 因此, 组 MCP 模型通过减小对目标变量回归系数的压缩来实现近似无偏估计<sup>[22]</sup>. 当  $\gamma \rightarrow \infty$  时, 求解的回归系数向量的稀疏性逐渐变小; 当  $\gamma \rightarrow 1$  时, 回归系数向量的稀疏性越来越大.

构建方法如下: 首先, 使用协调下降算法实现组 MCP 的稀疏线性回归; 然后, 将每一个 ROI 表示为一个节点, 通过变化  $\lambda$  的值来构建超边, 根据  $\alpha_k$  中的非零项构建超边, 通过在指定范围内变化  $\lambda$  值可以产生一定数目的超边; 最后, 将不同  $\lambda$  值构建出来的网络合并为一个, 最终的超网络是一个具有 90 个节点, 810 条超边的  $90 \times 810$  的矩阵.  $\gamma$  采用默认值 3.

### 2.2.2 复合 MCP 方法构建超网络

组 MCP 方法只能反映组级选择, 未体现组间选择, 因此引入复合 MCP<sup>[15]</sup> 方法, 它也是近似无偏稀疏模型. 通过将 MCP 惩罚同时用作内部惩罚和外部惩罚, 可以实现双级选择, 既能在组间选择变量, 又能选择组内的重要变量. 复合 MCP 方法的这种惩罚方式意味着它既可以在组间产生稀疏度, 又能在组内保持稀疏性, 因而可以同时选择组间变量和该组内的变量.

复合 MCP 方法定义如下<sup>[16]</sup>:

$$\min_{\alpha_k} \|x_k - A_k \alpha_k\|_2^2 + \sum_{p=1}^P \varphi_{\lambda, \gamma_2} \left( \sum_{j=1}^p \varphi_{\lambda, \gamma_1}(\alpha_k^{pj}) \right) \quad (3)$$

式中各参数与上述公式 (1) 中的相应参数有相同的含义.  $\varphi_{\lambda, \gamma}(\cdot)$  为 MCP 惩罚:

$$\varphi_{\lambda, \gamma}(\beta) = \begin{cases} \lambda |\beta| - \frac{|\beta|^2}{2\gamma}, & |\beta| \leq \gamma\lambda \\ \frac{1}{2}\gamma\lambda^2, & |\beta| > \gamma\lambda \end{cases} \quad (4)$$

$\eta$  可以是  $\gamma_1$  或  $\gamma_2$ .  $\gamma_1, \gamma_2$  分别是内部惩罚和外部惩罚的调整参数,  $\lambda \geq 0$  是惩罚参数.  $\alpha_k^{pj}$  表示的是第  $p$  组的第  $j$  个变量的回归系数,  $\beta = \alpha_k^{pj}$ .

复合 MCP 方法首先把 MCP 惩罚应用于各个组间变量选择上, 然后在组级加入 MCP 惩罚, 变量是否进入模型取决于模型自身以及它所属的组. 因此, 该变量也许进入模型, 也可能不进入模型而被消除. 实验过程与组 MCP 相似, 不同之处在于进行稀疏线性回归模型求解时的惩罚函数为复合 MCP, 通过在一定范围内变化  $\lambda$  的值可以构建一定数目的超边.  $\gamma_1$  和  $\gamma_2$  均采用默认值 3.

### 2.3 特征提取

构建功能连接超网络之后, 选择超网络中每个顶点的属性值作为特征. 对已构建的超网络进行指标计算, 聚类系数如今已广泛用于衡量网络的局部属性, 这 3 类指标统称为 HCC 指标, 该指标下有 3 种聚类系数, 分别具有不同的定义<sup>[23]</sup>, 分别记作 HCC1、HCC2 以及 HCC3.

$$HCC1(v) = \frac{2 \sum_{u, t \in N(v)} I(u, t, v)}{|N(v)| (|N(v)| - 1)} \quad (5)$$

$$HCC2(v) = \frac{2 \sum_{u, t \in N(v)} I'(u, t, v)}{|N(v)| (|N(v)| - 1)} \quad (6)$$

$$HCC3(v) = \frac{2 \sum_{e \in S(v)} (|e| - 1) - |N(v)|}{|N(v)| (|S(v)| - 1)} \quad (7)$$

公式 (5) 计算与节点  $v$  不存在连接的相邻节点的数目. 节点用  $u, t, v$  表示,  $N(v)$  是集合, 指超边中除了节点  $v$  以外, 还包括其它节点的集合. 如果  $E$  表示边集,  $e_i$  表示某一条超边, 则  $I(u, t, v) = 1$  当且仅当  $\exists e_i \in E, u, t \in e_i$ , 但  $v \notin e_i$ ; 否则  $I(u, t, v) = 0$ . HCC1 查找不包含  $u$  的邻居之间的连接, HCC1 的优点是, 在这个集合中发现的任何交互都可能表示邻居之间的真实连接. 它的缺点在于可能过分关注于那些次

要的共享连接,这些连接与  $u$  的交互没有什么关系。

公式(6)计算与节点  $v$  存在连接关系的相邻节点的数目,式中各相同参数的含义同上。如果  $\exists e_i \in E$  且  $u, t, v \in e_i$ , 则  $I'(u, t, v) = 1$ 。HCC2 查找那些包含  $u$  的邻居的连接,这种方式找到的边真实地反应了  $u$  和邻居之间的聚集。但是需要注意这种连接可能只是与  $u$  共享连接的人工数据。

公式(7)计算超边之间的重叠量,具体指节点  $v$  的相邻超边,式中各相同参数的含义同上。 $|e|$  表示超边中包含的节点数目; $S(v)$  指超边集合,这些超边中都含有节点  $v$ ,且  $S(v) = \{e_i \in E: v \in e_i\}$ 。通过邻域超边的重叠量来衡量邻域的密度。其分子表示与  $u$  相关联的超边的顶点个数。分母表示这种重叠可能的数目。HCC3 则通过邻域超边的重叠量来衡量邻域的密度。与以上两种定义都不同,它从节点的角度来定义重叠量。

超网络的局部聚类属性可由 HCC1、HCC2 以及 HCC3 3 种属性从不同的角度反映得出。在两种方法构建的超网络中,将每一种聚类系数作为特征,分别提取。每个被试都有 90 个脑区,为了简便,计算出被试的平均聚类系数(对 90 个脑区进行平均,分别记为平均 HCC1,平均 HCC2 以及平均 HCC3)。将这 3 种不同定义下的聚类系数应用于两种方法所构建的超网络,最后分别把 3 种聚类系数进行融合。

## 2.4 特征选择和分类

本文利用统计学方法进行特征选择。在样本量较小的情况下,使用 Kolmogorov-Smirnov (KS)<sup>[24]</sup> 非参数置换检验来分析两组数据是否具有差异是一种常用的方法。在本文中,对抑郁患者和正常人的超网络在计算 HCC 聚类系数后获得的 270 个节点属性进行组间 KS 非参数检验,并进行 FDR 校验( $q = 0.05$ ),通过非参数检验分别得出每个被试的 270 个脑区的  $p$  值大小。选取抑郁患者和正常人的  $p < 0.05$  的显著性特征,并将这些显著性差异特征作为分类特征,进行分类。

分类性能通过留一交叉验证进行评估。训练集使用 K 折交叉验证(K-fold Cross Validation, K-CV)对参数  $c$  和  $g$  进行寻优<sup>[25]</sup>。在训练集验证中, $c$  和  $g$  的最优值取分类准确率最高的那组值,并由此建立  $N$  个不同的模型。在标准化分类特征的平均值和标准偏差后,最后的分类结果取不同模型分类准确率的平均值。

## 3 结果和分析

### 3.1 超网络结构分析

为确定显著差异是否存在于不同方法构建的超网络中,对两种方法下的超网络结构进行了比较,并进行了如下的实验:

首先分别分析两种方法的正常人和抑郁患者的超网络的超边,计算它们的超边的边度。结果表明,两种方法下的超网络结构存在差异,具体说来,复合 MCP 方法构建的超网络的超边度大多数分布在 2-13 范围内;组 MCP 方法构建的关于正常人和抑郁患者的超网络的超边度分布范围较广,分布比较离散,超边度分布情况如图 1 所示。

在每一个被试的超网络中,计算出它们(对 90 个脑区进

行平均)的平均聚类系数,分别记作平均 HCC1,平均 HCC2 和平均 HCC3。对抑郁组和正常对照组进行非参数置换检验,

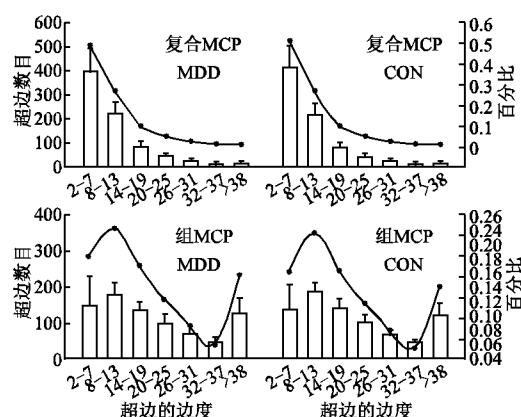


图 1 抑郁组和正常对照组的超边度分布图

Fig. 1 Hyper-edge distribution of depression group and normal control group

对于平均 HCC1,平均 HCC2 以及平均 HCC3,分别比较两种方法构建的超网络的差异。在两种方法所构建的超网络中,平均聚类系数的分布情况如图 2 所示。

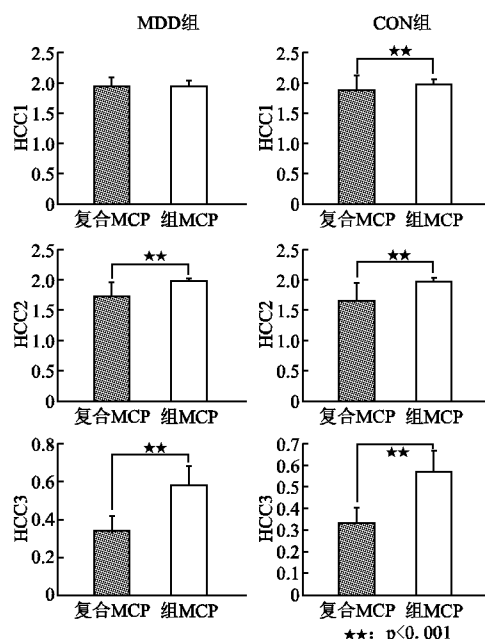


图 2 抑郁组和正常对照组的平均聚类系数

Fig. 2 Average clustering coefficient of depression group and normal control group

结果表明,无论是正常对照组还是抑郁组,两种方法下的平均聚类系数关于 HCC2 和 HCC3 分别具有显著性差异,而关于 HCC1 的显著性差异只存在于正常对照组中,抑郁组中则未发现显著性差异( $p > 0.05$ , FDR 校验,  $q = 0.05$ )。因此,两种方法构建的超网络存在结构上的差异。分析其潜在的原因,组 MCP 方法是在组级上进行变量选择,而复合 MCP 是双级选择,既在组级上进行变量选择,又选择组内重要变量,从而造成了超网络结构上的不同。

### 3.2 差异性脑区

在两种不同的方法下分别进行超网络构建并提取特征,对于每一个特征,进行非参数置换检验对所有被试评估抑郁症患者和正常人之间的差异.表2和表3分别列出了两种不同超网络构建方法下所得到的具有显著差异的大脑区域.

表2 组 MCP 方法下具有显著差异的大脑区域  
Table 2 Brain regions with significant differences under the group MCP method

脑区	p 值		
	HCC1	HCC2	HCC3
左侧背外侧额上回	<b>0.0142</b>	0.9582	0.6780
左侧内侧额上回	<b>0.0241</b>	0.5730	<b>0.0470</b>
右侧内侧额上回	<b>0.0135</b>	0.9000	0.0658
右侧脑岛	<b>0.0181</b>	0.7929	<b>0.0066</b>
右侧前扣带和旁扣带脑回	<b>0.0451</b>	0.0837	0.2521
左侧海马	<b>0.0165</b>	0.9990	0.5473
左侧海马旁回	0.4493	0.9973	<b>0.0432</b>
左侧舌回	<b>0.0413</b>	0.4610	0.2145
左侧梭状回	0.3219	0.9995	<b>0.0491</b>
左侧角回	<b>0.0331</b>	0.7431	0.6253
左侧中央旁小叶	0.5219	0.9635	<b>0.0241</b>
左侧颞中回	0.7683	0.5095	<b>0.0135</b>
左侧颞极:颞中回	0.0685	0.4610	<b>0.0264</b>

表3 复合 MCP 方法下具有显著差异的大脑区域  
Table 3 Brain regions with significant differences under the composite MCP method

脑区	p 值		
	HCC1	HCC2	HCC3
左侧中央前回	0.7929	<b>0.0142</b>	0.3033
右侧中央前回	<b>0.0181</b>	0.8394	0.8049
右侧额中回	0.6253	<b>0.0451</b>	0.5473
右侧补充运动区	<b>0.0362</b>	0.5473	0.7431
左侧内侧和旁扣带脑回	0.8504	<b>0.0413</b>	0.7302
右侧杏仁核	0.7302	<b>0.0181</b>	0.8609
左侧舌回	0.2601	<b>0.0142</b>	<b>0.0181</b>
左侧枕上回	<b>0.0362</b>	0.8049	0.5859
左侧顶上回	0.5729	0.0534	<b>0.0276</b>
左侧顶下缘角回	0.3033	0.0606	<b>0.0331</b>
右侧中央旁小叶	0.5473	<b>0.0451</b>	0.2855
左侧豆状壳核	0.8812	<b>0.0149</b>	0.8167
左侧颞极:颞上回	<b>0.0229</b>	0.8394	0.3033
左侧颞极:颞中回	<b>0.0451</b>	0.7043	0.7929

表中黑体字表示对应的聚类系数的 P 值  $< 0.05$ . 两种方法利用统计分析获得的差异性脑区也不相同,其中不重叠的脑区数目较多,主要包括左侧中央前回、右侧中央前回、右侧额中回、右侧补充运动区、左侧内侧和旁扣带脑回、右侧杏仁核、左侧枕上回、左侧顶上回、左侧顶下缘角回、右侧中央旁小叶、左侧豆状壳核和左侧颞极:颞上回.在后面的研究中,复合 MCP 方法得出了最高的分类准确率,因此本文对该方法下的差异性脑区进行讨论,且验证了复合 MCP 方法下的很多脑区都与已有文献一致,证明它们和抑郁症病理有着密切联系,具

体见表4.

表4 已有文献中关于抑郁症的脑区  
Table 4 Brain areas of depression in the literature

ROI	引文
右侧额中回	[郁仁强,2019] <sup>[26]</sup>
右侧补充运动区	[潘梦洁,2018] <sup>[27]</sup>
左侧内侧和旁扣带脑回	[郁仁强,2019] <sup>[26]</sup>
左侧舌回	[张帆,2018] <sup>[28]</sup>
左侧豆状壳核	[陈诚,2020] <sup>[29]</sup>
右侧中央旁小叶	[蔡秋艺,2018] <sup>[30]</sup>
双侧中央前回	[邹小平,2018] <sup>[31]</sup>
左侧枕上回	[李鹏,2018] <sup>[32]</sup>
右侧杏仁核	[Damborská,2020] <sup>[33]</sup>
左侧顶上回	[Lai,2016] <sup>[34]</sup>
颞极	[马平平,2019] <sup>[35]</sup>

### 3.3 分类表现

为评估两种超网络构建方法的分类性能,本文使用了分类准确率、敏感度以及特异度3个特定指标.除此之外,由于不平衡数据集所带来的膨胀性能,将特异度和敏感度的算术平均值定义为 BAC.

同时,将文中所提出的方法与广泛应用至脑疾病中的 LASSO 方法和组 LASSO 以及稀疏组 LASSO 方法进行了对比.同样也使用分类准确性、敏感度以及特异度3个指标.结果证明本文所提出的两种方法的分类准确度均优于已有方法.比较结果如表5所示.

表5 不同方法的分类表现

方法	准确率 (%)	灵敏度 (%)	特异度 (%)	BAC (%)
传统 LASSO 方法	83.33	84.21	82.14	83.18
组 MCP	86.36	82.14	86.84	84.49
复合 MCP	89.39	82.14	94.74	88.44
组 LASSO <sup>[36]</sup>	80.30	84.22	75.00	79.61
稀疏组 LASSO <sup>[36]</sup>	84.85	92.11	82.14	87.13

组 MCP 和复合 MCP 方法的分类准确率都能达到 85% 以上,其中复合 MCP 方法的分类准确率最高,能达到 89.39%.两种方法对变量是否进入模型采取了不同的解决办法,组 MCP 方法将惩罚函数运用在组级层面,解决了组效应问

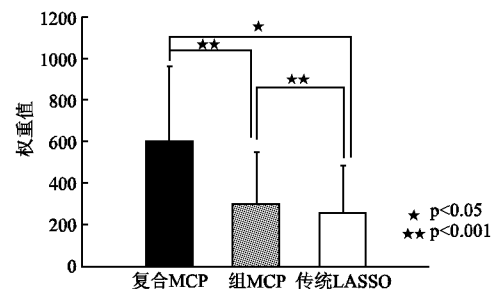


图3 两种方法的分类权重

Fig. 3 Classification weight of the two methods

题;而复合 MCP 方法则在组级和组内都应用了惩罚函数,既

选择了组间重要变量,又选择出组内重要的变量,因而是比组 MCP 更为严格的方法模型.以上结果表明在脑功能超网络构建中,组级选择方法能更好地体现人脑中固有的组结构,且双级选择能较好地体现人脑中复杂的层次关系.

此外,使用 Relief 算法<sup>[37]</sup>对特征的有效性进行验证.该方法会根据每个类别和功能相关性分配不同的权重,特征的分类能力越强,分配的权重越大(反之亦然).本文对两种方法以及传统 LASSO 方法对应的分类权重进行了比较,结果表明,两种方法的分类权重均高于 LASSO 方法,复合 MCP 方法具有最高的分类权重,结果显示在图 3 中.结果同样表明基于双级选择的复合 MCP 方法能构建合理有效的超网络.

## 4 参数讨论

### 4.1 分组参数 k 的讨论

两种方法都在预定义分组下进行超网络构建,实验中使用了 k-means<sup>[38]</sup>聚类算法,k 表示分组大小,可以将数据预先分成 k 个组,是预定义分组的重要基础.本实验以 3 为步长,进行组数不大于最大分组 90 的分组过程,分别在每一个 k 值下进行 50 次实验,选取正确率的算术平均值作为最后的分类结果.结果显示当组数 k=48 时,复合 MCP 方法表现出最高的正确率 89.394%.当 k=24 时,组 MCP 方法表现出最高正确率达到 86.364%.图 4 为两种方法在不同的 k 值下的正确率.

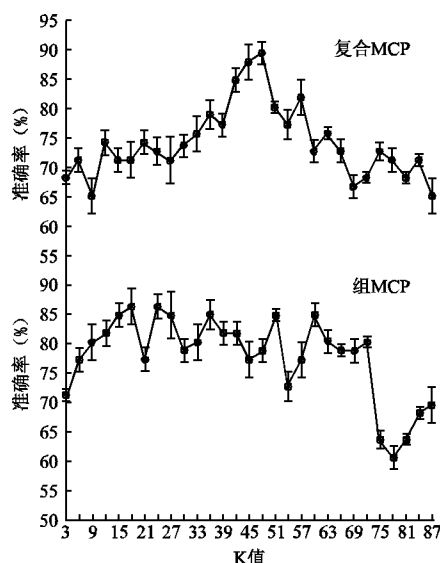


图 4 两种方法的不同 k 值下的准确率

Fig.4 Accuracy of the two methods under different k values

### 4.2 SVM 分类参数 c 和 g 的影响

SVM 分类器在各领域中的应用都比较广泛,而分类时则经常涉及到核函数的选取问题.由于 RBF 核函数应用广泛,无论是小样本还是大样本均适用,因此,在分类中选取 RBF 核函数. SVM 模型中有两个参数对分类影响较大,即惩罚因子 c 和核参数 g. c 表示调节优化方向中两个指标(间隔大小,分类准确度)偏好的权重,即对误差的宽容度, c 过大或过小,其泛化能力都会变差; g 隐含地决定了数据映射到新的特征空间后的分布, g 越大,支持向量的个数越少,反之则越多,而支持向量的个数影响训练与预测的速度.最优的 c 和 g 能使

SVM 的分类性能达到最佳.对于如何找出最佳的 c 和 g,首先选取某一组给定的 c 和 g 的值,利用 K-CV 方法得到在该组 c 和 g 下的分类准确率,然后不断更换 c, g 的值,最终取分类准确率最高的那组 c 和 g 作为最佳参数. c, g 两参数可在  $[2^{-8}, 2^8]$

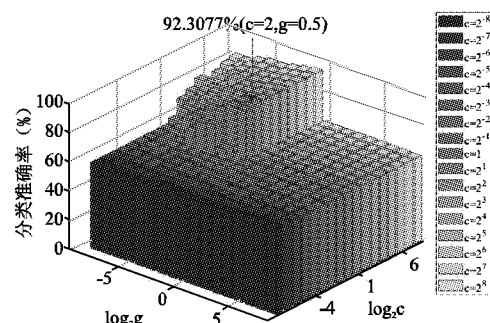


图 5 c 和 g 的参数寻优结果

Fig.5 Parameter optimization results of c and g

内变化,并将 1 设置为步长大小.结果显示,分类准确率最高时的 c, g 参数的值分别为 2 和 0.5,最高准确率是 92.308%,如图 5 所示.

## 5 结论

如今静息态功能磁共振成像 (R-fMRI) 技术已被广泛应用于预测抑郁症等精神疾病中.超网络因能充分反应各个脑区之间的互动而被广泛应用.现有的超网络构建多是基于稀疏表示方法,传统的基于 LASSO 的稀疏表示方法缺乏解释脑区之间的组效应的能力,而且惩罚函数对系数的过强压缩,导致模型中目标变量回归系数的有偏估计.考虑到这些问题,本文提出两种超网络构建方法:组 MCP 方法和复合 MCP 方法,前者只能在组级上进行变量选择;后者既能在组级上进行变量选择,又能选择组间的重要变量,称为双级选择.

实验结果表明,本文所提出的两种方法均优于 LASSO 方法.在超网络的结构上,两种方法之间存在差异,复合 MCP 构建的超网络的超边分布更为集中,能更好地体现人脑中复杂的层次关系;在分类性能上,本文所提出的两种方法均优于传统 LASSO 方法以及组 LASSO 和稀疏组 LASSO 方法,并能够获得较高的分类权重.

尽管在目前的研究中,实验已取得较好的结果,但仍存在局限性.首先,在实验中假设组间没有重叠的信息交互,但是在很多情况下,重叠组是存在的;其次,两种方法由于聚类初始种子点的随机选取以及聚类数 k 的不同会造成网络结构以及分类结果的不唯一,建立更加稳定的超边被期望以进一步改善超网络.

## References:

- [1] Zeng Ling-li, Shen Hui, Liu Li, et al. Identifying major depression using whole-brain functional connectivity: a multivariate pattern analysis[J]. Brain, 2012, 135(5): 1498-1507.
- [2] Zhang Qi, Wang Bin. Research progress on functional magnetic resonance imaging brain network for depression[J]. Magnetic Resonance Imaging, 2018, 70(4): 58-62.

- [3] Ni Bing-jie, Li Wei, Chen Xi. Modeling the evolution of brain networks in Alzheimer's disease[J]. Journal of Intelligent Science and Technology, 2019, 1(4): 54-63.
- [4] Chen Rong, Herskovits E H. Graphical-model-based multivariate analysis of functional magnetic resonance data[J]. NeuroImage, 2007, 35(2): 635-647.
- [5] Wu Xiao-jing. Research on the neural mechanism of face recognition based on the analysis of the functional connectivity of the resting brain[D]. Beijing: Beijing Jiaotong University, 2018.
- [6] Huang Shuai, Li Jing, Sun Liang, et al. Learning brain connectivity of Alzheimer's disease by sparse inverse covariance estimation[J]. NeuroImage, 2010, 50(3): 935-949.
- [7] Wee Chong-yaw, Yap Pew-thian, Zhang Dao-qiang, et al. Group-constrained sparse fMRI connectivity modeling for mild cognitive impairment identification[J]. Brain Structure & Function, 2014, 219(2): 641-656.
- [8] Alex Cayco-gajic N, Zylberberg Joel, Eric Shea-brown. A moment-based maximum entropy model for fitting higher-order interactions in neural data[J]. Entropy, 2018, 20(7): 489-505.
- [9] Li Yao, Sun Chao, Li Peng-zu, et al. Hypernetwork construction and feature fusion analysis based on sparse group lasso method on functional fMRI dataset[J]. Frontiers in Neuroscience, 2020, 14: 60-85, doi:10.3389/fnins.2020.00060.
- [10] Souly Nasim, Shah Mubarak. Visual saliency detection using group lasso regularization in videos of natural scenes[J]. International Journal of Computer Vision, 2016, 117(1): 93-110.
- [11] Piepho Hans-peter, Ogutu Joseph O. Regularized group regression methods for genomic prediction: bridge, MCP, SCAD, group bridge, group lasso, sparse group lasso, group MCP and group SCAD[J]. BMC Proceedings, 2014, 8(5): 1-9.
- [12] Furqan Mohammad-shaheryar, Siyal Mohammad-yakoob. Elastic-net copula granger causality for Inference of biological networks[J]. PLoS ONE, 2016, 11(10): e0165612- e0165628.
- [13] Teipel Stefan J, Grothe Michel J, Metzger Coraline D, et al. Robust detection of impaired resting state functional connectivity networks in Alzheimer's disease using elastic net regularized regression[J]. Frontiers in Aging Neuroscience, 2016, 8(1): 318-327.
- [14] Li Peng-xiang, Ye Wan-zhou. Sparse group variable selection based on two non-convex penalty functions[J]. Journal of Applied Mathematics and Computation, 2018, 32(3): 437-446.
- [15] Huang Jian, Breheny Patrick, Ma Shuang-ge. A selective review of group selection in high-dimensional models[J]. Statistical Science, 2012, 27(4): 481-499.
- [16] Li Chun-hong, Zhong Xiao-min, Zong Rui-xue. Several types of group variable selection methods and their block coordinate descent algorithm[J]. Advances in Applied Mathematics, 2019, 8(8): 1478-1486.
- [17] Chen Qiang, Ursini Gianluca, Romer Adrienne L, et al. Schizophrenia polygenic risk score predicts mnemonic hippocampal activity[J]. Brain, 2018, 141(4): 1218-1228.
- [18] Hu Ying, Wang Li-jia, Nie Sheng-dong. A review of brain function division in resting state functional magnetic resonance imaging[J]. Chinese Journal of Image and Graphics, 2017, 22(10): 1325-1334.
- [19] Peng Yao, Zu Chen, Zhang Dao-qiang. Multi-modal feature selection algorithm based on hypergraph and its application[J]. Journal of Computer Science and Exploration, 2018, 12(1): 112-119.
- [20] Li Ning, Yang Hu. Nonnegative estimation and variable selection under minimax concave penalty for sparse high-dimensional linear regression models[J]. Statistical Papers, 2019, 62(2): 661-680.
- [21] Breheny Patrick, Huang Jian. Group descent algorithms for nonconvex penalized linear and logistic regression models with grouped predictors[J]. Statistics and Computing, 2015, 25(2): 173-187.
- [22] Zhang Cun-hui. Nearly unbiased variable selection under minimax concave penalty[J]. The Annals of Statistics, 2010, 38(2): 894-942.
- [23] Gallagher Suzanne, Goldberg Debra. Clustering coefficients in protein interaction hypernetworks[C]//Proceedings of the International Conference on Bioinformatics, Computational Biology and Biomedical Informatics, ACM, 2013: 552-560.
- [24] Zhang Bin-bin, Wang Li-ping, Wu Jun, et al. Spindle radial accuracy evaluation and fault diagnosis based on a normal distribution[C]//12th International Conference on Reliability, Maintainability, and Safety(ICRMS), Shanghai, China, IEEE, 2018: 307-311.
- [25] Mishra Shailendra-kumar, Deepthi V Hima. Brain image classification by the combination of different wavelet transforms and support vector machine classification[J]. Journal of Ambient Intelligence and Humanized Computing, 2020, 12(3): 1-9.
- [26] Yu Ren-qiang, Zhang Zhi-wei, Lv Fa-jin, et al. The default network function connection of patients with first-onset major depression who have not been treated by resting state functional magnetic resonance studies[J]. Western Medicine, 2019, 31(4): 122-127.
- [27] Pan Meng-jie, Chen Feng, Lin Ming-fang, et al. Study of resting state functional magnetic resonance imaging based on low-frequency amplitude in major depression[J]. Chinese Journal of Modern Nervous Diseases, 2018, 18(3): 171-176.
- [28] Zhang Fan, Chen Jun-jie, Guo Hao. Multi-feature fusion classification method based on brain function super network[J]. Computer Engineering and Applications, 2018, 54(21): 120-127.
- [29] Chen Cheng, Wang Hui-ling, Wang Gao-hua, et al. Differences between resting state functional magnetic resonance imaging and cognitive function in patients with depression and schizophrenia[J]. Chinese Journal of Psychiatry, 2020, 53(5): 377-383.
- [30] Cai Qiu-yi, Xie Jun-shi, Zou Yu, et al. Research on functional connectivity of resting state fMRI in patients with depression and hypomania[J]. Chinese Journal of Medical Computer Imaging, 2018, 24(1): 14-20.
- [31] Wu Xiao-ping, Liu Chun-ying, Zhang Hong, et al. Study on functional connection of the banded brain network in major depression under resting MRI[J]. Clinical Medical Research and Practice, 2018, 3(12): 5-7.
- [32] Li Peng, Wang Chang-ming, Li Feng, et al. Study on the correlation between the resting state of the lower brain area and the severity of the disease by functional magnetic resonance imaging in patients with premature depression[J]. China Medical Equipment, 2018, 15(6): 61-66.
- [33] Damborska Alena, Rubega Maria, Honzirkova Eliska, et al. Altered directed functional connectivity of the right amygdala in depression: high-density EEG study[J]. Scientific Reports, 2020, 10(1): 4398-4422.
- [34] Lai Chien-han, Wu Yu-te, Chen Cheng-yu, et al. Gray matter increases in fronto-parietal regions of depression patients with aripiprazole treatment[J]. Journal of Affective Disorders, 2019, 245: 112-119.

- prazole monotherapy[J]. *Medicine*, 2016, 95(34): e4654-e4660.
- [35] Ma Ping-ping. Comparison of rs-fMRI of depression, bipolar disorder, and schizophrenia and its relationship with cognitive function and clinical features[D]. Kunming: Kunming Medical University, 2019.
- [36] Li Yao, Zhao Yun-peng, Li Xin-yun, et al. Construction of brain function super network and feature fusion analysis based on sparse group Lasso method[J]. *Journal of Computer Applications*, 2020, 40(1): 62-70.
- [37] Palma-Mendoza Raul-Jose, Rodriguez Daniel, DE-Marcos Luis. Distributed ReliefF based feature selection in spark[J]. *Knowledge & Information Systems*, 2018, 57(19): 1-20.
- [38] Hartono, Sitompul O S, Tulus, et al. Optimization model of K-means clustering using artificial neural networks to handle class imbalance problem[J]. *Iop Conference*, 2018, 288: 012075-012083, doi:10.1088/1757-899x/288/1/012075.
- 附中文参考文献:**
- [2] 张琪, 王滨. 抑郁症功能磁共振成像脑网络研究进展[J]. *磁共振成像*, 2018, 70(4): 58-62.
- [3] 倪冰洁, 李炜, 陈曦. 阿尔兹海默症脑网络演化建模[J]. *智能科学与技术学报*, 2019, 1(4): 54-63.
- [5] 吴筱菁. 基于静息态脑功能连接性分析的面孔识别神经机制研究[D]. 北京: 北京交通大学, 2018.
- [14] 李鹏翔, 叶万洲. 基于两种非凸惩罚函数的稀疏组变量选择[J]. *应用数学与计算数学学报*, 2018, 32(3): 437-446.
- [16] 李春红, 钟小敏, 宗瑞雪. 几类群组变量选择方法及其块坐标下降算法[J]. *应用数学进展*, 2019, 8(8): 1478-1486.
- [18] 胡颖, 王丽嘉, 聂生东. 静息态功能磁共振成像的脑功能分区综述[J]. *中国图象图形学报*, 2017, 22(10): 1325-1334.
- [19] 彭瑶, 祖辰, 张道强. 基于超图的多模态特征选择算法及其应用[J]. *计算机科学与探索*, 2018, 12(1): 112-119.
- [26] 郁仁强, 张志伟, 吕发金, 等. 静息态功能磁共振研究未治疗首次发作重症抑郁症患者默认网络功能连接[J]. *西部医学*, 2019, 31(4): 122-127.
- [27] 潘梦洁, 陈峰, 林明方, 等. 重症抑郁症基于低频振幅的静息态功能磁共振成像研究[J]. *中国现代神经疾病杂志*, 2018, 18(3): 171-176.
- [28] 张帆, 陈俊杰, 郭浩. 基于脑功能超网络的多特征融合分类方法[J]. *计算机工程与应用*, 2018, 54(21): 120-127.
- [29] 陈诚, 王慧玲, 王高华, 等. 抑郁症与精神分裂症患者静息态功能磁共振成像与认知功能的差异[J]. *中华精神科杂志*, 2020, 53(5): 377-383.
- [30] 蔡秋艺, 谢俊诗, 邹瑜, 等. 静息态 fMRI 对抑郁症伴嗅觉功能减退患者功能连接的研究[J]. *中国医学计算机成像杂志*, 2018, 24(1): 14-20.
- [31] 邹小平, 刘春英, 张宏, 等. 静息态磁共振下重度抑郁症带状盖脑网络功能连接的研究[J]. *临床医学研究与实践*, 2018, 3(12): 5-7.
- [32] 李鹏, 王长明, 李峰, 等. 首发抑郁患者功能磁共振成像对下脑区静息态自发活动与疾病严重程度的相关性研究[J]. *中国医学装备*, 2018, 15(6): 61-66.
- [35] 马平平. 抑郁症, 双相情感障碍, 精神分裂症的 rs-fMRI 比较及与认知功能和临床特征的关系[D]. 昆明: 昆明医科大学, 2019.
- [36] 李瑶, 赵云芃, 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.