

一、选题理由

脑机接口 (BCI) 领域中，通过非侵入式脑电信号 (EEG) 解码人类视觉感知并重建图像，是神经科学与人工智能交叉的核心研究方向。EEG 凭借便携性、低成本和高时间分辨率的优势，相比功能磁共振成像 (fMRI) 更具实际应用潜力，但低信噪比、低空间分辨率等特性导致其视觉解码与重建性能长期落后于 fMRI 和脑磁图 (MEG)。

现有研究存在三大核心挑战：一是 EEG 特征与视觉模态特征的跨域对齐难度大，缺乏高效的特征映射机制；二是生成模型难以同时捕捉 EEG 中蕴含的视觉语义信息与低层次结构特征；三是传统微调方式适配 EEG 数据时易过拟合，且计算成本高。

本选题聚焦 EEG 多模态图像生成，以 Stable Diffusion XL (SDXL) 大模型为核心生成引擎，通过特征投影对齐替代微调、多模态特征融合策略，解决上述挑战。SDXL 作为当前领先的文本 - 图像生成大模型，具备强大的高分辨率图像合成能力和灵活的模态扩展接口，其与 EEG 信号的结合不仅能充分发挥大模型的特征学习与生成优势，更能推动 EEG 在实时视觉重建、辅助沟通等 BCI 场景中的落地应用，兼具重要的理论价值与实践意义。

二、相关工作对比

图像生成相关研究

单阶段生成与传统模型方案：现有研究多采用轻量级生成模型或单一模态引导的生成方式，如直接将 EEG 特征输入基础扩散模型，因模型容量有限，难以还原复杂视觉细节。部分工作尝试微调小尺寸生成模型，但存在训练成本高、生成分辨率低（多为 512×512 以下）等问题。

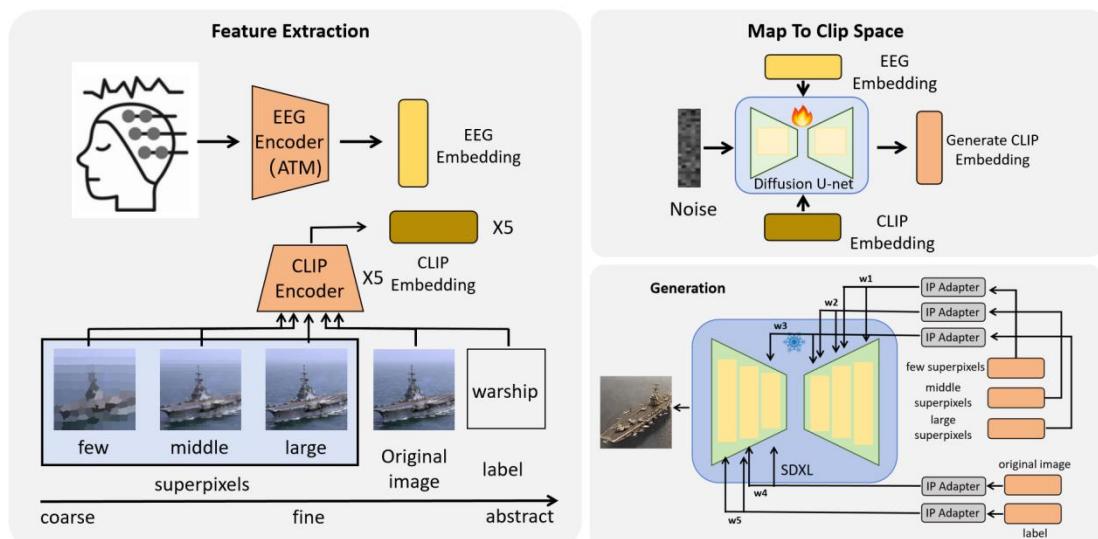
本研究方案：以 SDXL 大模型为核心，构建多模态特征融合生成框架。SDXL 通过 2B 参数量的双 U-Net 架构、改进的 VAE 编码器和高效的文本 - 图像对齐机制，原生支持 1024×1024 高分辨率生成，其跨模态适配能力为 EEG 信号的视觉重建提供了强大基础。本研究通过 IP-Adapter 插件将 EEG 对齐特征、多尺度超像素特征及文本特征注入 SDXL，既保留了大模型的预训练知识，又实现了对 EEG 信号的精准响应，生成质量远超传统小模型方案。

特征适配与大模型应用方式对比

传统微调方式：现有大模型在神经信号生成任务中的应用多依赖全参数微调，需调整 SDXL 等大模型的数百万甚至数十亿参数以适配 EEG 数据，存在计算成本极高、易过拟合、训练周期长等问题。

本研究方案：采用“特征投影对齐 + 插件式注入”策略，避免对 SDXL 大模型进行任何参数微调。通过专用投影层将 ATM 提取的 EEG 特征映射至 SDXL 的 CLIP 特征空间，再通过 IP-Adapter 插件将多模态融合特征注入 SDXL 的 U-Net 结构，引导 denoising 过程。该方式既发挥了 SDXL 大模型的强大生成能力，又降低了训练成本（单 GPU 即可完成），同时避免了过拟合风险，为大模型在小样本、高噪声的神经信号领域的应用提供了高效范式。

三、方法设计



整体框架

系统以 SDXL 大模型为核心生成单元，包含三大核心模块：EEG 特征提取与投影对齐模块、多模态特征融合模块、SDXL 驱动的图像生成模块，整体流程如下：

EEG 信号经 ATM 编码器提取时空特征，通过专用投影层对齐至 SDXL 的 CLIP 预训练特征空间；

从图像数据中提取原始图像的 CLIP 视觉特征、文本描述的 CLIP 文本特征，以及通过 SLIC 算法生成的大 / 中 / 小三个尺度超像素特征，构建多模态特征集；

以 EEG 特征为条件向量，多模态特征集为目标向量，指导 Unet Prior 生成重构的多模态特征。

重构的多模态特征经加权融合后，通过 IP-Adapter 插件分层注入 SDXL 的 U-Net 结构，利用 SDXL 的高分辨率生成能力和跨模态对齐机制，输出目标图像。

四、实验设置

数据集：采用 THINGS-EEG 数据集，包含 10 名被试的 63 通道 EEG 数据（采样率 250Hz）。训练集涵盖 1654 类图像的 4 次重复记录，测试集包含 200 类图像的 80 次重复记录；图像数据分为训练集与测试集，分别用于特征提取与生成结果验证。

评估指标：采用 FID (Frechet Inception Distance)、IS (Inception Score)、LPIPS (Learned Perceptual Image Patch Similarity) 和 SSIM (Structure Similarity Index Measure) 四项指标，全面评估 SDXL 生成图像的质量、多样性与相似度。

实验结果与分析：

表 1. 图像生成指标对比

Method	SSIM↑	LPIPS ↓	FID↓	IS↑
EEG-Conformer	0.125	0.785	135.92	26.27
EEG-Net	0.124	0.788	146.93	22.31
MB2C	-	-	153.37	12.85
BrainVis	-	-	126.66	30.99
ESG-ADA	-	-	174.13	10.82
Ours	0.217	0.684	88.03	31.41



图 1.SDXL-turbo 生成结果



图 2.SDXL-base1.0 生成结果

五、贡献与创新点

大模型应用范式创新：提出“特征投影对齐 + 插件式注入”的大模型适配策略，避免对 SDXL 大模型进行任何参数微调，解决了大模型在 EEG 小样本、高噪声数据中的应用难题，为大模型在神经科学领域的落地提供了高效可行的范式；

多模态融合与大模型协同创新：有效结合图像语义特征、文本描述特征及三个尺度超像素结构特征，通过加权融合与分层注入机制，与 SDXL 大模型的 U-Net 特征提取节奏精准匹配，充分发挥大模型的高分辨率生成能力与跨模态对齐优势；

工程实现优化：基于 SDXL+IP-Adapter 构建轻量化生成框架，无需复杂硬件配置，单 GPU 即可完成全流程训练与推理，同时提供 1024×1024 高分辨率生成结果，为 EEG 图像生成的工程化落地奠定基础。