

深層拡散生成モデルを用いたヒト脳活動からの動画再構成

大阪大学 大学院生命機能研究科 氏名 高木優

目的

自然な状態下における画像・意味・音情報に関する脳情報解析の基盤技術開発

内容

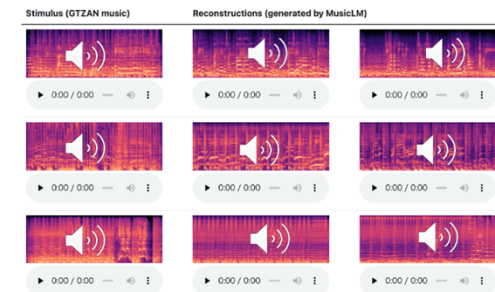
ヒトの活動の多くは視覚的な体験を伴い、その情報は幅広い脳領域で処理されている。ヒト脳内の視覚情報処理を非侵襲的に探るために、これまでの認知神経科学研究では主に機能的核磁気共鳴画像法（fMRI）を用いてきた。具体的には、被験者がfMRI内で画像を観ている際に、機器から得られた信号値を解析することで、脳のどこにどのような視覚情報が表現されているかを明らかにしてきた。一方で、fMRIを用いた先行研究では静止画を視聴中の脳活動解析が主に行われており、より自然な動画画像や音声を視聴中の脳活動を用いた解析は行われてこなかった。

結果

本研究では、次の3つに取り組み、それぞれの結果は以下の通りである。

1. 深層拡散生成モデルを用いた脳活動からの動画画像再構成に取り組んだものの、fMRI脳活動からの動的画像再構成は時間解像度やノイズの観点から難しく、静的な画像に比して大幅に精度が低い再構成となることがわかった。
2. 深層生成モデルを用いた脳活動からの音楽再構成に取り組み、高次の音情報については一定程度的内容再構成に成功した（右図; Denk and Takagi et al., arXiv 2023）。
3. 大規模言語モデルを用いた自然な動画視聴時の脳活動からの意味内容再構成に取り組み、表層的な情報だけでなく、背景のストーリーを一定程度解読可能なことを示した（Nakagi,..., and Takagi, bioRxiv 2024）。

利用した計算機 SQUID 汎用GPUノード群
ノード時間 1,430 時間



図（脳活動からの音楽再構成の結果。一番左が被験者が聴いている音楽のスペクトログラム。右二つが脳活動から再構成したスペクトログラム。）