

2. 项目的研究内容、研究目标，以及拟解决的关键科学问题（此部分为重点阐述内容）；

2.1 研究内容

本项目主要研究包括：基于脑功能影像的亚型划分评估方法评估；基于多中心脑影像、多模板 AD 的亚型划分稳定性、可重复性和临床意义；早期 AD 及高危人群 MCI 不同亚型的纵向变化表征刻画。

2.1.1 基于脑功能影像的亚型划分方法评估框架

已有的研究表明 AD 是一个复杂的疾病，包含多种不同的亚型，目前各种亚型算法较多，但是得到的结论依赖于数据、算法和特征。预实验的结果证明我们可以根据异常网络的改变将 AD 划分成 4 个不同的亚型(具体参见预实验的亚型划分研究)。

在本项目中，团队将继续研究不同亚型的算法，寻找最优的亚型计算方法。本项目考虑构建不同的仿真数据、接近真实数据的仿真数据和接近 AD 情况的仿真数据(具体请参见 3.1.1 研究方法和实验手段)中方法，基于此，研究团队备选的亚型分析方法的稳定性，敏感性和特异性。备选的算法包括：层次聚类，隐狄利克雷分布算法，基于图数据的社区发现算法(Louvain)，贝叶斯决策法，无监督随机森林算法，非负矩阵分解算法(NMF)，K-Medians，多专家 mixture-of-experts, Distribution Matching of Imaging Patterns, Heterogeneity of imaging and genetic patterns through a multiple max-margin Discriminative Analysis (HYDRA)，Subtype and Stage Inference (SuStain)等方法。

2.1.2 利用不同尺度的脑图谱、基于多中心磁共振影像准确刻画 AD 脑影像异常表征，研究亚型划分的稳定性和临床意义

团队初步计划利用不同粒度的图谱，以脑网络组图谱为基础，关注不同分辨率的脑图谱，刻画基于不同脑图谱的 AD 亚型。

项目组针对单个中心样本量小，容易受到样本方差大的影响，本项目重点关注第一个研究内容中确定的最优的前 3 种算法，利用不同尺度的脑图谱，精确刻画 AD 的亚型，研究高危人群 MCI 的亚型稳定性，研究纵向变化模式下的亚型异常变化模式。

重点关注融合结构网络和功能网络表征探寻 AD 异常网络变化模式的算法：结构网络和功能网络反应的信息有较大的差异，拟采用多元回归分析、多元独立成分分析和典型相关分析的方法，将功能像特征和弥散张量图像特征，异常网络特征互相融合，同时选取特定网络如默认网络的结构和功能为参考展开研究，通过加入先验信息来约束，有可能更加精确和稳定的提取出 AD 的特异性网络特征。从而实现从单模态、单尺度过度到多模态、多尺度精准刻画。

2.1.3 早期 AD 及高危人群 MCI 不同亚型的纵向变化表征刻画，提出基于亚型的早期精准预测模型

探索如何将脑影像的特征表示和 AD 已知的特征融合得到稳定的具有普适性的特征；如何通过这些特征构建合理的学习模型，融合脑区、脑网络以及行为等不同尺度的特征发现不同亚型 AD 的脑结构和脑功能表征，建立 AD“生物亚型”和早期高危人群的预测模型。

探索采用稀疏学习算法来充分利用多中心大样本的不完整多模态信息；提出适用于高维磁共振影像的深度学习模型，提高数据利用效率并提高最终分类统计功效和准确率；构造具备新类别自动发现能力、自适应迁移学习能力、以及半监督学习能力的模式识别模型和算法用于小样本学习模型建立。利用目前的对抗生成网络(GAN)模型生成部分样本，使得学习样本更加规范，加速机器学习中的参数优化学习。

项目组计划使用 ADNI 的纵向变化数据，使用自有的纵向变化数据，研究基线水平单个个体的亚型，将个体纵向变化的数据进行分析，研究不同亚型的个体纵向变化的脑结构、脑功能活动和脑网络变化的特征，以及这些变化模式与脑认知能力改变之间的关系。然后利用迁移学习、集成学习等深度学习方法，将已有的多中心数据库上训练人工智能方法，以识别有助于检测个体的疾病的模式，从而训练更加高效的模型，利用已有和新采集的纵向数据评估基于亚型的早期预测的准确性、敏感性和特异性。

2.2 研究目标

项目的总体目标是建立利用多模态脑影像对早期 AD 及高危人群的亚型个体化刻画，在此基础上为临床提供个体化干预方案的重要影像学基础。具体目标如下：

- 1) 系统评估已有算法的优缺点，提出 1-2 种新的亚型研究方法。
- 2) 明确功能亚型的稳定性、可重复性、泛化性和生物学意义，融合多模态脑影像数据精确刻画 AD 的个体化异常表征。
- 3) 明确早期 AD 及 MCI 的基于亚型的个体化变化表征，构建基于“生物亚型”的早期识别个体化表征。

2.3 关键科学问题

2.3.1 如何建立融合影像融合的方法研究不同亚型的特异性表征

AD 是一种复杂的神经退行疾病，已有的文献和项目团队的预实验提示 AD 可以根据不同的影像模态和计算方法细分成不同的亚型。另外，AD 也是一个影像多个脑区的活动和交互模式的复杂疾病，因此，本项目的科学问题之一就是如何确立早期 AD 的不同亚型和这些细分亚型所影响的脑网络，进而根据网络亚型的特异性，如何进行多中心综合评估管理也是要解决的问题之一。

2.3.2 如何建立适合脑影像数据多中心、小样本，高维异构特征的特点的学习算法，实现早期的 AD 个体化表征刻画

针对脑影像小样本、高维度、低信噪比等特点，该项目将采用先进的协调方法，整合各个中心 AD 诊断的认知、影像和生物标志物数，提出针对性的分析方法如何建立适合脑影像数据多中心、小样本高维异构特征的特点的学习算法，充分利用多中心大样本的不完整多模态信息；提高数据利用效率并提高最终分类统计功效和准确率，从而从网络上精确刻画 AD 的脑异常模式，做到特征具有生物学基础，临床可解释，实现早期个体化的可视化影像学表征。

2.3.3 如何将多图谱高维特点融合纳入统一的框架刻画 AD 脑个体化表征

目前的异常连接模式受到数据站点、采集参数和不同的图谱划分方式的影响，需要建立一套方法，将基于不同脑图谱的异常动态变化、静态连接、结构连接等融合，刻画 AD 的异常连接模式，研究其稳定性和泛化性表征等。

3. 拟采取的研究方案及可行性分析（包括研究方法、技术路线、实验手段、关键技术等说明）；

3.1 研究方案

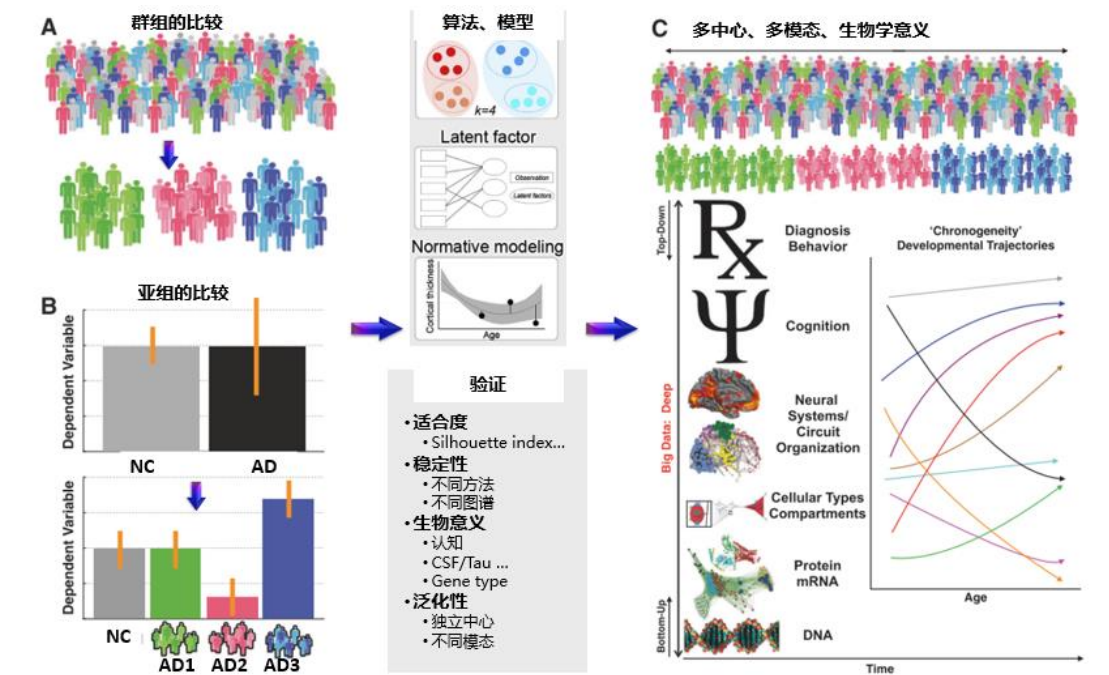


图 3. 本项目的主要研究框架

3.1.1 研究方法和实验手段

1) 数据库更新和数据采集

仿真数据：本项目计划采用 3-4 种不同类型的仿真数据进行亚型划分。第一步采取已有文献中的方法(*IEEE Trans Med Imaging* **35**(2): 612-621.)中的方法。第二个是对于 HCP 的数据，我们可以对其单个个体 4 次采集的功能影像数(单词时间序列长度 1200 个时间点)，从中每次采集连续的 170 个时间(团队已有数据的时间长度)，那么单个个体可以仿真出来 200 个以上的随机的数据，考虑到每个个体的数据不同时间点的差异<<不同个体的之间的差异，这样可以评估不同的算法的稳定性、敏感性和特异性。第三种仿真数据是可以对于我们已有的数据单个个体按照相关的采样办法采集 120-150 个时间长度的数据，这样可以验证不同的时间长度下，不同方法的稳定性、特异性等。

真实数据：本项目到目前为止已经完成整理 12 个中心的多模态影像数据(具体请参见研究基础部分)。因此，项目继续在中国人民解放军总医院收集已有数据的纵向跟踪数据，既可以作为新的数据集对比使用，又可以作为中国人民解放军总医院已有数据的纵向跟踪研究。本项目计划收集早期 AD，aMCI 和正常对照数据 200 例。使团队可用数据可以达到 15 个中心 (即已有 12 个，新采集 1 个，继续整理新加入多中心研究的数据)，近 3000 例样本。为保证实验的连续性和对原有数据的良好兼容性，被试主要完成临床和行为学测评、静息状态 fMRI、高分辨率结构 MRI 和弥散张量成像。数据采集主要计划和参数遵循已有的研究入组，本研究所有受试者自愿参加此项实验，并签署知情同意书，自愿参加并认真完成实验任务，在完成实验后给予一定报酬。概述如下：

2) 研究对象的入组标准和排除标准

所有被试均需满足：①右利手；母语为汉语；一般状况良好，无其他会影响到研究的疾病；无视听觉障碍；有合格的知情者，并且每次随访知情者都可以陪同；②汉密尔顿焦虑量表< 29 分；汉密尔顿抑郁量表< 24 分；③Hachinski 缺血量表小于等于 4 分；④不影响研究的药物服用已大于等于 4 周；⑤同意做 3.0 T MRI，无 MRI 检查禁忌症；⑥同意长期随访；⑦同意进行 APOE 基因检测；同意进行血和尿样本的采集。

正常被试-1：纳入标准：①年龄 60-80 岁；②智力粗测正常、可进行神经心理测试的操作；③两系三代内无符合 DSM-IV 诊断标准的精神障碍者；④无药物或酒精滥用史、无重大躯体疾病、颅脑外伤、神经与精神疾病史；⑤无磁共振扫描禁忌，经影像科医生判断无脑结构异常者；⑥在知情同意的情况下同意参加影像学研究。⑦根据美国国立神经病学、语言障碍和卒中研究所、AD 及相关疾病协会(NINCDS-ADRDA)及诊断与统计手册第 4 次修订版(DSM-WR)中的 AD 诊断标准进行诊断，认知功能评价正常，FAQ< 6；并同意采集静脉血。

正常被试-2(超高龄)：纳入标准：①年龄 80-95 岁；②智力粗测正常、可进行神经心理测试的操作；③两系三代内无符合 DSM-IV 诊断标准的精神障碍者；④无药物或酒

精滥用史、无重大躯体疾病、颅脑外伤、神经与精神疾病史；⑤无磁共振扫描禁忌，经影像科医生判断无脑结构异常者；⑥在知情同意的情况下同意参加影像学研究。⑦根据美国国立神经病学、语言障碍和卒中研究所、AD 及相关疾病协会(NINCDS-ADRDA)及诊断与统计手册第 4 次修订版(DSM-WR)中的 AD 诊断标准进行诊断，认知功能评价正常，FAQ< 6；并同意采集静脉血。

早期 AD 临床前：依据 Frank Jessen 2014 年所发表的 SCD 诊断框架、Jak/Bondi 等 2014 和 2015 年入组标准进行制定：①年龄 60-79 岁；②认知下降的主诉，或由知情者报告的认知下降；③总体认知检查正常：MoCA-B：小学> 19 分，中学>22 分，大学> 24 分；MMSE：小学> 20 分，初中及以上>24 分；CDR：0 分；④有或者没有轻微的认知下降(下降 1SD 以上)或 FAQ=6-8 分。

aMCI 患者入组标准：根据 2011 年 NIA-AA 诊断标准及 Jak/Bondi 于 2014 年提出的量化标准/ADNI 数据库入组标准制定如下：纳入标准：①符合美国 NINCDS-ADRDA 可能的 aMCI 诊断标准；②年龄、性别不限；③轻中度 aMCI：MMSE 总分为 16-28 分。

AD 患者入组：纳入标准：①符合美国 NINCDS-ADRDA 可能的轻度 AD 诊断标准；②年龄、性别不限；③轻度 AD：MMSE 总分为 6-20 分。

排除标准包括：①有脑卒中病史者；②血管性痴呆(Hachinski 缺血指数量表,HIS>5 分者)或其他原因的痴呆；③抑郁症、精神分裂症、先天精神发育迟缓等精神疾病患者；④可引起脑功能障碍的其他神经系统疾病(如脑肿瘤、帕金森病、脑炎等)；⑤因脑外伤致认知损害者；其他可引起认知损害的疾病，如甲状腺功能异常、严重贫血、梅毒、HIV 等；⑥合并严重心、肝、肾及造血系统疾病；⑦与感染性或血管性损伤相一致的颞中回 MRI 的 FLAIR 或 T2 信号异常；体内有起搏器，动脉瘤夹，人工心脏瓣膜，耳植入体，金属碎片；眼睛，皮肤或体内有异物等 MRI 扫描禁忌症。

B. 整体认知功能测评

①简易智力量表(MMSE)、焦虑评价量表(ASA)和汉密尔顿抑郁量表(HAMD)、蒙特利尔认知评估量表(MOCA)、日常生活能力评估(ADL)、词语记忆、图形自由回忆及领悟力、Hachinski 缺血指数量表等量表。②采用修订韦氏成人智力量表手册(Wechsler Adult Intelligence-Revised in China, WAIS-RC)、修订韦氏记忆量表手册(Wechsler Memory Scale - Revised Administration and scoring manual, WMS-R)、词汇流畅测试等维度评定神经认知功能。

3) 磁共振数据采集

本项目的数据采集将在中国人民解放军总医院第二临床中心放射科西门子 3.0T 成像设备上完成，为保持和已有数据的一致性和可比性，拟采集的数据包括常规 T1 结构像、全脑 3D 高分辨率结构像、静息状态功能磁共振成像(fMRI)数据、弥散张量成像(DTI)数据。主要扫描参数如下：

- 常规 T2 像：轴状位扫描，重复时间(TR)=1924ms，回波时间(TE)=75ms，20 层，

层内分辨率=256 × 256, 视野(FOV)=240 mm×240mm, 层厚=5mm, 层间距=1mm。

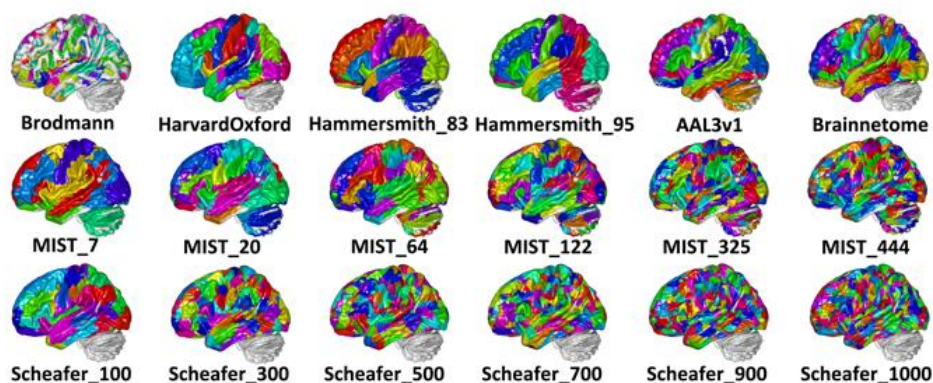
- 全脑 3D 高分辨率结构像: 轴状位扫描, TR=2530ms, TE=4.2ms, TI=900ms, FOV=256 mm×256mm, 层厚=1mm, 分辨率=256 × 256 × 176, 翻转角度=80 度。。
- 功能磁共振成像: 轴状位扫描, TR=2000ms, TE=40ms, 层内分辨率=64 × 64, FA=90 度, FOV=240 mm×240mm, 36 层, 层厚=3mm, 层间距=1mm, 扫描时间 6 分钟(180 个时间点)。
- 弥散张量成像(DTI): 轴位扫描, TR = 9000ms, TE = 85ms, FOV=256 mm×256mm, 层内分辨率=128×128, 层厚=2mm, 无间隔, 30 个方向弥散梯度方向, b 值=1000s/mm², b0 图像采集 4 次, 共 75 层。

第一部分: 多中心、多模态研究数据的综合使用

仿真数据的使用, 更加上一步的方法, 构建不同类型的仿真数据库, 评估亚型计算方法。**扩展已有的数据库:** 计划在本项目支持下, 在中国人民解放军总医院第二临床中心新收集 AD, aMCI, SCD 和正常对照数据的纵向跟踪数据共计 200 例, 加上我们在中国人民解放军总医院第二临床中心已有的近 400 例样本, 这样单个中心的样本可以接近 600 例, 有效的降低异质性。为保证实验的连续性和对原有数据的良好兼容性, 计划采用原参数采集新数据。**公共数据库的使用:** 为了验证我们的算法, 将以美国 ADNI 项目和 AIBL 数据库所提供的脑结构图像、PET 图像和纵向行为认知数据为基础验证我们的算法。目前已经取得该数据库的使用权, 完成了数据的预处理工作, 并发表了相关的研究论文。

第二部分: 基于多粒度脑图谱的 AD 亚型分析

- **构建基于脑影像的AD脑功能连接图谱及异常模式**
- **多尺度刻画AD脑功能连接图谱异常表征**



单模态
单尺度



多模态
多尺度

多个尺度的脑图谱：基于上述多模态特征的刻画，利用不同粒度的图谱，以脑网络组图谱为基础，关注不同分辨率的脑图谱，刻画基于不同脑图谱的 AD 亚型。

算法库构建及评估：利用构建的各个特征的常模，对特征进行归一化处理，然后用多元检验分析产生分类的特征图像。

亚新的生物学意义：借助不同的影像基因组研究，如 AHBA 数据库，研究不同亚型是否有其影响的特定的基因分布，对高权重基因进行富集分析，利用自有数据的血样，分析 AD 超高基因风险分数进行关联分析，探索不同亚型是否有其特定的基因特点。利用 ADNI 和本项目新采集的纵向数据评估不同亚型的纵向变化模式。利用 ADNI 的 PET 和 Tau 数据验证不同亚型是否有特定的代谢异常模式。

第三部分：基于多模态影像 AD 影像亚型个体化分析

本项目的科学问题之一就是如何确立早期 AD 的不同亚型和这些细分亚型所影响的脑网络，进而根据网络亚型的个体化表征，拟重点关注适合脑影像数据多中心、小样本，高维异构特征的特点的深度学习算法。脑影像具有样本量相对少，维度高，异质性强，易受多种因素干扰而信噪比相对较低等特点，过去的研究集中在传统的支持向量机（SVM），多元回归等方面。随着深度学习技术的发展，项目组拟提出采用规范化模型，利用单机器的大样本正常人学习到正常人群的特点和异质变化特点，然后对于任何一个单个个体，利用机器学习中常用的弱学习方式，探索将多个小的差异组合成稳定的、敏感性好的差异特征，继续融合脑功能和结构的异常，更进一步深入挖掘 AD 异常的连接模式与认知能力之间的关系；项目提出融合脑区、脑网络以及行为等不同尺度的特征发现 AD 所影响的脑结构和脑功能，建立 AD “生物亚型”和早期高危人群的个体化表征，促进脑影像在临床应用研究中特征的生物学“可解释”，加快真正的临床应用。融合多层次特征进行亚型的划分，最后再进行一站式辅助个体化诊断。

3.1.2 技术路线

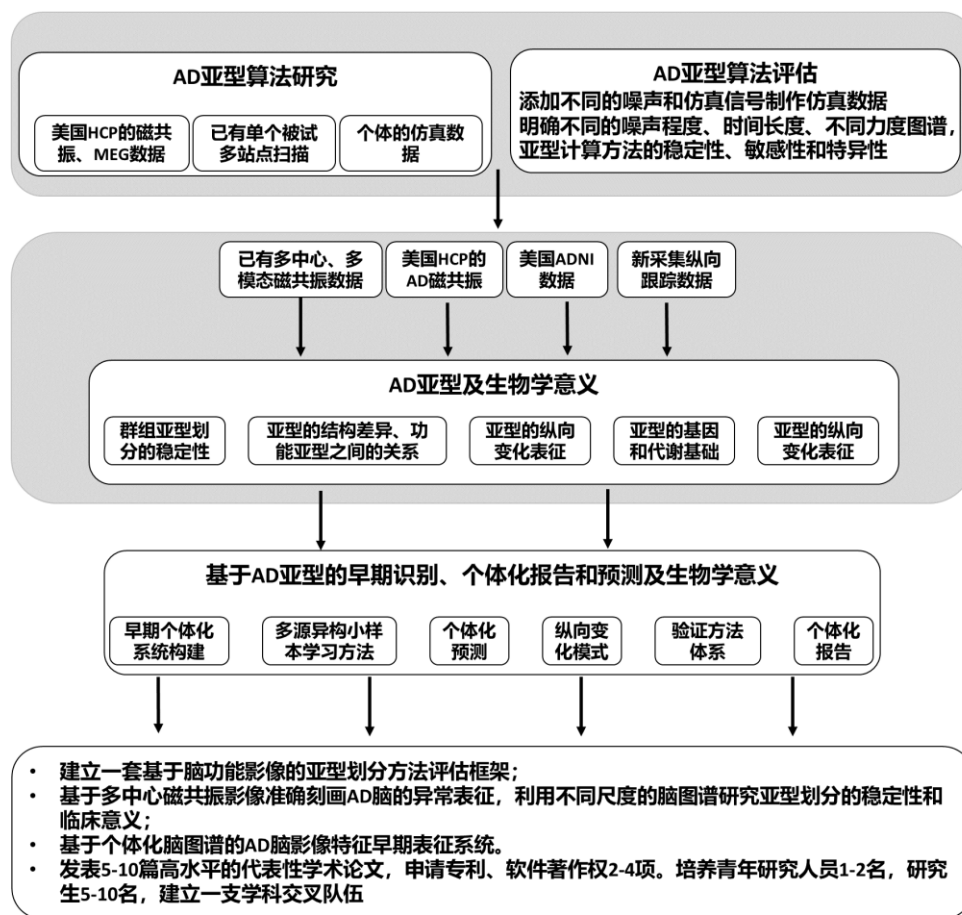


图 4. 本项目的主要技术路线

3.1.3 关键技术及可能的解决方案

本项目数据包括多个中心的结构磁共振、功能磁共振、弥散磁共振数据及这些数据特征的融合。

1) 如何融合多个中心的影像学测度

由于机器噪声、机器场强、采集参数等原因稍有不同，如何评估不同中心影像指标、如何归一化单中心的差异也是一个难点。我们计划借鉴基因组学的多中心统计方法，先对每个中心单独统计，然后对统计后的结果进行再统计、寻找公共的差异区域。预实验的结果已经验证了该方法的有效性。另外，最新的 PNAS 的文章也提出一种通过矫正单个中心的样本统计分布的方法进行多中心的融合(Zhou et al., Proc Natl Acad Sci U S A. 2018;115(7):1481-6.)，并在 ADNI 数据上取得了较好的效果，本项目计划尝试将该方法作为备选之一，对我们的影像组学数据进行统计矫正。

2) 多模态数据特征的融合问题

项目采取多模态磁共振影像，提取多个重要脑区的影像组学信息，丰富的图像信息无疑将有助设计更为高效的学习算法，但同时也增加的样本过学习（over-fitting）的风险。因此，本项目的另一个关键问题是如何从多模态图像中提取出有效信息，并剔除冗余信息并提高机器学习算法的泛化能力。我们将对多模态、多时间点图像特征在体元水平上以及区域特征水平进行融合。对多模态图像特征，我们将采取对不同模态特征进行加权平均，以及提取压缩性的统计学特性等方式对多模态图像特征进行融合处理。在图像特征融合中，将考虑图像特征的频域与空域特性，保证融合后新特征同样具有空间平滑性。将结合机器学习算法，以提高机器学习算法性能为评价标准，选择最优图像特征。

3) 纵向数据收集较为困难

在这个方面，合作单位中国人民解放军总医院已经有丰富的研究经验，病人的配合度较高，另外拟给与所有参与被试给予一定的报酬、报销往返的车费、免除挂号、检查费用等激励方式，对已有数据库的人员动员参加纵向跟踪的研究。在这方面，合作单位中国人民解放军总医院、北京邮电大学校医院还通过多次义诊、在微信、电视台等媒体上多次科普，提供患者及家属对该疾病的认知，使得由邀请参加变为积极主动参加。

3.2 可行性分析

3.2.1 研究目标明确，理论上可行

设定的研究目标源于对基于医学图像进行 AD 早期识别研究的深刻认识和前期项目的研究基础和积累，提出评估目前主流的亚型分析方法，发展新的适用于多中心磁共振影像的亚型研究算法，并基于此建立适用于磁共振影像数据特征的学习算法，基于多中心多模态磁共振影像，明确不同亚型的 AD 及高危人群的脑连接异常模式，然后利用模式识别以及机器学习技术，实现高效准确的 AD 亚型划分，为未来 AD 的早期智能精准识别、精准干预打下坚实的基础。

3.2.2 关键步骤的预实验结果表明研究方案切实可行

项目组已经进行了部分初步研究，探索了基于功能影像、PET 影像的 AD 亚型划分，关于模式分类，团队提出基于注意力机制的深度学习算法的 AD 早期识别系统构建等新的研究方法，实现了技术路线的几个关键步骤。所有的预实验的结果表明：项目提出的研究方案切实可行，可保证项目的顺利实施(具体的预实验结果请详见研究基础的预实验部分介绍)。