

▼ 결과 내러티브

▼ Part 1: Zoom-in to a large-scale pre-training with rsfMRI

주요 키워드: scalability (확장성) / emergence (창발성) / generalizability (일반화)

▼ Part 1-1: scalability check

- UKB의 단일 데이터셋을 이용해서 다음의 세가지 요소에 대한 scalability를 확인하는 것이 목적
 - model size scalability
 - data size scalability
 - compute budget scalability
- Held-out set에서의 age/sex/intelligence prediction의 성능들을 종합하여 "error rate"라는 통합 지표를 구성하고, 이 통합 지표를 통해 scalability 확인.
 - AUC와 MSE를 스케일링하여 classification과 regression을 "error rate"라는 단일 지표로 통일
 - 0.5에서 1사이 값을 지니는 AUC는 $2 \times (1 - \text{AUC})$ 를 하여 0~1 사이의 값을 지니는 error rate로 변환 (AUC < 0.5 일 때에는 0.5로 간주)
 - MSE는 Root Mean Square로 변환하고 데이터의 min과 max 값을 이용하여 min-max scaling을 하여 error rate로 변환. 즉, $\text{RMSE}/(\text{max}-\text{min})$ 을 통해 error rate를 구함
- 추가적으로, pre-train loss와 통합 지표 간의 correlation 등을 통해서, pre-train loss가 다양한 downstream task에서의 error rate를 얼마나 잘 반영할 수 있는지에 대한 추가 정보 제공.

▼ Part 1-2: A large-scale pre-training & Exploratory Data Analysis

- part 1-1에서의 scalability 확인 이후, 대규모의 오픈 소스 데이터를 모아서 구성한 life-span 데이터셋(i.e., ABCD, UKB, HCP)을 활용하여 사전 학습

을 진행하고, 사전학습된 모델에서 추출한 임베딩으로 exploratory data analysis를 진행함.

- exploratory data analysis중 사전학습 iteration에 따라 임베딩이 표상하는 바가 어떻게 달라지는지(e.g., sex, age, 뿐만 아니라 scanner/site effect 등 다양한 변수들을 t-sne같은 시각화 방식으로 분석)에 대한 분석을 통해, masked image modeling이라는 self-supervised learning 방식으로 학습을 진행하더라도 다양한 정보들이 담겨있는 표상을 모델이 학습할 수 있게 되는 emergence가 나타남을 보여주는 것이 목적.
- A large-scale pre-training
 - 사전 학습에 사용된 데이터셋들 중에서 held-out test set에서의 sex/age/intelligence의 성능 구하기.
- Exploratory Data Analysis
 - 사전 학습된 모델로부터 얻어낸 임베딩을 시각화하여 다양한 demographic 및 각종 변수(e.g., sex, age, 뿐만 아니라 scanner/site effect 등 다양한 변수들)들이 representation에 어떻게 표상되어 있는지에 대한 분석을 진행하기 -> 사전학습 iteration에 따라서 임베딩이 표상하는 바가 어떻게 달라지는지 변수들마다 보여주어도 좋을 듯
 - 임베딩의 Autocorrelation function을 통해서, 우리 모델이 전체 fMRI를 chunk로 쪼개서 넣는 방식이긴 하더라도 임베딩을 concatenation했을 때 time-series long-term dependency를 잘 표상하고 있다는 것을 보여줌. 즉, 우리 모델이 뇌의 동적 연결성(Dynamic Functional Connectivity)의 시간적 패턴을 잘 학습한다는 것을 보여줌. -> group-level에서 어떻게 수치화하고 시각화할지에 대한 고민 필요. (Resting-state fMRI는 데이터의 특성상 subject-specific characteristics들을 담고 있기 때문에 group-level analysis를 하게 되면 필연적으로 가장 중요한 특성을 살펴볼 수 없음. Subject-wise raw data-embedding act correlation 혹은 error를 구한다음에 이 수치를 group-level로 발전시키는 것이 하나의 접근이 될 수 있음.)

▼ Part 2: 다양한 site의 각종 independent 데이터셋에서의 downstream task

주요 키워드: Performance (성능) / **Robustness (견고성)** / **Generalization (일반화)**

- ROI-based fMRI foundation model (i.e., BrainLM and Brain-JEPA)와 4D fMRI foundation model (i.e., NeuroSTORM, SwiFT-v2)의 성능을 다

양한 downstream task에서 comprehensive 하게 비교.

- 기존의 fMRI foundation model들이 generalist model의 가능성을 보여주기는 했지만, 다양한 도메인의 independent dataset에 대한 comprehensive evaluation을 진행하지 않았다. 우리는 최초로 "Unified Brain Representation(통합된 뇌 표상)"을 구축했다. 이를 증명하기 위해 Neurodegenerative, Psychiatric, Cognitive, Functional의 4대 축(Pillars)을 설정하고 모두 검증했다. 특히, fMRI 연구의 특성상 가용한 데이터셋의 숫자가 매우 제한적인 상황이 많은데, 이런 상황에서 모델이 활용될 수 있는지에 초점을 맞추어서 검증함으로써 모델의 utility를 검증하고자 하였다.
- Downstream Task 분류
 - Neurodegenerative Domain
 - Mild Cognitive Impairment (MCI) to Alzheimer's Disease (AD) conversion prediction → ADNI 데이터셋
 - Psychiatric Domain
 - Obsessive-Compulsive Disorder (OCD) diagnosis → SNUH-OCD 데이터셋
 - Major Depressive Disorder (MDD) diagnosis → 김재원 교수님 데이터셋
 - Anti-depressants Response (OCD patient, MDD patient) → SNUH-OCD, EMBARC 데이터셋
 - Cognitive Domain
 - Attention → YooAttn 데이터셋
 - gradual-onset continuous performance task (gradCPT)
 - multiple object tracking (MOT)
 - visual short-term memory (VSTM)
 - Cognitive tasks
 - Word Order: Episodic Memory
 - Letter Sets: Fluid Reasoning
 - Digit Symbols: Perceptual Speed
 - Synonyms: Vocabulary
 - Functional Domain

- Pain-evoked state prediction → ToPS 데이터셋

▼ Part 3: Independent 데이터셋을 활용한 post-hoc analysis

주요 키워드: Post-hoc analysis / Biomarker

- part 2가 다양한 시나리오 및 domain에서 SwiFT-v2의 성능을 수치적으로 보여줬다면, part 3에서는 SwiFT-v2의 연구적인 implication을 보여주는 것이 목적
- 각 independent 데이터셋의 downstream task 중에서, 테스트셋에서 성능이 잘 나온 test split의 결과만을 사용해서 분석. → post-hoc 분석이기 때문에 모든 데이터셋을 사용하는 것이 오히려 부적절. 다만, "성능을 과시하기 위해서"가 아니라 "성공적인 예측이 일어날 때(When the model works), 뇌의 어떤 정보를 쓰는지 규명하기 위해서" 특정 Split을 사용했다고 명분(Justification)을 확실히 해야 함.
- Anatomical validity: In-silico Functional Lesioning
 - 주요 Brain Network(DMN, Visual, Salience 등)를 하나씩 지웠을 때 성능이 얼마나 떨어지는지 측정.
 - 모델이 질환의 핵심 병변(Core pathology)을 정확히 보고 판단함을 증명하는 것이 목적.
 - "모델이 발견한 Subtype이 기존 뇌과학 문헌에서 제시했던 가설과 일치한다."는 것을 잘 보여주어야 함.
 - 적용 Task:
 - MCI to AD Conversion prediction: 기억 관련 회로(DMN, Limbic)의 중요도가 높게 나와야 함.
 - OCD diagnosis: 인지 조절 회로(Frontoparietal, Salience)의 중요도가 높게 나와야 함.
 - MDD diagnosis
- Clinical Validity: Representational Similarity Analysis (RSA) & Latent Space Visualization
 - Linear Head 통과 전, 임베딩 벡터 공간에서 환자군(반응군 vs 비반응군)의 군집화 정도 분석.
 - 모델이 단순 진단을 넘어, "치료에 반응하는 뇌의 생물학적 특징(Biotype)"을 내재적으로 학습했음을 증명함으로써 정밀 의료(Precision

Medicine) 가능성 시사.

◦ 적용 task:

- Anti-depressant Response (MDD & OCD): 치료 전(Baseline) fMRI 만으로도 치료 예후가 좋은 환자와 나쁜 환자가 임베딩 공간에서 분리됨을 시각화.