机器学习驱动的帕金森病静息态脑电图检测：一项多中心研究

Anna Kurbatskaya[1], Alberto Jaramillo-Jimenez[2,3,4]John Fredy Ochoa-Gomez[5], Kolbjørn Brønnick[2,3] 和 Alvaro Fernandez-Quilez[1,4]

摘要— 已有研究表明，静息态脑电图（静息态脑电图）有助于帕金森病（帕金森病）诊断。尤其是，低频频带（δ 和 θ）与高频频带（α 和 β）的功率谱密度（PSD）在帕金森病患者与非帕金森病受试者（非帕金森病）之间存在显著差异。然而，静息态脑电图特征提取及其解读可能耗时较长，并且容易受到检查者差异的影响。机器学习（ML）有望实现静息态脑电图记录分析的自动化，并为临床医生提供支持工具，从而减轻其工作负担。在本研究中，我们汇总了来自不同中心、四个数据集中的84名帕金森病患者和85名非帕金森病受试者的静息态脑电图记录。我们提出了一套端到端流程，包括预处理、从经临床验证的频带中提取功率谱密度特征，以及在通过机器学习算法评估这些特征区分帕金森病与非帕金森病受试者的分类能力之前进行特征选择。此外，鉴于数据集具有多中心性质，我们还评估了特征协调的影响。验证结果显示，平均而言，在进行特征协调和单变量特征选择（k = **202** 特征）后，逻辑回归的帕金森病检测能力有所提高（准确率由69.6%提升至75.5%）。最终结果显示，平均全局准确率为72.2%；研究纳入的所有中心的平衡准确率分别为60.6%、68.7%、77.7%和82.2%。

临床相关性—我们提出了一种端到端流程，用于从静息态脑电图记录中提取具有临床相关性的特征，从而促进帕金森病的分析与检测。此外，我们提供了一种机器学习系统，该系统在检测帕金森病方面表现良好，即使存在采用不同采集协议的中心也是如此。我们的结果表明，统一特征具有重要意义，并为未来聚焦于静息态脑电图分析和多中心数据的研究提供了良好起点。

一、引言NTRODUCTION

帕金森病（PD）的临床特征包括运动症状，如运动迟缓、强直和静止性震颤，以及非运动症状，如认知缺陷 [1]。从病理学角度看，PD 的主要特征是黑质纹状体多巴胺能神经元的丧失，其原因是 α-突触核蛋白聚集 [2]。在 PD 早期阶段作出准确诊断，可提高神经保护性疾病修饰治疗的成功率。然而，基于临床检查的诊断准确率并不理想[3], , 尤其是在疾病早期，此时对多巴胺能治疗的反应较弱 [4]。另一方面，目前尚无法通过正电子发射断层扫描（PET）无创测量 α-突触核蛋白聚集；而使用单光子发射计算机断层扫描（SPECT）定量测量黑质纹状体多巴胺能神经元，成本高昂且因辐射剂量而受到限制 [5]。

定量脑电图（qEEG）是一种便捷、低成本且无创的技术，可用于分析大脑的电活动。一些定量脑电图特征已被证实是帕金森病的潜在生物标志物 [6]。例如，研究表明，δ 和 ϑ 频带的功率谱密度（PSD）增加，以及 α 和 β 频带的PSD降低，通常与帕金森病的进展相关 [7], [8]。为了从脑电信号中提取关键特征，必须应用先进的信号处理技术。此外，自动化诊断系统的应用有望克服脑电图目视分析的局限性 [9], , 而后者仍然是临床解读的金标准。

在过去几年中，多项研究旨在将定量脑电图与机器学习（ML）技术相结合，以实现帕金森病的自动检测。多种机器学习方法已在 [10], 中进行了测试，除α/ϑ 比值外，还采用ϑ、 δ、 α和β 频带中的相对功率、中位频率和峰值频率作为特征，以区分帕金森病受试者与非帕金森病受试者。此外，采用LASSO的惩罚逻辑回归（LR）用于获得表现最佳的特征子集。在[11], 中，还开展了利用高阶谱特征进行帕金森病自动诊断的研究，其中，使用t 值进行特征选择，对获得的特征进行排序，并将高度相关的特征作为多个分类器的输入。

在应用机器学习的脑电图研究中存在若干陷阱[12]。近期发表的大多数研究并未明确说明采用的是受试者级分类还是按记录分类。此外，由于训练数据与测试数据之间的交叉污染，所报告的性能指标通常过于乐观 [13]。此外，许多研究重复使用了先前发表的数据集来开发新模型，例如，相同的数据集一直在不同研究团队之间流传并被使用。该

*本研究未获得任何组织的资助 [1]Anna Kurbatskaya 和 Alvaro Fernandez-Quilez 就职于挪威斯塔万格斯塔万格大学科学技术学院电气工程与计算机科学系。电子邮箱：anna.kurbatskaya@uis.no 和 alvaro.f.quilez@uis.no [2,3] Alberto Jaramillo-Jimenez 和 Kolbjørn Brønnick 就职于年龄相关医学中心（SESAM）、斯塔万格大学医院精神病学系，以及挪威斯塔万格斯塔万格大学健康科学学院公共卫生系。电子邮箱：alberto.jaramilloj@udea.edu.co 和 kolbjorn.s.bronnick@uis.no [4] Alberto Jaramillo-Jimenez 就职于安蒂奥基亚大学医学院安蒂奥基亚神经科学小组。哥伦比亚麦德林。[5] John Fredy Ochoa-Gomez 就职于哥伦比亚麦德林安蒂奥基亚大学医学院神经心理学与行为小组。电子邮箱：john.ochoa@udea.edu.co [6] Alvaro Fernandez-Quilez 就职于斯塔万格大学医院斯塔万格医学影像实验室（SMIL）放射学系。

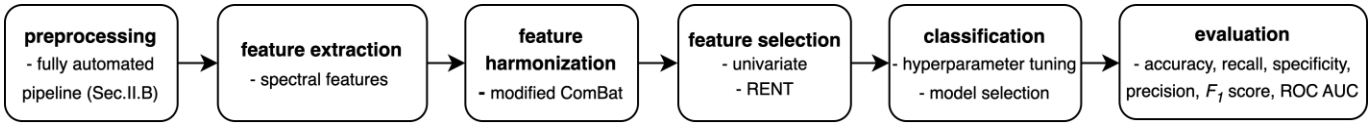

图 1：由预处理、特征提取、特征协调、特征选择、分类和评估组成的端到端流程。

反复使用相同的数据集而不进行外部评估（即使用其他数据集进行评估），会限制结果的可重复性。与此同时，采用由不同设备在不同研究中心针对不同人群采集的数据集，有助于提高泛化能力，并提供更可靠的性能。据我们所知，近期研究中只有两项同时使用了两个不同的数据集 [14], [15]。

据报道，即使对所有记录都采用标准预处理流程，将来自多个数据集的脑电图记录合并，也会导致组水平分析中的第一类错误率上升 [16], [17]。为克服多中心数据汇集的局限性，遗传学和脑成像领域中采用的解决方案已被广泛用于对从原始信号中提取的特征进行统计建模 [18], [19]。然而，脑电图分析中协调方法的应用还很少见 [17], [20]。

本研究旨在利用从不同研究中心采集的静息态脑电图（rs-EEG）记录，应用机器学习技术对帕金森病患者和非帕金森病受试者进行分类。我们提出了一个端到端流程（图1），包括预处理、特征提取、特征协调、特征选择、分类和基于验证结果的评估（代码可在 GitHub 上获取）。我们还考察了数据协调和不同特征选择技术如何影响分类性能，包括将所有中心的受试者合并进行分类（全局分类）以及分别对各中心的受试者进行分类。

II. 方ATERIALS 法与ETHODS材料

A. 数据

本研究共纳入169名受试者的静息态脑电图记录（84名帕金森病患者和85名非帕金森病受试者）。这些记录来自三个国家四个不同中心开展的四项横断面研究：哥伦比亚（麦德林）、芬兰（图尔库）和美国（爱荷华城与圣迭戈）。表 I 列出了各数据集的部分相关研究详情。

表 I：各纳入数据集的相关研究详情。

	受试者	诊断标准	匹配
爱荷华城 [14]	14例帕金森病，14 名非帕金森病患者	英国脑库 [21]	年龄、性别、教育程度相当
麦德林 [22]	36 例帕金森病，36 名非帕金森病患者	MDS [23]´	年龄、性别、ȳ受教育年限
圣迭戈 [24]	15例帕金森病，16 非帕金森病	未说明	年龄、性别，认知表现，利手
图尔库 [25]	19例帕金森病，19 名非帕金森病患者	英国脑库 [21] 和 MDS [23]	age

麦德林、圣迭戈和图尔库数据集的脑电图记录是在闭眼状态下采集的，而爱荷华城数据集的脑电图记录是在睁眼状态下采集的。此外，所有帕金森病受试者的爱荷华城、麦德林和圣迭戈数据集中的受试者在记录期间处于左旋多巴治疗的 ON 期。图尔库数据集中，13 名帕金森病受试者在 OFF 期接受评估，其余受试者则处于 ON 期。所有纳入分析的数据集共有29个通道：AF3、AF4、C3、C4、CP1、CP2、CP5、CP6、Cz、F3、F4、F7、F8、FC1、FC2、FC5、FC6、Fp1、Fp2、Fz、O1、O2、Oz、P3、P4、P7、P8、T7 和 T8。

B. 脑电图预处理

所有记录均转换为脑影像数据结构（BIDS）[26] ，并使用以 Python 实现的处理流程进行预处理，该流程此前已由 [27]验证。该流程包括稳健平均重参考、自适应电源线噪声相关性处理，以及通过以 Python 实现的 [28] PREP处理流程检测并插值噪声通道；该流程最初基于 MATLAB 实现 [29]。后续步骤包括在 1 Hz 下进行有限脉冲响应（FIR）高通滤波，以消除低频漂移；以及采用小波独立成分分析（ICA）抑制肌肉、眨眼和眼动伪迹 [30]。随后，对信号进行 30 Hz 的 FIR 低通滤波，并分割为长度为 5 秒的时段。对于仍受伪迹影响的时段，则依据一套流程予以剔除，该流程评估信号参数（例如极端振幅和谱功率值）以及统计特征（例如线性趋势、联合概率和峰度）。纳入分析的最终时段数根据每个数据集中所有受试者共有的时段数确定。爱荷华城、麦德林、圣迭戈和图尔库数据集的可用时段数分别为 17、44、28 和 19。

C. 分类框架

— – –___

1) 特征提取：特征提取时，将频谱划分为临床研究推荐的四个频段：δ [1 4 Hz)、ϑ [4 8 Hz)、α [8 13 Hz) 和 β [13 30 Hz) [6]。此外，ϑ 频段被进一步划分为慢-ϑ [4 5.5 Hz)和快-ϑ（又称“前-α”）子频段[5.5 8 Hz) [31]。对于每个时间段，通过将某一频率（子）频段内的 PSD 除以 1–30 Hz 范围内的总信号功率，计算功率谱密度值。PSD 使用多窗谱估计法 [32]计算数值。此外，我们使用相对 PSD α 和 ϑ 值 [33]计算了相对 α/ϑ PSD比值。最后，我们对所有时段的 PSD 功率谱密度值取平均，从而为给定通道的每个频谱特征得到一个代表值 [34], [35]。总计为每位受试者提取了 203 个特征（7 个特征 × 29 个通道）。

2）特征协调：按照固定频带谱分析的常规流程，对所有平均相对功率谱密度值进行对数变换。

表 II：全局验证性能（均值 [95% 置信区间]），未进行特征选择。

	模型	准确率,%	召回率,%	特异度,%	精确率,%	F_1,%	ROC曲线下面积
未协调	LR	69.6 [55.5 83.7]	64.7 [43.0 86.4]	74.9 [53.3 96.4]	72.9 [56.3 89.5]	67.3 [49.9 84.7]	0.70 [0.56 0.84]
	SVM	65.4 [50.9 79.8]	58.2 [31.5 84.9]	72.9 [57.7 88.0]	67.6 [55.2 80.0]	61.2 [42.0 80.3]	0.66 [0.51 0.80]
协调	LR	73.0 [59.2 86.9]	69.7 [45.8 93.6]	76.4 [69.0 83.7]	73.1 [60.6 85.6]	70.7 [51.4 90.1]	0.73 [59.2 86.9]
	SVM	73.0 [60.7 85.3]	69.7 [56.1 83.3]	76.2 [58.3 94.1]	75.8 [62.3 89.2]	72.1 [59.8 84.4]	0.73 [0.61 0.85]

表 III：特征选择下的全局验证性能（均值 [95% 置信区间]）。m 是单变量特征选择中要选取的顶级特征数量；τ 是 RENT 的选择标准。

	模型	特征选择	准确率，%	召回率，%	特异度，%	精确率，%	F_1,%	ROC曲线下面积
未协调	LR	单变量（m = 68）	73.7 [65.3 82.2]	77.9 [72.3 83.5]	69.2 [52.8 85.7]	73.2 [61.8 84.6]	75.1 [69.0 81.2]	0.74 [0.65 0.82]
		RENT（τ = 0.26）	72.2 [58.0 86.4]	70.0 [41.8 98.2]	75.0 [58.9 91.0]	74.0 [54.2 93.7]	70.0 [50.8 89.3]	0.73 [58.0 87.0]
	SVM	单变量（m = 65）	73.8 [66.2 81.3]	72.9 [56.4 89.4]	74.8 [57.8 91.9]	76.3 [64.5 88.1]	73.2 [64.7 81.6]	0.74 [0.67 0.81]
		RENT (τ = 0.12)	70.4 [57.0 83.9]	63.2 [35.2 91.2]	78.3 [62.8 93.8]	76.6 [58.4 94.8]	66.2 [47.3 85.2	0.71 [0.57 0.84]
协调	LR	单变量（m = 202）	**75.5 [66.3 84.7]**	**73.0 [56.7 89.3]**	**78.0 [70.0 86.1]**	**76.7 [68.9 84.4]**	**74.3 [62.6 86.0]**	**0.76 [0.66 0.85]**
		RENT (τ = 1.0)	69.7 [48.4 91.0]	66.4 [34.5 98.2]	73.0 [61.3 84.7]	67.4 [42.1 92.7]	66.5 [37.3 95.8]	0.70 [0.48 0.91
	SVM	单变量 (m = 193)	75.5 [65.4 85.5]	69.7 [53.2 86.2]	81.4 [71.3 91.5]	79.1 [70.2 87.9]	73.5 [61.4 85.7]	0.76 [0.66 0.85]
		RENT (τ = 1。0)	69.7 [52.8 86.5]	63.0 [33.6 92.4]	76.2 [58.3 94.1]	70.8 [52.1 89.4]	65.1 [38.4 91.8]	0.70 [0.53 0.86]

这些对数变换后的特征后续称为未协调化特征。随后，使用改进的 ComBat 协调模型 [18] 对这些对数变换后的特征进行协调，设置 1000 次自助法重复，并将性别、年龄和诊断（帕金森病和非帕金森病）作为协变量。麦德林数据集因其是规模最大的单个数据集，被用作参考中心。“批次”变量包含四个水平，分别对应每个数据集。

3）分类器：为进行按受试者分类，训练并比较了多种分类器，包括 LR、k近邻（k-NN）算法、决策树（DT）和支持向量机（SVM），比较指标为其验证准确率。

为验证、训练和测试这些算法，数据按中心和诊断属性分层，并以 70/30 的比例划分为训练/测试子集。为避免交叉污染，每位受试者的特征仅归入训练子集或测试子集之一。随后，使用训练子集，通过包含两个 5 折交叉验证（CV）循环的嵌套交叉验证（CV）进行超参数调优和模型选择。嵌套交叉验证的内层循环采用网格搜索方法，在整个搜索空间内进行超参数调优。超参数列表及其取值见所提供的代码。外层循环用于估计最终分类器的泛化性能。

4）性能评估：我们从准确率、召回率、特异性、精确率、F_1 分数以及受试者工作特征（ROC）曲线下面积（曲线下面积）等方面评估分类器在验证集和测试集上的表现。根据所有中心的平均全局验证准确率，选出最佳分类器。除报告各项指标的均值外，我们还报告了 95%置信区间（CI），以表示估计值的精确度 [36]。为获得最终的测试性能指标并估计性能的不确定性，我们对测试子集进行有放回自助抽样，抽样大小与原始测试分区相同，共迭代100次。

5）特征选择：鉴于数据的高维特性（每名受试者有203个特征），我们分别测试了两种特征选择方法，以确定能够提升验证准确率的特征子集对验证准确率产生积极影响的特征。

第一种方法是基于方差分析F 值的单变量特征选择，并将要选取的排名靠前特征数量作为参数。第二种方法是重复弹性网络技术（RENT），该方法依赖于广义线性模型集成，并使用三个标准评估各模型中特征的权重分布，将其作为参数 [37]。为降低过拟合导致的偏差，我们在嵌套交叉验证的外层循环中对每一折执行特征选择 [38]，，随后在内层循环中使用选定的特征进行超参数调优。两种特征选择方法均针对不同的参数值重复执行。由于每次迭代选出的特征集各不相同，我们对选定的特征集进行了合并，并去除了重复项。最后，根据验证性能确定的预测模型使用剩余特征在测试子集上进行测试。

III. 结果与讨论

表II展示了两种表现最佳的分类器——逻辑回归和支持向量机——在未进行特征选择时的全局验证性能（均值 [95% 置信区间]）。与未协调化特征相比，两种模型使用协调化特征时，各项指标的表现均更好。

表 III 展示了采用特征选择后的分类结果。可以看出，与未进行特征选择相比，对于逻辑回归和支持向量机，采用单变量特征选择后，无论使用未协调化特征还是协调化特征，各项指标的表现均有所改善；唯一例外是逻辑回归的特异度（74.9% vs. 69.2%）。另一方面，采用 RENT 时，未协调化特征的分类结果倾向于改善，但协调化特征的结果则倾向于变差。

总体而言，全局验证性能提升并不总意味着每个单独中心都能取得良好表现。协调化和特征选择带来的结果因中心而异。尽管协调化的全局准确率显示出良好前景，仍需在研究中纳入更多中心的数据集，以验证最终结果。因此，如果新的中心希望在研究中实施所提出的处理流程，应考虑其他方法，以确保其特定人群的个体表现良好。

表 IV：采用单变量特征选择方法，基于协调化特征训练的逻辑回归所取得的测试结果（均值 [95% 置信区间]）。

	准确率，%	召回率，%	特异度，%	精确率，%	F_1，%	AUC
全局	**72.2 [64.5 80.0]**	**68.4 [56.1 80.6]**	**76.1 [68.0 84.2]**	**73.2 [63.0 83.3]**	**70.3 [60.9 79.7]**	**0.72 [0.65 0.80]**
爱荷华城	60.6 [40.6 80.7]	36.1 [14.9 57.3]	100 [100 100]	100 [100 100]	51.1 [30.9 71.3]	0.59 [0.46 0.73]
麦德林	68.7 [55.8 81.6]	73.5 [56.5 90.4]	63.9 [45.8 81.9]	67.5 [50.2 84.7]	69.5 [55.1 84.0]	0.69 [0.56 0.82]
圣迭戈	77.7 [61.0 94.3]	74.5 [48.5 100]	81.6 [59.0 104]	78.8 [53.4 104]	74.2 [53.4 104]	0.78 [0.60 0.95]
图尔库	82.2 [67.8 96.6]	82.1 [62.2 103]	82.5 [62.1 103]	82.6 [62.5 103]	81.3 [64.2 98.5]	0.82 [0.66 0.98]

由于采用单变量特征选择的协调化特征逻辑回归模型取得了最有希望的验证结果，因此将其选用于最终分类评估

（表 IV）。所有中心的平均准确率为 72.2%，其中爱荷华城、麦德林、圣迭戈和图尔库数据集的平均准确率分别为 60.6%、68.7%、77.7% 和 82.2%。值得注意的是

爱荷华城数据集的表现不如所有其他数据集。一种可能的解释是，该数据集规模最小，因此训练和测试都只有有限数量的受试者

训练和测试。

我们工作的部分局限在于，在嵌套交叉验证方法中用于模型选择的测试数据量有限，未来的工作可以对此进行探索。

IV. 结论ONCLUSION

我们的结果表明，协调化能够提升全局分类性能。我们提出的端到端框架可以
支持帕金森病的临床诊断。我们的研究表明，在多中心环境中进行协调化十分重要，同时
受限于仅测试了一种我们认为相关的协调化技术。未来的研究可以测试
不同的协调化方法。

ACKNOWLEDGMENT

我们谨向加州大学圣迭戈分校、图尔库大学和图尔库大学医院、爱荷华大学以及安蒂奥基亚大学的所有研究人员和工作人员表示感谢。我们

我们尤其感谢 A. P. Rockhill 博士、H. Railo 博士和 N. S. Narayanan 博士公开提供他们的数据集。

REFERENCES

[1] O.-B. Tysnes 和 A. Storstein，“帕金森病流行病学”，神经传递杂志，第124卷，第8期，第901–905页，2017年。[2] J. Gratwicke、M. Jahanshahi 和 T. Foltynie，“帕金森病痴呆：神经网络视角”，《脑》，第138卷，第6期，第1454–1476页，2015年。[3] G. Rizzo、M. Copetti、S. Arcuti、D. Martino、A. Fontana 和 G. Logroscino，“帕金森病临床诊断的准确率”，《神经病学》，第86卷，第566–576页，2016年6月。[4] C. H. Adler、T. G. Beach、J. G. Hentz 等，“早期与晚期帕金森病临床诊断准确率较低：临床病理学研究”，《神经病学》，第83卷，第5期，第406–412页，2014年。[5] G. Pagano、F. Niccolini 和 M. Politis，“帕金森病影像学”，《临床医学》，第16卷，第4期，第371页，2016年。[6] C. Babiloni、R. J. Barry、E. Bas¸ar 等，“国际临床神经生理学联合会（IFCN）脑电图研究工作组：关于静息态脑电节律频率和地形图分析的建议。第1部分：临床研究中的应用”，《临床神经生理学》，第131卷，第1期，第285–307页，2020年。[7] M. Neufeld、S. Blumen、I. Aitkin、Y. Parmet 和 A. Korczyn，“痴呆与非痴呆帕金森病患者的脑电图频率分析”，《痴呆与老年认知障碍》，第5卷，第1期，第23–28页，1994年。[8] L. Pezard、R. Jech 和 E. Růžˇicˇka，“帕金森病早期阶段多通道脑电图非线性特性的研究”，《临床神经生理学》，第112卷，第1期，第38–45页，2001年。[9] W. O. Tatum IV，《脑电图解读手册》。Springer Publishing Company，2021年。[10] M. Chaturvedi、F. Hatz、U. Gschwandtner 等，“定量脑电图（QEEG）指标可区分帕金森病（PD）患者与健康对照组（HC）”，《衰老神经科学前沿》，第9卷，2017年1月。[11] R. Yuvaraj、U. Rajendra Acharya 和 Y. Hagiwara，“利用脑电信号高阶谱特征构建新型帕金森病诊断指数”，《神经计算与应用》，第30卷，第4期，第1225–1235页，2018年。

[12] M. de Bardeci、C. T. Ip 和 S. Olbrich，《深度学习在精神障碍脑电图数据中的应用[14]系统综述》，《生物心理学》，第 162 卷，108 117 页，2021年。[13] R. Sugden 和 P. Diamandis，《可泛化的脑电图分类

[15] 帕金森病的深度学习检测》，medRxiv，2022年。M. F. Anjum、S. Dasgupta、R.

Mudumbai、A. Singh、J. F. Cavanagh 和 N. S. Narayanan，《线性预测编码可区分帕金森病的脑电频谱特征》，《帕金森症及相关疾病》，第 79 卷，第 79–85 页，2020年。L. Qiu、J. Li 和 J. Pan，《基于多模式分析和多尺度卷积神经网络的帕金森病检测》，《神经科学前沿》，第 16 卷，2022年。

[16] N. Bigdely-Shamlo、J. Touryan、A. Ojeda、C. Kothe、T. Mullen 和 K. Robbins，《自动化脑电图大规模荟萃分析 I：不同研究中的频谱和振幅特征》，NeuroImage，第 207 卷，第 116 361 页，2020年。[17] M. Li、Y. Wang、C. Lopez-Naranjo 等，《协调的多国 QEEG

[19] 规范（harmnqeeg）》，NeuroImage，第 256 卷，第 119 190 页，2022年。R. Da-Ano、I.

Masson、F. Lucia 等，《面向多中心研究的影像组学特征协调：改良 Combat 方法的性能比较》，Scientific Reports，第 10 卷，第 1 期，第 1–12 页，2020年。J.-P. Fortin、N. Cullen、Y. I. Sheline 等，《跨扫描仪和采集地点协调皮层厚度测量》，NeuroImage，第 167 卷，第 120、2018年 104pp. –。

[20] S. Moguilner、A. Birba、S. Fittipaldi 等，《代表性不足场景中痴呆综合特征的多特征计算框架》，《

[21] 神经工程》，第19卷，第4期，第046 048页，2022年。C. E. Clarke、S. Patel、N. Ives 等，《英国帕金森病协会脑库诊断标准》，2016年。

[22] J. A. Carmona Arroyave、C. A. Tobon Quintero、J. J. Suarez Revelo 等，《帕金森病中的静息态功能连接与轻度认知障碍：一项脑电图研究》，《未来神经病学》，第14卷，第2期，

[23] FNL18, 2019. R. B. Postuma, D. Berg, M. Stern, 等，“帕金森病的MDS临床诊断标准”，《运动障碍》，第30卷，第12期，第1591–1601页，2015年。

[24] A. P. Rockhill, N. Jackson, J. George, A. Aron, 和 N. C. Swann，“加州大学圣迭戈分校帕金森病患者静息态脑电图数据”，2021年。[25] H. Railo, I. Suuronen, V. Kaasinen, M. Murtojärvi, T. Pahikkala, 和 A. Airola，“静息态脑电图作为帕金森病的生物标志物：测量的影响——

[26] ment 条件，” bioRxiv，2020年。C. R. Pernet、S. Appelhoff、K. J. Gorgolewski 等人，“EEG-BIDS：脑电图学脑影像数据结构的扩展”，《科学数据》，第 6 卷，第 1 期，第 1–5 页，2019年。

[27] J. X. Suarez-Revelo、J. F. Ochoa-Gomez 和 C. A. Tobon-Quintero，“通过重测信度验证脑电图预处理流程”，载于《工程应用研讨会论文集》，Springer，2018年，第290–299页。[28] S. Appelhoff、A. J. Hurst、A. Lawrence 等，“PyPREP：用于脑电图数据的预处理流程（PREP）的 Python 实现”，2022年。[29] N. Bigdely-Shamlo、T. Mullen、C. Kothe、K. M. Su 和 K. A. Robbins，“PREP处理流程：大规模脑电图分析的标准化预处理”，《神经信息学前沿》，第9卷，第1–19页，2015年6月。[30] N. P. Castellanos 和 V. A. Makarov，“恢复脑电图脑信号：采用小波增强独立成分分析进行伪迹抑制”，《神经科学方法杂志》，第158卷，第2期，第300–312页，2006年。[31] L. Bonanni、A. Thomas、P. Tiraboschi、B. Perfetti、S. Varanese 和 M. Onofrj，“对早期阿尔茨海默病、路易体痴呆和帕金森病痴呆患者进行脑电图比较，并随访2年”，《脑》，第131卷，第3期，第690–705页，2008年。[32] M. J. Prerau、R. E. Brown、M. T. Bianchi、J. M. Ellenbogen 和 P. L. Purdon，“透过多窗谱分析的视角观察睡眠神经生理动态”，《生理学》，第32卷，第60–92页，2017年1月。[33] A. Jaramillo-Jimenez、J. X. Suarez-Revelo、J. F. Ochoa-Gomez 等，“帕金森病静息态脑电图α/θ比值与神经心理测验表现的相关性”，《临床神经生理学》，第132卷，第756–764页，2021年。[34] G. Rodriguez、F. Copello、P. Vitali、G. Perego 和 F. Nobili，“用于阿尔茨海默病分期的脑电图频谱特征”，《临床神经生理学》，第110卷，第10期，第1831–1837页，1999年。[35] J. Schumacher、J.-P. Taylor、C. A. Hamilton 等，“定量脑电图作为路易体轻度认知障碍的生物标志物”，阿尔茨海默病研究与治疗，第12卷，第1期，第1–12页，2020年。[36] A. Claridge-Chang 和 P. N. Assam，“估计统计学应取代显著性检验”，《自然·方法》，第13卷，第2期，第108–109页，2016年。[37] A. Jenul、S. Schrunner、K. H. Liland、U. G. Indahl、C. M. Futsaether 和 O. Tomic，“Rent——用于特征选择的重复弹性网络技术”，《IEEE开放获取》，第9卷，第152 333–152 346页，2021年。[38] S. Raschka，“机器学习中的模型评估、模型选择与算法选择”，arXiv预印本 arXiv:1811.12808，2018年。