

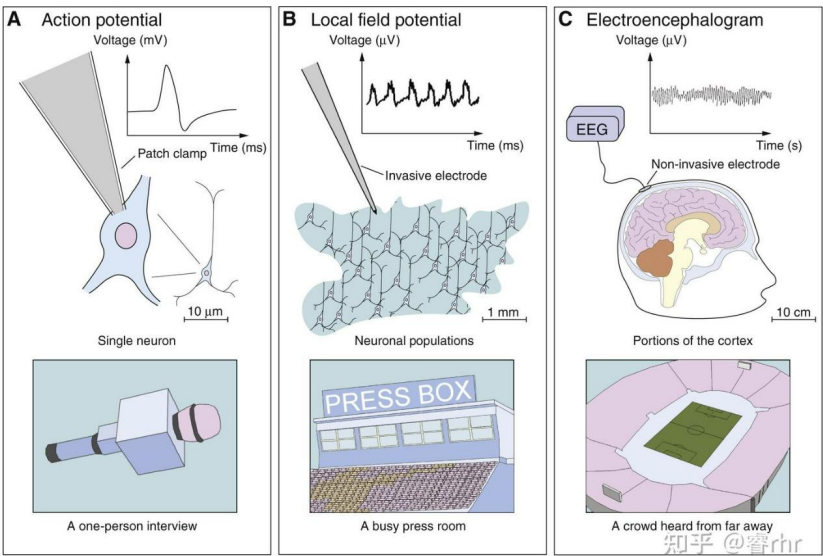
EEG 文本图像生成

1. 背景知识

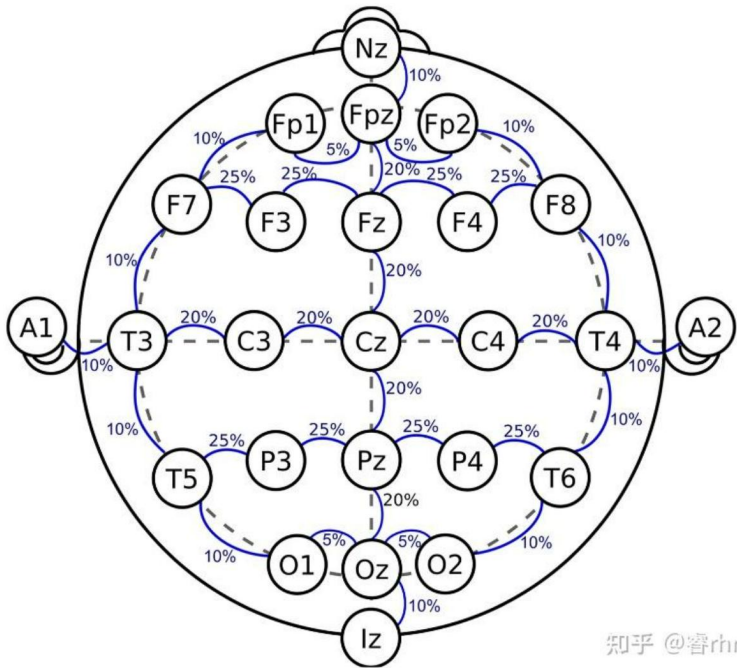
[\(91 封私信 / 82 条消息\) EEG 总结 \(脑电图\) - 知乎](#)

脑电图 (EEG) 是对大脑电场的非侵入性测量。放置在头皮上的电极记录神经元内和周围的电流产生的电压电位。中枢神经系统(大脑)的关键功能是调节全身身体或心理活动参与情感过程。脑电图可以记录中枢神经系统的生理活动。

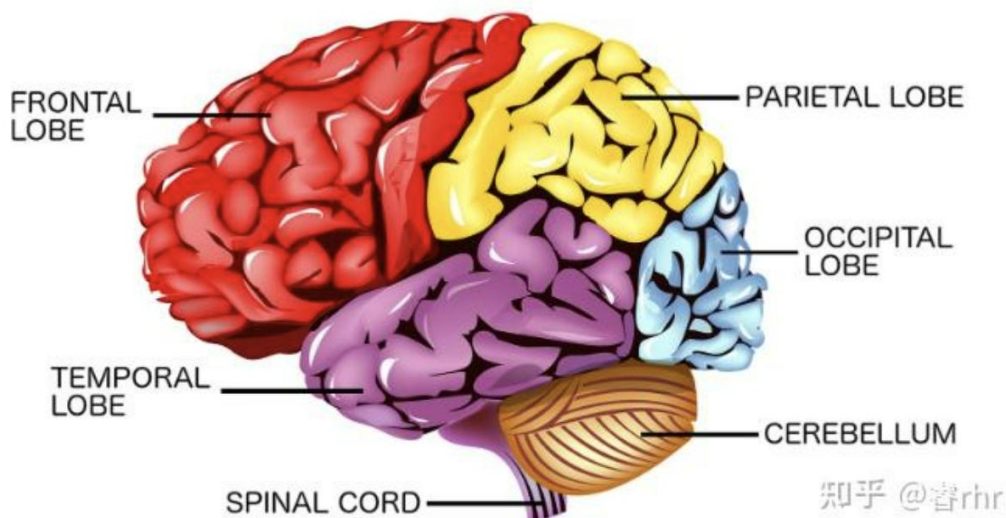
三种 eeg 采集信号



EEG 电极位置



大脑分区



eeg 信号采集相较其他设备来说成本较低，配置较为方便，穿戴式的采集方式又较植入式场景更加广泛，并且穿戴式的脑机接口没有植入排异等安全风险。

然而 eeg 信号从每个神经元发出然后传导到大脑皮层，然后又在大脑皮层经过几个通道记录下来，存在较大的信息流失。头盖骨的厚度较厚，传导到皮层后电信号微弱，由于脑中存在大量的生理活动（如眼动电信号等）还有外部的电信号干扰，造成脑电信号中掺杂了大量的无用信息。

近年来尤其是 2023 年以后，深度学习飞速发展逐步成熟，在 2025 年是人工智能的大落地时代，鉴于深度学习在特征提取上的突出优势，相较传统的信号处理更具优势，能够挖掘出更多的语义。

现在 eeg 信号已被应用于 eeg 情绪分类，语音识别，语音生成，图像识别，图像生成，文本生成等任务，应用广泛但是仍需探索和研究。

2. 研究现状

脑电图（EEG）与生成式人工智能（GenAI）的融合正在推动脑机接口（BCI）技术的革新，实现了从神经信号到图像、文本及语音的多模态生成。这一领域的研究主要围绕 EEG-to-Image、EEG-to-Text 和 EEG-to-Sound/Speech 三大方向展开，涉及生成对抗网络（GANs）、扩散模型、Transformer 架构等多种技术。以下将按场景分类归纳关键研究论文、作者及其方法。

2.1. EEG-to-Image 生成研究

1. 《DreamDiffusion: Generating High-Quality Images from Brain EEG Signals》

作者：Yunpeng Bai, Xintao Wang, Yan-pei Cao, Yixiao Ge, Chun Yuan, Ying Shan

方法：采用扩散模型（Diffusion Model）和掩码信号建模（Masked Signal Modeling），通过 EEG 信号直接生成高质量图像，避免了传统文本中间表示的需要。

2. 《Brain2Image: Converting Brain Signals into Images》

作者：Isaak Kavasidis, Simone Palazzo, Concetto Spampinato, Daniela Giordano, Mubarak Shah

方法：结合 LSTM 提取时序特征，并利用 GANs 和变分自编码器（VAE）生成类别判别性图像。

3. 《NeuroGAN: Image Reconstruction from EEG Signals via an Attention-based GAN》

作者：Rahul Mishra, Krishan Sharma, Ranjeet Ranjan Jha, Arnav Bhavsar

方法：基于注意力机制的 GAN，通过关注关键 EEG 通道提升图像重建质量，生成类特定编码作为潜在表示。

4. 《EEG2Image: Image Reconstruction from EEG Brain Signals》

作者：Prajwal Singh, Pankaj Pandey, Krishna Miyapuram, Shanmuganathan Raman

方法：使用 LSTM 与对比学习结合，在小规模 EEG 数据集上合成图像，并通过条件 GAN 优化生成效果。

5. 《Visual Decoding and Reconstruction via EEG Embeddings with Guided Diffusion》

作者：Dongyang Li, Chen Wei, Shiyong Li, Jiachen Zou, Haoyang Qin, Quanying Liu

方法：采用 Sinc-EEGNet 提取特征，结合引导扩散模型和注意力机制，重构多粒度视觉内容。

2.2. EEG-to-Text 生成研究

1. 《Open Vocabulary Electroencephalography-to-Text Decoding and Zero-Shot Sentiment Classification》

作者：Zhenhailong Wang, Heng Ji

方法：基于 Transformer 架构和 BART 模型，实现开放词汇的 EEG 到文本解码，支持零样本情感分类。

2. 《EEG2Text: Open Vocabulary EEG-to-Text Decoding with EEG Pre-training and Multi-View Transformer》

作者：Hanwen Liu, Daniel Hajialigol, Benny Antony, Aiguo Han, Xuan Wang

方法：通过多视图 Transformer 和掩码信号预训练，学习句子级 EEG 表征，提升开放词汇生成能力。

3. 《Deep Representation Learning for Open Vocabulary Electroencephalography-to-Text Decoding》

作者：Hamza Amrani, Daniela Micucci, Paolo Napoleone

方法：结合双向 GRU 和主题依赖表征学习模块，动态处理变长 EEG 信号，优化文本生成准确性。

4. 《Thought2Text: Text Generation from EEG Signal Using Large Language Models (LLMs)》

作者：Abhijit Mishra, Shreya Shukla, Jose Torres, Jacek Gwizdka, Shounak Roychowdhury

方法：通过视觉刺激规避语言处理复杂性，利用图像-文本跨模态对齐，微调 LLMs 直接从 EEG 生成文本。

5. 《Aligning Semantic in Brain and Language: A Curriculum Contrastive Method for EEG-to-Text Generation》

作者：Xiachong Feng, Xiaocheng Feng, Bing Qin, Ting Liu

方法：采用对比学习和课程学习策略，校准主题依赖的 EEG 表征，提升语义一致性。

3. EEG-to-Sound/Speech 生成研究

1. 《Towards Voice Reconstruction from EEG During Imagined Speech》

作者：Young-Eun Lee, Seo-Hyun Lee, Sang-Ho Kim, Seong-Whan Lee

方法：提出 NeuroTalk 框架，使用 GRU 序列模型生成梅尔频谱图，并通过 HiFi-GAN 声码器转换为语音波形。

2. 《An Investigation on the Speech Recovery from EEG Signals Using Transformer》

作者：Tomoaki Mizuno, Takuya Kishida, Natsue Yoshimura, Toru Nakashika

方法：基于 Transformer 架构，从 EEG 信号中重建连续语音，包括说话人特征。

3. 《Naturalistic Music Decoding from EEG Data via Latent Diffusion Models》

作者：Emilian Postolache, Natalia Polouliakh, Hiroaki Kitano, Akima Connelly, Emanuele Rodola, Luca Cosmo, Taketo Akama

方法：利用 ControlNet 适配器引导 AudioLDM2 扩散模型，从 EEG 解码自然音乐，无需手动预处理。

4. 《EEG-Based Talking-Face Generation for Brain Signal-Driven Dynamic Communication》

作者：Ji-Ha Park, Seo-Hyun Lee, Seong-Whan Lee

方法：结合 NeuroTalk 语音合成与 Wave2Lip 唇同步技术，生成与语音同步的虚拟人脸动画。

5. 《EEG-Driven Automatic Generation of Emotive Music Based on Transformer》

作者：Hui Jiang, Yu Chen, Di Wu, Jinlin Yan

方法：通过 DBSCAN 聚类提取 EEG 时序特征，利用 Transformer 模型生成情感音乐。

3.1. 技术趋势与评估方法总结

1. 共性技术

生成模型：GANs、VAEs、扩散模型广泛用于图像和语音生成。

注意力机制：提升 EEG 通道和频带的重要性加权。

对比学习：用于跨模态对齐（如 EEG-文本、EEG-图像）。

Transformer 架构：在文本和语音生成中处理长序列依赖。

2. 评估指标

图像生成：Inception Score (IS)、Frèchet Inception Distance (FID)、SSIM、LPIPS。

文本生成：BLEU、ROUGE、METEOR、BERTScore。

语音/音乐：Mel Cepstral Distortion (MCD)、Word Error Rate (WER)、Frèchet Audio Distance (FAD)。

4. Our method

[脑电脑网络分析代码使用流程介绍](#)

■ 调研方向—前置研究

深度学习方向

- 1. 直接使用深度学习网络进行跨模态映射

检索: EEG - invasive EEG - mapping
invasive EEG reconstruction from EEG
EEG - 2 - SEEG
deep learning model

数学建模方向

- 1. 使用数理方程构建脑部模型

检索: Mathematical equation -
Head modeling -
iEEG - Reconstruction

■ 调研方向—数理方程

微分方程向

- 1. 可用于头部电流模型计算的微分方程

检索: Differential Equations - Current conduction
Head modeling - iEEG - EEG

神经动力学向

- 1. 可用于头部模型构建的神经动力学方程

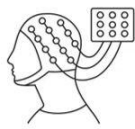
检索: Neural Dynamics - Current conduction -
Head modeling - iEEG - EEG

■ 调研方向—数据集

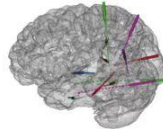
训练向

1. 同时采集的EEG-iEEG数据对 + 对应被试的T1 MRI

检索: dataset - EEG - invasive EEG/SEEG - T1 MRI
Simultaneous collection



EEG



iEEG

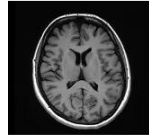
2. 同时采集的EEG-iEEG数据对

检索: dataset - EEG - invasive EEG/SEEG
Simultaneous collection

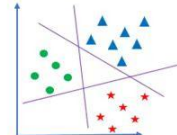
下游任务向

1. 包含 T1 MRI 的EEG多分类任务

检索: dataset - EEG - T1 MRI - Multi classification



T1 MRI



Multi classification

2. 同时采集的EEG-iEEG数据对

检索: dataset - EEG - iEEG
Simultaneous collection

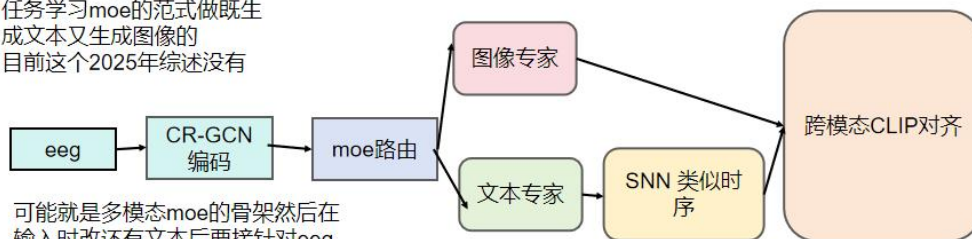
深圳先进技术研究院

1. CausalBrainMoE: 基于因果推理混合专家的脑电大模型及其多模态生成验证 (可解释)
2. 基于 moe 的脑机双路交叉验证 (生成)

我们的方法是以多任务学习的范式, 通过下游生成文本和图像实现互相促进, 当前的研究几乎集中在单分支架构, 忽略了文本图像与脑机的内在联系, 我们的 pipeline 是将 eeg 和文本和图像先接入 eeg 编码器, 然后过 moe 然后下游接 LLM 和 diffusion, 利用混合专家实现语义级解耦, 在末尾加上文本图像环路, 形成交叉验证, 一篇是针对可解释方向做 eeg 脑机可解释, 一篇是针对文本图像生成。

idea_1:

使用多任务的范式尤其是多任务学习moe的范式做既生成文本又生成图像的
目前这个2025年综述没有



可能就是多模态moe的骨架然后在输入时改还有文本后要接针对eeg的架构最后跨模态对齐其实大致应该不用改 CR这个就是脑区的划分失语者的医学报告生成

idea_2:

可解释方案 可以解决eeg输入收到噪声 脑区的输出贡献 也就是图像生成的路径
多模态数据 --> 协变量--> 支持因果分析
目前的baseline CR只有关系建模也就是哪个统计的皮尔逊相关系数但是可以有个因果 失语者的可解释

idea_3:

复杂信号就是从穿戴式采集最大特定就是信号复杂混到一起并且受到眼镜电的影响 所以如果能解耦出来

idea_3:

正常人是交互训练的也就是有监督的方式

5. 可行性分析

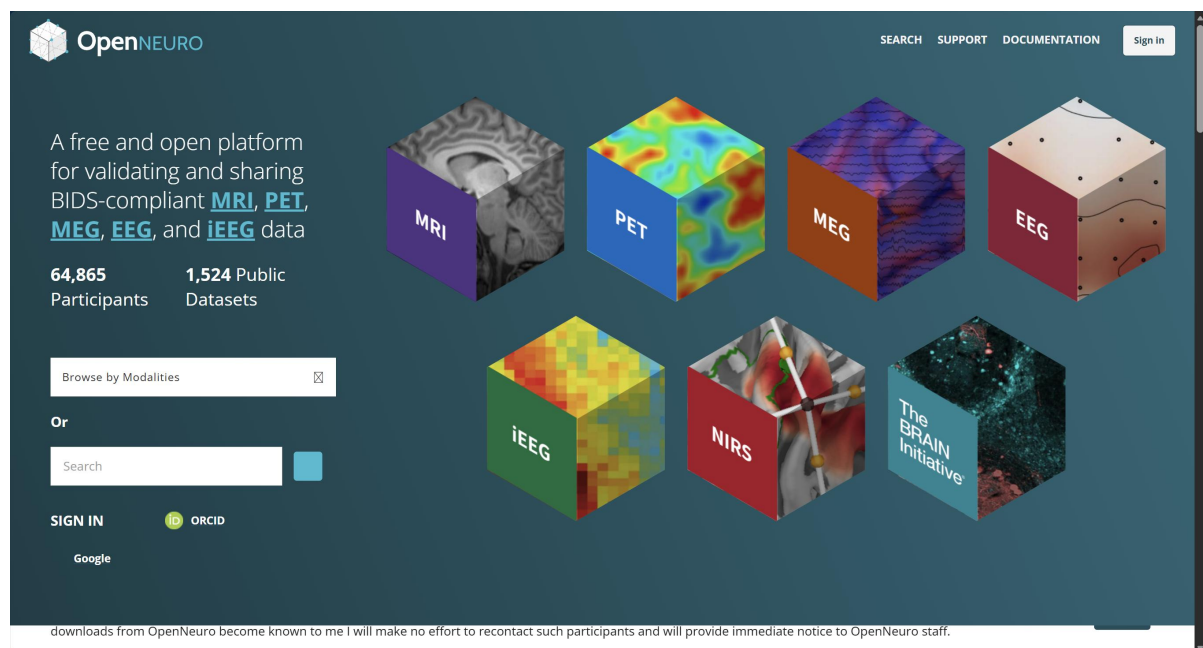
5.1. 算力

当前有充足的算力支持（保密）

5.2. 数据集

公开数据集：

<https://openneuro.org/>





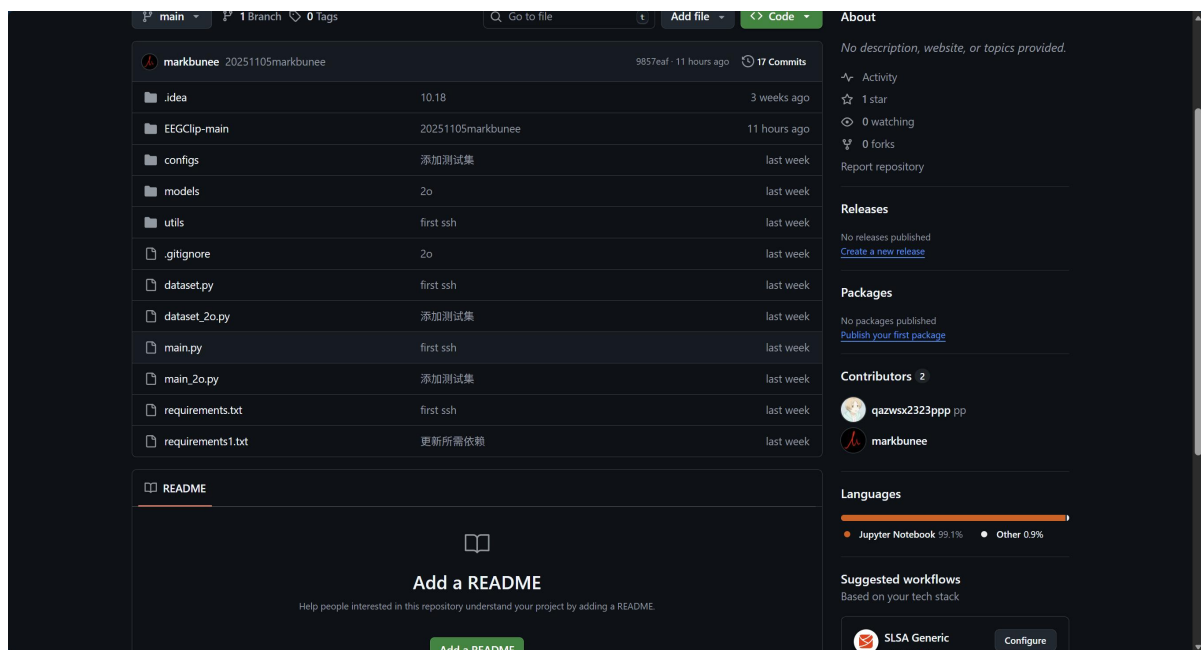
数据集名称	采集模态	受试者数量	受试者时长 / 图片数量	总时长 / 图片数量	任务描述
SMN4Lang 2024	fMRI/MEG	12	6 h	72 h	听觉刺激
MOUS 2019	MEG	96	0.85h	81.6h	听觉刺激
Gwilliams 2022	MEG	27	2h	54h	听觉刺激
Armeni 2022	MEG	3	10h	30h	听觉刺激
NOD 2025	fMRI/MEG/EEG	30	1900pic	57000pic	图像刺激
Things 2022	fMRI/MEG	7	16000pic	112000pic	图像刺激(3个fMRI,4个MEG)
Things 2021	EEG	58	22000pic	1270000pic	图像刺激
BABA 2025	fMRI/MEG	60	0.35h	21h	视频刺激(30个fMRI,30个MEG)



数据集名称	采集模态	受试者数量	总采集时长	任务
SMN4Lang 2024	fMRI/MEG	12	72h	听觉刺激
MOUS 2019	MEG	96	73.5h	听觉刺激
Broderick2018	EEG	19	19h	听觉刺激
SparrKULee 2023	EEG	85	170h	听觉刺激
Gwilliams 2022	MEG	27	54h	听觉刺激
Armeni 2022	MEG	3	30h	听觉刺激
NOD 2025	fMRI/MEG/EEG	30	57120pic	图像刺激
NSD 2021	fMRI	8	73000pic	图像刺激
Things 2022	fMRI/MEG	7	112000pic	图像刺激
BABA 2025	fMRI/MEG	60	21h	视频刺激(连续片段)
Algonauts2021	fMRI	10	9h	视频刺激(多个视频片段)
Rybář 2025	EEG-fNIRS	19	180item	默念单词/想象图片、声音、触觉
Chisco 2025	EEG	5	75h	想象句子
Wilson 2023	EEG	12	8h	想象句子
YOTO 2023	EEG	26	27h	想象画面、听觉、试听结合

6. 项目仓库 (在进行)

https://github.com/qazwsx2323ppp/EEG_encoder



7. 时间规划（20250908--20260101--20260301）

- 在 11 月基本搭建和测试
- 在 12 月整理实验结果并且继续修改
- 在 1 月完成大半开始论文阶段