Towards A General-purpose Foundation Model for fMRI analysis

Poster No:

3818

Submission Type:

Late-Breaking Abstract Submission

Authors:

Cheng Wang[1], Yu Jiang[1], Zhihao Peng[1], Chenxin Li[1], Changbei Bang[2], Carl Yang[3], Lifang He[4], Daniel Barron[5], Randy Hirschtick[6], Byung-Hoon Kim[2], Xiang Li[6], Yixuan Yuan[1]

Institutions:

[1]The Chinese University of Hong Kong, Hong Kong, Hong Kong, [2]Yonsei University, Seoul, Seoul, [3]Emory University, Atlanta, GA, [4]Lehigh University, Bethlehem, PA, [5]Brigham and Women's Hospital, Boston, MA, [6]Massachusetts General Hospital, Boston, MA

First Author:

Cheng Wang
The Chinese University of Hong Kong
Hong Kong, Hong Kong

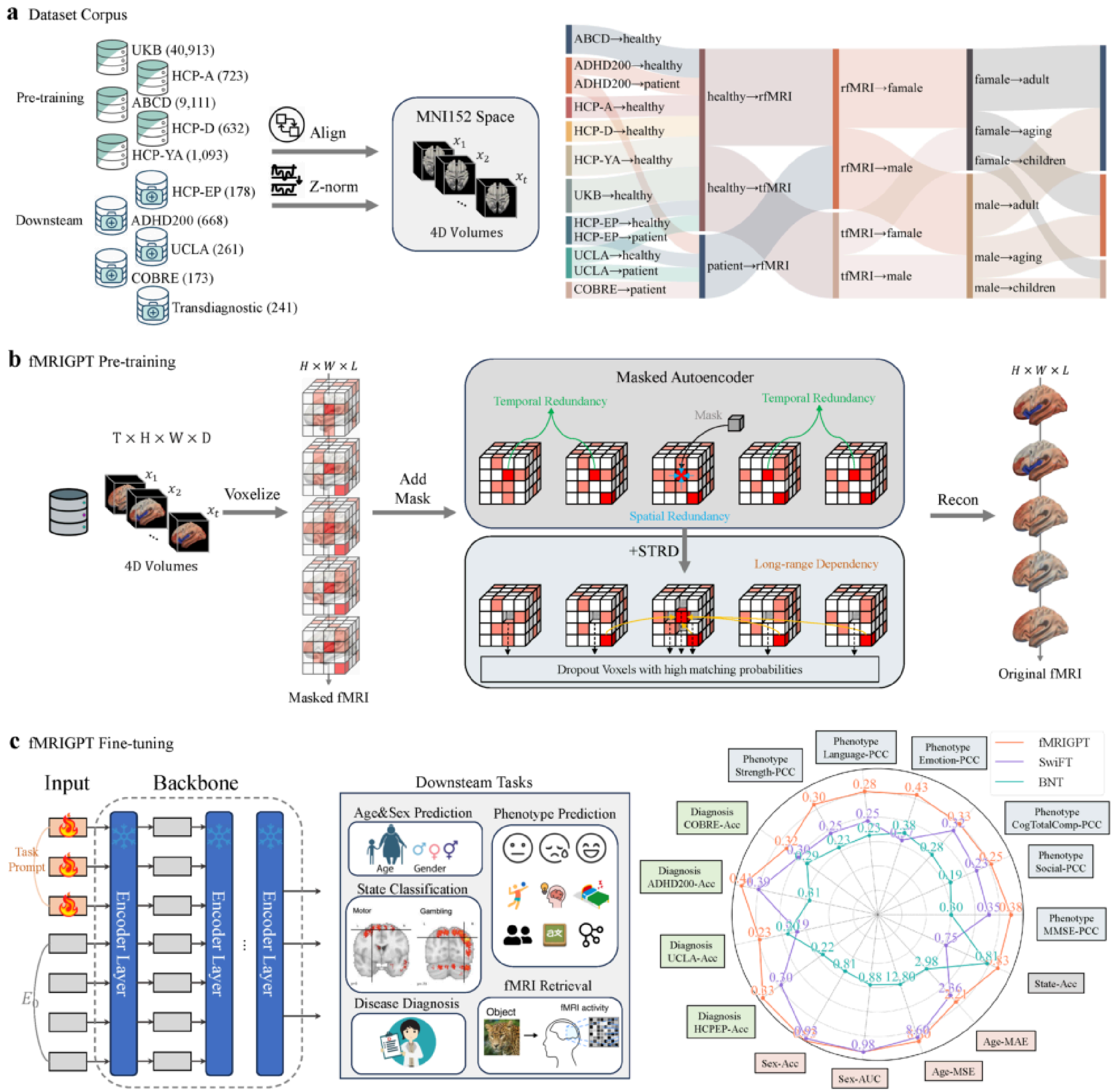
Co-Author(s):

Yu Jiang
The Chinese University of Hong Kong
Hong Kong, Hong Kong
Zhihao Peng
The Chinese University of Hong Kong
Hong Kong, Hong Kong
Chenxin Li
The Chinese University of Hong Kong
Hong Kong, Hong Kong
Changbei Bang
Yonsei University
Seoul, Seoul
Carl Yang
Emory University
Atlanta, GA
Lifang He
Lehigh University
Bethlehem, PA

Daniel Barron
Brigham and Women's Hospital
Boston, MA
Randy Hirschtick, MD, PhD
Massachusetts General Hospital
Boston, MA
Byung-Hoon Kim
Yonsei University
Seoul, Seoul
Xiang Li
Massachusetts General Hospital
Boston, MA
Yixuan Yuan
The Chinese University of Hong Kong
Hong Kong, Hong Kong

Introduction:

Functional Magnetic Resonance Imaging (fMRI) is essential for studying brain function and neurological disorders, yet its analysis is hindered by low signal-to-noise ratio, test-retest variability, complex preprocessing, and limited dataset sizes. These challenges contribute to a reproducibility crisis and limit model transferability across tasks and populations. In response, we introduce fMRI-GPT, a foundation model for fMRI analysis, pre-trained on 55,000+ multi-site fMRI 4D volumes. Evaluation on a diverse array of downstream tasks shows that fMRI-GPT achieves state-of-the-art performance while requiring fewer training samples. By learning whole-brain voxel-wise representations, fMRI-GPT provides a scalable and generalizable framework for fMRI analysis. It enhances brain function decoding for perception, memory, emotion, and decision-making, while improving reproducibility across studies. Bridging deep learning with neuroimaging, fMRI-GPT advances precision psychiatry, brain-computer interfaces, and early disease detection.



·Overview of the proposed fMRI-GPT model.

Methods:

fMRI-GPT is pre-trained on 55,000+ fMRI sequences from multi-center datasets, including UK Biobank, ABCD, and HCP, covering diverse demographics and clinical conditions. Data preprocessing includes motion correction, slice timing correction, spatial normalization to MNI152 space, and Z-score intensity normalization. The model employs a Masked Autoencoder (MAE) framework to learn latent representations of brain activity. A Spatiotemporal Redundancy Dropout (STRD) module enhances noise resilience by filtering redundant information, improving test-retest reliability. The Shifted-Window Mamba Backbone optimizes 4D volume processing, reducing GPU memory usage while preserving long-range dependencies in brain signals.
Pretraining follows a self-supervised approach, where the model reconstructs masked fMRI sequences to learn intrinsic brain activity patterns. Fine-tuning is performed using Task-specific

Prompt Tuning, which updates only a small subset of parameters, allowing efficient adaptation to new tasks with minimal labeled data. fMRI-GPT is evaluated across five key fMRI tasks: age and gender prediction, phenotype prediction, disease diagnosis, fMRI retrieval, and task-state classification. Performance is compared against state-of-the-art ROI-based and volume-based models using accuracy, Pearson correlation, mean squared error (MSE), and area under the curve (AUC).

Results:

fMRI-GPT achieves state-of-the-art (SOTA) performance across five key fMRI tasks while requiring significantly fewer training samples. For age prediction, it reduces mean squared error (MSE) by 12.5% compared to SwiFT (Kim et al., 2023), the best-performing volume-based method. In gender classification, it outperforms BrainGNN (Li et al., 2021), achieving 93.3% accuracy (vs. 85.6% for BrainGNN). For phenotype prediction, fMRI-GPT surpasses existing models, achieving a 0.429 Pearson Correlation Coefficient (PCC) in emotion-related phenotype estimation, 15.2% higher than previous SOTA models such as BrainLM (Ortega Caro et al., 2023). In disease diagnosis, it outperforms BrainLM and Com-BrainTF (Kan et al., 2022) on schizophrenia classification (HCP-EP dataset), reaching 75.2% accuracy, a 15.3% improvement over ROI-based approaches. Notably, fMRI-GPT is the first model to perform 4D volume-based fMRI retrieval, achieving 79.4% accuracy. In comparison, the SOTA method MindEyeV2 (Scotti et al., 2024) relies on pre-defined vision-related ROIs. In task-state classification, it attains 92.6% accuracy, exceeding SwiFT's 88.1%.

Conclusions:

fMRI-GPT introduces a new paradigm in fMRI analysis by leveraging large-scale data and a powerful pre-trained model. It provides a scalable and versatile solution with broad applications in cognitive neuroscience, precision psychiatry, and brain-computer interfaces.

Modeling and Analysis Methods:

Activation (eg. BOLD task-fMRI)
fMRI Connectivity and Network Modeling [2]

Neuroinformatics and Data Sharing:

Workflows [1]

Keywords:

FUNCTIONAL MRI
Modeling
Other - Foundation Model

[1|2]Indicates the priority used for review

Abstract Information

I accept

The Open Science Special Interest Group (OSSIG) is introducing a **reproducibility challenge** for OHBM 2025. This new initiative aims to enhance the reproducibility of scientific results and foster collaborations between labs. Teams will consist of a “source” party and a “reproducing” party, and will be evaluated on the success of their replication, the openness of the source work, and additional deliverables. Click here for more information. Propose your OHBM abstract(s) as source work for future OHBM meetings by selecting one of the following options:

I am submitting this abstract as an original work to be reproduced. I am available to be the “source party” in an upcoming team and consent to have this work listed on the OSSIG website. I agree to be contacted by OSSIG regarding the challenge and may share data used in this abstract with another team.

Please indicate below if your study was a "resting state" or "task-activation” study.

Other

Healthy subjects only or patients (note that patient studies may also involve healthy subjects):

Patients

Was this research conducted in the United States?

No

Were any human subjects research approved by the relevant Institutional Review Board or ethics panel? NOTE: Any human subjects studies without IRB approval will be automatically rejected.

Not applicable

Were any animal research approved by the relevant IACUC or other animal research panel? NOTE: Any animal studies without IACUC approval will be automatically rejected.

Not applicable

Please indicate which methods were used in your research:

Functional MRI

For human MRI, what field strength scanner do you use?

If Other, please list - varies

Which processing packages did you use for your study?

Free Surfer

Provide references using APA citation style.

Casey, B. J., Cannonier, T., Conley, M. I., Cohen, A. O., Barch, D. M., Heitzeg, M. M., … & Garavan, H. (2018). The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. Developmental Cognitive Neuroscience, 32, 43–54.
Kan, X., et al. (2022). Brain network transformer. Advances in Neural Information Processing Systems, 35, 25586–25599.
Kim, P., et al. (2023). SwiFT: Swin 4D fMRI transformer. Advances in Neural Information Processing Systems, 36, 42015–42037.
Li, X., et al. (2021). BrainGNN: Interpretable brain graph neural network for fMRI analysis. Medical Image Analysis, 74, 102233.
Miller, K. L., Alfaro-Almagro, F., Bangerter, N. K., Thomas, D. L., Yacoub, E., Xu, J., … & Smith, S. M. (2016). Multimodal population brain imaging in the UK Biobank prospective epidemiological study. Nature Neuroscience, 19(11), 1523–1536.
Ortega Caro, J., et al. (2023). BrainLM: A foundation model for brain activity recordings. bioRxiv.
Scotti, P. S., et al. (2024). MindEyeV2: Shared-subject models enable fMRI-to-image with 1 hour of data. arXiv preprint arXiv:2403.11207.
Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E., Yacoub, E., & Ugurbil, K. (2013). The WU-Minn Human Connectome Project: An overview. NeuroImage, 80, 62–79.

UNESCO Institute of Statistics and World Bank Waiver Form

I attest that I currently live, work, or study in a country on the UNESCO Institute of Statistics and World Bank List of Low and Middle Income Countries list provided.

No