

哈尔滨工业大学（深圳）

大一年度项目立项报告

项目名称：____基于脑磁图的信号处理及脑疾病辅助诊断应用____

项目负责人：____刘晰月____学号：____2024311461____

联系电话：____15969472667____电子邮箱：____3409778214@qq.com____

学 院：____基础学部____

指导教师：____叶辰飞____职称：____助理研究员____

联系电话：____13751184856____电子邮箱：____yechenfei@hit.edu.cn____

学 院：____信息科学与技术学院____

填表日期：2024 年 11 月 19 日

一、项目团队成员（包括项目负责人、按顺序）

姓名	性别	所在学院	学号	联系电话	本人签字
刘晰月	女	基础学部	2024311461	15969472667	刘晰月
寇家辉	男	基础学部	2024311312	19879921069	寇家辉
黄家祺	男	基础学部	2024311148	17796007256	黄家祺
马樾程	男	基础学部	2024311383	13778805009	马樾程
盛彦杰	男	基础学部	2024311173	13312984339	盛彦杰

二、立项报告（字数在 2000 字以上，篇幅不够可附页）

基于脑磁图的信号处理及脑疾病辅助诊断应用

（一） 立项背景

1.1 脑磁图的基本原理及优势

脑磁图（Magnetoencephalography, MEG）是一种高时空分辨探测大脑神经元电活动所产生磁场的神经影像学技术，通过分布在头皮表面的高灵敏度磁场传感器捕捉大脑神经元活动产生的微弱磁场（100fT 量级）^[1]，可实现对大脑皮层下活动模式的无创实时监测，在脑科学研究和临床诊断方面具有重要的作用和广泛的应用前景。作为少数几种直接观测人脑神经电磁信号的技术，它能够精确地对脑内活动神经元进行实时定位。

MEG 可以精确地对脑内活动神经元进行实时定位，且信号不受头皮软组织、颅骨等人体组织阻抗的影响，这是因为大脑不同组织的磁导率一致。而同样作为直接测量神经信号的技术脑电图（Electroencephalograph, EEG），由于大脑不同组织的电导率不同，使得由 EEG 来精确定位活动神经元变得较为困难^[2]。

MEG 的测量数据都是脑内神经元电活动的直接反映，所以同时具有毫米级的空间分辨率和毫秒级的时间分辨率。fMRI（磁共振成像）则通过检测与大脑活动相关的血液动力学变化，从而间接反映大脑活动。虽然 fMRI 能够提供非常详细的大脑结构图像，具有较好的空间分辨率，但是由于血流增加比神经活动迟延，fMRI 的时间分辨率明显受到“血液动力学”因素的制约^[3]，时间分辨率约为 1s。

如图 1 所示，可以明显看出 MEG 具有毫米级的空间分辨率和毫秒级的时间分辨率，优于其他影像学技术。

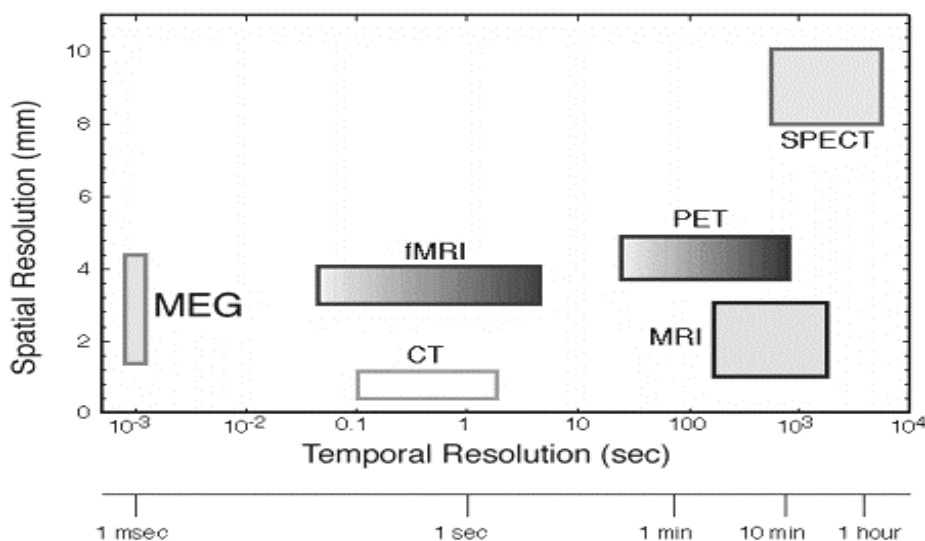


图 1 MEG 与其他成像技术的时空解像度比较

除了上述优势外，MEG 还是非侵入式的，完全无创的。它不需要注射同位素或暴露于 X 射线，适用于包括儿童孕妇在内的所有人群。近年来，MEG 技术在神经系统疾病的诊疗中取得了显著进展。可为神经系统疾病无创诊断、脑功能区定位等临床和科研领域提供全新技术手段。^[2]

1.2 帕金森病的临床诊断困难

帕金森病（Parkinson's disease, PD）是仅次于阿尔兹海默症(Alzheimer's disease)的世界第二大常见的神经退行性疾病。在生物学上其黑质纹状体细胞的进行性损失和细胞内路易斯小体的聚集是该疾病的主要病理，并且当相应的神经元细胞死亡到 50%-60%之后才会显现可被临床发现的症状。有研究表明只有在脑内损伤 60% 的黑质纹状体神经元细胞，耗尽 80% 纹状体多巴胺神经元的时候才会被诊断患有帕金森疾病^[4]。PD 早期易与多系统萎缩、进行性核上性麻痹等帕金森叠加综合征混淆，在早期帕金森病的临床试验中，高达 15% 的疾病患者被错误诊断^[5]。一项 Meta 分析提示，对帕金森病专家而言，初始诊断的准确率仅为 79.6%，进一步随访后的诊断准确率才达到 83.9%^[6]。

目前，国内外对于帕金森疾病的诊断大多是根据临床显现的症状、家族病史及抗帕金森药物的治疗反应，如 PD 患者病症所述典型患者一般不难做出诊断。但是该疾病起病缓慢，研究表明从黑质纹状体神经元细胞改变到损伤再到患者表现出临床症状，这需要经历一个较漫长的临床前过程。一般地，黑质纹状体神经元细胞丢失小于 50%的患者临床症状不明显，在此阶段医生尚无法做出患有帕金森疾病的诊断^[7]。然而当出现可以做出诊断的 PD 症状依据时，患者脑内多巴胺能神经元、黑质纹状体细胞丢失率已达 70%—80%^[8]。由此可见，此时做出确诊却早已处于疾病的中后期，已经错过了最佳的治疗时机。因此若能够在早期或亚临床期就可以确诊 PD，就能使患者尽早得到相应的神经保护治疗，无论是对于患者本身还是其家庭都至关重要。目前虽然目前已经有部分药物在疾病的早期可以帮助患者减轻负担，然而这些药物本身带有强刺激性，是一把双刃剑，如果长时间依赖便会对大脑神经造成极大的损伤。随着医学技术水平地不断提高，国内外学者一直在探寻新的、高效地、准确的诊断方法。

MEG 作为一种非侵入性的脑成像技术，通过监测大脑中的磁场活动，揭示帕金森病患者大脑神经活动和功能连接变化，可以无创便捷的辅助诊断 PD。且兼备高时间和空间分辨率的特性，能够从微观时域尺度上对 PD 患者的脑部异常变化进行观测。目前，MEG 应用于 PD 的辅助诊断已经取得了一些阶段性成果：基于 MEG，以运动网络为重点的研究和以全脑为重点的研究将脑活动变化与 PD 临床运动症状与

认知功能下降相关联^[9]；脑皮质和皮层下 MEG 信号变化与疾病进展有较强关联，有希望作为非侵入性标志物辅助诊断^[10]。

本次课题研究将对利用 MEG 辅助诊断帕金森症的应用进行探究，旨在利用 MEG 影像数据识别异常的大脑模式，最终建立能够有效区分早期 PD 患者和正常人的机器学习模型，为临床研究提供有力参考。

1.3 脑功能的研究指标

在 MEG 的研究中，我们常用的研究指标有特定频段下的功率谱（Power Spectrum）和幅度耦合指标（AEC（Amplitude Envelope Correlation）、AEC-c（corrected Amplitude Envelope Correlation））等。功率谱作为描述信号在不同频率上的功率分布的指标，常用于分析特定频谱段下的脑活动，常用的频段有 delta（0.5-4 Hz）、theta（4-8 Hz）、alpha（8-12Hz）、beta（12-30 Hz）和 gamma（30-100 Hz）等。AEC 和 AEC-c 用于评估不同脑区之间的幅度耦合，能够反映大脑各部分之间的功能连接。

ALFF（低频振幅）、fALFF（分数低频振幅）和 ReHo（局部一致性）是静息态 fMRI 研究中常用的三种指标。ALFF 能够反映大脑在静息状态下的功能活动强度，fALFF 则相较于 ALFF 具有更高的准确性，ReHo 值可以反映大脑局部功能活动的同步性或协调性从而表征大脑局部功能活动的异常。它们各自具有独特的定义和应用，是基于 BOLD fMRI 技术提取的生理变量，在实际应用中，这三个指标可以相互补充，共同揭示大脑在静息状态下的功能活动特征。

由于信号容易受到各种噪声的干扰，包括呼吸、心跳等生理噪声以及扫描仪噪声等，这些指标的准确性也会受到影响，导致结果出现偏差。为了减少噪声的影响，研究者们通常采取多种预处理措施，如去噪、平滑等。然而，这些预处理措施仍可能会在一定程度上改变 BOLD 信号的原始特征，从而影响指标的计算结果。

为更好的分析影像数据以及识别 PD 异常的大脑模式，我们在传统的指标外使用加权排列熵（Weighted Permutation Entropy, WPE）为度量进行信息论复杂性分析。使用 WPE 具有多种优势：（i）能够量化时间序列的复杂性，捕捉神经信号的动态变化和不规则性。这种方法对神经信号的细微变化具有高度敏感性，能够揭示大脑活动的深层次特征，（ii）具有较强的抗噪能力，通过计算时间序列的排列模式及其概率分布来量化复杂性，这一过程有助于减少噪声对结果的影响，（iii）不仅提供了时间序列复杂性的量化指标，还可以通过分析不同时间尺度上的排列熵来揭示时间序列的动态特征。这有助于研究人员更深入地了解大脑活动的时空特性。^[11]

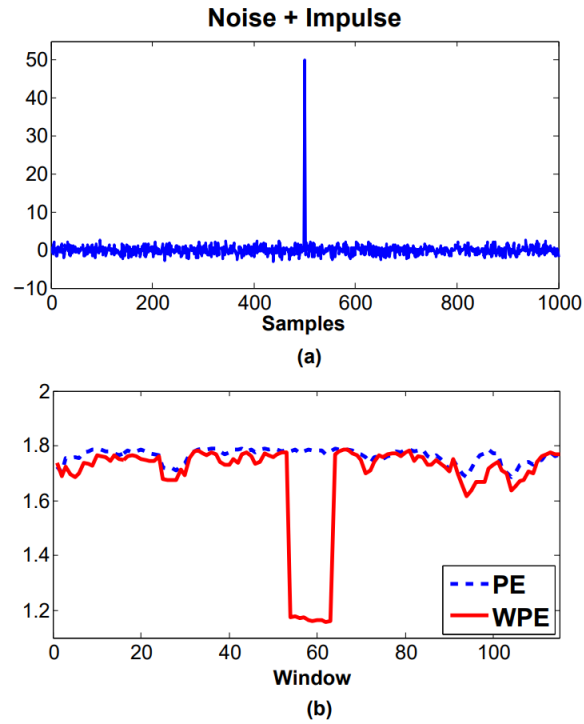


图 2 脉冲情况下的 PE 与 WPE; (a) 均值和单位方差为零的加性白色高斯噪声的脉冲;
(b) 窗口为 80 个样本时计算的 PE 和 WPE 值滑动了 10 个样本。

同时，与常用于描述非线性时间序列复杂性的排列熵（PE）相比，WPE 能够解决 PE 的一些局限性。如图 2 所示，我们注意到在存在较高噪声水平的情况下，在 PE 值没有显示任何明显变化的脉冲区域中注意到 WPE 值的显著下降，WPE 从具有尖峰特征或幅度突然变化的数据中提取复杂性信息的独特能力。

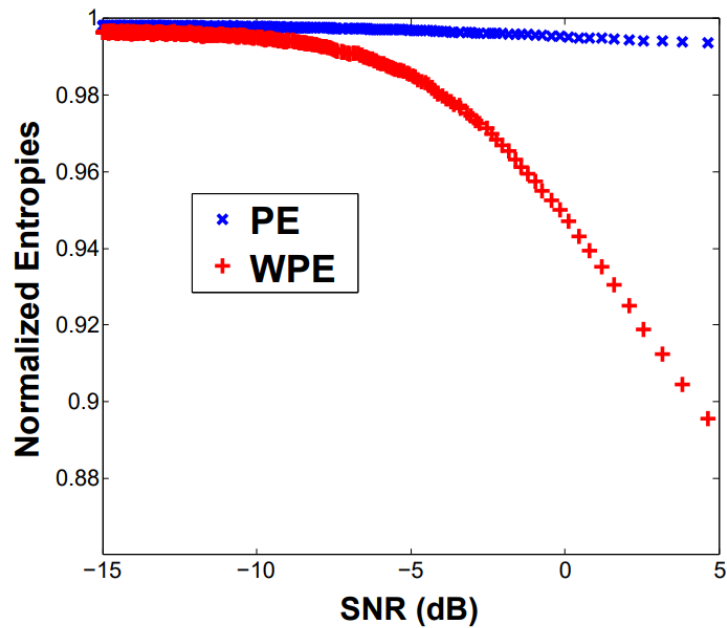


图 3 不同级别的信噪比（SNR）的 PE 和 WPE 值

如图 3 中所示，两种熵度量都随着 SNR 的增加而减小。因为噪声对复杂性的影响变得不那么显著，WPE 的下降速度高于 PE，这反映了 WPE 对噪声的鲁棒性更好。因此使用 WPE 可以十分有效的降低噪音对于数据的影响，从而使分析结果更加精确。^[11]

1.4 机器学习模型辅助疾病诊断的应用

医学诊断确定解释一个人症状和体征的疾病或病症。通常，诊断信息是从患者的病史和检查中收集的，由于许多适应症和症状不明确，只能由训练有素的专家进行诊断，因此通常很困难。

近年来，随着人工智能（artificial intelligence, AI）技术的发展，AI 在医学数据分析和神经影像自动化诊断研究中的应用日益广泛。机器学习（ML）是目前实现的主要方法，许多学者都使用机器学习（ML）方法进行疾病诊断，例如 Ansari et al.

（2011）提供了一种基于 Neuro-Fuzzy 集成系统的自动冠心病诊断系统，准确率约为 89%^[12]; Charleonnann et al.（2016）使用公开可用的数据集来评估四种不同的 ML 算法：K 最近邻（KNN）、支持向量机（SVM）、逻辑回归（LR）和决策树分类器，准确率分别为 98.1%、98.3%、96.55% 和 94.8%^[13]，准确度均处于较高水平。机器学习模型在乳腺癌、糖尿病等病症的辅助诊断方面也已经获成功应用。

机器学习在帕金森病的诊断中也已经有了较广的应用。表格 1 总结之前研究者基于 ML 的方法，开发的智能帕金森诊断系统，包括算法数据集和性能评估等方面指标。据表格 1 可以看到，五组研究使用了不同的算法模型，根据不同类型的的数据来实现辅助诊断，其诊断精度基本处于较高水平。^[14]

表格 1 基于影像学数据的机器学习辅助 PD 诊断实例

Author	Sample	Data	Methods	Main findings
Singh et al., 2018 ^[15]	408 PD 71 SWEDD (from PPMI); [128 AD 262 HC 447 MCI (from ADNI)]	Discretized Voxel Intensity Changes extracted from MRI images by SOM	Feature selection: SOM; Classification:SVM;Validation: tenfold cross-validation	Classification accuracy: 92.63 ± 0.06% (PD vs. HC); 94.63 ± 0.05% (PD vs. SWEDD); 92.65 ± 0.08% (SWEDD vs. HC); [94.29 ± 0.08% (AD vs. HC); 85.43 ± 0.08% (AD vs. MCI); 95.24 ± 0.05% (MCI vs. HC);]
Zeng et al., 2017 ^[16]	45 probable PD, 40 HC	GM in the cerebellum from MRI	Featureselection: Recursive feature elimination; Classification: SVM; Validation: Leave-one-out; fivefold (twofold,632-fold) cross-validation	Classification accuracy: 97.7% for leave-one-out validation, 97.2% and 96.9% for twofold and fivefold cross-validation respectively

Gu et al., 2016 ^[17]	52 PD (19 PIGD, 25 TD, 8 mixed subtype), 45 HC	GM, WM, CSF volumes from MRI; FA, MD, RD, AD from DTI; ReHo and ALFF from resting-state fMRI	Feature selection: Recursive feature elimination; Classification: SVM; Validation: Leave-one-out	Classification accuracy: PIGD vs. non-PIGD 92.3%
Haller et al., 2012 ^[18]	17 PD, 23 other Parkinsonism (5 MSA; 1 PSP; 17 other types)	TBSS from DTI	Feature selection: select the most discriminative features with RELIEF; Classification: SVM; Validation: tenfold cross-validation	Classification accuracy: 97.5 ± 7.54% (PD vs. other Parkinsonism)

ML 技术的发展为 PD 计算机辅助诊断与健康测评赋予了智能决策技术，从而有效地解决了特殊情况下的误诊或无法确诊的问题。ML 模型构建中的分类器训练旨在挖掘出海量数据中潜在的规律，构建一套相应规律的决策模型，从而能实现对各种未知数据特征的快速准确分析判别。

本课题也将基于处理分析后影像数据和 MEG 脑活动时空动态表征模式，建立相应的机器学习模型，以能够提取 PD 患者脑结构或功能的形态学特征，从而进一步挖掘出与疾病过程直接相关的有效病理信息。希望机器学习方法通过对全脑影像特征进行特征选择及模型构建，实现 PD 的自动分类及预测，最终实现基于 MEG 的个体化 PD 辅助诊断 AI 模型，并验证模型性能，从而提高诊断的准确性。

本课题研究目的的核心，是将 fMRI 数据分析的时序动态脑功能信号复杂度指标 WPE 迁移到 MEG 数据分析中，探索 PD 患者脑功能活动在微观时域尺度下是否也会出现病理特异性的复杂度异常变化，并检验该指标作为 PD 生物标志物实现 PD 影像学辅助诊断的真实性能。

（二） 项目研究内容及实施方案

2.1 研究内容

2.1.1 帕金森病的脑功能时域复杂度

大脑功能的实现依赖于不同功能的大脑系统之间的有效沟通^[19]。帕金森病往往导致病人脑网络的功能活动区域发生异质化，使不同脑区域的功能连接作用受损，导致病人出现不同程度的临床症状。分析信号的复杂度动态特征，表征脑区的功能活动特征，构建脑网络功能活动架构，有利于理解和识别帕金森病所引起的功能活动异质化。

为了研究帕金森病的脑功能时域复杂度，重点在于将 fMRI 数据分析的时序动态脑功能信号复杂度指标 WPE 迁移到 MEG 数据分析中，探索 PD 患者脑功能活动在微观时域尺度下是否也会出现病理特异性的复杂度异常变化，使我们能够将各个区域的复

杂度时间序列与这些区域之间的功能连接联系起来。

进一步探索解释得到的数据指标与帕金森症的关系，可以计算其与临床表现之间相关性，以便于后续辅助机器学习模型的构建与训练。

2.1.2 辅助诊断的机器学习模型

本课题建立机器学习的步骤如下:对 MEG 数据进行分析 and 处理，识别动态表征 模式；其次，利用 WPE 刻画人脑静息状态下，脑功能活动的动态变异性，解析 PD 脑功能信号可变性特征；选择适合 PD 分类任务的算法，将数据集分为训练集和测试集，用训练集来训练选定的模型;进行评估与优化。采用测试集评估已完成训练的模型，计算准确率(accuracy, ACC)、召回率(recall, R)、F1 等评估指标;根据评估结果，进行模型的优化和调整，以提高模型的性能,并尝试调整特征选择的方法、模型参数的设置等，对模型进行迭代优化。

其中，模型选择与训练是构建机器学习模型的核心步骤之一，常见的机器学习模型如逻辑回归 (Logistic Regression)、随机森林(random forest, RF)、支持向量机 (Support Vector Machine,SVM) 或极端梯度上升 (XGboost)。

逻辑回归是一种广泛使用的统计模型，逻辑回归通过使用逻辑函数（也称为 Sigmoid 函数）将线性回归模型的输出映射到 0 和 1 之间，从而预测一个事件发生的概率。尽管名字中有“回归”二字，但它主要用于分类问题，特别是二分类问题。即样本分为两个互斥的类别，每个样本都属于两个类别中的一个，且只能属于其中一个。例如预测患者是否患有帕金森病就属于一个二分类问题，可以适用于本课题。^[20]

随机森林是一种集成学习方法，由多个决策树组成，每个决策树都依赖于一个随机向量的值，这些值对于森林中的所有树都是独立且同分布的。随着森林中树木数量的增加，泛化误差会收敛到一个极限。随机森林的泛化误差取决于森林中单个树的强度以及它们之间的相关性。由于随机森林集成多个决策树，降低模型的方差，因此具有较好的鲁棒性，适用于具有噪声的数据。^[21]

支持向量机 (Support Vector Machine, SVM) 是一种监督学习算法，主要用于分类问题。SVM 算法的核心思想是找到一个超平面（在二维空间中是一条直线，在三维空间中是一个平面，在更高维空间中是一个超平面），这个超平面能够将不同类别的数据点尽可能地分开，并且具有最大的间隔。SVM 在解决小样本、非线性问题以及高维空间问题上表现出色，被广泛应用于模式识别、图像识别、生物信息学等领域，是一种非常强大的分类器。^[22]

XGBoost 类似于梯度上升框架，但是兼具线性模型求解器和树学习算法，作为一种

新型高效的 Boosting 算法，在基础算法上加以优化改进，提高了精度，专门设计用于处理大规模数据集并优化树模型性能。XGBoost 因其出色的预测性能和易用性，在包括疾病辅助诊断在内的多个领域得到了广泛应用，特别是在处理结构化数据和提高模型可解释性方面表现出色。^[23]

2.2 技术路线

本课题的技术路线如图 4 所示

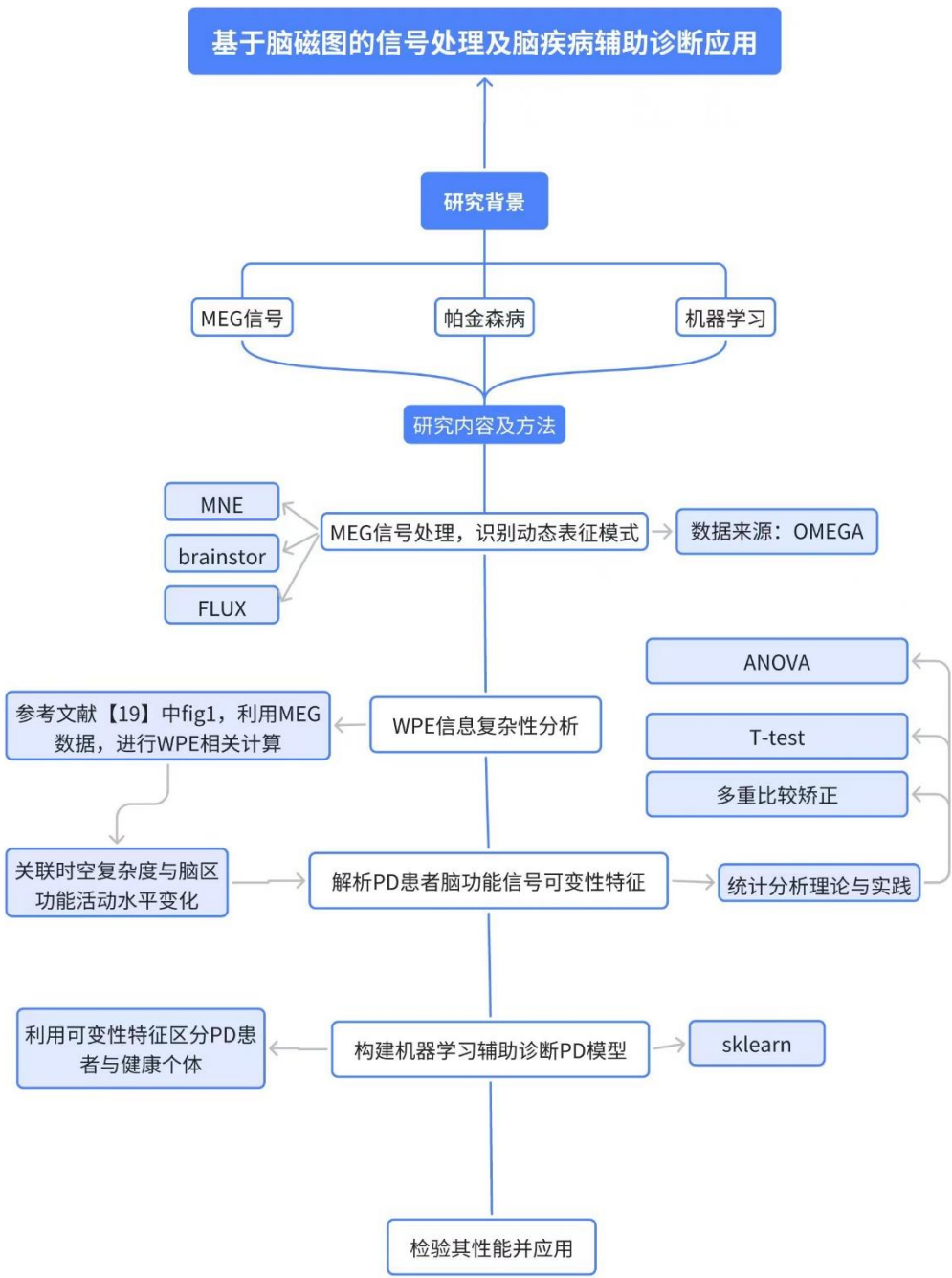


图 4 技术路线

2.3 实施方案

2.3.1 数据收集与预处理

本课题中相关的 MEG 数据主要来自于 OMEGA（目前可用的最大的 MEG 数据存储库之一）^[24]。

MEG 数据的频谱丰富而复杂。多个过程同时发生，涉及不同空间、时间和频率尺度的神经群体。数据预处理的目的是提高感兴趣的信号水平，同时衰减或甚至消除被伪影损坏的记录中的某些片段。在本课题中，以 Brainstorm 作为主要的处理软件^[25]，主要进行如下的预处理步骤：

(1) 寻找不良信号通道：在数据采集过程中，由于传感器可能损坏或不稳定等，往往会使部分通道出现不良信号影响整体数据。一些坏的通道很容易被检测到，仅查看信号轨迹，某些通道通常看起来比其他通道更嘈杂，或查看 2D 传感器地形，如果一个传感器在很长一段时间内显示与其相邻传感器截然不同的值，其质量也往往不能保证。而对于一些较难识别的通道，我们通常使用功率谱密度（PSD）来检测不良信号通道，这个过程使用 Welch 方法评估不同频率的 MEG/EEG 信号的功率^[26]。它将信号分割到给定长度的重叠窗口中，计算每个短段的傅里叶变换（FFT），并平均所有重叠窗口的 FFT 系数的功率，最终输出 PSD。通过在对应的 PSD 文件中查看并检查其频率分辨率来确定不良通道。

(2) 信号过滤：由于信号采集过程中生理和外界环境的影响，会在收集到的数据中产生噪声。为消除在特定频率下连续发生的噪声，可使用频率滤波器来消除特定频率范围的噪声。高通滤波器常用于去除在低频（ $< 1\text{Hz}$ ，例如呼吸或眼球运动）下出现的伪影；低通滤波器常用于去除高频时发生的强烈噪音（例如肌肉收缩、刺激器）；带通滤波器是低通滤波器和高通滤波器的组合，能够去除了目标频带之外的所有频率；陷波滤波器用于删除特定频率的正弦信号（电源线噪声、头部跟踪线圈）^[27]。根据存在的噪声和研究范围，调整滤波器长度和过渡带大小，选用适宜的滤波器能够有效的让信号更加纯净，便于后续分析。

(3) 信号空间分离技术：由于心脏活动、动脉搏动、肌肉运动以及眼动等有时跨越较大的频率范围或与感兴趣的大脑信号的频率重叠，时间传感器中周期性出现生物伪影，不便于利用频率滤波器消除，采用信号分离技术（如信号空间投影（SSP）^[28]）来处理这些伪影。生物伪影的产生有很强的可重复性并且始终发生、在同一位置，在处理软件中通过对串联的伪影运行主成分分析（PCA）实现空间分解，接着为空间分量设置 SSP 投影仪。这样每当从此文件中读取内容时，SSP 投影仪都会快速高效的应用于清除伪影。

2.3.2 WPE 计算与分析

WPE 定义如下，

$$H_{WPE}(m, \tau) = - \sum_{i: \pi_i^{m, \tau} \in \Pi} p_w(\pi_i^{m, \tau}) \ln p_w(\pi_i^{m, \tau})$$

其中 m 是时间序列中符号模体的长度， τ 是一个滞后参数，表示沿时间序列移动的时间点数， $\pi_i^{m, \tau}$ 表示给定模体长度的可能模体集合 Π 中的第 i 个符号模体， $p_w(\pi_i^{m, \tau})$ 是基序 i 的方差加权相对频率，由 Fadlallah 及其同事详细说明^[11]。

通过对具有相同序数模式但幅度变化不同的相邻向量进行加权，WPE 还可用于检测噪声或多分量信号的突然变化。值得注意的是，通过利用模式多样性（由 PE 评估）和振幅信息（由 SD 通过 BOLD 矢量评估），WPE 还捕获了在仅考虑这些特征之一时仍未检测到的信号动态，如图 5 和图 6 中所示。

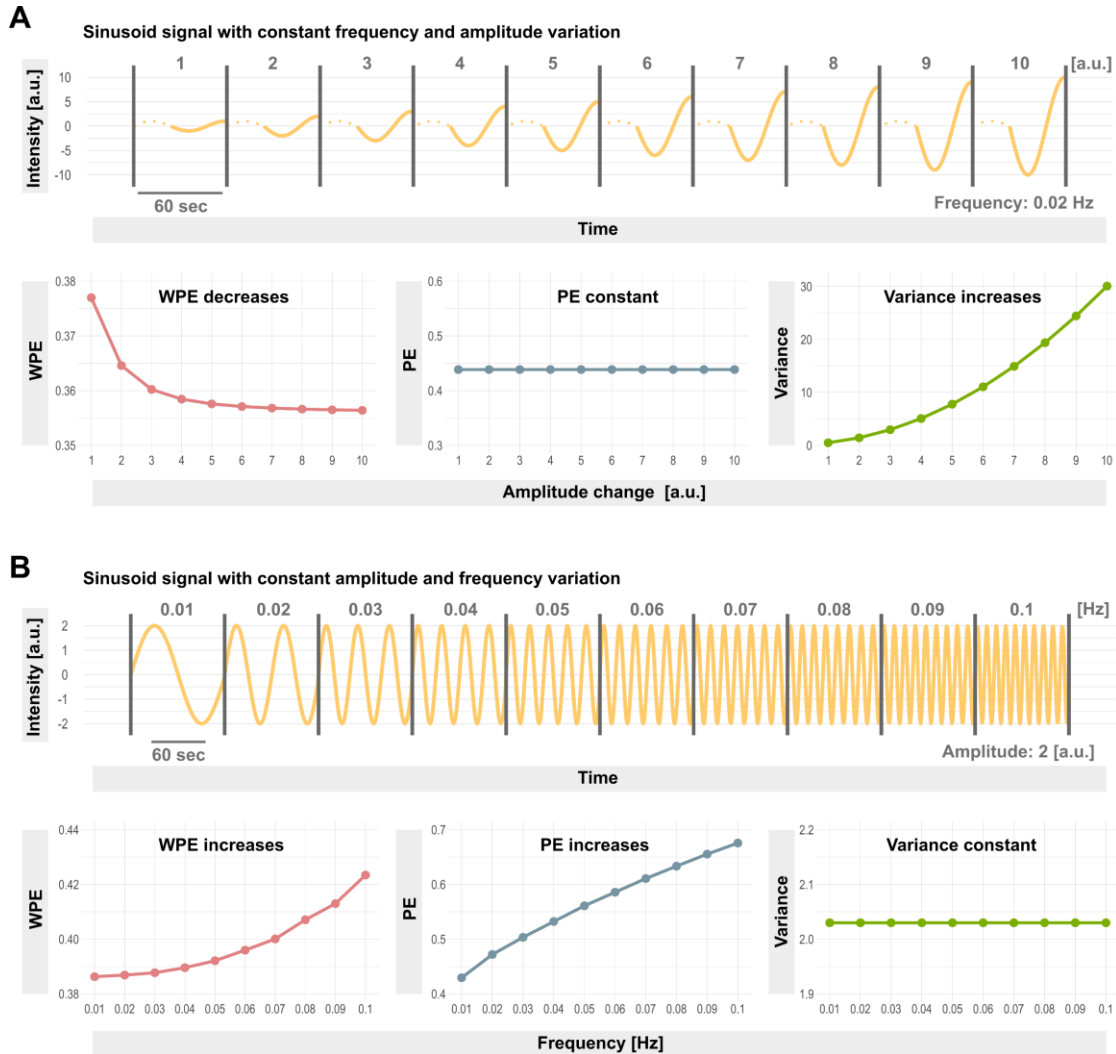


图 5^[19] 加权置换熵反映了在仅考虑模式分集或信号方差时仍未检测到的幅度和频率变化；（A）模拟的正弦信号，在实际 BOLD 信号范围（0.02 Hz）内具有恒定频率，并且振幅变化系统增加，

导致加权排列熵（WPE）和振幅方差发生变化，而单独的模式分集（排列熵，PE）保持不变。（B）模拟的正弦信号，具有恒定幅度和系统性增加的频率变化，导致加权排列熵（WPE）和模式分集（PE）发生变化，而信号方差保持不变。请注意，WPE 值随着频率含量的增加而单调增加，在 0.1 Hz 时达到约 0.42 个归一化位（与 BOLD 预处理中应用的低通滤波器设置相同），与根据经验观察到的从真实 BOLD 信号得出的复杂度时间序列的上限一致

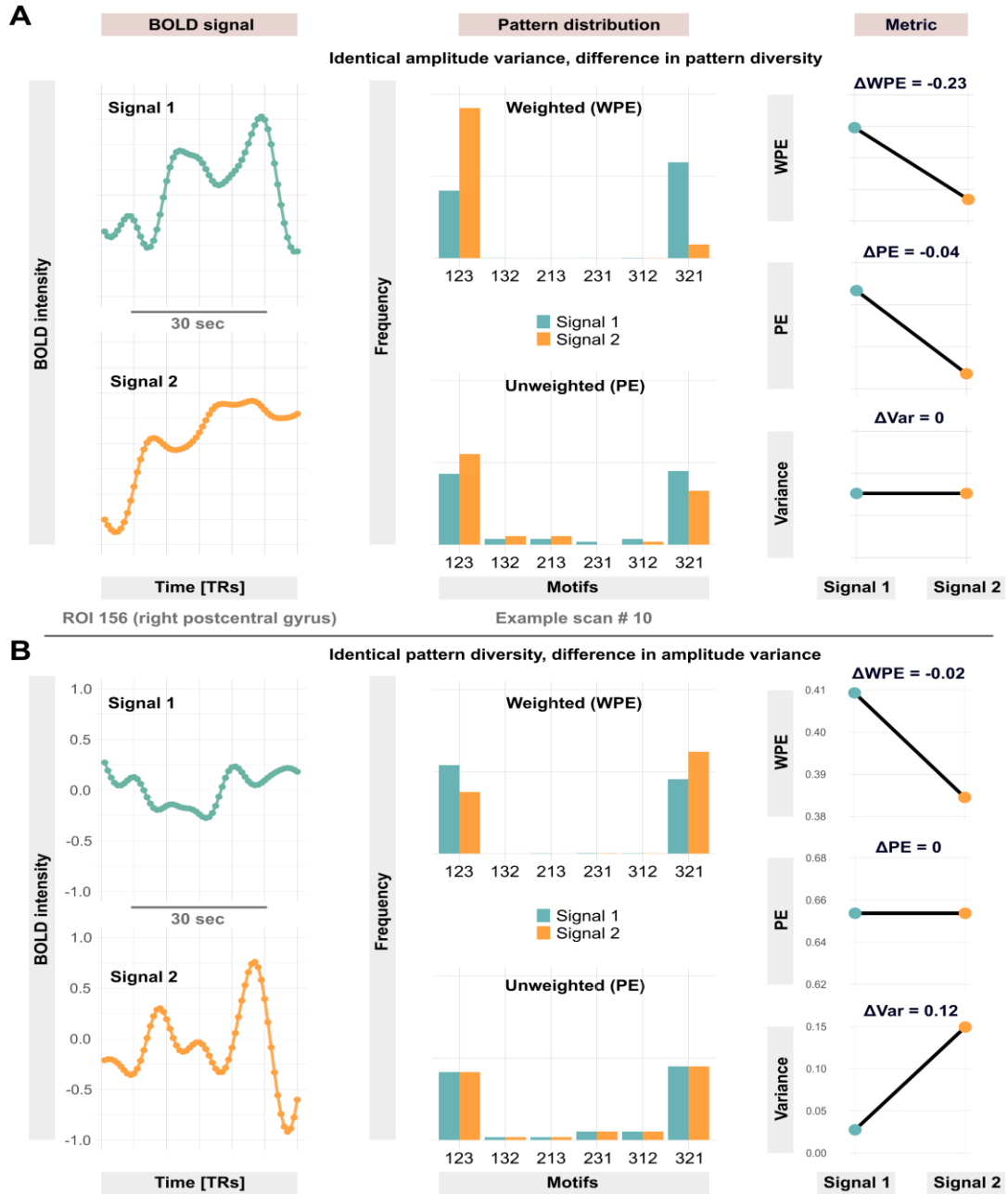


图 6^[19] 加权排列熵跟踪在仅考虑模式多样性或信号方差时仍未检测到的 BOLD 信号动态；（A）具有相同幅度方差但模式多样性（排列熵，PE）不同的 BOLD 信号；（B）具有相同模式分集但幅度变化的 BOLD 信号；所有信号片段都对应于来自代表性静息状态录音的右中央后回的非重叠 BOLD 窗口；柱形显示 BOLD 信号（左）、相关的方差加权模式分布（中）和信号量度的变化（右）

2.3.3 针对帕金森病的脑功能时域复杂度的分析

使用信息论复杂性分析，（i）将单个大脑区域的局部活动与网络的全局属性相关联，（ii）表示标准化空间中神经波动的内在可变性，以及（iii）允许对神经可变性进行时间分辨说明，以捕获单个信号中的时间动态。这种方法基于血氧水平依赖性（BOLD）活动的符号编码，并利用振幅信息和信号中抽象模式的多样性来量化给定时刻的不规则程度（即信号复杂性）（图 5 和图 6）。因此，我们获得了每个大脑区域的复杂性时间序列，它通过神经变异性的元表示来捕捉局部动态，并允许在网络中的节点（即单个区域）和网络中的边缘（即区域到区域的功能连接）之间建立链接。

该分析方法已经被用于过人类连接组计划（HCP）的静息态 fMRI 分析，如参考文献^[19]中，揭示了人脑活动的高度典型复杂性结构，结果如图 7。

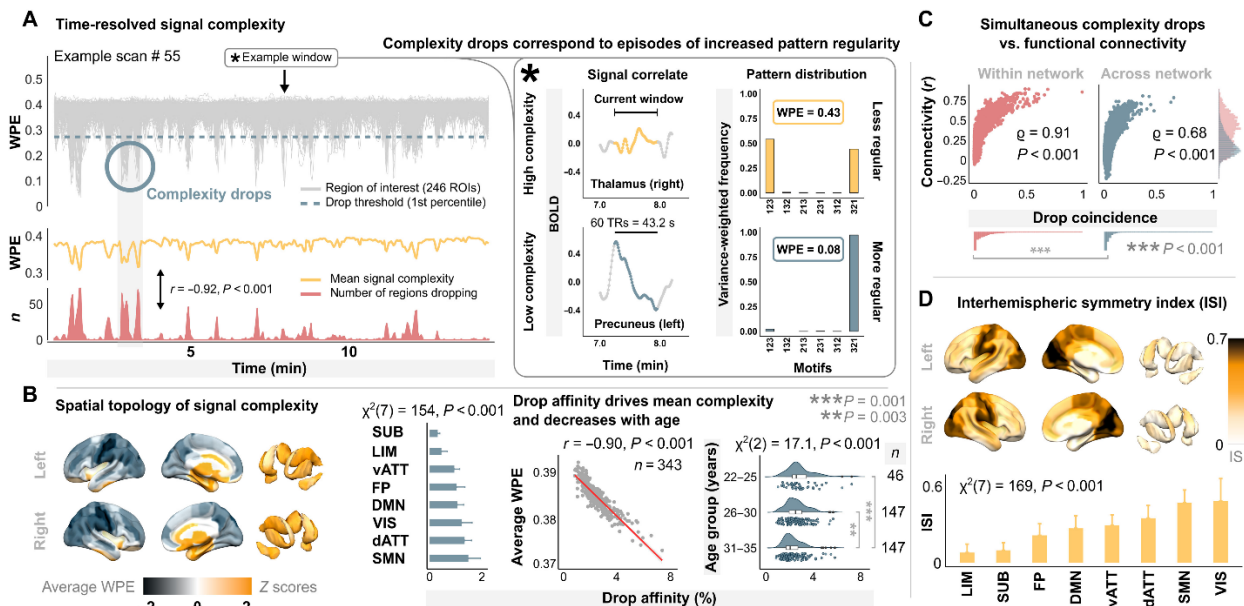


图 7^[19] 人脑活动的特点是神经规律的关键时刻，随着瞬态复杂性的下降而可见 (A)在示例性静息态记录中对来自每个 ROI 的神经信号进行时间分辨 WPE（Brainnetome parcellation;窗口长度为 60 TRs, 95% 重叠），插图显示两个具有代表性的高复杂度和低复杂度信号的 BOLD 相关；(B) 平均信号复杂度的空间分布：对 RSN 的 drop 亲和力，链接到总平均信号复杂性，以及与年龄相关的 drop 亲和力降低；(C) FC 作为典型 RSN 内和之间的下降重合的函数；(D) 半球间复杂性对称性的空间拓扑和网络分布；dATT，背侧注意力；DMN，默认模式；FP，额顶叶；LIM，边缘；SMN，躯体运动；SUB，皮质下；vATT，腹侧注意力；VIS，可视网络

这种 BOLD 动态的时间分辨表示随后使我们能够将各个区域的复杂度时间序列与这些区域之间的功能连接联系起来。分析大脑中复杂性下降的共现情况表明，任意两个区域之间的功能连接强度与它们表现出复杂性一起下降的程度密切相关。我们期望在本课题中利用 MEG 数据复现参考文献^[19]中相关计算以识别帕金森患者的脑功能区

域的复杂度水平和脑区的功能活动水平变化。

处理分析得到的数据指标，我们应用一般线性模型评估影像学-行为关联，评估 WPE 复杂度及相关指标与 PD 的关联。综合使用 T-test、ANOVA、多重比较矫正等分析方法增强结果的统计学显著性和可靠性。[29]

2.3.4 辅助诊断的机器学习模型训练与评估

本课题将使用 Scikit-learn 来实现相关建模。Scikit-learn 是一个开源的机器学习库，用于 Python 编程语言。它建立在 NumPy、SciPy 和 matplotlib 这些科学计算库之上，提供了简单有效的数据挖掘和数据分析工具，和一系列完善的机器学习算法。我们将基于已经得到的相关数据类型与指标，选择适宜的机器学习模型。

使用 Scikit-learn 进行模型训练时，首先，导入必要的 Scikit-learn 模块，包括数据集加载、模型选择、数据预处理、模型训练和评估等关键组件，这些模块构成了机器学习工作流的基础架构。接着，通过 Scikit-learn 提供的数据集加载函数，获取并理解数据集。随后，将数据集划分为特征集 x 和标签集 y ，并进一步分割为训练集和测试集，通常按照 80% 训练集和 20% 测试集的比例进行，以确保模型在未见数据上的泛化能力，防止过拟合。接着，进行模型选择，可以采用交叉验证和网格搜索等技术，对不同模型和参数组合进行评估，以选择最优模型。最后，将数据集的特征 x 和标签 y 传递给模型，模型通过算法学习数据中的模式和关系，调整内部参数使预测误差最小化。

在训练过程中，应当确保测试集独立于训练集，未在训练过程中使用来模拟模型在实际应用中的表现。模型训练完成后，利用先前分割好的测试集评估模型性能，通常涉及调用模型的 `predict` 方法。对于分类问题，常用的评估指标包括准确率、精确率、召回率等。分析计算好的性能指标，确保模型满足项目需求，若未达到预期性能，需要回到模型选择与训练阶段，尝试不同模型或调整参数。在模型部署前，使用 `Pipeline` 将数据预处理和模型训练步骤串联起来，构建一个完整的数据分析流程。这不仅简化了模型训练和预测过程，还提高了工作流的可重复性和可维护性。最后，将训练好的模型部署到实际应用中，对新数据进行预测和分析，实现机器学习项目的价值转化。

（三） 进度安排

2024 年 11 月至 2025 年 2 月

1) 熟悉 MEG 基础背景。查找学习 MEG 相关专业文献、学术期刊，理解 MEG 扫描的基本概念，以及 MEG 信号处理背景、理论、核心软件和方法。

2) 学习并理解人脑神经系统信号分析理论。掌握神经科学的基本概念和原理以及信号的基本概念、信号的分类、信号的传输和处理等基础知识。学习使用图像处理技术

和机器学习算法来处理和分析神经信号。

3) 熟练掌握课题相关编程计算环境。明确课题的具体需求,学习和掌握所选编程语言的语法、数据类型、控制结构。学习和掌握与课题相关的算法和模型,在编程环境中实现课题目标。

4) 制定并细化围绕课题目标的技术路线。深入学习研讨课题目标,并以此为出发点设计具体可行的技术路线,并在课题实施过程中不断进一步细化具体路线。

2025 年 3 月至 2025 年 5 月

1) 表征人脑活动 MEG 时空动态模式。使用 MEG 仪器记录大脑在特定任务或静息状态下的磁场信号,通过算法来增强 MEG 信号的信噪比,实现对大脑活动的精准测量和分析,并进行 dFC 可视化,进而表征人脑活动 MEG 时空动态模式。

2) 帕金森病 MEG 时空动态标志识别。利用 MEG 数据的时空特性,识别出与帕金森病相关的 MEG 时空动态标志。

3) 面向帕金森病辅助诊断的 AI 模型训练与性能验证。基于各项数据建立基于脑功能活动复杂性特征的帕金森病机器学习分类模型,对模型进行训练以提高诊断准确性,并进行性能验证。

(四) 预期目标

1. 提出一种表征人脑静息态功能活动的局部复杂性的方法;在至少两种脑图谱中评估不同脑区(皮层及皮层下区域)的神经活动复杂性特点,并验证该方法的跨图谱稳定性(DK 图谱和 HCPMMP 图谱)。

2. 建立一个基于脑功能活动复杂性特征的帕金森病机器学习分类模型,分类准确率不低于 80%。

3. 发现与脑功能活动复杂性特征具有统计显著性相关且相关系数 >0.3 的帕金森病认知行为功能表型(如 UPDRS、MMSE 等)不少于 2 种。

4. 对比 MEG 和静息态 MRI 的 PD 特异性 WPE 模式。

(五) 经费计划

1. 知识传播费:论文出版、专利申请(70%)

2. 图书费、资料费、打印费、耗材费(30%)

(六) 参考文献

[1] 杨宇佳,孙畅,丁铭.脑磁图技术在神经系统疾病诊疗中的应用[J].中国临床医生杂志,2024,52(08):887-891.

- [2] 霍小林. 脑磁源定位技术研究 [D]. 浙江大学,2001.
- [3] 苏敏, 尧德中. 脑功能多种成像方式整合技术的研究及其进展 [J]. 生物医学工程学杂志,2005,(02):385-388.
- [4] K Ray Chaudhuri;;Daniel G Healy;;Anthony HV Schapira.Non-motor symptoms of Parkinson's disease: diagnosis and management[J].Lancet Neurology,2006(3):235-245.
- [5] Bloem, Bastiaan R et al., Parkinson's disease, The Lancet, Volume 397, Issue 10291, 2284 - 2303
- [6] RAZZO G, COPETTI M, ACUTI S, et al. Accuracy of clinical diagnosis of Parkinson disease. A systematic review and metaanalysis[J]. Neurology, 2016, 85(1): 1-11
- [7] 武泉州. 基于多模态脑影像的帕金森早期诊断研究 [D]. 西安电子科技大学,2020.DOI:10.27389/d.cnki.gxadu.2020.001195.
- [8] 刘军;. 中国帕金森病的诊断标准 (2016 版)[J]. 中华神经科杂志,2016(04)
- [9] Boon LI, Geraedts VJ, Hillebrand A, Tannemaat MR, Contarino MF, Stam CJ, Berendse HW. A systematic review of MEG-based studies in Parkinson's disease: The motor system and beyond. Hum Brain Mapp. 2019 Jun 15;40(9):2827-2848. doi: 10.1002/hbm.24562. Epub 2019 Mar 7. PMID: 30843285; PMCID: PMC6594068.
- [10] Boon, L.I., Hillebrand, A., Schoonheim, M.M. et al. Cortical and Subcortical Changes in MEG Activity Reflect Parkinson's Progression over a Period of 7 Years. Brain Topogr 36, 566–580 (2023).
- [11] Fadlallah, Bilal H. et al. “Weighted-permutation entropy: a complexity measure for time series incorporating amplitude information.” Physical review. E, Statistical, nonlinear, and soft matter physics 87 2 (2013): 022911 .
- [12] Abdul Quaiyum Ansari and Neeraj Kumar Gupta. Automated diagnosis of coronary heart disease using neuro- fuzzy integrated system. In 2011 World Congress on Information and Communication Technologies, pages 1379 - 1384. IEEE, 2011.
- [13] Anusorn Charleonnann, Thipwan Fufaung, Tippawan Niyomwong, Wandee Chokchueypattanakit, Sathit Suwan-nawach, and Nitat Ninchawee. Predictive analytics for chronic kidney disease using machine learning techniques.In 2016 management and innovation technology international conference (MITicon), pages MIT–80. IEEE, 2016.
- [14] Zhang, J. Mining imaging and clinical data with machine learning approaches for the diagnosis and early detection of Parkinson's disease. *npj Parkinsons Dis.* 8, 13 (2022).
- [15] Singh, G., Samavedham, L. & Lim, E. C. Alzheimer's disease neuroimaging initiative; Parkinson progression marker initiative. Determination of imaging biomarkers to

- decipher disease trajectories and differential diagnosis of neurodegenerative diseases (Disease TreND). *J. Neurosci. Methods* 305, 105–116 (2018).
- [16] Zeng, L. L. et al. Differentiating patients with Parkinson’s disease from normal controls using gray matter in the cerebellum. *Cerebellum* 16, 151–157 (2017).
- [17] Gu, Q. et al. Automatic classification on Multi-Modal MRI data for diagnosis of the postural instability and gait difficulty subtype of Parkinson’s disease. *J. Parkinsons Dis.* **6**, 545–556 (2016).
- [18] Haller, S. et al. Individual detection of patients with Parkinson disease using support vector machine analysis of diffusion tensor imaging data: initial results. *AJNR Am. J. Neuroradiol.* **33**, 2123–2128 (2012).
- [19] Stephan Krohn et al. ,A spatiotemporal complexity architecture of human brain activity.Sci. Adv.9,eabq3851(2023).
- [20] Wu Y, Zhang L, Bhatti UA, Huang M. Interpretable Machine Learning for Personalized Medical Recommendations: A LIME-Based Approach. *Diagnostics (Basel)*. 2023 Aug 15;13(16):2681. doi: 10.3390/diagnostics13162681. PMID: 37627940; PMCID: PMC10453635
- [21] Donick, D., Lera, S.C. Uncovering feature interdependencies in high-noise environments with stepwise lookahead decision forests. *Sci Rep* 11, 9238 (2021).
- [22] Guido, R.; Ferrisi, S.; Lofaro, D.; Conforti, D. An Overview on the Advancements of Support Vector Machine Models in Healthcare Applications: A Review.*Information* 2024, 15, 235.
- [23] QI Qiaona, LIU Yan, CHEN Jihui, LIU Xinzhu, YANG Rui, ZHANG Jinyuan, CUI Mengxuan, XIE Yimeng, WANG Zeyuan, YU Ze, GAO Fei, ZHANG Jian. Research progress on machine learning XGBoost algorithm in medicine[J]. *Journal of Molecular Imaging*,2021, 44(5): 856-862.
- [24] G. Niso, C. Rogers, J.T. Moreau, L-Y. Chen, C. Madjar, S. Das, E. Bock, F. Tadel, A.C. Evans, P. Jolicoeur, S. Baillet, OMEGA: The Open MEG Archive, *NeuroImage*, 124B(1) Jan 2016, pp 1182-1187, ISSN 1053-8119.(04):265-266+298.
- [25] Tadel F, Baillet S, Mosher JC, Pantazis D, Leahy RM (2011) Brainstorm: A User-Friendly Application for MEG/EEG Analysis Computational Intelligence and Neuroscience, vol. 2011, ID 879716.
- [26] P. Welch, "The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," in *IEEE Transactions on Audio and Electroacoustics*, vol. 15, no. 2, pp. 70-73, June 1967.

- [27] Coffman, B.A., Salisbury, D.F. (2020). MEG Methods: A Primer of Basic MEG Analysis. In: Kubicki, M., Shenton, M. (eds) Neuroimaging in Schizophrenia . Springer, Cham.
- [28] Uusitalo MA, Ilmoniemi RJ. Signal-space projection method for separating MEG or EEG into components. *Med Biol Eng Comput.* 1997 Mar;35(2):135-40. doi: 10.1007/BF02534144. PMID: 9136207.
- [29] Ye C, Huang J, Liang L, Yan Z, Qi Z, Kang X, Liu Z, Dong H, Lv H, Ma T, Lu J. Coupling of brain activity and structural network in multiple sclerosis: A graph frequency analysis study. *J Neurosci Res.* 2022 May;100(5):1226-1238. doi: 10.1002/jnr.25028. Epub 2022 Feb 20. PMID: 35184336.

三、指导教师意见

该组同学以 MEG 脑活动时空动态表征模式为研究对象，研究如何利用信息理论复杂性分析来刻画人脑静息状态下功能活动的动态变异性，解析帕金森病患者的脑功能信号的可变性特征，从全脑功能活动时空复杂性分析的角度揭示帕金森病的脑机制；最终实现基于 MEG 的个体化 PD 辅助诊断 AI 模型，并验证模型性能。立项目标清楚，研究路线合理，具有一定的科学意义，同意立项。

签 名：

年 月 日

四、评审专家组意见

批准经费： 元 组长签名：

年 月 日