

重度抑郁症患者脑功能网络的分类研究*

温洪, 郭浩, 李越, 陈俊杰[†]

(太原理工大学 计算机科学与技术学院, 太原 030024)

摘要: 为了构建辅助诊断模型, 为抑郁症的诊断提供一个新的方法, 以提高抑郁症诊断的准确率。在连续的阈值空间(8%~32%)内构建所有被试的功能脑网络并使用复杂网络理论对抑郁症患者的脑网络进行分析, 力求提取出可以从各个维度来表征抑郁症患者的脑网络的特征值, 采用不同的属性组合并使用 SVM 分类算法对所有被试进行分类研究, 结果发现将全局属性与局部属性组合作为分类特征得到的分类正确率最高, 因此该方法可以用于抑郁症的辅助诊断中。

关键词: 重度抑郁症; 复杂网络; 特征选择; 脑功能网络; 分类

中图分类号: TP18

文献标志码: A

文章编号: 1001-3695(2013)08-2304-04

doi:10.3969/j.issn.1001-3695.2013.08.016

Research on classification of brain functional network in major depressive patient

WEN Hong, GUO Hao, LI Yue, CHEN Jun-jie[†]

(College of Computer Science & Technology, Taiyuan University of Technology, Taiyuan 030024, China)

Abstract: In order to construct a computer-aided diagnosis model and propose a new method for depressive diagnosis to improve the accuracy of depression diagnosis, this article constructed the whole-brain functional networks of all subjects and analyzed the brain networks of depressive patients from different perspectives to extract characteristic values which could represent brain networks of depressive patients from different dimensions. And it classified all subjects by SVM classification algorithm research. Using different property group, results show that the rate of correct is highest when combination of global and local properties used as classification characteristics, so this method can be used in the auxiliary diagnosis of depression.

Key words: major depressive; complex network; feature selection; brain functional network; classification

在竞争压力巨大的当今社会, 抑郁症发病率居高不下。而目前在医学上还没有一种检查或者化验可以作为抑郁症诊断的依据, 对抑郁症的诊断还是以临床诊断为主, 这种诊断方式很大程度上依赖于医师的临床经验, 诊断准确率偏低。因此, 如何提高精神脑疾病诊断的准确率是目前医师所关心的一个重要问题。

在以往的研究中已经证明, 无论是抑郁症患者还是正常人的脑功能网络都明显具有小世界属性, 并且抑郁症患者脑网络与正常人脑网络的拓扑结构是有些差异的^[1]。因此, 可以使用复杂网络理论对抑郁症患者和正常人的脑功能网络的拓扑结构进行研究。虽然功能影像学已经表明重度抑郁症与海马、海马旁回、中央前回、尾状核等大脑区域的异常有关^[1], 但是目前对重度抑郁症患者脑网络的研究都是在一个特定的单一的阈值下进行的, 在一个连续阈值空间内对抑郁症患者和正常人全脑网络的拓扑结构的差异变化研究屈指可数。

本文的研究重点是使用复杂网络理论分析抑郁症患者与正常人脑网络在一个连续的阈值空间内的差异变化情况, 进而选取可以从不同维度来表征抑郁症患者脑网络特点的分类特征, 运用支持向量机(support vector machine, SVM)分类算法,

对抑郁症进行分类研究, 以此来构建准确率较高的抑郁症辅助诊断模型, 提高抑郁症诊断的准确率。

1 基础理论

1.1 复杂网络

衡量复杂网络拓扑结构的指标包括为全局属性和局部属性。全局属性包括聚合系数(clustering coefficient)、特征路径长度(characteristic path length), 节点属性包括节点的度(degree)、中间中心度(betweenness centrality)、节点效率(node efficiency)^[1]。

a) 聚合系数。在一个脑网络中, 假设某个节点有 k 个边, 则这 k 条边连接的节点(k 个)之间最多可能存在的边的个数为 $k(k-1)/2$, 用实际存在的边数除以最多可能存在的边数得到的分数值, 定义为这个节点的聚合系数。所有节点的聚合系数的均值定义为网络的聚合系数。网络的聚合系数定义如下:

$$C = \frac{1}{N} \sum_{i=1}^N C_i = \frac{1}{N} \sum_{i=1}^N \frac{2E_i}{k_i(k_i-1)}$$

其中: 脑区节点 i 的度为 k_i , E_i 为节点 i 的邻居脑区节点间存在的边数, C_i 为脑区节点 i 的聚类系数(若 $k_i < 2$, 则 $C_i = 0$)。

收稿日期: 2012-10-31; **修回日期:** 2012-12-06 **基金项目:** 国家自然科学基金资助项目(61170136, 61070077, 60975032, 30971054, 81171290); 山西省科技厅资助项目(2011011015-4, 2010011020-2)

作者简介: 温洪(1986-), 女, 硕士, 主要研究方向为智能信息处理、脑信息学(wenh002@163.com); 郭浩(1981-), 男, 讲师, 博士研究生, 主要研究方向为智能信息处理、脑信息学、脑影像学、语义网; 李越(1985-), 男, 硕士研究生, 主要研究方向为智能信息处理; 陈俊杰(1956-), 男(通信作者), 教授, 博导, 主要研究方向为数据库与智能信息处理。

b)特征短路径长度。一个网络中两点之间特征路径长度的平均值。网络特征短路径长度定义如下:

$$L = \frac{1}{N} \sum_{i=1}^N L_i = \frac{1}{N} \sum_{i=1}^N \frac{\sum_{j=1, j \neq i}^N d_{ij}}{N-1}$$

其中: d_{ij} 为两个脑区节点 i 和 j 的最短路长。一般认为,对于随机图模型,网络的平均路径长度 $L_{\text{random}} \approx \ln N / \ln k$ 。

c)节点的度。度被定义为一个节点的边数,其公式表示为

$$k_i^B = \sum_{j \in G} a_{ij} \quad \text{或者} \quad k_i^W = \sum_{j \in G} w_{ij}$$

其中: a_{ij} (w_{ij}) 表示在二值网络中第 i 个节点与第 j 个节点之间的连接数。

d)中间中心度。节点在图中中心性的度量(同样也有边介数),出现在许多其他节点最短路径中的节点有更高的中间中心度值。其公式表示为

$$b_i = \frac{\sum_{m \neq i \neq n \in G} \sigma_{mn}(i)}{\sigma_{mn}}$$

其中: σ_{mn} 表示从节点 m 到 n 之间的最短路径的数目, $\sigma_{mn}(i)$ 表示从 m 到节点 n 经过节点 i 的最短路径数目。

e)节点效率。其定义如下:

$$e_i = \frac{1}{N-1} \sum_{j \neq i \in G} \frac{1}{d_{ij}}$$

其中: d_{ij} 表示在网络 G 中节点 i 与 j 之间的最短特征路径长度。这里所说的特征路径长度就是在网络 G 中从一个节点到另一个节点所有可能的路径中距离最短的那条路径的长度。

f)AUC(area under the curve),它提供了一种测量网络节点属性在不同稀疏下总的变化强度的方法^[1]。本文所求的 AUC 是在一个特定阈值空间内网络某属性的线下面积。定义为

$$Y^{\text{AUC}} = \sum_{k=1}^{n-1} [Y(S_k) + Y(S_{k-1})] \times \Delta S / 2$$

其中: S_1 与 S_n 表示阈值空间的上下边界, ΔS 表示两个稀疏度之间的间隔。而在本文中 S_1 取值为 0.08, S_n 取值为 0.32, ΔS 取值为 0.01。

1.2 分类算法

SVM 是 Cortes 和 Vapnik 于 1995 年提出来的,与其他算法相比,该算法应用在解决小样本、非线性及高维模式识别,特别是在 fMRI 数据分析中表现出很大的优势,并且得到了广泛的应用^[2,3]。该算法是基于统计学习理论和结构风险最小原理而提出来的,为了得到较好的推广能力,该算法在模型的学习能力和复杂性中间力求找到的一个最佳的平衡点。由于很难把低维空间的集合划分开来,所以只能将低维空间向量集向高维空间映射,然而这种解决方法又提高了计算的复杂度。SVM 利用核函数很好地解决了这个问题,因此 SVM 的关键就在于核函数,只有使用了适当的核函数才能更好地得到高维空间划分函数。在 SVM 方法中,只要定义不同的内积函数,就可以实现多项式逼近、贝叶斯分类器、径向基函数(radial basic function)方法、多层感知器网络等许多现有学习算法。

2 数据的采集与预处理

2.1 被试

本实验共收集 60 名被试,其中有 34 名首发、无用药、重度

抑郁症患者作为抑郁组,26 名身体健康的志愿者作为对照组。所有对照组被试均无精神或神经紊乱史并以精神卫生筛选表与神经症筛选表为依据;抑郁组中的所有被试均为重度抑郁症病人并且都是以美国精神障碍诊断与统计学手册第四版(DSM-IV)和 17 项汉密尔顿评定量表的抑郁症(HRSD)为依据。所有被试均排除以下情况:重大躯体疾病史,特别是与大脑组织变化有关的疾病;不稳定的躯体疾病,如重大头部创伤、重症哮喘等;IQ 低于 70;任何磁共振禁忌情况。所有被试都为右利手。实验前同每位参与者均达成了书面协议,数据采集与预处理同山西医科大学第一医院合作。所有被试的详细信息如表 1 所示。

表 1 被试基本信息表

项目	正常组	抑郁组	P 值
年龄	17 ~ 63 (27 ± 68.5)	17 ~ 54 (28.9 ± 87.8)	0.4466
性别(男/女)	14/12	14/20	
利手(右/左)	26/0	34/0	-
HAMD	无	15 ~ 34 (22.8 ± 13.3)	-

2.2 磁共振数据的获得

本实验采用 fMRI(functional magnetic resonance imaging,功能磁共振)功能图像采集数据,对静息态下脑功能连接模式进行考证。数据的采集工作是由山西医科大学第一医院完成的,所有的扫描工作由熟悉磁共振操作的放射科医生来完成。在扫描的过程中,要求被试闭眼、放松、不去想特定的事情但要保持清醒不能睡着。

扫描参数设置如下:33 axial slices, repetition time(TR) = 2000 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。

2.3 数据的预处理

对得到的核磁共振原始数据进行预处理,预处理时使用 MATLAB 7.8 软件的 statistical parametric mapping 8(SPM8)^[4]: a)层间时间校正,以校正图像在获取时间上的差异;b)头动校正;c)空间标准化,将得到的 fMRI 图像标准化到 SPM 标准 MNI 空间的 EPI 模板上,以便于不同被试进行分析;d)低频滤波,以降低低频漂移及高频的生物噪声。

3 脑功能网络的构建

3.1 节点的定义

本文将全脑按照国际脑成像领域广泛使用的标准 automated anatomical labeling(AAL)^[5]模板分成 90 个脑区作为感兴趣区(regions of interest, ROI),每一个区域定义为一个节点。该区域内所有体素的时间序列的平均值作为该节点的时间序列。

3.2 边的定义

a)对每个节点的平均时间序列作多元线性回归去除一些由头动所引起的伪差异;b)用每个节点多元线性回归得到的残差两两作偏关,求得任意两节点平均时间序列的相关系数,所采用的方法为

$$r_{ij} = \frac{\sum_{t=1}^T [x_i(t) - \bar{x}_i][x_j(t) - \bar{x}_j]}{\sqrt{\sum_{t=1}^T [x_i(t) - \bar{x}_i]^2} \sqrt{\sum_{t=1}^T [x_j(t) - \bar{x}_j]^2}}$$

其中: x_i, x_j 分别表示体系 i 和体系 j 的时间序列, $\bar{x}_i$ 和 $\bar{x}_j$ 表示时间序列的均值。由此得到一个 90×90 的时间序列相关矩阵 A 。最后,根据事先设定好的阈值将关联矩阵 A 转换成二值矩阵,即当 r_{ij} 大于该阈值时对应二值矩阵的对应元素为 1,否则为 0。在本文中采用一系列连续的阈值,在一个阈值空间内构建脑网络。

3.3 阈值选择

对于所有的关联矩阵要选定一个阈值才能将关联矩阵转换成二值矩阵。本文选用的不是单一的阈值而是一个连续阈值范围,在该阈值空间内构建对应于每个阈值的脑功能网络。确定阈值空间时遵循以下原则:

a) 对于所有被试来说,在该阈值空间内构建的所有网络的平均度值 K 要大于 $2 \ln N \approx 9$,本文中 $N=90$ 。

b) 对于所有被试来说,在该阈值空间内构建的所有网络的小世界属性都要大于 1.1。

根据以上原则,可以计算阈值空间为 8% ~ 32%,而在该阈值空间内以步长为 0.01 的所有稀疏度下,构建所有被试的脑功能网络。

4 分类研究

4.1 特征选择

4.1.1 全局属性选择

求取抑郁组和正常组的所有被试在该阈值空间内所构建的所有脑网络的聚合系数和特征路径长度;再对抑郁组和正常组的所有被试在本文所确定的阈值空间内所构建的所有脑网络的聚合系数和特征路径长度作非参数置换检验,结果如图 1 所示。

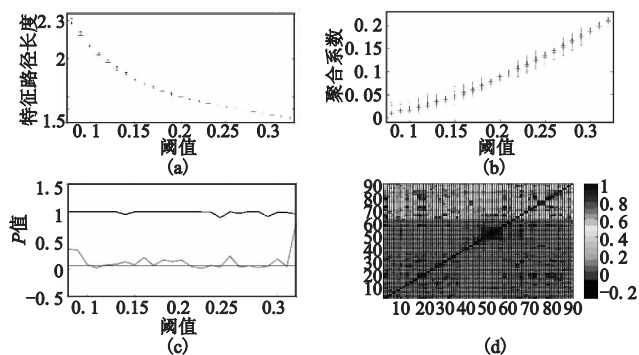


图1 全局属性统计检验结果

图 1 中,横坐标表示阈值,0.1 表示在稀疏度为 10% 下;纵坐标表示所求得的脑功能网络的聚合系数和特征路径长度。图 1(见电子版)中,(a)(b)的蓝色曲线表示正常组,粉红色的曲线表示抑郁组;(c)中的蓝线和绿线分别表示特征路径长度和聚合系数作统计检验后的结果。

在以往的抑郁症患者研究中得到结论:抑郁症患者脑网络的聚合系数与正常人脑网络的聚合系数没有明显差异,抑郁症患者脑网络的聚合系数与正常人脑网络的特征路径长度有明显差异,这一结论在本实验中也得到了印证^[6]。由此可以选

取脑功能网络的特征路径长度作为分类特征来构建分类模型。

4.1.2 局部属性选择

对抑郁组和正常组的所有被试在阈值空间内所构建的所有脑网络的每个节点的度的 AUC 作非参数置换检验。结果如表 2 所示,结果表明,一些节点(如丘脑、楔叶、海马等)在抑郁组脑网络中的度的 AUC 与正常组有明显差异。这一结果与 He 等人^[9]研究所得出的结论基本一致。

表2 节点度的 AUC 作非参数置换检验的结果

脑区名称	P 值		
	度	中间中心度	节点效率
正常组大于抑郁组(Con > MDD)			
梭状回	0.016 7	0.151 5	0.047 9
眶部额下回	0.017 8	0.179 2	0.034 2
楔叶	0.025 6	0.000 9	0.093 8
内侧额上回	0.035 2	0.495 1	0.151 5
眶部额中回	0.049 9	0.557 1	0.025 6
岛盖部额下回	0.050 0	0.295 9	0.141 6
正常组小于抑郁组(Con < MDD)			
海马	0.001 5	0.005 5	0.003 3
角回	0.008 2	0.284 3	0.011 5
后扣带回	0.008 8	0.167 1	0.007 1
丘脑	0.008 8	0.004 0	0.008 8
豆状壳核	0.014 7	0.213 4	0.023 5
枕中回	0.021 5	0.443 0	0.027 1
楔叶	0.031 8	0.510 0	0.008 2

表 2 是抑郁组和正常组的所有被试在整个阈值空间(8% ~ 32%)下,所构建的所有脑网络的每个节点的度的 AUC 作非参数置换检验后对所得到的 P 值进行排序,截取显著度 $P < 0.1$ 的前 13 个脑区的结果。

再将各个节点度的 AUC 作非参数置换检验的结果进行从小到大排序,可以选取差异最为明显的前五个节点的度、中间中心度、节点效率的 AUC 作为分类特征,结果如表 3 所示。

表3 节点选取结果

节点编号	P 值	中文名称	节点编号	P 值	中文名称
38	0.001 453 967	海马	55	0.016 722 988	梭状回
65	0.008 200 786	角回	15	0.017 869 16	眶部额下回
36	0.008 821 024	后扣带回	52	0.021 564 751	枕中回
78	0.008 821 024	丘脑	46	0.025 671 692	楔叶
74	0.014 780 407	豆状壳核			

对抑郁组和正常组在阈值空间下所构建的脑功能网络的每个节点的度作非参数置换检验,结果表明,在抑郁组脑网络中的度的 AUC 与正常组也有明显差异,而且在不同的阈值下,差异明显程度是有变化的。

图 2 表示在阈值空间内对每个稀疏度下,海马、楔叶、角回、丘脑、后扣带回、豆状壳核各脑区正常组所有被试功能脑网络的度与抑郁组所有被试功能脑网络的度作非参数置换检验所得到的结果。由图 2 所示,本文将节点在差异最明显的阈值下的度、中间中心度、节点效率作为节点在该阈值空间内度、中间中心度、节点效率的极值。由此结论,可以将前文选取的五个节点的度、中间中心度、节点效率的极值作为分类特征。

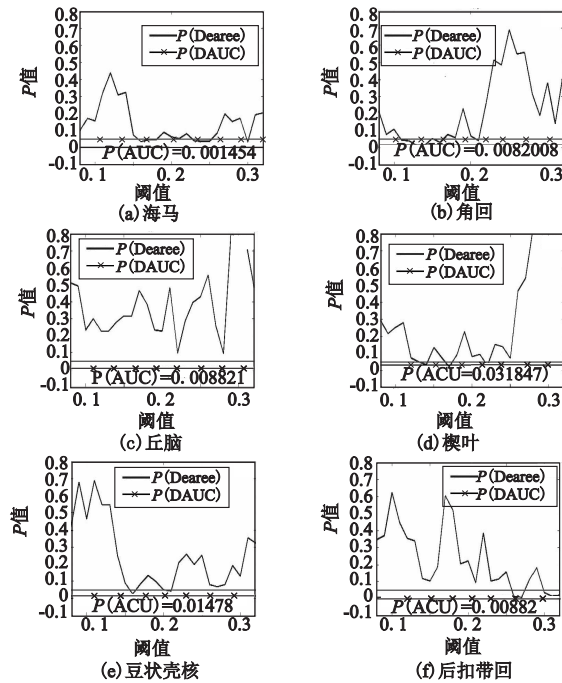


图2 节点度的统计检验结果

4.2 分类结果对比分析

4.2.1 节点属性特征组合分类结果对比

分别用所选取的五个节点的度的AUC、极值以及AUC和极值的组合作为分类特征,采用SVM方法对所有被试进行分类;与此同时,为了降低分类结果的随机性,对于每一种不同的分类特征都重复进行10次分类,得到的正确率如图3所示。

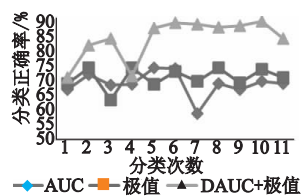


图3 节点的度的AUC和极值分类结果对比

由图3可知,使用节点属性的AUC和极值的组合作为分类特征的分类正确率要高于只使用AUC或者极值作为分类特征的正确率,这也说明了节点属性的整体特征与局部特征相结合作为特征值会提高分类正确率。

为了验证此结论,本文进一步用所选取的五个节点的中间中心度和节点效率的AUC、极值以及AUC和极值的组合作为分类特征,采用SVM算法对被试进行分类。与上述实验相同,对于每个特征组合都会重复10次,将这10次的平均值作为每个特征组合的分类正确率,所得到的节点的度、中间中心度以及节点效率的分类正确率如图4所示。

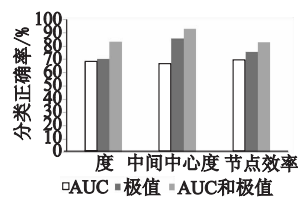


图4 节点各属性分类结果

4.2.2 全局属性特征组合分类结果对比

用特征路径长度与所选节点的各属性值的组合作为分类特征,采用SVM算法对被试进行分类并把分类结果与只用节点属性作为分类特征的结果进行比较,结果如图5所示。

由图5可知,使用节点属性和全局属性的组合作为分类特征的正确率要高于只使用节点属性或全局属性作为分类特征的正确率,这个结果进一步说明了将从不同的维度、不同角度选择的特征值相结合作为分类特征可以提高分类正确率。

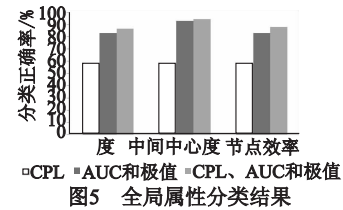


图5 全局属性分类结果

5 结束语

本文在一个连续的阈值空间内构建了抑郁症患者与正常人的脑功能网络,利用复杂网络理论和统计学方法进行差异分析发现,抑郁症患者脑功能网络的特征路径长度明显低于正常人,而且抑郁症患者的一些与情感相关的脑区的度、中间中心度等属性与正常人有明显差异,并且在不同阈值下差异明显程度不同。使用全局属性、节点属性的极值、节点属性的AUC的组合作为分类特征,从不同的维度来表征抑郁症患者的脑网络,以此构建的分类模型的准确率较高的,结果与现有研究结果相吻合^[1,8],可以用于抑郁症的辅助诊断中。

参考文献:

- [1] ZHANG Jun-ran, WANG Jin-hui, WU Qi-zhu, *et al.* Disrupted brain connectivity networks in drug-naive, first-episode major depressive disorder[J]. *Biological Psychiatry*, 2011, 70(4): 334-342.
- [2] MITCHELL T M, HUTCHINSON R, JUST M A, *et al.* Classifying instantaneous cognitive states from fMRI data[C]//Proc of American Medical Informatics Association Annual Symposium. 2003:465-469.
- [3] GOUTTE C, HANSEN L K, LIPTRON M G, *et al.* Feature-space clustering for fMRI Meta-analysis[J]. *Human Brain Mapping*, 2001, 13(3): 165-183; 465-469.
- [4] SPM8[EB/OL]. (2009-04). <http://www.fil.ion.ucl.ac.uk/spm/software/spm8>.
- [5] TZOURIO-MAZOYERN, 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.
- [6] HE Yong, WANG Jin-hui, WANG Liang, *et al.* Uncovering intrinsic modular organization of spontaneous brain activity in humans[J]. *PLoS ONE*, 2009, 4(4): e5226.
- [7] 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.
- [8] COX D D, SAVOY R L. Functional magnetic resonance imaging (fMRI) "brain reading": detecting and classifying distributed patterns of fMRI activity in human visual cortex[J]. *NeuroImage*, 2003, 19(2): 261-270.
- [9] HE Yong, CHEN Z J, EVANS A C. Small-world anatomical networks in the human brain revealed by cortical thickness from MRI[J]. *Cerebral Cortex*, 2007, 17(10): 2407-2419.