
华中科技大学

本科生毕业设计[论文]

(基于 rfMRI 的抑郁症预测研究)

院 系 电子信息与通信学院

专业班级 电信 1705 班

姓 名 杜咏琦

学 号 U201713356

指导教师 彭勤牧

2021 年 05 月 15 日

学位论文原创性声明

本人郑重声明：所呈交的论文是本人在导师的指导下独立进行研究所取得的研究成果。除了文中特别加以标注引用的内容外，本论文不包括任何其他个人或集体已经发表或撰写的成果作品。本人完全意识到本声明的法律后果由本人承担。

作者签名： 年 月 日

学位论文版权使用授权书

本学位论文作者完全了解学校有关保障、使用学位论文的规定，同意学校保留并向有关学位论文管理部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅。本人授权省级优秀学士论文评选机构将本学位论文的全部或部分内容编入有关数据进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

本学位论文属于 1、保密口，在 年解密后适用本授权书

2、不保密口。

(请在以上相应方框内打“√”)

作者签名： 年 月 日

导师签名： 年 月 日

摘 要

抑郁症是一种严重的精神疾病,当前包括抑郁症在内的精神疾病的诊断均是基于量表评分,这会带来误诊和治疗不客观等问题。抑郁症的科学客观的诊断是亟待研究的课题。神经影像学特别是脑功能成像手段的发展,为精神类疾病的诊断提供了极大的便利。同时,近年来机器学习和深度学习快速发展并且它在医学影像分析领域得到广泛应用。我们将开展基于 fMRI 的抑郁症智能诊断的研究。

本文采用大脑静息态功能磁共振 rfMRI 小数据集,基于动态功能连接(DFC)和静态功能连接(SFC),用机器学习算法 SVM 对抑郁症进行分类。实现了基于两种脑图集 AAL, Brainnetome 分区的四组实验,并对重要特征进行了生理上的解释和可视化。进一步,我们基于双向 LSTM(BiLSTM)的深度学习框架,采用更大的 REST-meta-MDD 数据集,实现了基于动态功能连接(DFC)和基于原始 BOLD 信号的分类模型,对抑郁症患者进行诊断。实验结果表明尽管 DFC 是原始信号的更高层次的表述,但提取动态功能连接特征并未有信息的丢失,而是表现出了更好的分类特性。

本文在两个数据集上,实现基于传统机器学习和深度学习对抑郁症进行智能诊断,预测精度基本满足要求。

关键词: 静息态功能磁共振; 抑郁症; 机器学习; 深度学习; 动态功能连接; 静态功能连接

Abstract

Major Depression Disorder(MDD) is a serious mental illness, and the diagnosis of mental diseases like major depression disorder is based on scale scores, which will cause problems such as misdiagnosis and delayed treatment. The accurate diagnosis of depression and the search for anatomical landmarks are urgently needed to be studied. Neuroimaging, especially the development of brain functional imaging technology, provides a huge opportunity for the diagnosis of mental diseases. In recent years, machine learning and deep learning algorithms have been utilized widely in medical problems. In this paper, we will do some research on the diagnosis of MDD using rfMRI data.

In this paper, on a small size dataset, four sets of experiments with two kinds of brain atlases AAL, Brainnetome, two kinds of functional connections dynamic functional connection (DFC) and static functional connection (SFC) are implemented, utilizing machine learning algorithm SVM classification, and the important features selected in the pipeline are explained physiologically and visualized. In the second set of experiments, on a larger size dataset, REST-meta-MDD dataset, two sets of classification experiments based on dynamic functional connection (DFC) and based on the original BOLD signal are implemented using deep learning framework based on bidirectional LSTM (BiLSTM). The experiment result shows that DFC, as a higher-level representation of the original signal, does not lose information of original signal, and even shows better classification characteristics.

In this paper, based on traditional machine learning and deep learning, the intelligent diagnosis of depression is realized, and the prediction accuracy basically meets the requirements.

Key Words: Resting state fMRI(rfMRI); Major depressive disorder (MDD); machine learning; deep learning; dynamic functional connectivity (DFC); static functional connectivity (SFC)

目 录

摘 要.....	I
Abstract.....	II
1 绪论.....	1
1.1 问题背景.....	1
1.1.1 问题来源.....	1
1.1.2 国内外研究现状.....	1
1.2 研究内容概述.....	5
2 相关原理	8
2.1 fMRI 数据.....	8
2.2 fMRI 数据预处理.....	9
2.3 功能连接计算.....	10
2.4 数据划分.....	11
2.5 样本均衡.....	13
2.6 SVM.....	13
2.7 LSTM.....	15
2.8 评价指标.....	16
3 基于 fMRI 的抑郁症诊断方法设计	18
3.1 基于传统机器学习 SVM 算法的抑郁症诊断	18
3.1.1 功能连接数据提取前处理.....	18
3.1.2 分类数据准备.....	19
3.1.3 数据划分和样本均衡.....	20
3.1.4 基于 SVM 的分类流程.....	21
3.2 基于 BiLSTM 深度学习方法的抑郁症诊断	27
3.2.1 数据前处理.....	27

3.2.2	数据准备.....	27
3.2.3	数据划分.....	29
3.2.4	基于 LSTM 的分类器模型.....	29
4	实验与分析	31
4.1	实验数据集描述.....	31
4.2	实验结果与分析.....	32
4.2.1	基于传统机器学习实验结果.....	32
4.2.2	基于 BiLSTM 深度学习实验结果:	35
5	结论与展望	36
6	致谢	38
7	参考文献	39
附录 1	44
附录 2	46
附录 3	54

1 绪论

1.1 问题背景

1.1.1 问题来源

抑郁症是一种严重的精神疾病，其核心特征在于持续（至少连续两周）的低落情绪，精神不振[3]，最严重的情况下可能会出现自杀的想法。根据世界卫生组织（WHO）的数据[1]，就残障人士而言，抑郁症是全球范围内致残的主要原因。全球约有 3.22 亿人患有抑郁症。从 2005 年到 2015 年，抑郁症患者的增长率为 18.4%。抑郁症已经是当今社会面临的一个严峻的问题，根据世界卫生组织（WHO）预测[2]，到 2030 年，抑郁症将会成为全球疾病负担的主要原因。

而包括抑郁症在内的精神疾病的诊断均是基于量表评分，如（DSM-5 诊断标准）[3]，贝克抑郁量表（BDI）和汉密尔顿抑郁量表（HAM-D）等，最终的诊断由专业人士结合量表的结果以及访谈过程的表现进行最终诊断，目前没有现成的黄金标准可以用于最终验证（如生物学或解剖学的标志），因此有时会出现误诊的情况，也因为缺乏相应的黄金标准，对诊断结果的质疑可能会影响及时的治疗[4]。

进几十年来神经影像学的发展，特别是脑功能成像手段的发展，为精神类疾病的诊断提供了巨大的机会，而近年来机器学习和深度学习算法的发展与广泛应用向我们展示出了机器学习算法，特别是深度学习算法良好的特征表示能力。将机器学习和深度学习相关算法应用于神经影像数据，自动提取具有标识性的疾病诊断特征[5][6][7]，对于精神类疾病的诊断以及后续对特征的解释中寻求具有标识性的生物学特征具有重要的意义，这也是本课题的研究意义所在。

1.1.2 国内外研究现状

1.1.2.1 数据模态

包括 fMRI[12]，PET，EEG，DTI 在内的大脑的功能以及结构影像技术已经成为大脑疾病诊断的重要工具，人们希望能够从这些功能影像数据中提取能够来

刻画大脑功能异常的特征，并以此进行大脑功能异常疾病的诊断[8][9]。包括抑郁症在内的精神疾病又与神经科学和大脑的功能息息相关[42]，借助这些成像手段，基于大脑的精神疾病诊断[36]引起了各界重视。

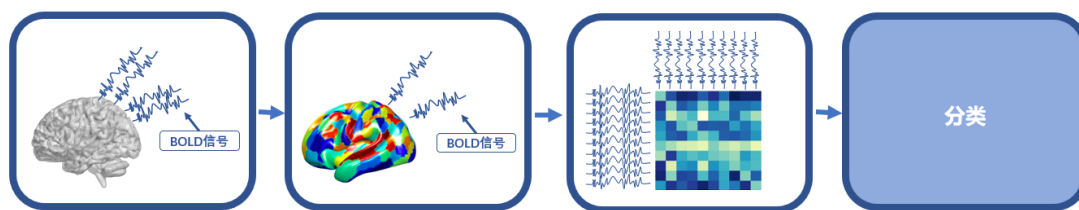
根据调查，基于这几个模态或者几种模态结合的大脑功能成像的抑郁症诊断研究都有许多，如[8]总结了 11 篇使用 EEG 数据进行重度抑郁症诊断的文章，[11]借助 PET 数据研究抑郁症患者神经回路的异常，然而在这些模态中，fMRI 由于其高空间分辨率以及非侵入式成像方式成为最为广泛使用的模态。

值得注意的是，fMRI 数据包括任务态(task state)和静息态(resting state)的数据，最早的时候对大脑机制的研究主要基于任务态的数据，即探究对应于特定任务的大脑激活区域研究大脑的作用机制，然而，任务态的数据的采集往往需要复杂的实验设计和实验装置，相较之下，静息态 fMRI，即被试处于一种无任务刺激的非睡眠静息状态时采集的 fMRI 数据，则更加容易获取，并且有着更加简单且高度标准化的数据采集程序使得采集变得更加简单[9]。更重要的是，研究[10]发现静息态的大脑并非是处于静止状态，而是有着一些自发的神经活动，并且这些活动对应的一些大脑区域与任务态 fMRI 的激活区域具有高度的相似性。基于这两点，静息态 fMRI 受到越来越广泛的关注。与此同时，因为采集程序的标准化，多中心的静息态数据也能够聚集起来用于共同研究，许多公开数据集逐渐构建起来，如用于 AD 诊断的 ADHD200，用于自闭症诊断的 ABIDE 数据集以及前几年公开的 REST-meta-MDD 数据集等，这也为后来的机器学习和深度学习方法的使用提供了数据基础。

1.1.2.2 基于功能连接的抑郁症诊断

目前存在的借助机器学习或深度学习的方式进行抑郁症诊断的文章大部分包括两个类别，分别是基于功能连接的预测和非功能连接的预测。可以理解为在分类前基本预处理后的 fMRI 数据经过了特定的数据初步特征提取的方法和直接端到端的输入经过基本预处理之后的 fMRI 数据的方法。

基于功能连接的预测的流程大致上可以分解为对原始采集数据的预处理，脑区分割 (parcellation)，功能网络的构建，功能连接特征的提取，对特征进行分类，如下图图片 1-1 所示：



图片 1-1 基于功能连接的抑郁症诊断流程

预处理

对于预处理的研究主要集中在过去,例如文章[43][44],在后来一些集成了这些研究成果的预处理工具逐渐发展起来,包括 FSL[45], SPM[47], AFNI[46], DPABI[49], DPARSF[13], 以及 python 的安装包 fMRIPrep[48]等工具或代码安装包

脑区分割

经过预处理后的 fMRI 数据具有很高的空间分辨率,在构建大脑功能网络以进行后续诊断研究时如果使用每一个体素作为网络的节点,计算会变得非常困难,对噪声敏感并且难以解释,此外,在体素尺度上存在冗余信息,并且消除这种冗余可以显着减小 fMRI 数据的维数[17],因此在进行后续脑功能连接分析,即脑网络分析时,首先进行脑区功能划分,之后将区域作为节点进行脑网络分析是很重要的一步。[18]介绍了一些常用的脑区分割的方法,主要集中在聚类方法以及一些公开的脑图集,如需了解更多,请参阅这篇文章。

分割方式有许多种,包括基于模板的分割和基于聚类的分割[18][17],为了保证分割区域的对应,一般使用大脑分割模板进行分割。

功能连接的度量

在脑区域划分后,就可以将区域作为节点进行网络分析,网络分析的关键是寻求网络节点功能连接的度量。

功能连接的度量有许多方式,可以分为静态功能连接(SFC)和动态功能连接(DFC):

静态功能连接(SFC)的获取包括通皮尔森相关系数计算,通过 ASR, NASR 获取,通过超图学习的方式获取等方式,最常用的是通过皮尔森相关系数计算获得。

研究表明[34][35][36],脑功能连接性可能在短时间内(例如数十秒)变化,

并且根据这个特性可以成功借助脑功能连接的动态变化捕获疾病人群中的功能连接异常。因此基于动态功能连接(DFC)的诊断逐渐引起重视。值得注意的是,[34]通过对比实验探究在同一个分类模型下,动态功能连接和静态功能连接对后续诊断效果的影响,发现动态 FC(分类准确度: 84.28%)明显优于静态 FC(分类准确度: 59.12%),这也证明了动态功能连接含有更加丰富的判别信息。

动态功能连接主要通过滑动窗格的方式获取。研究表明窗口长度需要在 30–60 s[39],对每一个窗格内的时间序列做功能连接分析,接着滑动窗格就可以得到多个短时功能连接特征,将得到的所有特征用于下一步的分类诊断。另有一些无窗的方式,如[40],最主要使用的方式是滑动窗格的形式。

分类模型

早几年人们大多采用机器学习的方法,如 SVM[21][22][35], ridge regression[24]等,可用于动态或静态功能连接特征的分类;近年来随着深度学习的发展,许多深度方法被引入到这一环节,包括 BrainnetCNN[20][23]借助 CNN 实现基于静态功能连接的预测,[21]对比了 SVM, CNN 和两种树结构的分类方法的分类效果基于静态功能连接进行分类;有一些研究借助图卷积神经网络(GCN)实现诊断,包括[25][26]将个体作为节点,功能连接矩阵作为每个节点的特征,输出诊断图(每个节点一个 label), [28][27]则是从单个个体的功能连接矩阵中提取图,再提取图的一维特征用来做分类,[31]则是通过度量学习学习图之间的相似度,基于此相似度做进一步的谱聚类得到诊断结果。这些都是基于静态功能连接的特征做的进一步分类。[29]结合动态功能连接并且利用 GCN 自动学习功能连接表示,在公开数据集上取得了优秀的结果。

1.1.2.3 基于非功能连接的抑郁症诊断

不使用功能连接方法的文章，大都直接将原始的 fMRI 数据或脑区分割后的 fMRI 数据以 2D 或 3D 或 4D 的数据维度输入到不同的深度学习的框架中，直接得到输出，模型框架如 deepAD[30]，3DCNN+LSTM[32]，ConvRNN[19]等文章。

1.1.3 趋势和存在的问题

基于动态功能连接（DFC）和借助深度方法进行诊断是抑郁症这类精神疾病的诊断方法的发展趋势，并且动态功能连接和深度方法结合的研究还不多，目前所知只有[29]。作为能够提取与时间相关特征的 RNN，直接用在原始 fMRI 数据上的研究有不少[19][32]，但是将其用于动态功能连接数据的目前还没有，同时也缺乏相关的对比实验，这也是本次想要研究的问题之一，希望能够研究 FC 这种基于原始信号的数据是否比原始信号更具有表述能力，还是会因为缺失了一些信息而失去了一些表述力。

1.2 研究内容概述

本研究使用两种类型的数据集，共设计并实现了两部分的对照试验内容。

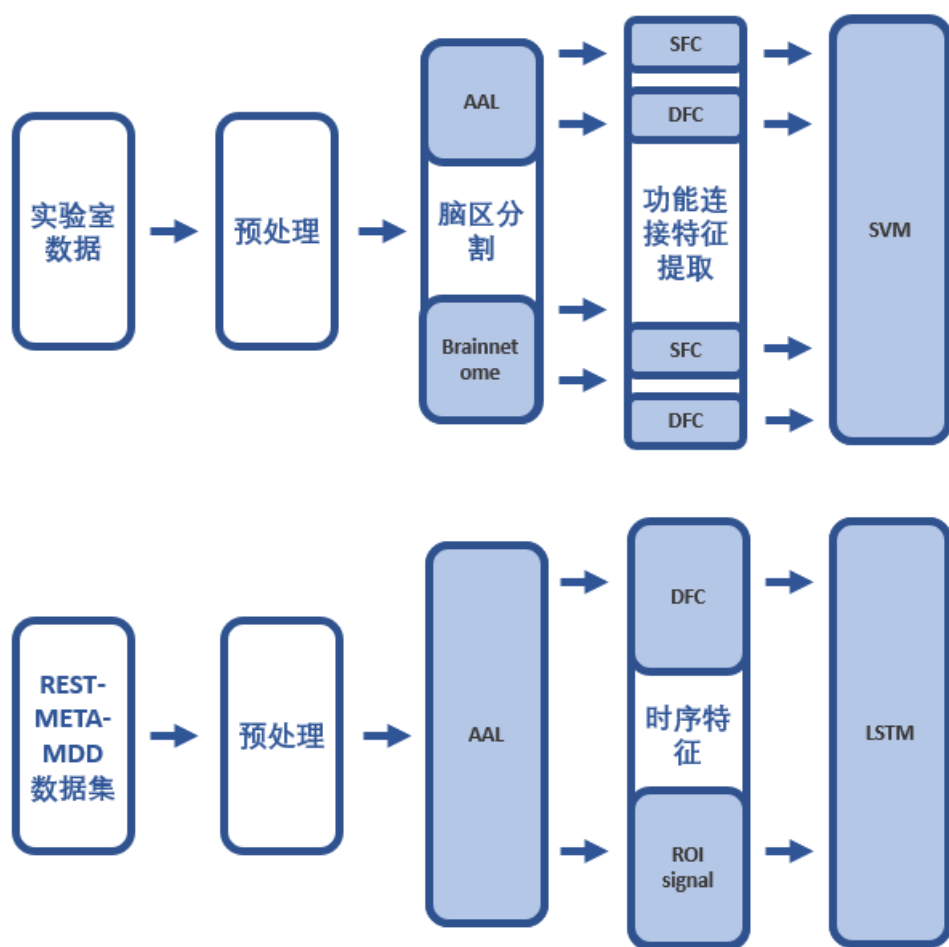
第一个是部分是使用实验室小数量的数据集，基于两种脑图集，AAL 和 Brainnetome，分别提取静态和动态的功能连接（SFC），（DFC），使用 SVM 分类器进行分类，如图图片 1-2 所示，图中上半部分代表该部分实验的内容，共四组对照实验。

第一组实验的主要目的是研究静态功能连接（SFC）和动态功能连接（DFC）特征的分类效果，验证先前的研究[34]表明的 DFC 会比 SFC 特征更具有分类效果。同时，因为功能连接的提取与脑区分割关系比较紧密，因此基于两种脑图集分别做 SFC 和 DFC 的对比实验，从而探究脑图集对于 SFC 和 DFC 特征的分类效果是否有影响。

实验的结果表明，在两种脑图集上，DFC 均比 SFC 数据具有更好的分类效果，并且 DFC 和 SFC 选择的特征对应的生理区域，均与先前的抑郁症相关的生理区域的先验知识一致，证明 DFC 数据比 SFC 数据具有更佳的特征表示能力，也意味着抑郁症患者和健康组之间在大脑的动态特性上很可能也会存在着差异。只是不同的脑图集会带来分类效果的影响，这与脑图集的分辨率有很大的关系。

第二个部分是使用 REST-meta-MDD 数据集，基于 AAL 脑图集，对于两种类型的时序特征，DFC 和 ROI 的区域信号，采用基于 LSTM 的分类器进行分类。

主要目的是研究功能连接这种建立在原始信号之上的特征，和直接使用原始信号进行分类相比，是否更具有表示意义还是会丢失某些信息从而不利于分类。本研究的实验结果表明 DFC 数据比 ROI 区域信号数据具有更好的分类效果。



图片 1-2 整体研究内容

其中动态功能连接（DFC）使用滑动窗格（Sliding window）进行提取，基于

SVM 的分类在分类流程中加上 t 检验和 RFE 进行特征选择降维, 并使用 10 折交叉验证选择最佳参数。

下述文章在第 0 部分介绍了原理和相关概念, 第**错误!未找到引用源。**部分介绍了数据和实验方法, 第**错误!未找到引用源。**部分介绍了实验结果和分析。

代码已上传到 GitHub 上。请访问:https://github.com/duyongqi/fmri-data_based-major-depressive-disorder-ML-DL-classification-/blob/main/README.md 获取实验代码。代码可实现指定文件目录后自动进行特征提取, 判断功能连接类别, 选择指定的地图集, 选择指定的分类方法对数据进行自助训练和测试, 并能够自动保存中间文件, 生成对应的训练和评估曲线图, 输出重要特征的生理意义

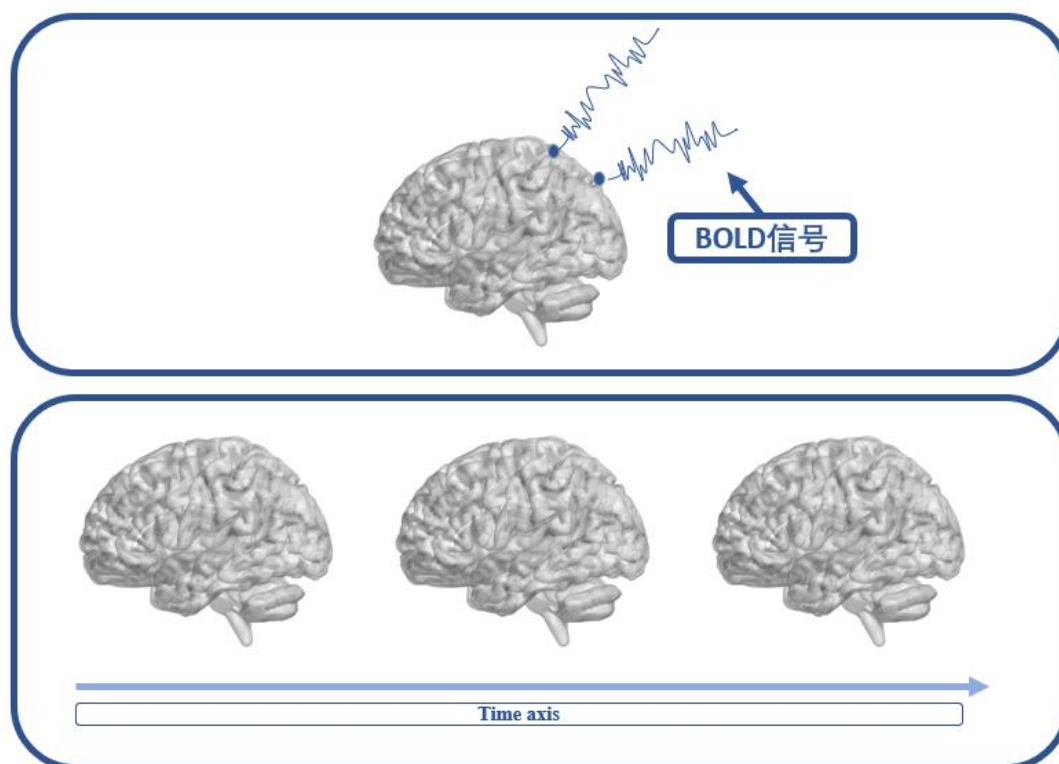
2 相关原理

2.1 fMRI 数据

医学图像可以主要分为两种模态，一种是结构模态的图像，一种是功能模态的图像，像我们常见的 CT 图像，MRI 图像，就是进行结构像采集的结果，而像 PET 图像，fMRI 图像等，是进行功能像采集的结果，功能像的采集一般是依据如某种与功能活动相关的物质的浓度，来进行成像，以此来反应功能的强弱或者功能是否存在异常。

fMRI 成像就是一种功能成像技术，fMRI 的全称是功能磁共振图像 (functional magnetically resonance image)，主要用来采集大脑的血氧浓度水平来反应大脑的功能。fMRI 的采集技术是磁共振成像技术。更加细节的解释是，当大脑的某一个区域比较活跃的时候，这个区域附近的血管会扩张，会使该区域的血氧浓度增加，而血氧浓度的增加会使 MRI 成像采集到的信号更强，从而使得区域的功能活跃度与成像的强度对应起来。也因此把 fMRI 数据中每个体素的采集信号的集合称为血氧水平依赖信号 (blood oxygen level dependent, BOLD)。

fMRI 数据是一个 4 维图像，其中一个维度是时间，另外三个维度是空间坐标维度，可以把 fMRI 数据看成是沿着时间采集的 3 维图像，也可以把 fMRI 数据看成是大脑三维空间中的每一个体素都附有一个 BOLD 信号，如下图图片所示。



图片 2-1 fMRI 数据的两种理解方式

2.2 fMRI 数据预处理

静息态功能磁共振数据的预处理是一个复杂的过程，在前人的不懈研究下，逐渐形成了一套预处理流程，主要包括以下步骤：

(1)首先去除前十个时间点的扫描，以消除磁场不稳定的影响

(2)接着对剩余的 fMRI 图像进行切片时间校正 (slice timing)，这一个步骤的目的是补偿切片之间的采集时间差异。在 fMRI 成像时，一次只能采集一个切片，在一个 TR 内采集所有的切片，这就造成了一个时间点上的 fMRI 数据的不同切片采集时间不同，当 TR 比较长的时候，切片之间的时间差就会比较大，而原始假设是这些切片在同一个时间采集，这个偏差可能会对后续分析产生影响，因此需要进行切片时间矫正。

(3)然后，进行头动矫正，通过平移和旋转执行重新对齐，以校正 fMRI 图像在不同时间点之间的头部运动。

(4)接着进行高分辨率的结构图像与功能图像之间的配准，得到配准参数，并

将 T1 图像分为白质，灰质和脑脊髓液（CSF）信号用于后续的回归协变量。

(5)然后使用从结构图像到 MNI 模板的变形参数将静态 fMRI 图像标准化到标准空间中。

(6)接下来，使用半峰全宽为 6 毫米的高斯滤波器对功能图像进行空间平滑处理。

(7)然后，去除线性漂移，回归出白质信号，CSF 信号和头部运动参数的影响。

(8)最后，使用 0.01~0.1Hz 的带通滤波器进行滤波以去除时间域的噪声。

抑郁症的数据预处理的整个过程已经有一些集成的软件可以使用，在上文中也有所介绍，如 FSL[45]，SPM[47]，AFNI[46]，DPABI[49]，DPARSF[13]，以及 python 的安装包 fMRIPrep[48]等工具或代码安装包。

2.3 功能连接计算

由于功能磁共振数据在空间维度上分辨率比较高，如果直接将所有体素对应的 BOLD 信号纳入考虑进行后续的分析，将会遇到维度灾难，并且将会受到严重的噪声干扰，同时这些体素之间是存在着相似性的，也就是采集到的信号事实上在空间维度上存在着冗余，因此，在进行后续的处理之前，首先将大脑划分成一些区域，从而减少空间冗余，更加方便后续实验。

大脑区域划分，即脑区分割也是一个被许多学者关注的领域，有许多基于不同算法实现的分割结果构建的脑图集，公开在网络上方方便研究人员获取和研究。如下图图片 2-2 脑区分割示意图所示。获取脑图集之后对预处理后的 fMRI 数据进行脑区分割，将分割的每个区域的 BOLD 信号进行平均，作为这个区域的 BOLD 信号，从而得到脑图集个数分辨率的 fMRI 数据。



图片 2-2 脑区分割示意图

在获取了分割后的 fMRI 数据后, 下一步是计算动态功能连接 (DFC) 和静态功能连接 (SFC)。功能连接矩阵事实上就是上一步分割后, 得到了每个区域的 BOLD 信号, 计算这些信号之间的相关性, 得到的相关矩阵, 就被成为功能连接矩阵。其中信号之间相关性的度量也有许多方法, 其中最常用的是皮尔森相关系数 (Pearson coefficient), 这也是本文中计算功能连接矩阵, 量度区域信号之间相关性的方法。

对于 SFC 的提取, 直接基于皮尔森相关系数构建相关矩阵即可获得, 而对于 DFC 的提取, 存在着许多方式, 其中最常用的是基于滑动窗格的动态功能连接特征的提取, 它是最广泛用于评估大脑动态功能连接的算法之一。

滑动窗格算法进行动态功能连接特征提取的过程是, 首先使用窗口内的数据, 计算这个窗口内部的数据的短时静态功能连接, 接着将窗格移动一个步长, 在新的窗口中接着计算短时静态功能连接, 不断移动窗格, 最终得到动态功能连接数据。在这个过程中很重要的步骤就是确定窗格的长度, 即窗格对应的时间点数, 先前研究建议, 在该任务中窗格的宽度应该不小于 $1/f_{min}$, f_{min} 代表信号的最小频率。由于在预处理的环节中进行了信号的滤波处理, 得到的信号的最小频率为 0.01HZ, 由此可得到对应得最小窗格个数。

以上即为 fMRI 相关知识的介绍, 下面将介绍深度学习和机器学习相关的知识。

2.4 数据划分

在做深度学习, 机器学习相关实验的时候, 都会有一个数据划分的步骤, 将整个数据划分为训练集, 验证集和测试集 3 个部分, 这一小节将详细的介绍什么是训练集, 验证集, 测试集, 以及为什么需要这些集合的划分。

一般情况下, 在建立一个机器学习或深度学习的模型之后, 训练的最终目的是让模型能够学习到数据的整体分布, 而非仅仅去拟合训练数据, 那么应该如何评估模型是否学习到了数据的分布。一个重要的办法是, 再取另一个与训练数据同分布的数据, 这些数据未必与训练数据相同, 只是与训练数据服从相同的分布, 类似于在一个分布下做采样, 一部分样本用来训练使得模型能够学习到分布, 另一部分样本输入到训练后的模型中, 如果模型能够学习到数据的分布, 那么模型

在面对这些与训练集不同却是同一个分布的采样的数据的时候,也能够有很好的效果,这个与训练集不同却又来自与同一个分布的数据,就是我们常说的测试集。用来训练的数据被称为训练集。

模型在训练集上的表现可以量度算法是否收敛,模型是否有很好的学习效果,模型在测试集上的表现则可以量度模型的泛化性,也就是模型能否学习到整体的分布。训练集和测试集一般来自于整个数据集的随机划分,因此在进行机器学习或深度学习的相关实验的时候,都会有一个数据划分的步骤,也就是将整个数据集按照一定的比例随机划分为训练集和测试集,一般训练集需要多于测试集,以更好的训练模型。

但是如果仅仅将数据划分为训练集和测试集,会出现一个问题。一般情况下,模型都含有一些可以调节的人工参数,比如深度框架中节点的个数,正则化参数,机器学习中特征选择的个数,分类器的参数等等,这些参数的改变也会使模型的效果发生改变,参数的调节,其实也就是依据模型的效果进行调节,我们希望模型的学习性能要好,也希望模型泛化性要好,如果基于测试集的评价指标值作为依据进行模型参数的调节,那么模型实际上在调节的过程中,在逐渐向测试集靠拢,也就是说,测试集的信息实际上泄露给了模型,模型也会向着拟合测试集的方向靠近,因为信息的泄露,测试集在这个时候作为模型泛化性的评价已经没有任何意义了,因此我们还需要一个验证集。

验证集就是在之前所说的数据集划分时划分为三份,其中的一份被称为验证集,如图所示。验证集的作用,一个是在训练的时候,通过验证集的效果评估模型的泛化性,一个是基于验证集的效果进行参数的调节,而模型最终的效果评价,是来自于网络训练和参数调节过程中未曾出现过的测试集,以测试集的效果作为模型最终的评价,



图片 2-3 数据集的划分

在数据样本数量比较小的时候,划分成三个部分很容易导致划分的集合具有

不同的分布,因此在遇到比较少的数据的时候,也可以仅划分为训练集和测试集,但是需要通过交叉验证的方式进行测试集上效果的评估。交叉验证的意思是



图片 2-4 十折交叉验证示意图

本实验为分类任务,为了保证训练集和测试集分布尽量相同,在划分时选择按照类别取样,而非随机划分。

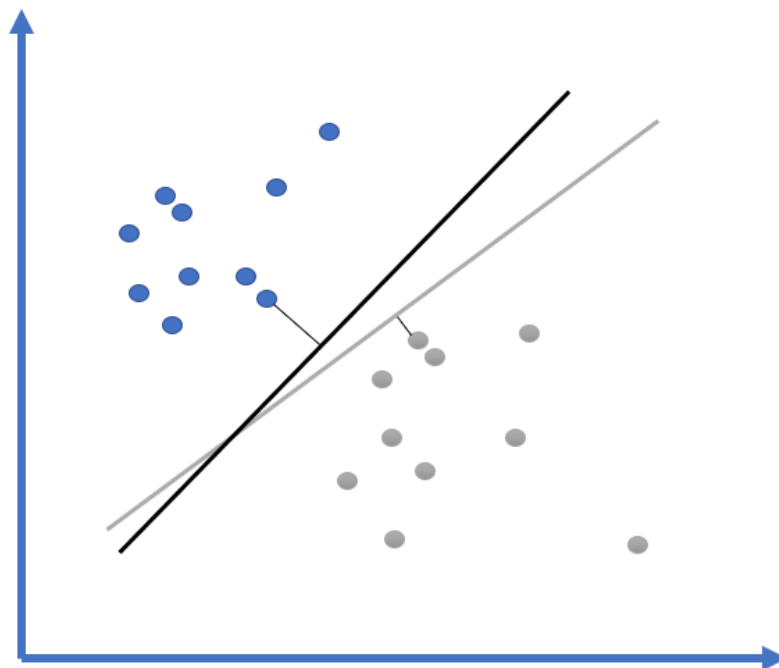
2.5 样本均衡

正如上一小节所述,机器学习和深度学习实际上是想要学习到数据的分布,如果在分类问题中训练数据中的样本不均衡,那么对于数量较多的类别,模型应能更好的学习数据的分布,对于数量较少的类别,模型学习到的分布会不够稳定,同时可能会受到数据较多类别的影响,因此在进行训练的时候,首先进行样本的均衡。在该实验中,样本均衡选择的方式是对样本较少的类别数据进行上采样,选用的上采样算法是 SMOTE 算法。SMOTE 算法主要基于 KNN 和随即线性插值实现的,除此以外还有一些其他的如下采样的方法均可以用来进行样本均衡。

2.6 SVM

支持向量机(SVM)是一种监督式机器学习模型,全称为支持向量机(support vector machine)。SVM 算法是试图寻找能将两类样本分类,且与两类样本的距离最小值最大的超平面,超平面的意思是指特征空间的一个平面,如图所示灰色的超平面虽然能将样本分隔开,但是和一个样本之间的距离非常小,而另一条黑色

的超平面，和两类样本之间的最近距离均大于灰色的超平面，因此黑色的超平面优于灰色的超平面，而使超平面与两类样本之间的距离最小值最大的超平面就是 SVM 算法最终寻找的超平面。其中离最佳超平面距离最近的点被称为支持向量，注意支持向量最少需要两个。



图片 2-5 SVM 示意图

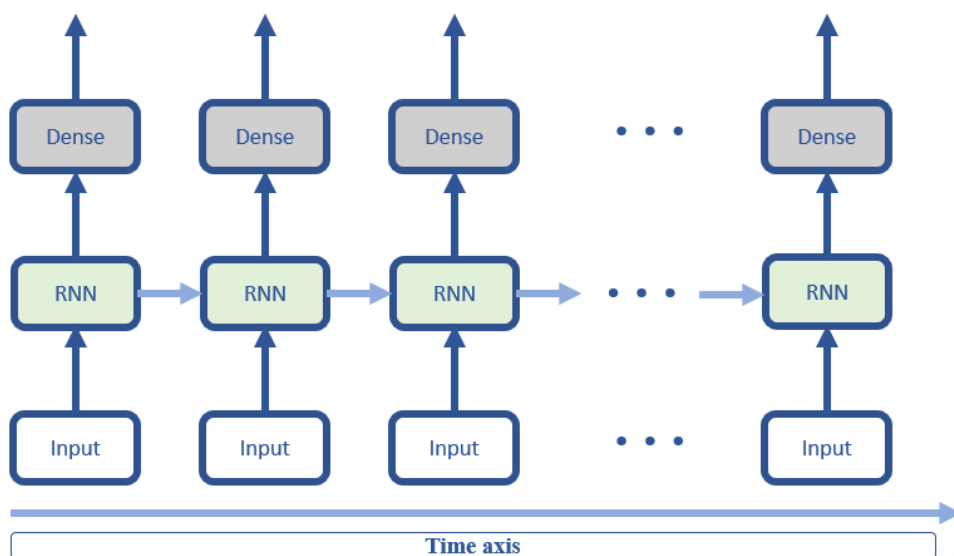
SVM 可以分为线性 SVM 和非线性 SVM。最初的 SVM 是线性分类器，并且有很好的分类效果，线性 SVM 又可以分为硬间隔分类器和软间隔分类器，它们的含义分别是样本需要严格线性可分和样本可以存在一些在支持向量撑起的间隔之间分布的样本点从而使分类器更加稳定而不易受极端样本数据的影响。

而非线性 SVM 则可以对线性不可分的样本数据进行分类，主要的实现方式是通过核技巧，将非线性分布的样本映射到更高维的空间，使其具有线性可分的分布，实现这一映射的函数被称为核函数，有多种可选择的核函数，包括线性核函数，多项式核函数以及 RBF 核函数，一般情况下，当数据本身就处于比较高的维度空间的时候，这个时候再用核函数进行高维空间的映射会使得计算复杂同时可能效果提升不大，因此在样本主要处于高维空间，也就是特征的个数比较多的情况下（这个特征个数比较多是相对于样本量来说的）一般选择线性核函数，

基于线性核函数的分类器也就是上述中提到过的线性 SVM。而当样本分布的空间维度比较低，也就是样本特征相对较少的时候，主要选择 RBF 核函数。

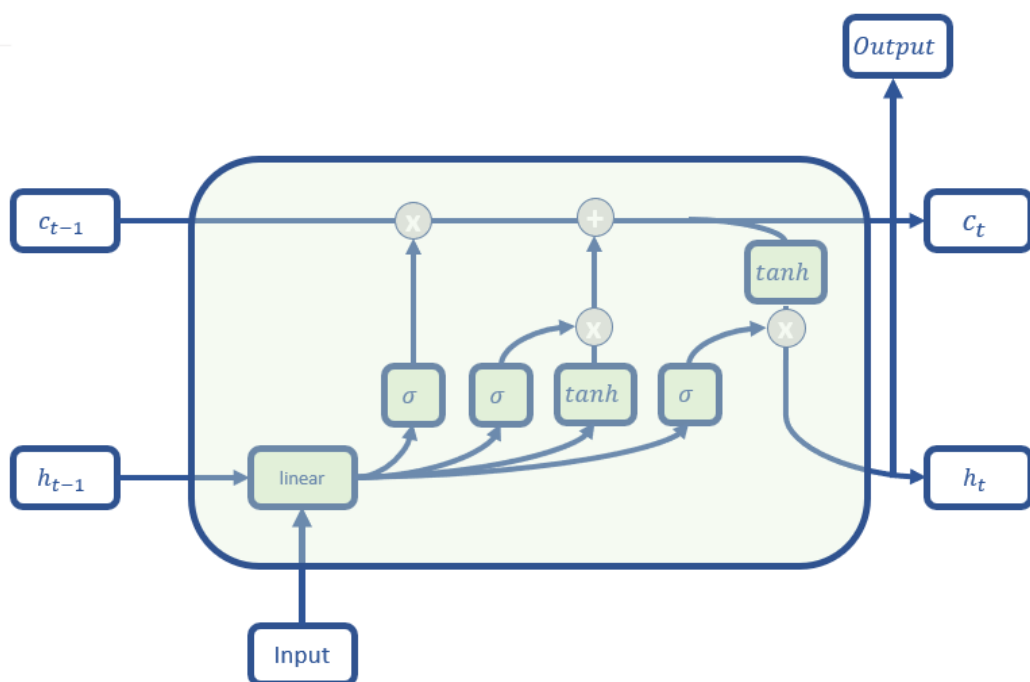
2.7 LSTM

LSTM 是一种考虑时间特征的深度学习基础单元，是 RNN 的一种变体，与传统 RNN 相比能够记忆更长的序列特征。RNN 全称为循环神经网络，其示意图如下图片所示。网络的 RNN 层在上一个时间点的输出值，除了接着传给下一层，例如图中所示的 Dense 层，还会将与之相关的值，也被称为隐藏值，传给下一个时间点的 RNN 层，使得 RNN 层提取的特征能够考虑整个输入序列的时间特性。



图片 2-6 RNN 示意图

LSTM 是 RNN 的一种变体，由于 RNN 在面对比较长的序列的时候，会存在梯度消失和梯度爆炸的情况，为了解决这些问题，前人提出了如 LSTM，GRU 等一系列的改进结构，在本次研究中为了量度长时间的序列选择了 LSTM，并使用双向 LSTM 更好的量度长序列的动态特性。其原理可以参考图片：



图片 2-7 LSTM 结构示意图

LSTM 作为深度学习的一层的时候,向下一层的输出 $output$,一般与 h 相同,如图中所示。而 h 和存储器 c 一同输给下一个时刻的 LSTM 层,这个存储器就是重要的能够实现长序列记忆的结构。可以将内部的每一个需要双曲正切激活的函数看作是一个权值系数,而每一个 sigmoid 激活函数,也就是图中的 σ 代表着一个开关,可以将其看作是一个开关电路,基于这些就可以理解 LSTM 的原理实际上是通过一些开关控制是否将上一个时刻的隐藏值和存储单元的存储值保留下来,进而计算输入到下一个时刻的存储值和隐藏值,以及输入到下一层网络中的输出值。

2.8 评价指标

在基于机器学习和深度学习处理分类问题的时候,主要的评价指标包括准确率,基于混淆矩阵的一些计算量,如 PR 曲线的面积,精确率,召回率, AUC 值等,下面主要介绍本次实验中用到的评价指标 AUC 和 Recall。

Recall 代表的是实际的正样本被预测为正样本概率。

AUC 的全称是曲线下面积(area under curve),是 ROC 曲线的曲线下面积,ROC 曲线是基于混淆矩阵计算的,通过不断地调节分类的阈值,计算负样本被

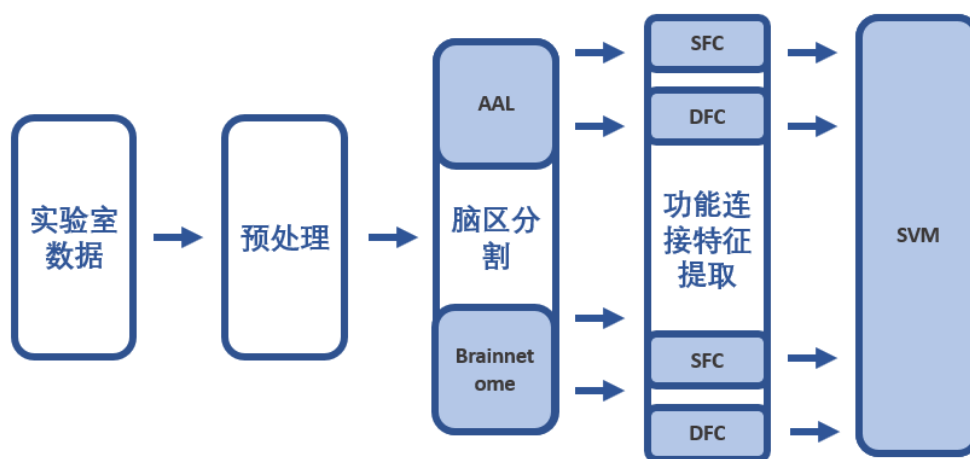
分为正样本的比例(FPR)以及正样本被分为正样本的比例(TPR),在图上绘制出以(FPR, TPR)为坐标的点,不断改变阈值,连接这些点就可以得到 ROC 曲线,曲线下的面积就是 AUC 值。

另外在第二组实验中,对照文献中使用的指标,也是医学中常用的指标,包括准确度,特异度和敏感度,第一个是分类的准确率,后面两个指标中的特异度为正常人不被误诊的概率,敏感度是患者能够被诊断出来的概率,这两者在医学中有重要的意义,因此一些医学领域的分类或者说诊断问题中,常常以这几个指标为诊断评价指标。

3 基于 fMRI 的抑郁症诊断方法设计

3.1 基于传统机器学习 SVM 算法的抑郁症诊断

第一组实验的总任务如下：



图片 3-1 第一组实验任务

3.1.1 功能连接数据提取前处理

借助 DPABI[13]工具在原始数据上实现包括去除前十个时间点，头动矫正，切片时间矫正，配准到标准空间，去除协变量以及时间和空间标准化等一系列预处理操作，最终得到四维 fMRI 预处理数据，包括时间一个维度和空间三个维度。根据预处理的结果和中间参数，检查被试数据的质量，排除头动参数大于 3mm 的被试。

借助 GREYNET[51]工具将经过预处理后的每个被试的 fMRI 数据进行脑区分割，将分割区域内的 BOLD 信号进行平均得到每个区域对应的 BOLD 信号，实现了原始数据中以体素为单位的 BOLD 信号降到区域单位的 BOLD 信号。其中脑区分割选择了两个脑图集，以进行后续的比较。分别是 AAL116[52]和 Brainnetome[53]脑图集，AAL116 脑图集是脑科学专家根据已有经验知识对大脑进行的解剖学标记，将大脑划分为 116 个区域。Brainnetome 脑图集是基于结构

像和功能像的多模态脑区分割, 基于聚类的方法对大脑进行的区域划分, 共划分为 246 个区域。两个脑图集均可在网上进行下载, 同时脑图集对应的区域生理解释在网上的一些论坛中分享, 通过 Python 进行网页爬取解析以及存储, 以待后续生理解释所需。

借助 GREYNET[51]工具对脑区分割后的数据进行 SFC 和 DFC 的提取, 其中提取 DFC 的窗格长度为 100s, 共记 125 个时间点, 步长为 1, 共得到 125 个短时 FC, 以待后续分析。(在后续分析中, 考虑到特征过多带来的计算资源的消耗和噪声的干扰, 在得到步长为 1 的 DFC 后, 可以进行时间维度的采样获得其他更长步长的 DFC 数据, 实验中为加快计算速度, 最终采样间隔为 4 个时间点, 等价于将步长设置为 5)。

3.1.2 分类数据准备

3.1.2.1 SFC 数据准备

经过预处理流程和脑区分割以及功能连接的计算, 对于静态功能连接(SFC)数据, 每个被试对应一个二维的对称矩阵, 如图 3-2 中的输入矩阵; 为了后续的分类, 将对称矩阵的上半部分提取出来, 减少数据冗余, 接着将上半矩阵按行进行拼接之后, 得到最终用来分类的一维向量, 经过 SFC 数据分类前处理流程之后, 每个被试的特征是一个一维的向量, 将这些向量输入分类 pipeline 中进行相应的训练和测试。如图 3-2 所示。

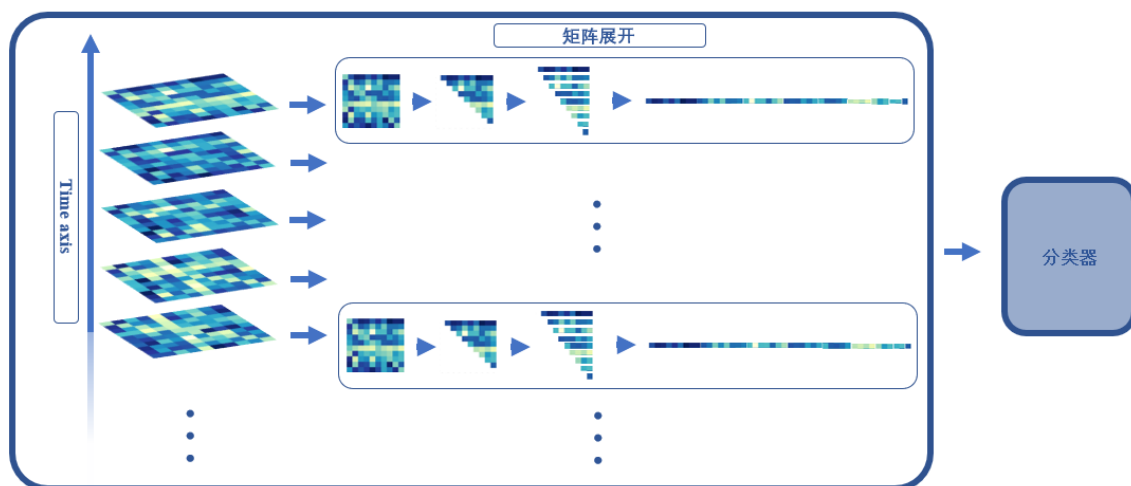


图片 3-2 SFC 数据分类前处理流程

3.1.2.2 DFC 数据准备

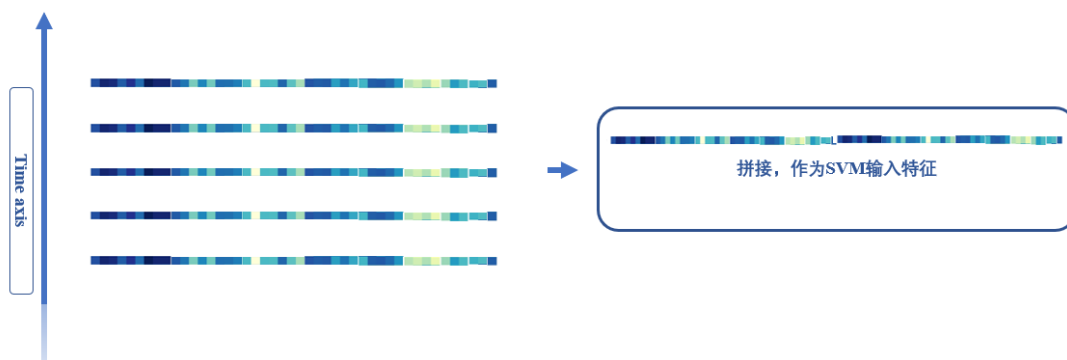
对于动态功能连接(DFC)数据, 每个被试对应一个三维矩阵, 见图片 3-3 中的输入三维矩阵, 动态功能连接(DFC)可以看作是沿着时间轴的短时 SFC 的集合, 对每个时间点对应的短时静态功能连接(SFC)进行静态功能连接(SFC)

数据前处理流程,进行上半矩阵的提取以及上半矩阵的拆分和拼接,最后在每个时间点,得到一个一维的特征向量,沿着时间轴的这些一维特征向量组成一个被试的最终用来分类的特征向量,是一个二维矩阵,包括时间和向量空间两个维度。如图图片 所示



图片 3-3 DFC 数据分类前处理流程

在将 DFC 数据进行矩阵展开后,需要将展开后的向量进行拼接用于后续的分类输入,如图片 所示,将展开后的矩阵经过拼接,这样每个被试的 DFC 数据就可以转换为一个一维的向量,可以用 SVM 对这个一维向量进行分类。



图片 3-4 DFC 经过矩阵展开之后的操作

3.1.3 数据划分和样本均衡

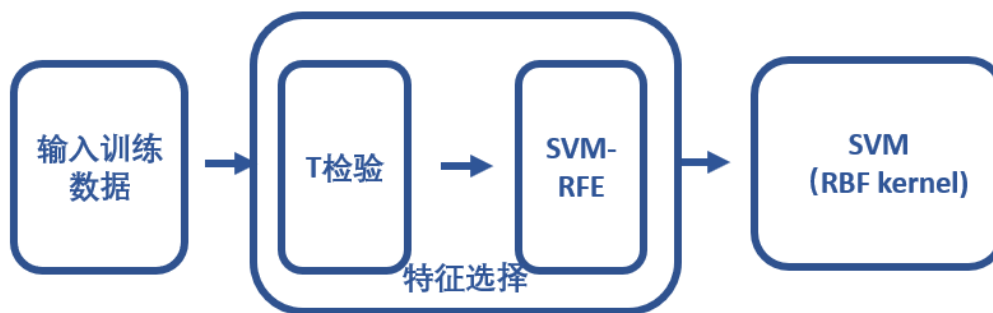
在进行了数据准备,将数据转换到对应的结构之后,就可以将这些特征输入到对应的分类器里做分类了。将数据集按照测试集 0.2 的比例划分为训练集和测

试集，并未继续划分验证集，因为实验数据量过少，划分成多个集合会进一步减少学习数据的数量，同时即使在验证集上选择了最佳的模型，划分的测试集因为数量过少也有可能存在一些比较偏离原始数据分布的样本，这样的评价是没有意义的。因此在本次实验中仅仅划分为训练集和测试集合，并计算十折交叉验证的分数作为测试集的分数。

同时为了防止因为类别的数目的差异带来的学习误差，在划分了训练集和测试集之后，对训练集进行样本均衡，以消除类别数目的差异带来的影响，虽然数据集中的两种类别数目比较接近，但是因为数据总体个数太少，很少的差距可能会有比较大的影响，因此在实验中增加了这一个步骤。

3.1.4 基于 SVM 的分类流程

基于 SVM 的的分类的流程包括特征选择过程和分类过程。对于经过数据准备后的功能连接特征，由于分割时的区域一般在百数量级，一个 SFC 矩阵展开成一维向量后就有几千个特征，而对于 DFC 矩阵对应的数据准备后的特征则更多，为了减少特征中存在的噪声对后续分类的影响，同时为了使特征维度尽量降低以便特征的存储，在分类前首先进行特征的选择。特征选择分为两个步骤分别是 t 检验和 SVM-RFE，t 检验进行粗略的特征选择，在 t 检验进行之后，SVM-RFE 进行更加精细的特征选择，并且将待选择的特征个数作为超参数在后续进行调节和最优参数的选择。在特征选择后，将选择出来的特征作为最终进行分类的特征，采用基于 RBF 核的非线性 SVM 对特征提取后的特征进行分类，RBF 核的两个参数同样作为待调节的超参数在后续进行调节和最优参数的选择。整个分类流程如下图所示：



图片 3-5 SVM pipeline

3.1.4.1 t test 特征初步选择

在将功能连接特征作为分类特征时，由于脑区分割的分割区域个数较大，而功能连接矩阵的尺寸时脑区域个数的平方，尽管只提取了上半矩阵，经过 SFC 数据准备流程和 DFC 数据准备流程处理后的 SFC 数据，特别是 DFC 数据的特征维数会非常高，为了避免噪声的干扰，消除信息的冗余以及便于在后续分类中选择 RBF 核函数做特征分类，在分类 pipeline 的第一步，首先先对数据进行独立双样本 t test，基于显著性值进行特征选择，在 SFC 数据分类时，将 t test 的显著形值中的选择百分比定为 30，在 DFC 数据进行分类时，将 t test 的显著性值得选择百分比设定为 0.3，以达到特征初步选择得目的。

值得一提的是，这里进行特征降维的时候，并未选择像主成分分析（PCA）等一类的方法，是因为在后续需要对提取的特征进行生理上的映射和解释，直接做主成分分析等这一类的降维将不易于后续特征的生理映射与解释，因此在这里首先选择 t test，这种类似于基于可信的原因给特征加上一个遮罩的方法，在后续的解释中可以清楚的看到遮罩对应的特征，从而更易于解释选择出来的特征是哪些以及有什么意义。

3.1.4.2 SVM-RFE 特征选择

RFE 是一种特征选择方式，全称是递归特征消除（Recursive feature elimination），递归特征消除在进行的时候需要设定一个评估器，也就是能够评估特征重要程度的一个评估模型，递归特征消除之后就会根据这个评估器给出的特

征权重,逐次迭代删除最不重要的特征,知道特征个数达到想要的特征个数位置。整个流程具体化描述是,首先,在初始特征集上训练估计器,根据评估器的特征权重得到每个特征的重要性。然后,从当前的特征中删除最不重要的特征,删除的特征个数可以自己设定,一般被称为特征选择的步长,在进行了第一步的剔除特征后,接着在剩余的特征上继续进行特征的评估和删除,不断地迭代直到剩余特征的个数为想要选择的特征个数为止,最终根据删除的顺序,对特征进行排序,并取最后剩余的特征为选择出来的特征。而 SVM-RFE 是将 SVM 作为 RFE 的评估器的方法。

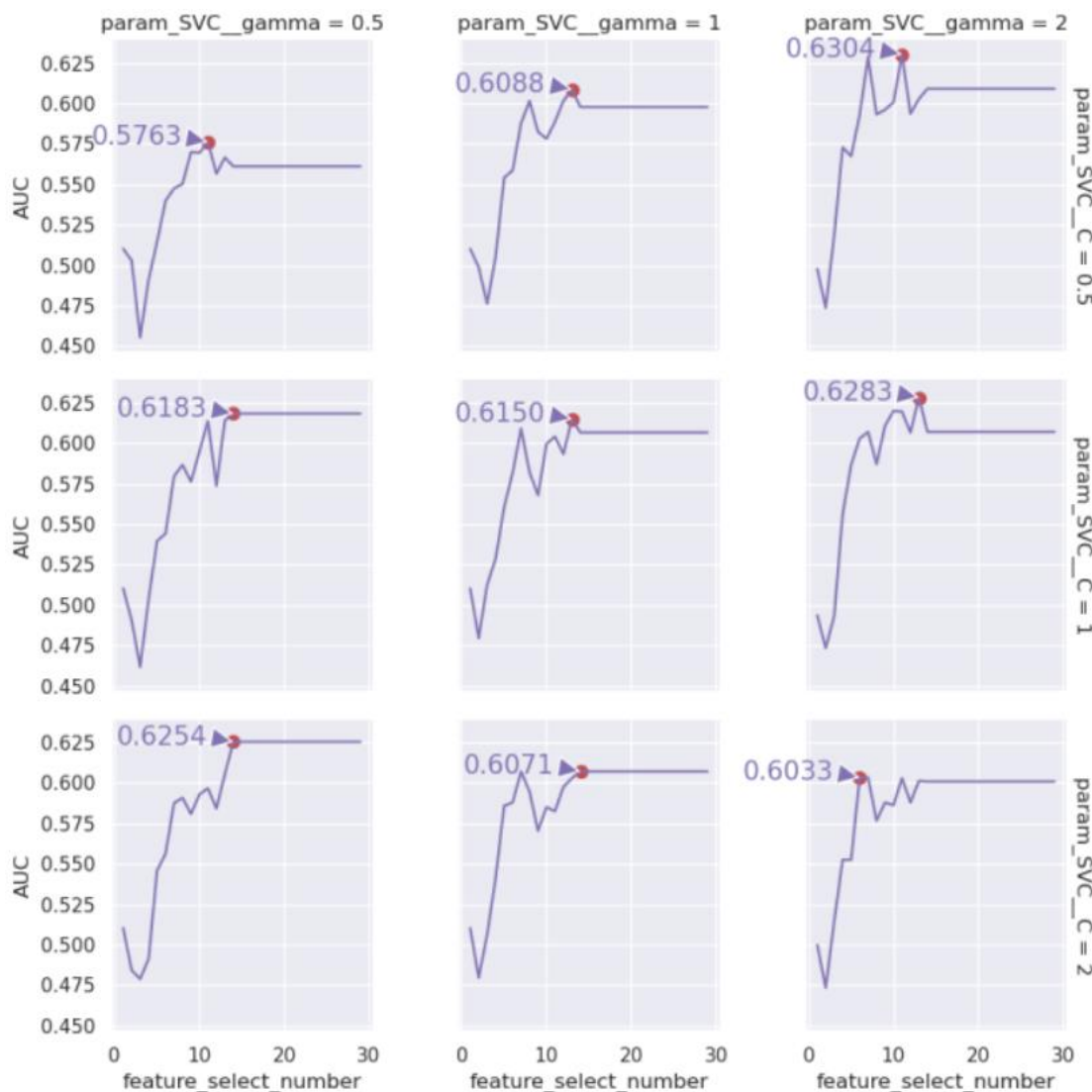
在经过 t test 进行初步得特征选择后,为了使得特征得数目能够尽量少,也使得特征得个数更加适合后续非线性 SVM 的训练,在训练前加上了 SVM-RFE 进行特征选择,特征选择的个数作为等待交叉验证的参数在后续与其他参数一同在 Grid search 中进行组合与泛化性验证,通过选择验证结果最好的一组参数作为最优参数,从而将模型最终确定下来,这样就可以避免人工调节参数带来的繁琐。

3.1.4.3 非线性 SVM 进行特征分类

经过以上两步特征选择之后,目前的特征数量已经很适合做 SVM 的非线性分类了,非线性 SVM 的核函数选择 RBF 函数,函数的两个待确定的参数 C 和 γ 同样作为等待交叉验证的参数在后续与其他参数一同在 Grid search 中进行组合与泛化性验证,通过选择验证结果最好的一组参数作为最优参数,从而将模型最终确定下来,其中验证过程选择十折交叉验证。

借助 scikit-learn 库进行相应 pipeline 的构建和 Gridsearch 以及十折交叉验证的实现。

在实验过程中为了进行特征选择,将参数与评价指标 AUC 的关系可视化为以下的曲线图便于进行进一步的参数选择,如图片 所示。



图片 3-6 可视化参数与验证结果之间的关系

图中每一行代表一个 C 参数的值，每一列代表一个 γ 参数的值，曲线绘制的是 RFE 特征选择的个数和十折交叉验证 AUC 值的关系，并将每一对 C 和 γ 下具有最佳效果的 RFE 特征选择个数标记出来，并标注 AUC 的值，以便后续进一步缩小参数搜索范围。

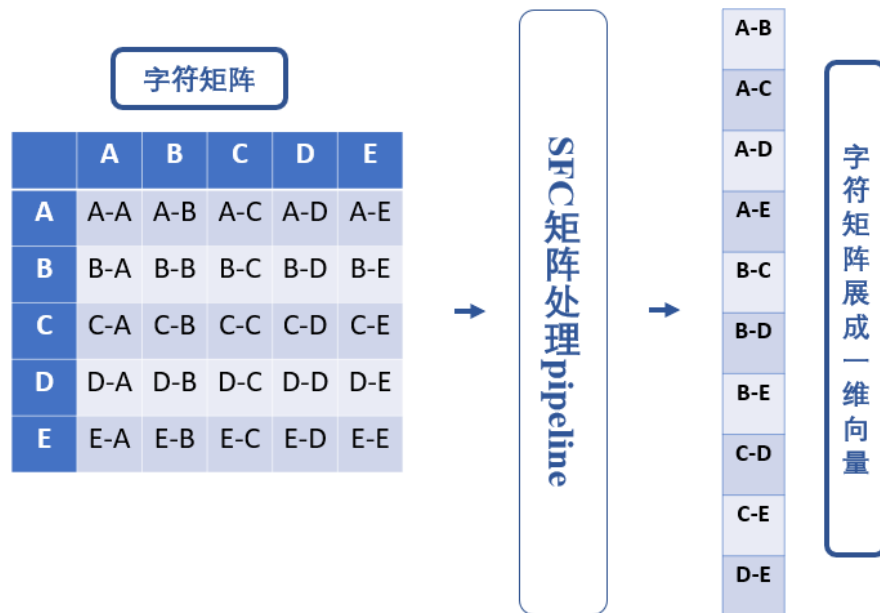
3.1.4.4 特征的生理理解

将机器学习等数据科学手段应用在医学问题里，一个很重要的部分就是特征或结果等数据的应用于医学上的生理理解。在上文中也提到过，模型设计事实上是带着可解释性的想法进行设计的，在这一步骤，根据前两个环节，这两个环节都是类似于给输入数据加上了遮罩再进行 SVM 分类的训练，因此在解释网络的

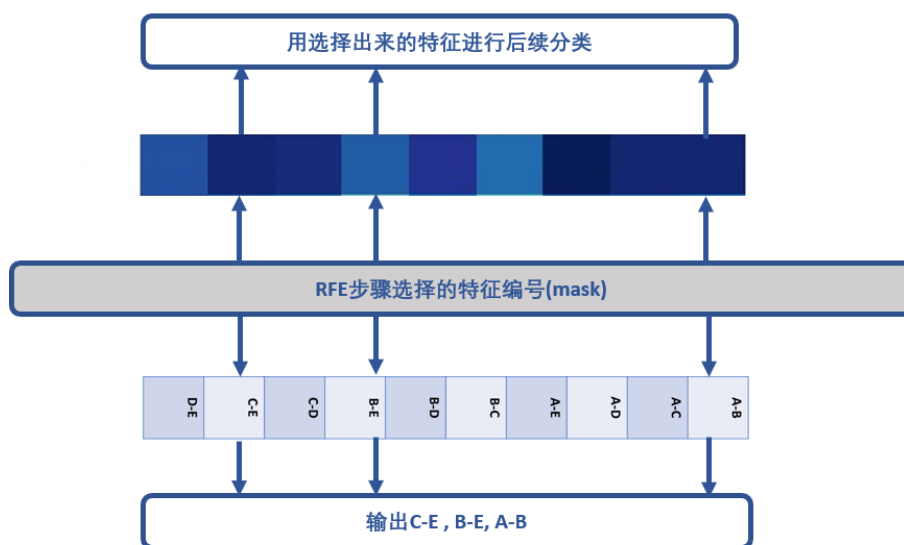
时候，基于 RFE 步骤中选择特征的“遮罩”寻找到特征原始的编号 (index)，接着寻找该编号对应的原始 FC 矩阵中的位置对应了哪两个脑区，从而得到选择的特征对应 FC 值的物理解释。

本文实现这个对应的方法是构建一个大脑区域对的字符类型的功能连接矩阵，，接着将这个字符矩阵做 SFC 数据准备流程同样的操作，取这个字符矩阵的上半矩阵，将上本矩阵的每一行拼接起来，既可以得到一个字符类型的向量，如图片 所示。

这个向量和 FC 矩阵拼接得到的向量同一个编号对应的是两个原始矩阵的同一个位置，而字符原始矩阵和 FC 原始矩阵是一一对应，见图图片 所示，因此，在得到 RFE 步骤的特征选择编号后，就可以根据特征编号输出特征的生理意义。



图片 3-7 脑区字符矩阵，用于解释选择的特征对应哪两个脑区的功能连接

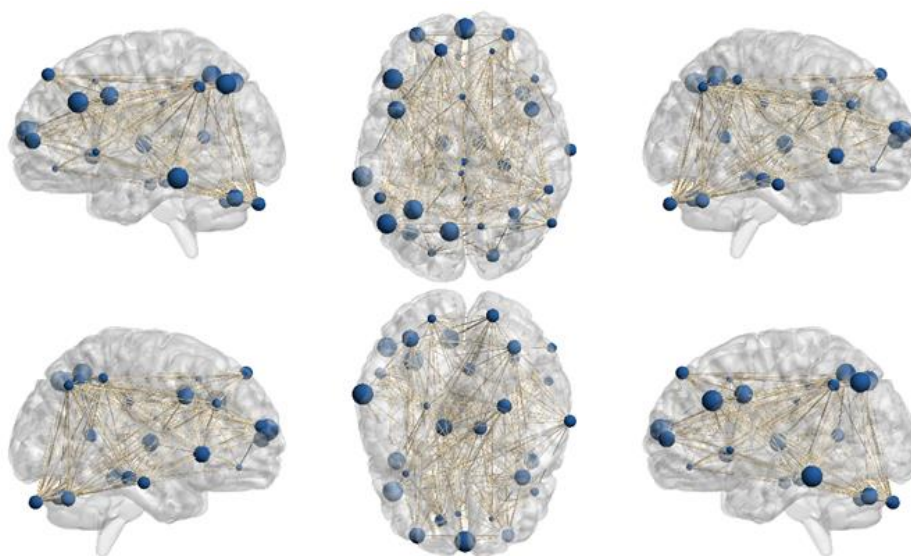


图片 3-8 字符矩阵展开后, 如何进行选择特征的生理解释

特别的是, 在 DFC 分类任务中的生理解释, 除了将编号和生理意义对应起来, 还会根据编号的范围确定该编号对应的特征是哪一個时间点的哪一個生理特征, 以便更好的理解动态特性。

3.1.4.5 特征的可视化

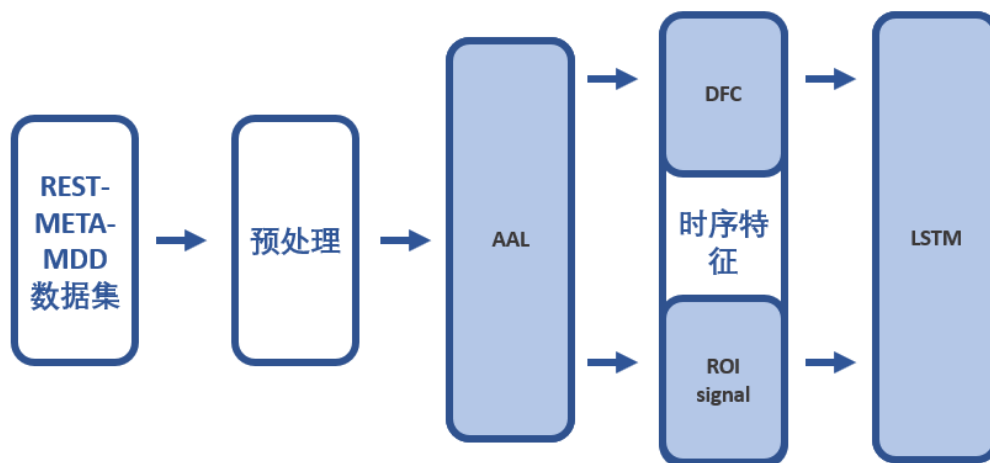
借助 BrainNet Viewer 工具, 将选择出来的特征对应的脑区节点和连接绘制在大脑模型上, 如图所示:



图片 3-9 BrainNet Viewer 可视化示意图

3.2 基于 BiLSTM 深度学习方法的抑郁症诊断

第二组实验的任务如下图所示：



图片 3-10 第二组实验任务

3.2.1 数据前处理

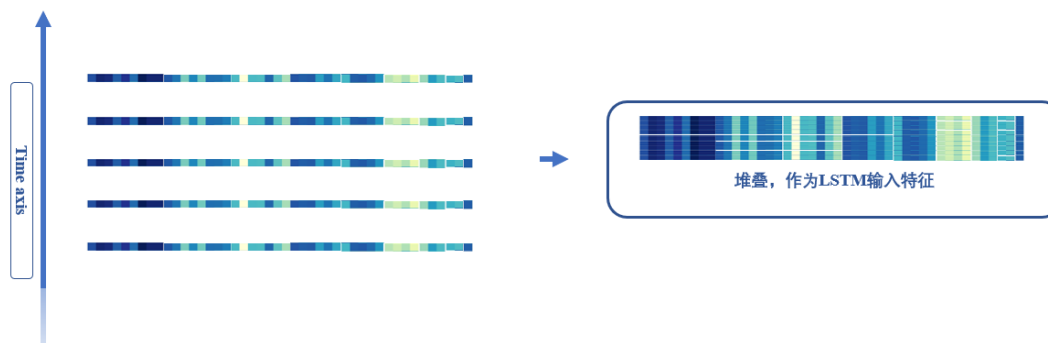
由于该数据集提供的是已经经过预处理和脑区分割的数据，因此前两个步骤在该实验中不需要另外实现，数据集提供了脑区分割后 ROI 区域的信号数据。

因此首先基于提供的 ROI 区域信号数据，手动编程实现基于滑动窗格的 DFC 特征的提取，与之前 SVM 中的实验类似，滑动步长选择为 5，得到 20 个 FC 矩阵。

3.2.2 数据准备

3.2.2.1 DFC 数据准备

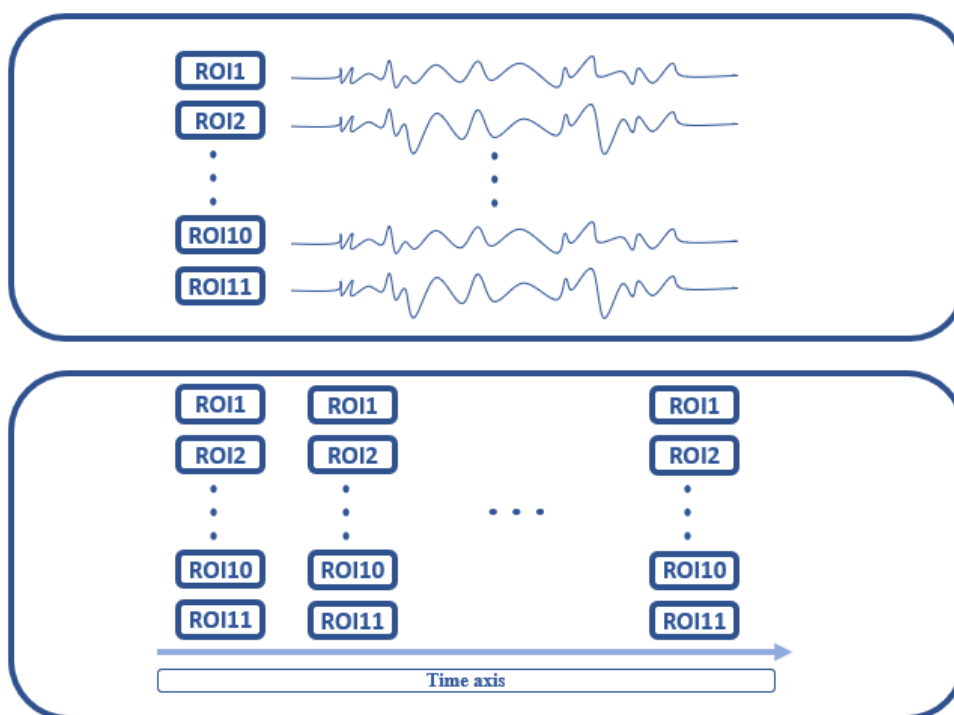
DFC 的准备阶段与上一个实验类似，首先将沿着时间的短时 FC 展开成一维向量，如图图片 所示。接着将展开后的向量进行堆叠，这样每个被试的 DFC 数据会转成一个二维矩阵，矩阵的一个维度是时间维度，可以用基于 LSTM 的深度学习框架对特征的时间序列进行分类，如图图片 所示：



图片 3-11 DFC 堆叠前处理

3.2.2.2 ROI 信号数据准备

由于该数据集提供的数据包括 ROI 信号数据, 这些信号在之前的预处理阶段就已经经过了时域的平滑, 因此不需要再对信号进行处理, ROI 信号数据的格式如图所示, 将数据看成沿着时间的区域信号强度组合, 即可输入到基于 LSTM 的分类器里。



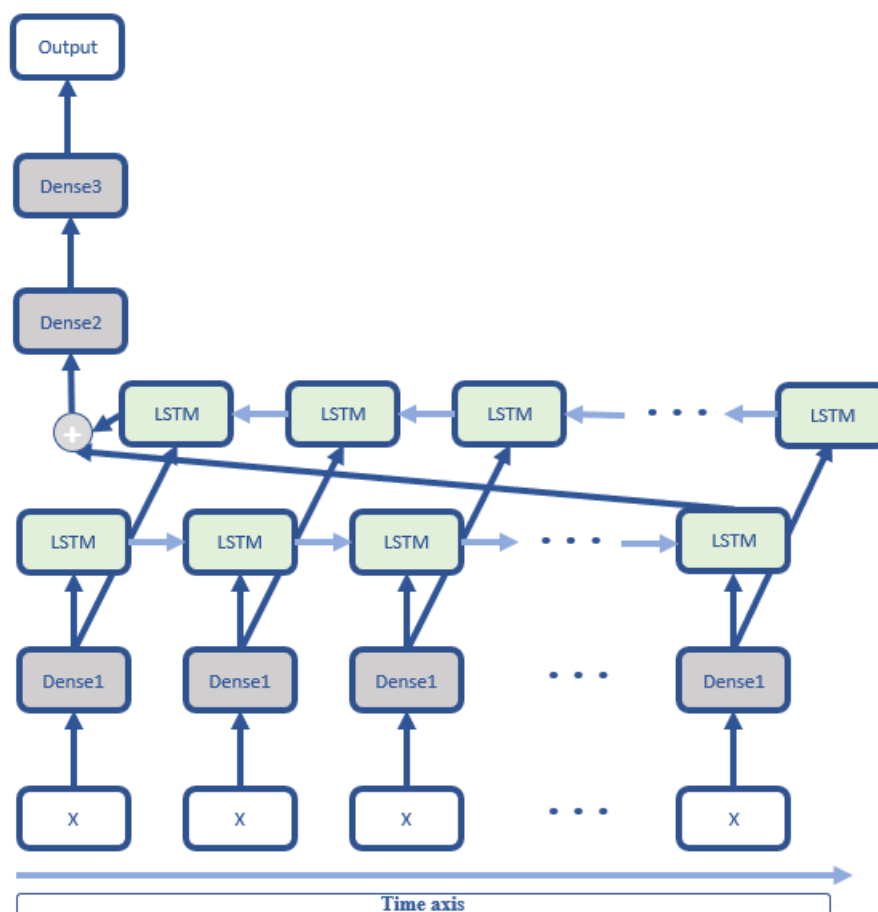
图片 3-12 ROI 数据结构示意图

3.2.3 数据划分

将数据集按照 4:1 的比例划分为训练集和测试集,同时将训练集按照 4: 1 的比例划分为训练集和验证集,基于验证集的评估效果调节模型。

3.2.4 基于 LSTM 的分类器模型

如下图所示代表了整个模型的框架,利用一个基于双向 LSTM 的架构实现分类器。分类器在输入与 LSTM 层之间加了一个全连接层做初步的特征提取,在输出前加上一个全连接层进行特征的整合,最终经过一个激活函数为 softmax 的含有两个单元的全连接层,得到最终的输出用来计算交叉熵损失。在 lstm 层前和层后加了 dropout 层和 batch normalize 层,另外在第一层加上了 L1 正则化防止过拟合。



图片 3-13 深度学习框架示意图

模型参数设置如表格 3-1 模型参数设置表所示。

对 DFC 数据进行分类时候,考虑到时间点数更少,特征维度更大,将第一层输出维度设置为 32, LSTM 层输出维度设置为 16, 第三层输出维度设置为 16.

对原始信号进行分类时,考虑到时间点数比较多,输入特征维度比较小,将第一层输出维度设置为 32, LSTM 层输出维度设置为 16, 第三层输出维度设置为 16.由于该实验在前四组实验之后做,因此参考了前四组基于特征选择和 SVM 分类的实验中,特征选择个数的先验知识,一般 DFC 的特征选择个数在 30 上下波动,因此将 DFC 数据分类的第一层输出维度设置为 32

表格 3-1 模型参数设置表

输入数据	Dense1	LSTM	Dense2	Dense3
SROI signal	32	16	16	2
DFC	32	16	16	2

4 实验与分析

4.1 实验数据集描述

表格 4-1 数据集描述(预处理剔除被试之后)

数据集	HC	MDD	总人数	时间点数	时间间隔
实验室数据	48	47	95	373	0.8s
REST-meta-MDD	894	1104	1998	190	2s

基于机器学习分类的四组实验使用实验室提供的数据集,包括 HC 正常组 49 人, MDD 抑郁症患者组 47 人, 共 96 个原始静息态功能磁共振 (rfMRI) 数据样本, 所有抑郁症患者的诊断均符合 DSM-5 抑郁症诊断标准。静息态功能磁共振数据的采集重复时间 TR 为 0.8s, 切片个数 (slice number) 为 60, 时间点数共 383 个时间点。

基于深度学习分类的两组实验使用从 REST-meta-MDD[50]项目中选择的时点点数在 190 个以上的数据, 包括正常组 HC 894 人, 抑郁症组 MDD 1104 人, 共计 1998 个被试的已经进行了预处理和脑区分割的数据。其中静息态功能磁共振的采集重复时间 TR 为 2s, 时间点数共 190 个时间点。

4.2 实验结果与分析

4.2.1 基于传统机器学习实验结果

4.2.1.1 实验评价指标

表格 4-2 实验结果

数据	Atlas	FC	分类模型	验证集 AUC
实验室数据	AAL	SFC	SVM	0.6729 (十折交叉验证)
实验室数据	AAL	DFC	SVM	0.8521 (十折交叉验证)
实验室数据	Brainnetome	SFC	SVM	0.6937 (十折交叉验证)
实验室数据	Brainnetome	DFC	SVM	0.6083 (十折交叉验证)

其中每一个的结果都是选择的最佳参数对应的十折交叉验证的 AUC 指标值。

这四组实验的参数选择过程,以及最佳参数下特征个数与 AUC 之间的关系,均放置在了附录 1 里,敬请参阅。

保持脑图集不变,变化 FC 类型,可以看到在 AAL 脑图集下,DFC 的分类效果要优于 SFC 的分类效果;在 Brainnetome 脑图集下,DFC 的分类效果却不如 SFC 的分类效果;

保持 FC 类型不变,变化脑图集,可以看到对于 SFC,Brainnetome 脑图集的分类效果优于 AAL 脑图集;对于 DFC,AAL 脑图集的分类效果优于 Brainnetome 脑图集。

4.2.1.2 特征选择个数与对应的生理解释

在基于 AAL 脑图集的实验中,基于 SFC 的实验,最终选择出来的为 11 个特征,这些特征共涉及到 18 个区域节点,主要分布在默认网络,视觉网络等网络中,具体选择列表见附录 2。

将这些特征对应的区域和存在的连接可视化参见附录 3。

基于 DFC 的实验,最终具有最佳分类效果的特征个数为 46 个,对应 54 个

大脑区域，主要包括默认网络中的一些区域，如内侧前额叶皮层，前突神经等区域，同时包括中央执行网络的外侧前额叶皮质等区域，由于区域个数过多，放置在了附录 2。

最终经过与各个脑区的生理区域进行对照，可视化参见附录 3。

可以看出 DFC 选择的特征与 SFC 选择的特征具有比较高的重合性，分布比较相似。

另外的两组基于 Brainnetome 脑图集的实验结果如下。

基于 SFC 的分类选择出来的特征共计 50 个，分布在 67 个区域内，参见附录 3；基于 DFC 的分类选择出来的特征共计 91 个，分布在 104 个区域，如果考虑同一个位置的特征在不同时间点被多次选择，不重复的特征区域共有 75 个，参见附录 3；这些特征与 AAL 特征选择的结果的分布一致，只是因为分辨率更高因此选择了更多的特征数目。

相应的选择出来的特征附在了附录里，参见附录 2 和附录 3

4.2.1.3 实验结果分析

许多研究表明 DFC 比 SFC 的分类效果要好[34][35][36]，在本实验的 AAL 图集下，也的确验证了这个结论。事实上只是在理论上也可以大致推断，DFC 的分类效果应该优于 SFC，因为增加了时间信息。

但是根据表格，在 Brainnetome 脑图集上的实验却和这个结论不一致。一种比较合理的，也是在实验中很容易观察到的原因是，基于 Brainnetome 脑图集的 DFC 分类特征维数过高，造成了信息的冗余，也带来了噪声。Brainnetome 的区域个数是 AAL 的二倍，这使得基于 Brainnetome 脑图集的 FC 特征维数是 AAL 脑图集中 FC 特征维数的四倍左右，而在 DFC 中，又将多个时间点的 FC 特征拼接在一起，特征维度达到了几十万个，尽管在机器学习 pipeline 中存在特征选择的步骤，但是过高的特征维度带来的噪声很可能会影响到后续分类，也就导致基于 Brainnetome 脑图集的分类实验中 DFC 的效果不如 SFC。

根据表格，在保持 FC 类型不变时，对于 SFC 的分类，Brainnetome 脑图集上的分类效果略优于 AAL 脑图集。这可能是因为 Brainnetome 脑图集是一种更高分辨率的大脑区域划分，并且结合了结构像和功能像，比起生理划分的 AAL

能够提供一些更有意义的划分；对于 DFC 的分类，Brainnetome 脑图集由于分类特征维度过高带来的噪声，分类效果与 AAL 脑图集下 DFC 的分类效果差距很大。

根据可视化的结果可以看到两种脑图集的两种方法选择的特征区域具有很高的相似性，同时也可以看到 AAL 脑图集下选择的特征是包括小脑区域的，但是 Brainnetome 脑图集是不包括小脑的，因此这也可以在一定程度上说明，Brainnetome 一方面提供了更加功能结构一致的区域，对 SFC 的分类效果有所影响，另一方面，由于脑图集的区域限制，后续的 FC 特征中缺失了一些有分类特性的重要特征，这也有可能是导致其 DFC 分类效果大幅度下降的原因之一。

由此我们可以得知，在考虑利用 DFC 能够提供动态信息的优势的时候，也应该考虑脑区分割时的分辨率的影响；同时，从包含功能数据的多模态数据中提取出来的大脑图集在一定程度上有利于提升对应功能连接特征的分类特性。

4.2.2 基于 BiLSTM 深度学习实验结果:

表格 4-3 深度学习实验结果

输入	test sensitivity	test specificity	Test auc	test accuracy
ROI signal	0.6244	0.6145	0.6356	0.6200
DFC	0.5814	0.4476	0.5415	0.5206

表中所示评价值为测试集上的评价指标值。

值得一提的是,在上一组实验中只用到了 AUC,因为上一组中样本比较均衡,而且 auc 最终的值也比较高,因此不会出现 AUC 很高,但是敏感度和特异度一个很大一个很小的情况,而这一组中本身分类精度不太高,样本也不够均衡,可能会出现 auc 或者 accuracy 比较高但是敏感度和特异度中有一个比较低的情况,因此这里用了四个评价指标。

根据实验结果可以看出,在基于 LSTM 的分类模型下,ROI signal 的分类效果相较于 DFC 更好一些,比较直观的原因是基于 DFC 的分类输入特征维度太高,导致第一层的参数太多,而原始信号的输入特征维度比 DFC 输入特征维度小了 50 倍左右,因此第一层的参数量适中,这就导致 ROI signal 的分类更适合这个模型一些。因为并未尝试其他模型,不能说明任何一种情况下,这两者中的某一种始终具有更好的分类特性,因此还无法得出两者哪一个更适合做分类的生理上的结论,只是可以得出两者相较于这个 LSTM 模型框架更适合的一方是 ROI signal。

5 结论与展望

本研究通过在两个数据集上设计的两个对比实验,一方面基于机器学习分类验证了前人在研究中表述的 DFC 比 SFC 在抑郁症的诊断实验中表现更优的结论,设计了 SFC 和 DFC 的对比实验发现在一个脑图集下 DFC 的诊断特性确实优于 SFC,但是可能在利用 DFC 进行诊断之前需要考虑脑区分割的分辨率,避免使用分辨率较高的脑图集。另一方面基于深度学习分类比较了对于一个基于 LSTM 的深度学习框架来说 DFC 和原始信号 ROI signal 相比,后者更具有分类特性。

本项工作仍然有一些需要进一步研究的问题,未来即将进行的工作包括以下几点:

首先,本研究由于时间和数据的限制,未进行基于 SVM 的机器学习分类和基于深度学习的分类方法的研究,因为本实验这两种方法使用了两种数据集,在未来会在同一组数据上增加对比实验。

其次,在前文中也提到过目前的诊断主要是基于 DSM-5 标准,值得注意的是,在标准中规定的九条诊断参考条件中,只要包括必须包含的两种,同时符合的条件总数不小于五条就会被诊断为抑郁症,导致该诊断的症状可能有 227 种组合。这就带来了一个问题,在被诊断为抑郁症患者的人之间,存在症状和病理生理上的多样性[41],因此未来的一个研究方向就是进行 MDD 亚组的划分。

这与其他精神疾病如阿尔兹海默症的更精细的分类不同,还没有一个对抑郁症亚组的诊断标准,因此 MDD 亚组的划分任务是无监督的,是希望数据科学能够从各种可利用的数据中挖掘出有助于亚组划分的特征,并且最好能够附加生理解释。而其他一些精神疾病如阿尔兹海默症的诊断问题中的精细化分是有监督的,存在着诊断标准,可以将疾病诊断为不同的阶段或者不同的亚组。

最后,在脑区分割的步骤,本次实验直接选取了脑图集进行脑区分割,但是如果直接用算法,如聚类算法对个体大脑进行划分,即使都是健康组的个体,脑区分割的结果也会不同,直接使用模板划分固然方便并且能够使不同个体之间对应起来,但是个体的差异会导致划分的区域本身就具有异质性,之后做分类的时

候就算是同样的群体,也可能因为大脑存在个体差异性导致分割后得到的 fMRI 序列存在差异,这也是未来需要解决的问题。

一个可能的方法是先得到群体的脑区分割结果,再将群体脑区分割结果根据算法如区域增长在每个个体上进行微调,进而使区域的划分更加准确。

另一个可能有效的办法是,有研究表明,也可以借助深度学习的方法,如 RBM, DBN, AE 等方法,对大脑进行区域的划分,并且也证明这些划分的区域是具有可解释的生理意义的,同时深度学习的特征提取能力,以及深层网络的应用可以用来研究大脑的分层特性,也就是网络的不同层对应不同分辨率的脑区和花粉,这些分层划分结果的组合可能更有利于后续的分类,这也为端到端的而非过长的处理流程的分类方法提供了理论依据,也是未来的一个研究方向。

6 致谢

感谢指导老师彭勤牧老师和学长洪梓铭提供的指导、帮助和鼓励，感谢家人朋友一如既往的陪伴与关心，感谢学校提供的平台与环境，也感谢自己跌跌撞撞但越来越坚定的往前走，祝自己毕业快乐！

7 参考文献

- [1] Funk, M. "Global burden of mental disorders and the need for a comprehensive, coordinated response from health and social sectors at the country level." Retrieved on 30 (2016).
- [2] World Health Organization. Depression and other common mental disorders: global health estimates. No. WHO/MSD/MER/2017.2. World Health Organization, 2017.
- [3] Regier, Darrel A., et al. "The conceptual development of DSM-V." American Journal of Psychiatry 166.6 (2009): 645-650.
- [4] Hasin, Deborah S., et al. "Epidemiology of adult DSM-5 major depressive disorder and its specifiers in the United States." JAMA psychiatry 75.4 (2018): 336-346.
- [5] Ju, Ronghui, Chenhui Hu, and Quanzheng Li. "Early diagnosis of Alzheimer's disease based on resting-state brain networks and deep learning." IEEE/ACM transactions on computational biology and bioinformatics 16.1 (2017): 244-257.
- [6] LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." nature 521.7553 (2015): 436-444.
- [7] Schmidhuber, Jürgen. "Deep learning in neural networks: An overview." Neural networks 61 (2015): 85-117.
- [8] Mahato, Shalini, and Sanchita Paul. "Electroencephalogram (EEG) signal analysis for diagnosis of major depressive disorder (MDD): a review." Nanoelectronics, Circuits and Communication Systems (2019): 323-335.
- [9] Sundermann, Benedikt, Mona Olde lütke Beverborg, and Bettina Pfleiderer. "Toward literature-based feature selection for diagnostic classification: a meta-analysis of resting-state fMRI in depression." Frontiers in human neuroscience 8 (2014): 692.
- [10] Smith, Stephen M., et al. "Correspondence of the brain's functional architecture during activation and rest." Proceedings of the national academy of sciences 106.31 (2009): 13040-13045.

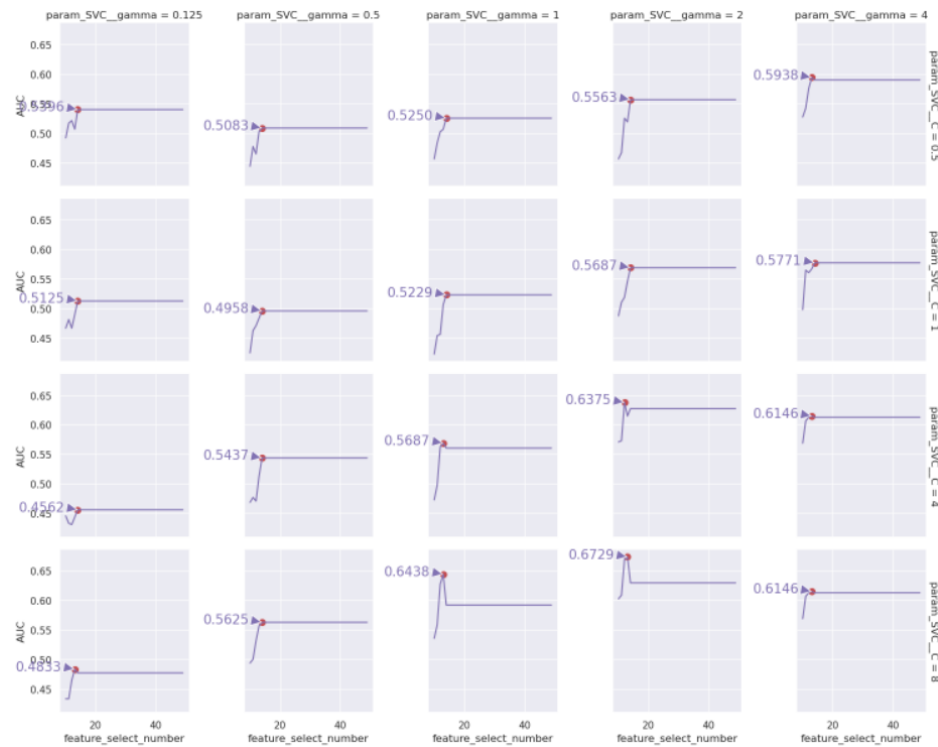
- [11] Price, Joseph L., and Wayne C. Drevets. "Neurocircuitry of mood disorders." *Neuropsychopharmacology* 35.1 (2010): 192-216.
- [12] Power, Jonathan D., Bradley L. Schlaggar, and Steven E. Petersen. "Studying brain organization via spontaneous fMRI signal." *Neuron* 84.4 (2014): 681-696.
- [13] Yan, Chaogan, and Yufeng Zang. "DPARSF: a MATLAB toolbox for" pipeline" data analysis of resting-state fMRI." *Frontiers in systems neuroscience* 4 (2010): 13.
- [14] Shen, Xilin, et al. "Groupwise whole-brain parcellation from resting-state fMRI data for network node identification." *Neuroimage* 82 (2013): 403-415.
- [15] Li, Wei, Xuefeng Lin, and Xi Chen. "Detecting Alzheimer's disease Based on 4D fMRI: An exploration under deep learning framework." *Neurocomputing* 388 (2020): 280-287.
- [16] Sarraf, Saman, Ghassem Tofghi, and Alzheimer's Disease Neuroimaging Initiative. "DeepAD: Alzheimer's disease classification via deep convolutional neural networks using MRI and fMRI." *BioRxiv* (2016): 070441.
- [17] Craddock, R. Cameron, et al. "A whole brain fMRI atlas generated via spatially constrained spectral clustering." *Human brain mapping* 33.8 (2012): 1914-1928.
- [18] Arslan, Salim, et al. "Human brain mapping: A systematic comparison of parcellation methods for the human cerebral cortex." *NeuroImage* 170 (2018): 5-30.
- [19] Wang, Lebo, et al. "Application of convolutional recurrent neural network for individual recognition based on resting state fMRI data." *Frontiers in neuroscience* 13 (2019): 434.
- [20] Kawahara, Jeremy, et al. "BrainNetCNN: Convolutional neural networks for brain networks; towards predicting neurodevelopment." *NeuroImage* 146 (2017): 1038-1049.
- [21] Chun, Ji Ye, et al. "Visualizing Functional Network Connectivity Difference between Healthy Control and Major Depressive Disorder Using an Explainable Machine-learning Method." 2020 42nd Annual International Conference of the

- IEEE Engineering in Medicine & Biology Society (EMBC). IEEE, 2020.
- [22] Yan, Baoyu, et al. "Quantitative identification of major depression based on resting-state dynamic functional connectivity: a machine learning approach." *Frontiers in neuroscience* 14 (2020): 191.
- [23] Pervaiz, Usama, et al. "Optimising network modelling methods for fMRI." *Neuroimage* 211 (2020): 116604.
- [24] Abraham, Alexandre, et al. "Deriving reproducible biomarkers from multi-site resting-state data: An Autism-based example." *NeuroImage* 147 (2017): 736-745.
- [25] Parisot, Sarah, et al. "Disease prediction using graph convolutional networks: application to autism spectrum disorder and Alzheimer's disease." *Medical image analysis* 48 (2018): 117-130.
- [26] Parisot, Sarah, et al. "Spectral graph convolutions for population-based disease prediction." *International conference on medical image computing and computer-assisted intervention*. Springer, Cham, 2017.
- [27] Wang, Lebo, Kaiming Li, and Xiaoping P. Hu. "Graph convolutional network for fMRI analysis based on connectivity neighborhood." *Network Neuroscience* 5.1 (2021): 83-95.
- [28] Hojjati, Seyed Hani, et al. "Predicting conversion from MCI to AD using resting-state fMRI, graph theoretical approach and SVM." *Journal of neuroscience methods* 282 (2017): 69-80.
- [29] Yao, Dongren, et al. "Temporal-Adaptive Graph Convolutional Network for Automated Identification of Major Depressive Disorder Using Resting-State fMRI." *International Workshop on Machine Learning in Medical Imaging*. Springer, Cham, 2020.
- [30] Sarraf, Saman, Ghassem Tofighi, and Alzheimer's Disease Neuroimaging Initiative. "DeepAD: Alzheimer's disease classification via deep convolutional neural networks using MRI and fMRI." *BioRxiv* (2016): 070441.
- [31] Ktena, Sofia Ira, et al. "Metric learning with spectral graph convolutions on brain connectivity networks." *NeuroImage* 169 (2018): 431-442.

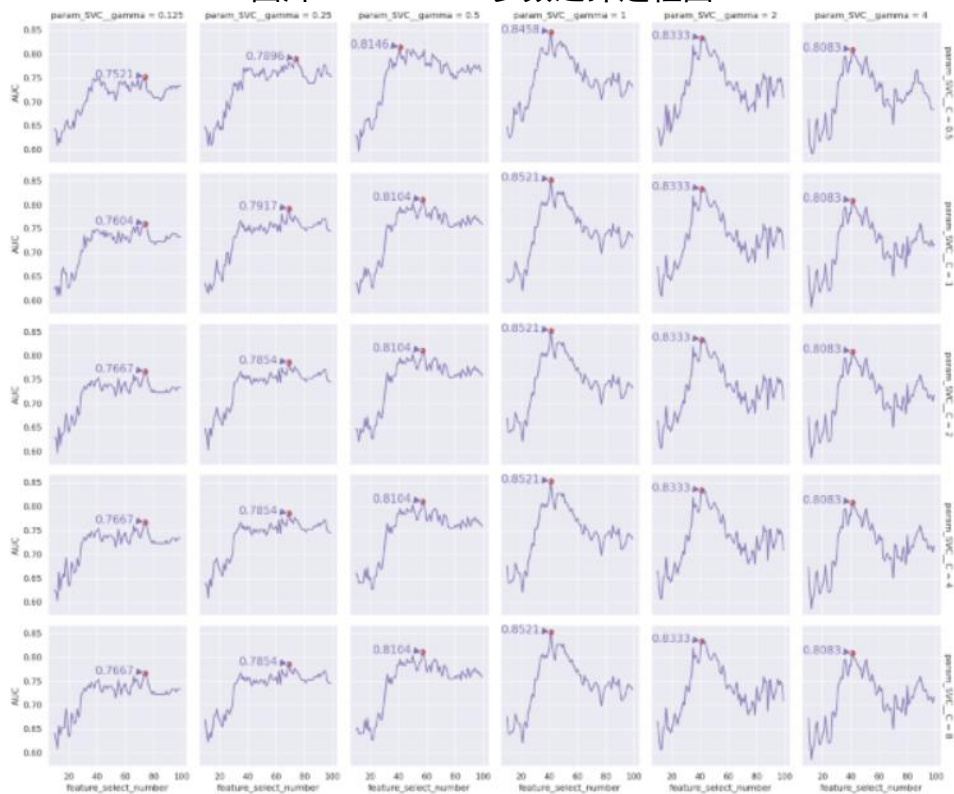
- [32] El-Gazzar, Ahmed, et al. "A hybrid 3DCNN and 3DC-LSTM based model for 4D spatio-temporal fMRI data: an ABIDE autism classification study." OR 2.0 Context-Aware Operating Theaters and Machine Learning in Clinical Neuroimaging. Springer, Cham, 2019. 95-102.
- [33] Riaz, Atif, et al. "Fcnet: a convolutional neural network for calculating functional connectivity from functional mri." International Workshop on Connectomics in Neuroimaging. Springer, Cham, 2017.
- [34] Rashid, Barnaly, et al. "Classification of schizophrenia and bipolar patients using static and dynamic resting-state fMRI brain connectivity." *Neuroimage* 134 (2016): 645-657.
- [35] Demirtaş, Murat, et al. "Dynamic functional connectivity reveals altered variability in functional connectivity among patients with major depressive disorder." *Human brain mapping* 37.8 (2016): 2918-2930.
- [36] Rashid, Barnaly, and Vince Calhoun. "Towards a brain-based predictome of mental illness." *Human brain mapping* 41.12 (2020): 3468-3535.
- [37] Gers, Felix A., Jürgen Schmidhuber, and Fred Cummins. "Learning to forget: Continual prediction with LSTM." (1999): 850-855.
- [38] Chandra, Mayank Arya, and S. S. Bedi. "Survey on SVM and their application in image classification." *International Journal of Information Technology* (2018): 1-11.
- [39] Du, Yuhui, Zening Fu, and Vince D. Calhoun. "Classification and prediction of brain disorders using functional connectivity: promising but challenging." *Frontiers in neuroscience* 12 (2018): 525.
- [40] Yaesoubi, Maziar, et al. "Dynamic coherence analysis of resting fMRI data to jointly capture state-based phase, frequency, and time-domain information." *Neuroimage* 120 (2015): 133-142.
- [41] Miller, Greg. "Beyond DSM: seeking a brain-based classification of mental illness." (2010): 1437-1437.
- [42] Insel, Thomas R., and Philip S. Wang. "Rethinking mental illness." *Jama* 303.19

- (2010): 1970-1971.
- [43] Strother, Stephen C. "Evaluating fMRI preprocessing pipelines." *IEEE Engineering in Medicine and Biology Magazine* 25.2 (2006): 27-41.
 - [44] Sladky, Ronald, et al. "Slice-timing effects and their correction in functional MRI." *Neuroimage* 58.2 (2011): 588-594.
 - [45] Smith, Stephen M., et al. "Advances in functional and structural MR image analysis and implementation as FSL." *Neuroimage* 23 (2004): S208-S219.
 - [46] Cox, Robert W. "AFNI: software for analysis and visualization of functional magnetic resonance neuroimages." *Computers and Biomedical research* 29.3 (1996): 162-173.
 - [47] Tzourio-Mazoyer, Nathalie, et al. "Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain." *Neuroimage* 15.1 (2002): 273-289.
 - [48] Esteban, Oscar, et al. "fMRIPrep: a robust preprocessing pipeline for functional MRI." *Nature methods* 16.1 (2019): 111-116.
 - [49] Yan, Chao-Gan, et al. "DPABI: data processing & analysis for (resting-state) brain imaging." *Neuroinformatics* 14.3 (2016): 339-351.
 - [50] Yan, Chao-Gan, et al. "Reduced default mode network functional connectivity in patients with recurrent major depressive disorder." *Proceedings of the National Academy of Sciences* 116.18 (2019): 9078-9083.
 - [51] Wang, Jinhui, et al. "GRETNA: a graph theoretical network analysis toolbox for imaging connectomics." *Frontiers in human neuroscience* 9 (2015): 386..
 - [52] Tzourio-Mazoyer, Nathalie, et al. "Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain." *Neuroimage* 15.1 (2002): 273-289.
 - [53] Fan, Lingzhong, et al. "The human brainnetome atlas: a new brain atlas based on connectional architecture." *Cerebral cortex* 26.8 (2016): 3508-3526.
 - [54] Xia, Mingrui, Jinhui Wang, and Yong He. "BrainNet Viewer: a network visualization tool for human brain connectomics." *PloS one* 8.7 (2013): e68910.

附录 1



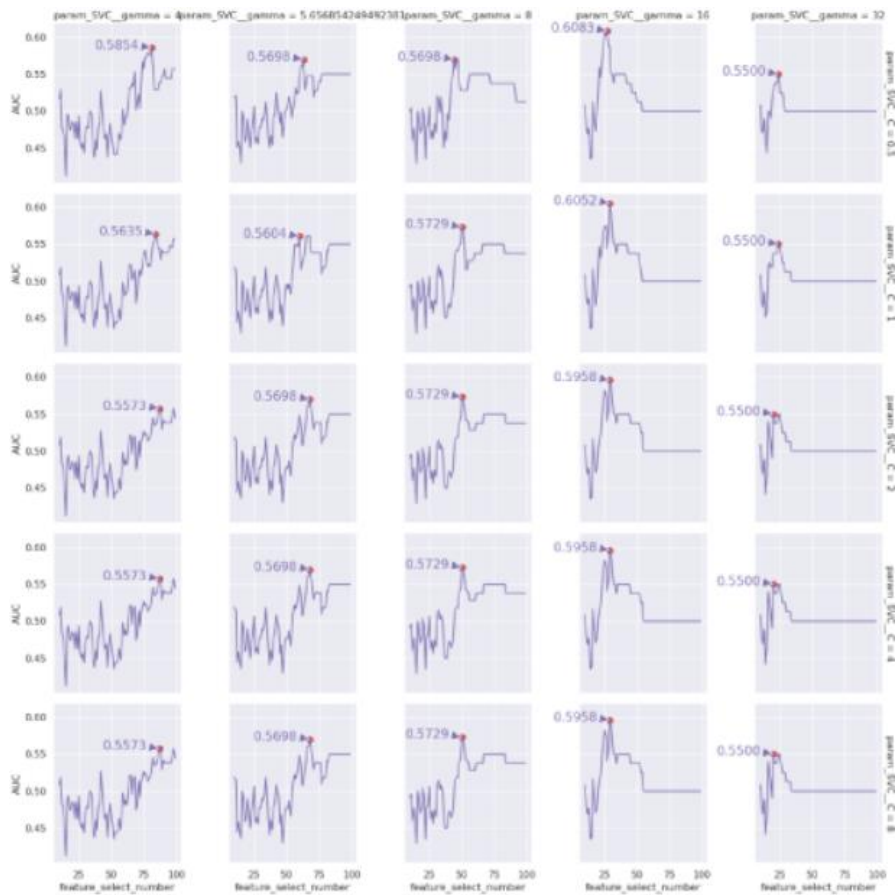
图片 1 AAL SFC 参数选择过程图



图片 2 AAL DFC 参数选择过程



图片 3 Brainnetome SFC 参数选择过程



图片 4 Brainnetome DFC 参数选择过程

附录 2

表格 1 AAL SFC 特征选择结果

Number	对应的区域连接 (ROI1_ROI2)
1	Frontal_Inf_Oper_R2302_Cingulum_Post_R 4022
2	Frontal_Inf_Tri_R 2312_Cingulum_Post_R 4022
3	Olfactory_L 2501_Vermis_1_2 9100
4	Frontal_Med_Orb_L 2611_Cerebelum_10_L 9081
5	Cingulum_Post_R 4022_Parietal_Inf_R 6202
6	Putamen_R 7012_Vermis_3 9110
7	Pallidum_R 7022_Cerebelum_Crus2_L 9011
8	Pallidum_R 7022_Cerebelum_Crus2_R 9012
9	Pallidum_R 7022_Cerebelum_7b_R 9052
10	Pallidum_R 7022_Vermis_7 9140
11	Pallidum_R 7022_Vermis_8 9150

表格 2 AALDFC 特征选择结果

Number	对应的区域连接 (ROI1_ROI2)
1	Frontal_Sup_L 2101_Calcarine_L 5001
2	Frontal_Sup_L 2101_Lingual_L 5021
3	Frontal_Sup_L 2101_Lingual_R 5022
4	Frontal_Sup_L 2101_Occipital_Sup_L 5101
5	Frontal_Sup_L 2101_Occipital_Sup_R 5102
6	Frontal_Sup_L 2101_Temporal_Pole_Mid_R 8212
7	Frontal_Sup_R 2102_Temporal_Pole_Mid_R 8212
8	Frontal_Inf_Oper_L 2301_Frontal_Inf_Tri_R 2312
9	Frontal_Inf_Orb_R 2322_Vermis_1_2 9100
10	Rolandic_Oper_L 2331_Vermis_3 9110
11	Rolandic_Oper_L 2331_Vermis_9 9160
12	Rolandic_Oper_R 2332_Vermis_3 9110
13	Supp_Motor_Area_L 2401_Vermis_6 9130
14	Frontal_Sup_Medial_R 2602_Temporal_Pole_Mid_R 8212
15	Frontal_Med_Orb_L 2611_Calcarine_R 5002
16	Frontal_Med_Orb_R 2612_Calcarine_R 5002
17	Insula_R 3002_Vermis_3 9110
18	Cingulum_Post_L 4021_Cuneus_R 5012
19	Amygdala_L 4201_Cerebellum_8_L 9061
20	Amygdala_R 4202_Cerebellum_7b_R 9052
21	Calcarine_R 5002_Cerebellum_10_R 9082
22	Lingual_L 5021_Vermis_7 9140
23	Lingual_R 5022_Vermis_7 9140
24	Occipital_Mid_L 5201_Vermis_7 9140
25	SupraMarginal_R 6212_Vermis_3 9110
26	Angular_R 6222_Cerebellum_4_5_R 9032

27	Caudate_L 7001_ Heschl_R 8102
28	Caudate_R 7002_ Cerebelum_8_L 9061
29	Putamen_L 7011_ Vermis_3 9110
30	Putamen_R 7012_ Cerebelum_Crus2_R 9012
31	Putamen_R 7012_ Cerebelum_7b_R 9052
32	Putamen_R 7012_ Cerebelum_9_L 9071
33	Putamen_R 7012_ Vermis_7 9140
34	Pallidum_L 7021_ Vermis_3 9110
35	Pallidum_R 7022_ Cerebelum_7b_R 9052
36	Heschl_L 8101_ Vermis_3 9110
37	Temporal_Sup_L 8111_ Vermis_3 9110
38	Temporal_Sup_R 8112_ Vermis_3 9110
39	Temporal_Pole_Sup_L 8121_ Vermis_9 9160
40	Temporal_Pole_Sup_L 8121_ Vermis_10 9170
41	Temporal_Pole_Mid_L 8211_ Cerebelum_3_L 9021
42	Cerebelum_Crus2_R 9012_ Cerebelum_4_5_R 9032
43	Cerebelum_6_L 9041_ Vermis_7 9140
44	Cerebelum_6_R 9042_ Vermis_7 9140
45	Vermis_4_5 9120_ Vermis_7 9140
46	Vermis_6 9130_ Vermis_7 9140

表格 3 brainnetome SFC 特征选择结果

Number	对应的区域连接 (ROI1_ROI2)
1	SFG_R_7_2_PCun_L_4_4
2	SFG_R_7_2_PCun_R_4_4
3	SFG_L_7_3_PhG_R_6_1
4	SFG_R_7_7_ITG_R_7_3
5	MFG_R_7_1_IPL_R_6_2
6	MFG_R_7_3_ITG_R_7_3
7	MFG_L_7_6_INS_L_6_5
8	MFG_L_7_7_SPL_L_5_3
9	IFG_L_6_1_ITG_L_7_3
10	IFG_L_6_3_MVOcC_L_5_1
11	IFG_L_6_6_Tha_R_8_8
12	OrG_L_6_4_PrG_L_6_4
13	OrG_L_6_4_SPL_R_5_3
14	OrG_L_6_4_PoG_L_4_1
15	OrG_L_6_4_PoG_L_4_3
16	OrG_L_6_4_PoG_L_4_4
17	OrG_L_6_5_SPL_R_5_1
18	OrG_L_6_5_SPL_R_5_5
19	OrG_L_6_5_IPL_R_6_3
20	OrG_L_6_6_PhG_R_6_4
21	PrG_R_6_1_PhG_R_6_6
22	PCL_L_2_1_INS_L_6_5
23	PCL_R_2_1_PhG_L_6_3
24	PCL_R_2_1_INS_L_6_5
25	PCL_R_2_2_PhG_R_6_6
26	PCL_R_2_2_PCun_L_4_3
27	STG_L_6_1_ITG_R_7_5

28	MTG_R_4_3_IPL_L_6_3
29	MTG_R_4_3_INS_L_6_1
30	PhG_L_6_1_IPL_L_6_6
31	PhG_L_6_3_PoG_L_4_4
32	PhG_R_6_6_BG_L_6_5
33	SPL_L_5_3_IPL_R_6_5
34	SPL_R_5_4_INS_L_6_5
35	IPL_L_6_6_INS_L_6_5
36	PCun_R_4_2_INS_L_6_5
37	PCun_R_4_3_LOcC_R_4_1
38	INS_L_6_5_CG_R_7_5
39	MVOcC_L_5_2_Tha_R_8_1
40	MVOcC_L_5_2_Tha_R_8_4
41	MVOcC_R_5_2_Tha_R_8_1
42	MVOcC_R_5_2_Tha_R_8_4
43	MVOcC_L_5_4_BG_R_6_4
44	MVOcC_L_5_5_Tha_R_8_4
45	LOcC_R_2_2_BG_R_6_6
46	BG_R_6_1_Tha_R_8_8
47	BG_L_6_2_BG_R_6_6
48	BG_R_6_2_BG_L_6_4
49	BG_L_6_4_Tha_R_8_2
50	BG_R_6_4_Tha_R_8_8

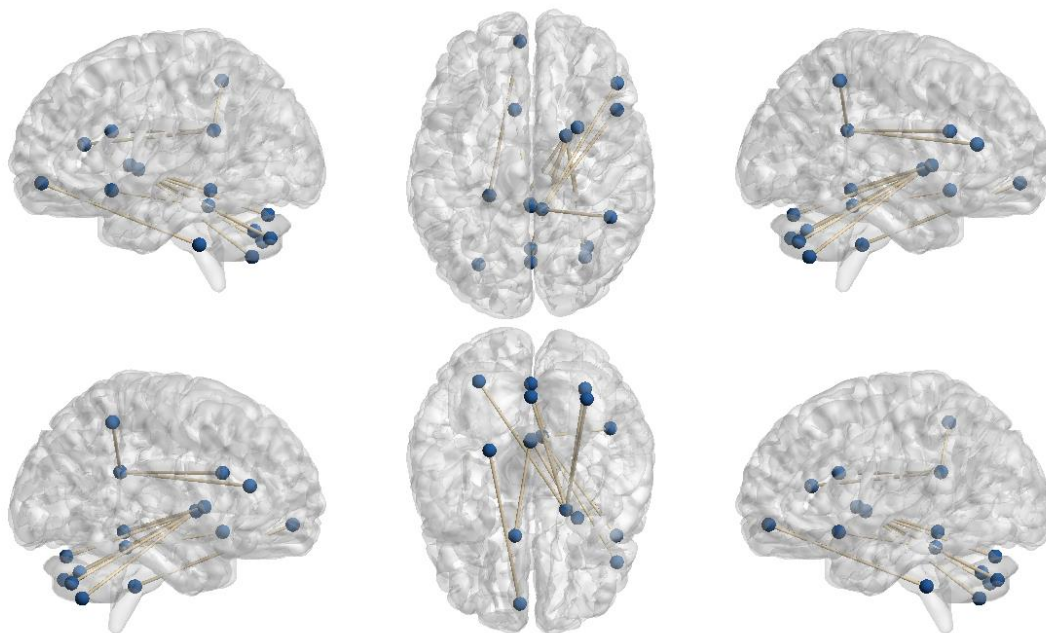
表格 4 Brainnetome DFC 特征选择结果

Number	对应的区域连接 (ROI1_ROI2)
1	SFG_R_7_2_SFG_L_7_7
2	SFG_R_7_2_MTG_L_4_2
3	SFG_R_7_5_MFG_L_7_6
4	SFG_R_7_5_IPL_L_6_6
5	MFG_L_7_2_MFG_L_7_6
6	MFG_L_7_3_ITG_R_7_3
7	MFG_L_7_3_Tha_L_8_2
8	MFG_L_7_3_Tha_L_8_3
9	MFG_R_7_3_ITG_R_7_3
10	MFG_R_7_4_PhG_L_6_5
11	MFG_R_7_5_PhG_L_6_4
12	MFG_L_7_6_OrG_L_6_5
13	MFG_L_7_6_PrG_L_6_2
14	MFG_R_7_6_IPL_L_6_2
15	MFG_L_7_7_ITG_R_7_1
16	MFG_L_7_7_SPL_L_5_3
17	MFG_R_7_7_ITG_R_7_1
18	IFG_L_6_1_IFG_R_6_2
19	IFG_L_6_1_OrG_L_6_5
20	IFG_R_6_2_PhG_L_6_5
21	IFG_R_6_3_PrG_L_6_6
22	IFG_L_6_5_MVOcC_L_5_1
23	IFG_L_6_5_MVOcC_R_5_1
24	IFG_L_6_6_Tha_R_8_2
25	IFG_L_6_6_Tha_L_8_8
26	IFG_R_6_6_PrG_L_6_6
27	OrG_L_6_2_PhG_R_6_1

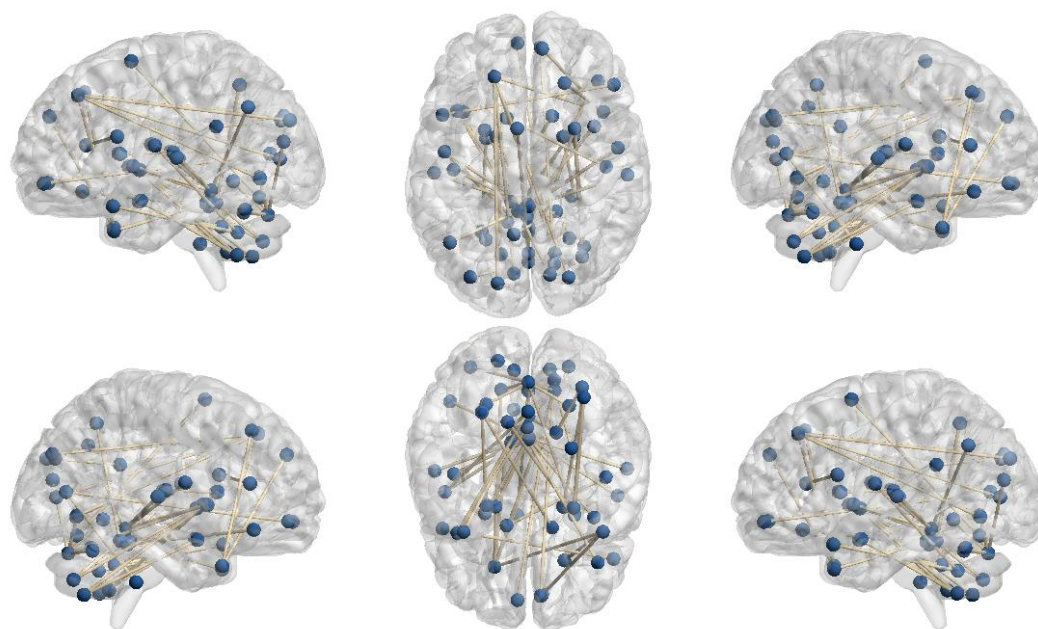
28	OrG_R_6_4_PrG_R_6_4
29	OrG_R_6_4_ITG_R_7_4
30	OrG_L_6_5_MTG_L_4_4
31	OrG_L_6_5_SPL_L_5_3
32	OrG_L_6_6_PhG_R_6_1
33	PrG_L_6_1_IPL_R_6_6
34	PrG_L_6_4_PhG_L_6_5
35	PCL_L_2_1_INS_L_6_1
36	STG_L_6_1_IPL_R_6_5
37	STG_L_6_3_PhG_L_6_1
38	STG_L_6_4_FuG_L_3_1
39	STG_L_6_5_SPL_L_5_3
40	STG_L_6_5_SPL_R_5_5
41	STG_L_6_6_INS_L_6_1
42	MTG_L_4_1_MTG_L_4_2
43	MTG_L_4_4_IPL_L_6_4
44	ITG_R_7_1_PCun_R_4_4
45	ITG_L_7_3_BG_L_6_2
46	ITG_L_7_4_BG_L_6_4
47	ITG_L_7_6_SPL_L_5_3
48	FuG_R_3_2_Amyg_L_2_1
49	PhG_L_6_1_INS_L_6_1
50	PhG_L_6_2_IPL_R_6_3
51	PhG_L_6_2_IPL_R_6_4
52	PhG_L_6_2_CG_R_7_6
53	PhG_R_6_5_Tha_L_8_3
54	PhG_L_6_6_INS_L_6_4
55	SPL_L_5_1_Tha_R_8_1

56	SPL_L_5_1_Tha_R_8_8
57	SPL_R_5_4_BG_R_6_2
58	SPL_R_5_5_CG_L_7_1
59	SPL_R_5_5_CG_R_7_1
60	IPL_R_6_4_BG_R_6_4
61	IPL_L_6_6_CG_L_7_5
62	PCun_L_4_4_PCun_R_4_4
63	PoG_L_4_1_CG_R_7_4
64	PoG_L_4_2_LOcC_R_4_2
65	INS_R_6_2_CG_R_7_3
66	INS_R_6_4_BG_R_6_1
67	INS_L_6_6_Tha_L_8_7
68	CG_L_7_1_LOcC_R_2_2
69	CG_R_7_7_Tha_L_8_2
70	MVOcC_R_5_5_Hipp_R_2_2
71	LOcC_R_4_1_Hipp_R_2_1
72	Amyg_L_2_2_Tha_R_8_5
73	Hipp_L_2_1_Tha_R_8_5
74	Hipp_R_2_1_Tha_L_8_5
75	Hipp_R_2_1_Tha_L_8_6

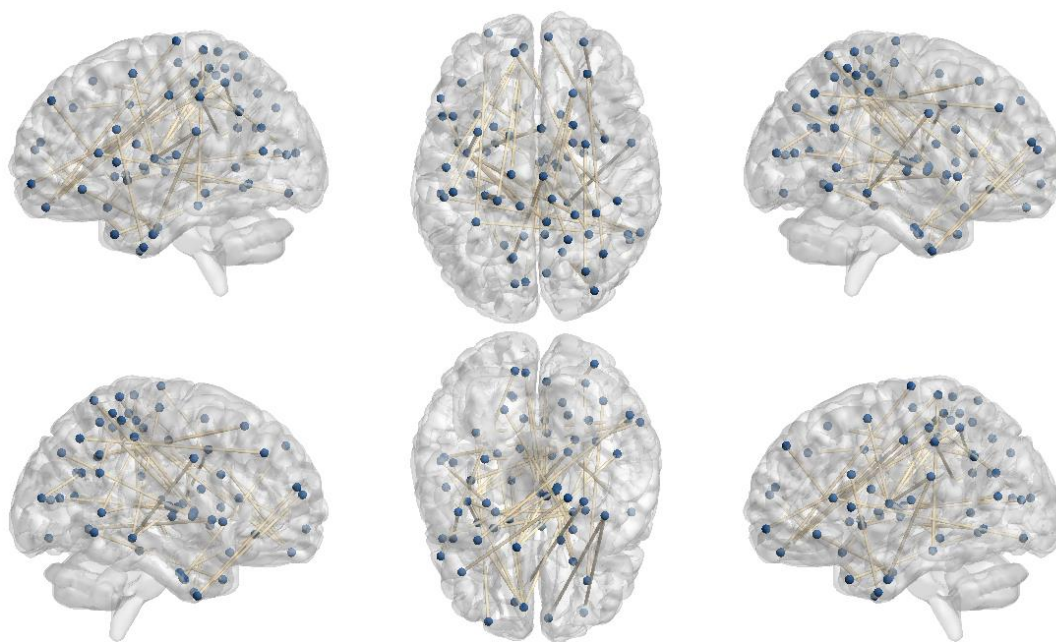
附录 3



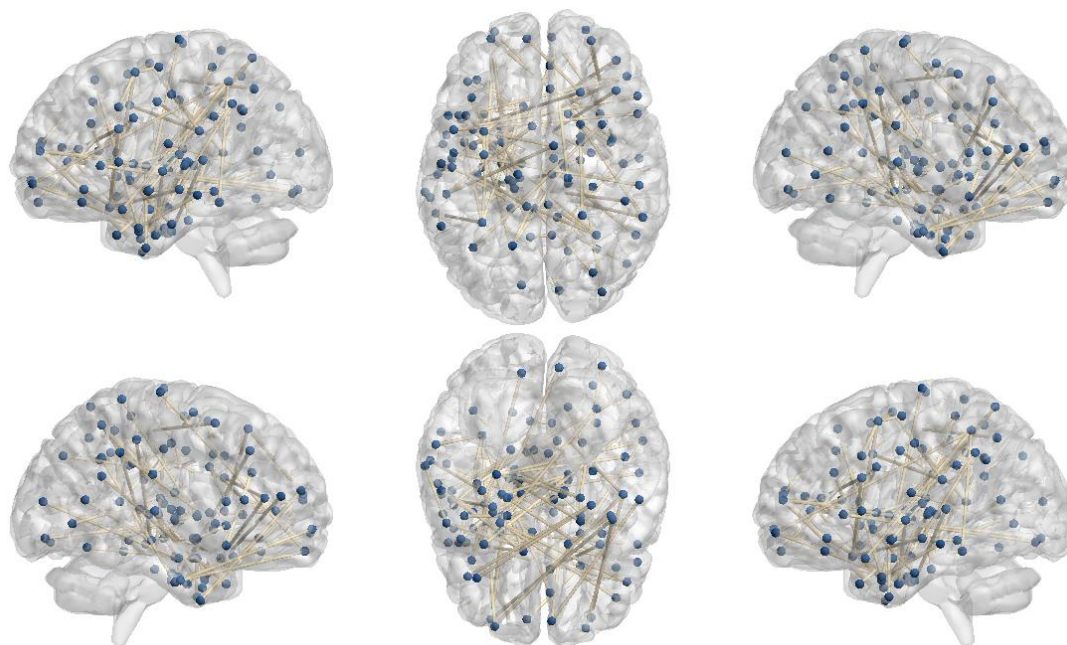
图片 1 AAL+SFC 特征选择可视化



图片 2 AAL+DFC 特征选择可视化



图片 3 Brainnetome+SFC 特征选择可视化



图片 4 Brainnetome+DFC 特征选择可视化