



文章编号:1001-9081(2019)07-1948-06

DOI:10.11772/j.issn.1001-9081.2018112421

基于多层次模板的脑功能网络特征选择及分类

吴浩¹, 王昕璨², 李欣芸¹, 刘志芬³, 陈俊杰¹, 郭浩^{1*}

(1. 太原理工大学 信息与计算机学院, 山西 晋中 030600; 2. 太原理工大学 艺术学院, 山西 晋中 030600;

3. 山西医科大学第一医院 精神卫生科, 太原 030000)

(*通信作者电子邮箱 feiyu_guo@sina.com)

摘要: 基于单一脑图谱模板的功能连接网络中提取的特征表示不足以揭示患者组和正常对照组(NC)之间的复杂拓扑结构差异,而传统的基于多模板的功能脑网络定义多采用独立模板,缺乏模板间的关联,从而忽略了各模板构建的功能脑网络中潜在的拓扑关联信息。针对上述问题,提出了一种多层次脑图谱模板和一种使用关系诱导稀疏(RIS)特征选择模型的方法。首先定义了具有关联的多层次脑图谱模板,挖掘模板之间潜在关系和表征组间网络结构差异;然后用RIS特征选择模型进行参数优化,进而提取组间差异特征;最后利用支持向量机(SVM)方法构建分类模型,并应用于抑郁症患者的诊断。在山西大学第一医院抑郁症临床诊断数据库上的实验结果显示,基于多层次模板的功能脑网络通过使用具有RIS特征的选择方法取得了91.7%的分类准确率,相比传统多模板方法的准确率提高了3个百分点。

关键词: 多层次模板; 功能脑网络; 关系诱导稀疏; 机器学习; 抑郁症

中图分类号: TP181 **文献标志码:** A

Brain function network feature selection and classification based on multi-level template

WU Hao¹, WANG Xincan², LI Xinyun¹, LIU Zhifen³, CHEN Junjie¹, GUO Hao^{1*}

(1. College of Information and Computer Science, Taiyuan University of Technology, Jinzhong Shanxi 030600, China;

2. College of Art, Taiyuan University of Technology, Jinzhong Shanxi 030600, China;

3. Department of Mental Health, First Hospital of Shanxi Medical University, Taiyuan Shanxi 030000, China)

Abstract: The feature representation extracted from the functional connection network based on single brain map template is not sufficient to reveal complex topological differences between patient group and Normal Control (NC) group. However, the traditional multi-template-based functional brain network definitions mostly use independent templates, ignoring the potential topological association information in functional brain networks built with each template. Aiming at the above problems, a multi-level brain map template and a method of Relationship Induced Sparse (RIS) feature selection model were proposed. Firstly, an associated multi-level brain map template was defined, and the potential relationship between templates and network structure differences between groups were mined. Then, the RIS feature selection model was used to optimize the parameters and extract the differences between groups. Finally, the Support Vector Machine (SVM) method was used to construct classification model and was applied to the diagnosis of patients with depression. The experimental results on the clinical diagnosis database of depression in the First Hospital of Shanxi University show that the functional brain network based on multi-level template achieves 91.7% classification accuracy by using the RIS feature selection method, which is 3 percentage points higher than that of traditional multi-template method.

Key words: multi-level template; functional brain network; relationship induced sparse; machine learning; Major Depressive Disorder (MDD)

0 引言

抑郁症(Major Depressive Disorder, MDD)是最常见的精神疾病之一。了解MDD的神经基础对于改善这种使人衰弱的疾病的预防、检测和治疗至关重要^[1]。大量神经影像学研

究已经发现并指出MDD患者在结构和功能脑区间存在异常^[2]。同时,基于脑网络的方法在精神障碍的机器学习分类中发挥着越来越重要的作用^[3]。例如,这些技术可以用于诊断或预测治疗结果。许多研究使用支持向量机(Support Vector Machine, SVM)分类器,以识别患有MDD的被试与正

收稿日期:2018-12-07;修回日期:2019-01-07;录用日期:2019-01-08。 基金项目:国家自然科学基金资助项目(61672374, 61741212, 61876124, 61873178);山西省科技厅应用基础研究项目青年面上项目(201601D021073);山西省教育厅高等学校科技创新研究项目(2016139);教育部赛尔网络下一代互联网技术创新项目(NGII20170712);国家留学基金资助出国留学项目(201708140216)。

作者简介: 吴浩(1994—),女,山西临汾人,硕士研究生,CCF会员,主要研究方向:人工智能、智能信息处理与脑影像学;王昕璨(1993—),女,山西太原人,硕士研究生,主要研究方向:人工智能、智能信息处理与脑影像学;李欣芸(1999—),女,山西太原人,主要研究方向:人工智能;刘志芬(1981—),女,山西太原人,副主任医师,博士,主要研究方向:情感障碍基础与临床研究;陈俊杰(1956—),男,河北定州人,教授,博士,主要研究方向:人工智能、智能信息处理与脑影像学;郭浩(1981—),男,山西祁县人,副教授,博士,CCF会员,主要研究方向:人工智能、智能信息处理与脑影像学。



常对照组 (Normal Control, NC) 的网络指标^[4]。

基于静息态功能磁共振成像的分类研究已经在 MDD 的诊断中被广泛研究^[5-8]。在传统的静息态功能磁共振成像分析中,通常利用单一脑图谱模板划分脑区进而构建功能连接网络,再进行之后的分类研究^[9]。然而,利用单个脑图谱模板提取的功能脑网络拓扑特征可能不足以揭示受疾病影响的患者组和 NC 组之间的潜在差异。具体地,不同的脑图谱模板对所构建网络的结构及其拓扑属性影响很大,而网络的不同分割节点数量实质上对网络的小世界属性、局部属性、功能连接强度以及网络连通性均存在显著影响^[10]。此外,脑图谱模板的不同对网络的影响还体现在以网络拓扑属性作为特征的分类中。

为了解决上述问题,有研究提出可以利用多种不同的脑图谱模板来更丰富地表示被试的分类特征^[11]。在减少误差、提取更丰富的分类特征表示以及提高分类准确率上取得了相比单模板更好的成效。研究表明,基于多图谱模板的方法通常可比单模板的方法实现更准确的诊断。

目前,这个领域仍存在许多值得探索的问题,其中最重要的问题便是多模板的选取。以往基于多模板脑网络的研究在选择模板时忽略脑图谱模板之间的关联性。基于此,本文通过定义多层次模板,建立了不同模板的层次关联性,挖掘出了多个模板间脑区的对应关系,表征了具有关联性的多模板对于网络拓扑属性的显著影响,以针对现有在脑网络领域中对多模板划分研究的不足,旨在突破神经影像学在多模板研究上的壁垒,为 MDD 数据的分类研究提供新的参考方案。

此外,该领域另外一个问题便是如何提取多个模板中的多组分类特征的相关关系,以便从多组特征中筛选出更加有效的分类特征,从而取得更高的分类准确率。然而,大多数现有的基于多模板的方法简单地平均或连接从不同模板提取的多组特征,或者仅仅对比分析不同模板所得到的特征有效性及特征间冗余性^[12-13]。这可能忽略多模板数据中包含的重要结构信息,如何更加有效地筛选多个模板中的多组特征是一个具有挑战性的问题。基于此,本文提出参数自寻优的关系诱导稀疏 (Relationship Induced Sparse, RIS) 多模板特征选择方法。这一方法明确模拟出多模板中的结构信息,降低了多模板特征间的冗余性,提高分类特征的有效性,并被广泛应用在同类研究中^[14]。值得注意的是,本文提出参数自寻优框架,优化了 RIS 模型,旨在获得参数的最优组合,防止了参数的过度拟合,提高了分类器的泛化性能,使特征选择和分类更加准确和有效。

本文定义多层次脑图谱模板,分析模板之间、被试之间的潜在关系,利用使用参数自寻优框架的 RIS 特征选择方法,提取组间差异特征。结果显示,本文所提方法具有更高的分类准确率。

1 材料与方法

1.1 数据采集及预处理

实验的采集工作,在遵照山西省医学伦理委员会的建议下,所有被试均知情同意,并根据赫尔辛基宣言签署了书面知情同意书。被试包括 38 名 MDD 患者和 28 名健康右利手志愿者,被试基本信息如表 1。其脑部静息态磁共振成像扫描工作使用西门子 3T 超导核磁共振成像扫描仪,所有扫描由山西大学第一医院熟悉磁共振的放射科医生进行。其间,要求

被试在扫描过程中闭眼,呈放松状态,不进行思考活动且不能进入睡眠状态。扫描参数设置如下:33 axial slices, Repetition Time (TR) = 2 000 ms, Echo Time (TE) = 30 ms, thickness/skip = 4/0 mm, Field Of View (FOV) = 192 × 192 mm, matrix = 64 × 64 mm, flip angle = 90°, 248 volumes。

表 1 被试基本信息

Tab. 1 Basic information of subjects

类别	对照组	MDD 组	P 值
年龄	17 ~ 51 (26.6 ± 9.4)	17 ~ 49 (28.4 ± 9.68)	0.44 ^a
性别(男/女)	13/15	15/23	0.57 ^b
利手(右/左)	28/0	38/0	—
HAMD	N/A	15 ~ 42 (22.8 ± 13.3)	—

表 1 中的数据范围是最小值 ~ 最大值(平均值 ± 标准偏差)。HAMD 为 24 个汉密尔顿量表,a 表示双值双尾 T 检验得到 P 值,b 表示通过双尾 Pearson 卡方检验得到 P 值。

本文使用 Matlab 平台下的统计参数映射 (Statistical Parametric Mapping, SPM) 对实验采集来的图像进行预处理,主要包括如图 1 所示处理步骤^[15]。

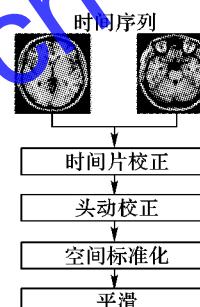


图 1 数据预处理流程

Fig. 1 Flow chart of data preprocessing

其中,时间片校正用来校正功能核磁共振成像 (functional Magnetic Resonance Imaging, fMRI) 影像扫描中各层之间采集时间的差异。头动校正即为对影像进行刚体变换并重排至同一被试的某一参考影像中,以形成新的影像序列。在此过程中,2 例 MDD 组及 NC 组数据由于校正过程中头动大于 3 mm 或转动大于 3° 而被弃除。在空间标准化步骤中,本文选择了蒙特利尔神经学研究所 (Montreal Neurological Institute, MNI) 标准的回声平面成像 (Echo Planer Imaging, EPI) 脑模板,体素大小选取为 3 mm。空间平滑则是为了进一步消除被试脑形态结构的细微差异。

1.2 网络构建和网络属性

1.2.1 多层次模板构建

本文挖掘出了不同脑图谱模板间脑区的对应关系,重点突出多个模板之间的关联性,构建了 AAL90、Parc256、Parc497、Parc1003、Parc1501 这五个多层次模板。

其中,AAL90 模板是自动解剖标记图谱模板 (Automated Anatomical Labeling, AAL)^[15],将大脑划分为 90 个脑区 (左右每个半球各 45 个脑区),每个脑区代表脑网络中的一个节点,且节点的值是通过计算各个脑区中所包含的所有体素的算术平均值来表示的。其余四个模板,是在 AAL90 模板基础上基于动态随机种子的方法细分大脑区域得到的。具体方法如下。



首先,将预期节点数用 N 表示,在此预设 N 为250。将原有AAL模板中的各脑区占所有脑区的体素比例用 V 表示。原有AAL脑区可划分的子区域的个数用 k 表示,其值可通过 $k = VN$ 计算得到。也就是说,在 N 个节点的规模下,脑区应该被划分为 k 个子区域,然后,针对该脑区,设置 k 个随机种子体素,用 $S = \{S_1, S_2, \dots, S_k\}$ 来表示。之后分别计算各种子体素 S_i 与其他剩余体素的距离。从中选取距离当前种子体素最小的体素,将其与种子体素 S_i 组合,形成包含该新体素的新子区域。接着计算这些新子区域的物理中心,并设置这个物理中心为下一个种子体素,然后计算体素间与种子体素的距离。循环往复此过程,直到脑区内所有未被划分的体素均被划分到其中。不同脑区的随机种子体素的设置及脑区划分都是各自独立完成的,也就是说,一个种子体素不可能同时属于两个脑区,即避免存在某个划分的脑区在空间上被其他划分的脑区分割的情况。待所有AAL独立脑区划分完成即得到一个新的脑图谱模板。

值得注意的是,可划分子区域的个数 k 是通过计算得到的,因此这个值在计算中可能不是一个整数。在这种情况下,本文通过四舍五入的方式来解决,所以,实际得到新的脑图谱模板的节点个数与预设的节点数 N 会存在细微的偏差。

AAL标准模板下划分的脑区有相应的名字和编号,以90个节点的脑图谱模板下的1号脑区(即左侧中央前回)为例,若在256个节点的模板下其被划分为了5个小脑区,则这5个小脑区均命名为左侧中央前回,与AAL模板中的命名一致,而没有对其重新命名。

同样地,497、1003、1501三个模板也是在此方法下得到的。这五个具有层次关联的脑图谱模板被标记为AAL90、Parc256、Parc497、Parc1003和Parc1501。前缀AAL旨在表示原始的AAL模板。前缀Parc表示使用上述算法确定的模板。模板AAL90划分的每个脑区对应到模板Parc256、Parc497、Parc1003和Parc1501脑区个数分别为6、5、10、15,不同模板的层次关联性如多层次模板节点关联示意图如图2所示。

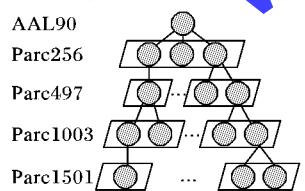


图2 多层次模板下节点关联示意图

Fig. 2 Schematic diagram of node association under multi-level template

以L(Left)、R(Right)表示左、右半脑,不同颜色代表不同脑区,则对AAL模板对应的脑区划分后得到多层次模板脑区的示意图如图3所示。

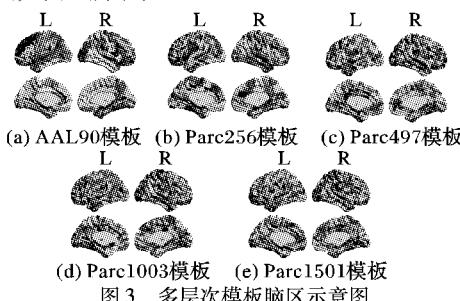


图3 多层次模板脑区示意图

Fig. 3 Schematic diagram of multi-level template brain area

连接即为节点间的边。连接的定义使用传统的方法,即

皮尔逊相关法。提取每个脑区血氧水平依赖(Blood Oxygenation Level Dependent, BOLD)的时间序列,即节点的值,对节点的值做皮尔逊相关,得出的值可以反映出节点之间的交互关系^[16]。

1.2.2 稀疏度的选择

本文使用稀疏度来构建二值网络^[17]。脑网络的节点数为 N 时,其对应的完全图的边数为 $N(N - 1)/2$,稀疏度就是脑网络中现存的边数与其完全图边数 $N(N - 1)/2$ 的比值。稀疏度可以控制脑网络的密度,与阈值的作用一样,但区别在于稀疏度是对网络整体的描述,而阈值控制的是网络中的边^[18]。

选取稀疏度时要保证最后的二值网络具有小的噪声,且没有孤立点的存在^[19]。有研究表明,当稀疏度大于40%时,脑网络会退化为随机网络^[20]。故为了使脑网络具有小世界属性,且无孤立点的存在,将稀疏度的范围设置为5%到40%,以0.05为步长将加权网络转化为二值网络。最后,所有被试的每个脑图谱模板下的脑网络都得到了8个不同稀疏度的二值网络。

1.3 特征提取和选择

1.3.1 特征定义

在本文中,选用三个局部指标作为分类特征,包括度、中间中心度和节点效率^[21]。针对不同的脑图谱模板,分别计算MDD组和NC组被试网络指标曲线下面积(Area Under the Curve, AUC)值^[20, 21-23],选取其经过T检验后具有统计显著性(即P值小于0.05)的作为特征。之所以计算网络指标的AUC值,是因为网络指标的计算并不是在某一确定稀疏度下,而是在一个选定的稀疏度空间。为了表征指标在完整稀疏度空间下的整体特性,本文计算每个指标的AUC值,以反映测量的网络节点属性在不同稀疏度下总的变化强度^[20]。AUC示意图如图4所示。

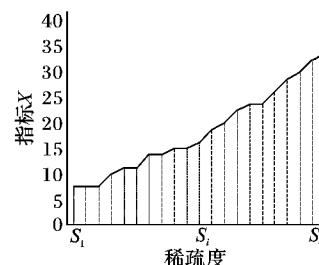


图4 AUC示意图

Fig. 4 Schematic diagram of AUC

为了找到特征的最优子集,避免过度拟合,提升模型性能,更快地训练分类器,就需要对特征进行筛选^[24-25]。本文采用RIS特征选择模型对特征进行筛选。

1.3.2 RIS特征选择模型

显而易见,在多个脑图谱模板下提取的特征数量要高于单一模板。然而,仅仅连接多个模板特征的方法并未取得更高的分类准确率。其原因可能是由于特征数量的增多,多模板特征间的冗余性提高,使得特征的有效性降低,因此,考虑到多模板特征之间的相似程度,使用RIS特征选择模型对提取的特征进行筛选。

其次,多模板脑网络中包含有很多重要的结构信息。然而,其他基于多模板脑网络的方法忽略了这一先验信息。RIS特征选择模型旨在明确地模拟多模板脑网络中的结构信息,并且在此基础上进行最优的参数设置,从而在MDD组与NC组分类中达到更高的分类准确度。具体而言,首先将从5个



脑图谱模板中的多组网络属性作为特征提取出来。接下来,通过引入两个正则化项来建立一种新颖的特征选择算法,以模拟模板之间和各个被试之间的关系。也就是说,通过 RIS 模型挖掘模板之间以及被试之间潜在的关系,以进一步筛选特征。最后,通过 SVM 分类器的输出,以实现最终结果。下面首先简要介绍 RIS 特征选择模型,然后介绍模型的推导过程。

RIS 特征选择模型主要由 4 个部分组成:1)训练数据经验损失;2)组稀疏正则化项;3)模板关系诱导正则化项;4)被试关系诱导正则化项。其中:1)、2)两个部分在于决定特征选择的数量,以保证在多组特征中选择出更加有效的分类特征;3)用于提取多个模板之间的相关关系;4)用于提取不同被试之间的相关关系。

RIS 特征选择模型的目标函数可写如下:

$$\begin{aligned} \min_{\mathbf{W}} \sum_{k=1}^K & \| \mathbf{Y} - \mathbf{X}^k \mathbf{w}^k \|_2^2 + \lambda_1 \| \mathbf{W} \|_{1,1} + \\ & \lambda_2 \sum_{n=1}^N \text{tr}((\mathbf{B}_n \mathbf{W})^T \mathbf{L}_n (\mathbf{B}_n \mathbf{W})) + \lambda_3 \sum_{k=1}^K (\mathbf{X}^k \mathbf{w}^k)^T \mathbf{L}^k (\mathbf{X}^k \mathbf{w}^k) \end{aligned} \quad (1)$$

首先,介绍式(1) 中的前两项,即:

$$\min_{\mathbf{W}} \sum_{k=1}^K \| \mathbf{Y} - \mathbf{X}^k \mathbf{w}^k \|_2^2 + \lambda_1 \| \mathbf{W} \|_{1,1} \quad (2)$$

式(2) 表示多任务特征学习模型。在这里,将每个模板看作一个任务,K 表示研究中有 K 个模板,即 K 个学习任务。 $\mathbf{X}^k = [\mathbf{x}_1^k, \mathbf{x}_2^k, \dots, \mathbf{x}_n^k, \dots, \mathbf{x}_N^k]^T \in \mathbf{R}^{N \times d}$ 表示包含全部 N 个被试的第 k 个学习任务(对应于第 k 个模板) 的训练数据,其中 $\mathbf{x}_n^k \in \mathbf{R}^d$ 表示第 k 个模板空间中的第 n 个被试的特征向量。同样地, $\mathbf{Y} = [y_1, y_2, \dots, y_n, \dots, y_N]^T \in \mathbf{R}^N$ 作为训练数据 \mathbf{X}^k 的响应向量。其中 $y_n \in \{-1, 1\}$ 是第 n 个被试的类别标签(即 MDD 组和 NC 组),在这里,-1 表示该被试为 MDD 患者,1 表示该被试为正常对照组。 $\mathbf{W} = [\mathbf{w}^1, \mathbf{w}^2, \dots, \mathbf{w}^k, \dots, \mathbf{w}^K] \in \mathbf{R}^{d \times K}$ 表示权重矩阵。其中, $\mathbf{w}^k \in \mathbf{R}^d$ 为第 k 个任务的参数化线性判别函数。式(2) 中的第一项是训练数据的经验损失,第二项 $\| \mathbf{W} \|_{1,1} = \sum_{i=1}^d \| \mathbf{W} \|_1$,是矩阵 \mathbf{W} 中行 ℓ_1 的范数的总和。该项有助于在不同模板中选择少量的特征。 λ_1 是用来平衡该项在式(1) 中的相对贡献,并且 λ_1 越大,选择的特征数量越少。

接下来,介绍表示模板关系诱导正则化项,即:

$$\sum_{n=1}^N \sum_{k_1=1}^K \sum_{k_2=1}^K (f(\mathbf{x}_n^{k_1}) - f(\mathbf{x}_n^{k_2}))^2 = \sum_{n=1}^N \text{tr}((\mathbf{B}_n \mathbf{W})^T \mathbf{L}_n (\mathbf{B}_n \mathbf{W})) \quad (3)$$

在式(3) 中,线性映射函数(也就是 $f(\mathbf{x}) = \mathbf{x}^T \mathbf{w}$) 被用来将原始高维特征空间的数据转换为一维标签空间。Liu 等认为,同一被试在不同模板中的特征在被映射到类标签后,它们应该彼此靠近。同样地,不同被试在同一模板中的特征 $\mathbf{x}_{n_1}^k$ 和 $\mathbf{x}_{n_2}^k$ 若相似,则说明这两个被试的估计类标签(即 $f(\mathbf{x}_{n_1}^k)$ 和 $f(\mathbf{x}_{n_2}^k)$) 也是相似的。在这里,tr(·) 表示方阵的轨迹, $\mathbf{B}_n = [\mathbf{x}_1^n, \mathbf{x}_2^n, \dots, \mathbf{x}_n^n, \dots, \mathbf{x}_N^n]^T \in \mathbf{R}^{N \times d}$ 表示从第 n 个被试的 K 个模板中导出的多组特征,并且, $\mathbf{L}_n \in \mathbf{R}^{K \times K}$ 是对角元素为 K-1,其他元素为 -1 的对角矩阵。

然后,介绍表示被试关系诱导正则化项,即:

$$\sum_{k=1}^K \sum_{n_1=1}^N \sum_{n_2=2}^N S_{n_1, n_2}^k (f(\mathbf{x}_{n_1}^k) - f(\mathbf{x}_{n_2}^k))^2 = \sum_{k=1}^K (\mathbf{X}^k \mathbf{w}^k)^T \mathbf{L}^k (\mathbf{X}^k \mathbf{w}^k) \quad (4)$$

这里的 \mathbf{x}^k 是如上所述的第 k 个任务(即第 k 个模板) 中的特征数据矩阵,并且, $S^k = \{S_{n_1, n_2}^k\}_{n_1, n_2}^N$ 表示第 k 个模板中的 N 个训练对象之间相似性的元素相似矩阵。在这里, $\mathbf{L}^k = \mathbf{D}^k - \mathbf{S}^k$ 代表任务 k 的 Laplacian(拉普拉斯) 矩阵。其中, \mathbf{D}^k 为对角

元素为 $D_{n_1, n_2}^k = \sum_{n_2=1}^N S_{n_1, n_2}^k$ 的对角矩阵。 S_{n_1, n_2}^k 被定义为:

$$S_{n_1, n_2}^k = \begin{cases} e^{-\frac{\|\mathbf{x}_{n_1}^k - \mathbf{x}_{n_2}^k\|^2}{\sigma}}, & \mathbf{x}_{n_1}^k, \mathbf{x}_{n_2}^k \text{ 与 } q \text{ 相邻} \\ 0, & \text{其他} \end{cases} \quad (5)$$

在这里,σ 是一个常数,q = 3。式(5) 旨在保留原始数据的局部相邻结构。

最后,讨论 $\lambda_1, \lambda_2, \lambda_3$ 三个参数的设置。式(1) 中这三个参数分别表示每项所占权重,这也直接影响分类结果。本文提出参数自寻优框架以选出最优参数,从而保证分类结果取得较高的准确度。

1.3.3 参数自寻优框架

在 RIS 模型特征选择的过程中,会涉及到一些参数的选择,分别是 $\lambda_1, \lambda_2, \lambda_3$ 三个参数的设置。这三个参数分别表示组稀疏正则化项、模板关系诱导正则化项以及被试关系诱导正则化项所占权重。 λ_1 的值直接影响筛选的特征个数, λ_1 值越小,选出的特征数越多。 λ_2, λ_3 的值则表征模板关系及被试关系对分类结果的影响程度。

选择不同的参数值,会对分类的结果造成一定的差异,为了使构建的分类器效果达到最优,本文提出一种参数自寻优框架,该框架可获得参数的最优组合,防止参数的过度拟合,使得 RIS 模型筛选的特征具有显著差异,从而使分类更加准确、有效,即用当前对模型影响最大的参数 λ_1 借助网络搜索(Grid Search, GS)^[26] 方法调优,通过循环遍历,直到最优化,并选取最优化过程中结果相对较好的几个参数值作为之后的候选值;剩下两个对模型影响相对较小的参数 λ_2 和 λ_3 ,借助随机搜索(Random Search, RS)^[27-28] 方法,随机选出若干候选值进行调优,得出这两个参数最优的参数组合。如此,所有的参数调整完毕。这个方法相对省时省力,但为避免可能会调到局部最优而不是全局最优,最后再次使用 GS 的方法,将上一步骤保存的相对数量较少的 λ_1 参数值作为候选值,再结合参数 λ_2 和 λ_3 的最优参数组合值,得到表现最好的参数组合作为最终结果。

结合本文使用的 RIS 特征选择方法和 SVM 分类器,具体而言,包括以下 2 个步骤:

1) 将输入的数据集 A,即提取出来待筛选的特征,按比例分为两组,分别是训练集 B 和测试集 C,划分比例约为 3:1。

2) 在步骤 1) 得到的训练集 B 上,选择不同的参数 $\lambda_1, \lambda_2, \lambda_3$ 的值,构成参数自寻优设置后的 RIS 模型。基于每一组参数组合在测试集 C 上的表现,得出效果最佳的参数组合,表现优劣通过 SVM 分类结果判定。

在框架寻优过程中可看出分类精确度随着三个参数 $\lambda_1, \lambda_2, \lambda_3$ 的改变而变化的情况,并且对于参数 λ_2 和 λ_3 ,分类结果一般是稳定的;对于参数 λ_1 ,分类精确度变化明显,表明 RIS 特征选择方法对参数 λ_1 特别敏感。参数自寻优结果得出,在 $\lambda_1 = 0.037, \lambda_2 = 0.4, \lambda_3 = 0.5$ 时,分类结果最佳。

1.4 分类

本文采用 SVM 分类器,其分类工作基于 Matlab 的 LIBSVM 工具包^[29] 进行,这是相似的研究中常用的方法^[30]。在特征选择阶段,首先利用 SVM 分类器与参数自寻优框架结



合,通过对分类准确率对比分析,选出最优参数组合。这部分工作是通过分类准确率来衡量参数优劣。最后,将使用最优参数的RIS模型筛选出的特征进行最后分类,并使用10折交叉验证的方法来评估分类器的泛化性能。

2 结果

2.1 不同方法对分类结果的影响

为了证明本文提出的方法具有更好的分类性能,本文更精确地比较在不同单一脑图谱模板以及不同于本文的多模板方法中的特征数量、敏感性、特异性以及准确率。本文在AAL90、Parc256、Parc497、Parc1003、Parc1501这五个多层次模板中进行研究,并计算每个单一模板的分类性能以作对比分析,如表2所示。多模板的方法相比单一脑图谱模板具有更好的分类性能。同时本文的方法,即经过参数自寻优后的RIS特征选择模型,相比仅仅连接不同模板特征的多模板方法取得了更好的分类结果。实验结果得出本研究方法的分类准确率为91.7%,特异性为92.3%,敏感性为88.4%。其原因主要是多层次模板之间具有关联性,其关联关系可作为研究的先验信息。RIS特征选择方法挖掘出来多层次模板之间的潜在关系,并且解决了由于特征数量增多而使得特征间冗余性高的问题,优化了特征子集,降低了特征数目,从而筛选出更有效的特征,在表2特征数量均有体现。

表2 基于不同脑模板的分类器性能

模板	特征数	敏感性/%	特异性/%	准确率/%
AAL90	16	79.3	66.3	74.3
Parc256	48	89.3	74.0	82.7
Parc497	95	92.0	75.2	82.5
Parc1003	186	87.3	88.2	87.5
Parc1501	240	91.9	83.6	88.5
多模板连接	585	81.5	91.7	88.7
优化后的RIS	185	88.4	92.3	91.7

2.2 RIS参数对分类结果的影响

RIS特征选择模型涉及到 λ_1 、 λ_2 、 λ_3 三个参数的设置,参数设置范围均为 $(10^{-10}, 10^0)$,选择不同的参数值,会对分类的结果造成一定的差异。为使分类器效果达到最优,本文提出参数自寻优框架。该框架运行结果表明,RIS特征选择方法对参数值 λ_1 最为敏感,其中 λ_1 在 $(10^{-3}, 10^{-1})$ 范围内分类效果较好。由此可见,在多模板方法中,对特征数量的控制,降低特征间冗余性,提高特征有效性尤为重要。 λ_2 与 λ_3 两个参数分别表示模板间相关关系和被试间相关关系的权重,均在 $(10^{-1}, 10^0)$ 范围内分类效果较好。可以得出,模板间的相关关系和被试间的相关关系对于分类性能同等重要。框架最终结果得出,在 $\lambda_1 = 0.037$, $\lambda_2 = 0.4$, $\lambda_3 = 0.5$ 时,分类结果最佳。

3 结语

本文解决了单一脑图谱模板提取脑网络拓扑特征可能不足以揭示患者组和NC组潜在的解剖学差异的问题,同时针对已有的多模板方法仅仅比较不同模板的分类性能或连接不同模板的特征,且选取的多个模板节点间不存在关联的问题,提供了解决办法。定义了多层次模板,突出多个脑图谱模板间的关联性。同时,提出经过参数自寻优的RIS多模板特征选择方法,模拟出多模板中的结构信息。其中,参数自寻优框

架是为了获得参数的最优组合,防止了参数的过度拟合,提高了分类器的泛化性能,使特征选择和分类更加准确和有效。分类结果显示,本文提出的方法达到更高的分类准确率。

本文工作仍存在一些潜在问题。首先,虽然本文的方法取得了比单个脑图谱模板更高的分类准确率;但是,需要注意的是,本次研究无法忽略构建、分析多个模板脑网络所带来的时间消耗。假设选取的模板数量很多,则需要的时间会较长。除了网络规模外,时间消耗同样受到算法、硬件配置等多种因素的影响,因此,接下来的研究着重平衡分类准确率和时间消耗。其次,在模板选择上,本文定义了AAL90、Parc256、Parc497、Parc1003、Parc1501这五个脑图谱模板。由于时间复杂度和空间复杂度对计算机资源要求的限制,没有尝试在更多的模板中进行实验,但这是未来研究的方向之一。再者,由于相关数据采集工作难度大,尤其抑郁症病人的功能核磁共振成像数据采集上可供采集的被试数量较少,且相关成本较高,使得本实验样本量较少的问题难以解决。因本研究样本量小,在参数自寻优方法中不足以将原始数据集A划分成训练集B、验证集C和测试集D,但相关数据的采集以及受试者招聘工作仍在持续进行中,这个问题将得到进一步解决。以上几点也作为未来工作的一部分,以确保本文的方法进一步完善。

参考文献 (References)

- [1] SACCHET M D, PPASAD G, FOLANDROSSL C, et al. Elucidating brain connectivity networks in major depressive disorder using classification-based scoring[C]// Proceedings of the 2014 IEEE International Symposium on Biomedical Imaging. Piscataway, NJ: IEEE, 2014: 246–249.
- [2] NIXON N L, LIDDLE P F, NIXON E, et al. Biological vulnerability to depression: linked structural and functional brain network findings[J]. British Journal of Psychiatry, 2014, 204(4): 283–289.
- [3] PEREIRA F, MITCHELL T, BOTVINICK M. Machine learning classifiers and fMRI: a tutorial overview[J]. Neuroimage, 2009, 45(1): S199–S209.
- [4] HAHN T, MARQUAND A F, EHLIS A C, et al. Integrating neurobiological markers of depression[J]. Archives of General Psychiatry, 2011, 68(4): 361–368.
- [5] MULDERS P C, van EIJDHOVEN P F, SCHENE A H, et al. Resting-state functional connectivity in major depressive disorder: A review[J]. Neuroscience and Biobehavioral Reviews, 2015, 56: 330–344.
- [6] ANAND A, LI Y, WANG Y, et al. Activity and connectivity of brain mood regulating circuit in depression: a functional magnetic resonance study[J]. Biological Psychiatry, 2005, 57(10): 1079–1088.
- [7] CRADDOCK R C, HOLTZHEIMER R P, HU X P, et al. Disease state prediction from resting state functional connectivity[J]. Magnetic Resonance in Medicine, 2010, 62(6): 1619–1628.
- [8] GREICIUS M D, FLORES B H, MENON V, et al. Resting-state functional connectivity in major depression: abnormally increased contributions from subgenual cingulate cortex and thalamus[J]. Biological Psychiatry, 2007, 62(5): 429–437.
- [9] 温洪,郭浩,李越,等.重度抑郁症患者脑功能网络的分类研究[J].计算机应用研究,2013,30(8):2304–2307.(WEN H, CUO H, LI Y, et al. Classification of brain function networks in patients with major depression [J]. Application Research of Computers, 2013, 30(8): 2304–2307.)



- [10] FORNITO A, ZALESKY A, BULLMORE E T. Network scaling effects in graph analytic studies of human resting-state fMRI data [J]. *Frontiers in Systems Neuroscience*, 2010, 4: 22–25.
- [11] MIN R, WU G, CHENG J, et al. Multi-atlas based representations for Alzheimer's disease diagnosis[J]. *Human Brain Mapping*, 2015, 35(10): 5052–5070.
- [12] OTA K, OISHI N, ITO K, et al. A comparison of three brain atlases for MCI prediction [J]. *Journal of Neuroscience Methods*, 2014, 221: 139–150.
- [13] JING B, LONG Z, LIU H, et al. Identifying current and remitted major depressive disorder with the Hurst exponent: a comparative study on two automated anatomical labeling atlases[J]. *Oncotarget*, 2017, 8(52): 90452–90464.
- [14] LIU M, ZHANG D, SHEN D. Relationship induced multi-template learning for diagnosis of Alzheimer's disease and mild cognitive impairment [J]. *IEEE Transactions on Medical Imaging*, 2016, 35(6): 1463–1474.
- [15] TZOURIOMAZOYER 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.
- [16] 牛会兰.局部信息方法在功能脑网络建模中的应用研究[D].太原:太原理工大学,2016:10–13.(NIU H L. Application research of local information method in functional brain network modeling[D]. Taiyuan: Taiyuan University of Technology, 2016: 10–13.)
- [17] KHAZAEI A, EBRAHIMZADEH A, BABAJANI-FEREMI A. Identifying patients with Alzheimer's disease using resting-state fMRI and graph theory [J]. *Clinical Neurophysiology*, 2015, 126(11): 2132–2141.
- [18] LANGER N, PEDRONI A, JÄNCKE L. The problem of thresholding in small-world network analysis[J]. *PLoS One*, 2013, 8: e53199.
- [19] KHAZAEI A, EBRAHIMZADEH A, BABAJANI-FEREMI A. Application of advanced machine learning methods on resting-state fMRI network for identification of mild cognitive impairment and Alzheimer's disease[J]. *Brain Imaging and Behavior*, 2016, 10(3): 799–817.
- [20] 郭浩.抑郁症静息态功能脑网络异常拓扑属性分析及分类研究[D].太原:太原理工大学,2013:5–10.(GUO H. Analysis and classification of abnormal topological attributes of resting brain function in depression[D]. Taiyuan: Taiyuan University of Technology, 2013: 5–10.)
- [21] ZHANG J, WANG J, WU Q, et al. Disrupted brain connectivity networks in drug-naïve, first-episode major depressive disorder[J]. *Biological Psychiatry*, 2011, 70(4): 334–342.
- [22] GUO H, LIU L, CHEN J, et al. Alzheimer classification using a minimum spanning tree of high-order functional network on fMRI dataset[J]. *Frontiers in Neuroscience*, 2017, 11: 639.
- [23] GUO H, ZHANG F, CHEN J, et al. Machine learning classification combining multiple features of a hyper-network of fMRI data in Alzheimer's disease [J]. *Frontiers in Neuroscience*, 2017, 11: 615.
- [24] CAO L, GUO S, XUE Z, et al. Aberrant functional connectivity for diagnosis of major depressive disorder: a discriminant analysis [J]. *Psychiatry and Clinical Neurosciences*, 2014, 68(2): 110–119.
- [25] MEIER T B, DESPHANDE A S, VERGUN S, et al. Support vector machine classification and characterization of age-related reorganization of functional brain networks[J]. *Neuroimage*, 2012, 60(1): 601–613.
- [26] HESTERMAN J Y, CAUCCI L, KUPINSKI M A, et al. Maximum-likelihood estimation with a contracting-grid search algorithm [J]. *IEEE Transactions on Nuclear Science*, 2010, 57(3): 1077–1084.
- [27] BERGSTRA J, BENGIO Y. Random search for hyper-parameter optimization[J]. *Journal of Machine Learning Research*, 2012, 13(1): 281–305.
- [28] BECSEY J C, BERKE L, CALLAN J R. Nonlinear least squares methods: a direct grid search approach[J]. *Journal of Chemical Education*, 1968, 45(11): 728.
- [29] CHANG C C, LIN C J. LIBSVM: a library for support vector machines[J]. *ACM Transactions on Intelligent Systems and Technology*, 2011, 2: 1–27.
- [30] YE F, LOU X Y, SUN J F. An improved chaotic fruit fly optimization based on a mutation strategy for simultaneous feature selection and parameter optimization for SVM and its applications[J]. *PLoS One*, 2017, 12(4): e0173516.

This work is partially supported by the National Natural Science Foundation of China (61672374, 61741212, 61876124, 61873178), the Youth Surface Program of Shanxi Provincial Science and Technology Department Applied Basic Research Project(201601D021073), the Shanxi Provincial Department of Education Science and Technology Innovation Research Project for Universities (2016139), the CERNET Next Generation Internet Technology Innovation Project of Ministry of Education (NGII20170712), the National Funded Overseas Studying Program (201708140216).

WU Hao, born in 1994, M. S. candidate. Her research interests include artificial intelligence, intelligent information processing and brain imaging.

WANG Xincan, born in 1993, M. S. candidate. Her research interests include artificial intelligence, intelligent information processing and brain imaging.

LI Xinyun, born in 1999. Her research interest includes artificial intelligence.

LIU Zhifeng, born in 1981, Ph. D., deputy chief physician. Her research interest includes basic and clinical research of affective disorder.

CHEN Junjie, born in 1956, Ph. D., professor. His research interests include artificial intelligence, intelligent information processing and brain imaging.

GUO Hao, born in 1980, Ph. D., associate professor. His research interests include artificial intelligence, intelligent information processing and brain imaging.