Novelty Assessment Report

Paper: Brain-Semantoks: Learning Semantic Tokens of Brain Dynamics with a Self-Distilled Foundation Model

PDF URL: https://openreview.net/pdf?id=ANkm27vNuk

Venue: ICLR 2026 Conference Submission

Year: 2026

Report Generated: 2025-12-30

Abstract

The development of foundation models for functional magnetic resonance imaging (fMRI) time series holds significant promise for predicting phenotypes related to disease and cognition. Current models, however, are often trained using a mask-and-reconstruct objective on small brain regions. This focus on low-level information leads to representations that are sensitive to noise and temporal fluctuations, necessitating extensive fine-tuning for downstream tasks. We introduce Brain-Semantoks, a self-supervised framework designed specifically to learn abstract representations of brain dynamics. Its architecture is built on two core innovations: a semantic tokenizer that aggregates noisy regional signals into robust tokens representing functional networks, and a self-distillation objective that enforces representational stability across time. We show that this objective is stabilized through a novel training curriculum, ensuring the model robustly learns meaningful features from low signal-to-noise time series. We demonstrate that learned representations enable strong performance on a variety of downstream tasks even when only using a linear probe. Furthermore, we provide comprehensive scaling analyses indicating more unlabeled data reliably results in out-of-distribution performance gains without domain adaptation.

Core Task Landscape

This paper addresses: **Learning Abstract Representations of Brain Dynamics from fMRI Time Series**

A total of **50 papers** were analyzed and organized into a taxonomy with **21 categories**.

Taxonomy Overview

The research landscape has been organized into the following main categories:

- **Graph-Based Spatiotemporal Representation Learning**
- **Foundation Models and Self-Supervised Pretraining**
- **Generative Models for Visual Stimulus Reconstruction**
- **Temporal Dynamics and State-Space Modeling**
- **Multi-Subject and Cross-Modal Representation Learning**
- **Task-Based and Predictive Representation Learning**
- **Specialized Decoding and Analysis Frameworks**
- **Dimensionality Reduction and Latent Feature Models**
- **Semantic and Cognitive Representation Models**
- **Survey and Review Literature**

Complete Taxonomy Tree

- Learning Abstract Representations of Brain Dynamics from fMRI Time Series Survey Taxonomy
- Graph-Based Spatiotemporal Representation Learning
  - Dynamic Graph Neural Network Architectures (4 papers)
  - [1] Learning dynamic graph representation of brain connectome with spatio-temporal attention (Kim Byung-Hoon, 2021) View paper
  - [2] Dynamic graph representation learning for spatio-temporal neuroimaging analysis (Rui Liu, 2025) View paper
  - [6] BrainTGL: A dynamic graph representation learning model for brain network analysis (Lingwen Liu, 2023) View paper
  - [23] Multi-scale dynamic graph learning for brain disorder detection with functional MRI (Yunling Ma, 2023) View paper
  - Graph Transformer Methods for Brain Connectome (2 papers)
  - [8] Large-scale Graph Representation Learning of Dynamic Brain Connectome with Transformers (Kim Byung-Hoon, 2023) View paper
  - [18] Learning Dynamic Brain Connectome with Graph Transformers for Psychiatric Diagnosis Classification (Byung Hoon Kim, 2024) View paper
  - Manifold and Geometric Learning on Brain Networks (2 papers)
  - [11] Neurospectrum: A Geometric and Topological Deep Learning Framework for Uncovering Spatiotemporal Signatures in Neural Activity (Yan-lei Zhang, 2025) View paper
  - [35] Learning brain dynamics of evolving manifold functional MRI data using geometric-attention neural network (Tingting Dan, 2022) View paper
- Foundation Models and Self-Supervised Pretraining
  - Masked Autoencoding and Predictive Architectures ★ (3 papers)
  - [0] Brain-Semantoks: Learning Semantic Tokens of Brain Dynamics with a Self-Distilled Foundation Model (Anon et al., 2026) View paper
  - [20] Brain-jepa: Brain dynamics foundation model with gradient positioning and spatiotemporal masking (Dong Zijian, 2024) View paper
  - [26] A Foundational fMRI Model for Representing Continuous Brain States (Li Yang, 2025) View paper
  - Graph Contrastive and Modular Pretraining (3 papers)

  - [22] A Brain Graph Foundation Model: Pre-Training and Prompt-Tuning for Any Atlas and Disorder (Wei Xinxu, 2025) View paper
  - [28] Leveraging brain modularity prior for interpretable representation learning of fMRI (Qianqian Wang, 2024) View paper
  - [39] BrainMAE: A Region-aware Self-supervised Learning Framework for Brain Signals (Yang, 2024) View paper
- Generative Models for Visual Stimulus Reconstruction
  - Latent Diffusion and VAE-Based Reconstruction (3 papers)
  - [7] High-resolution image reconstruction with latent diffusion models from human brain activity (Yu Takagi, 2023) View paper
  - [15] Optimized AI-based neural decoding from BOLD fMRI signal for analyzing visual and semantic ROIs in the human visual system (Lorenzo Veronese, 2025) View paper
  - [32] Mindldm: Reconstruct visual stimuli from fmri using latent diffusion model (Jun-Hao Guo, 2024) View paper
  - Contrastive and Adversarial Reconstruction Methods (3 papers)
  - [12] Reconstructing seen image from brain activity by visually-guided cognitive representation and adversarial learning (Ziqi Ren, 2021) View paper
  - [19] Decoding realistic images from brain activity with contrastive self-supervision and latent diffusion (Sun Jingyuan, 2023) View paper
  - [21] Reconstructing natural scenes from fmri patterns using bigbigan (Milad Mozafari, 2020) View paper
- Temporal Dynamics and State-Space Modeling
  - Recurrent and Sequential Deep Learning (3 papers)
  - [40] â¦ , highly accurate brain decoding of subtly distinct brain states from functional MRI using intrinsic functional networks and long short-term memory recurrent neural â¦ (H Li, 2019) View paper
  - [47] Learning deep temporal representations for fMRI brain decoding (Orhan Firat, 2015) View paper
  - [50] Temporal Graph Representation Learning for Autism spectrum disorder Brain Networks (Peng Cao, 2021) View paper
  - State-Space and Dynamical Systems Models (3 papers)
  - [25] Characterization of regional differences in resting-state fMRI with a data-driven network model of brain dynamics (Viktor Sip, 2023) View paper
  - [31] State-space model with deep learning for functional dynamics estimation in resting-state fMRI (Heung-Il Suk, 2016) View paper
  - [45] Reconstructing brain causal dynamics for subject and task fingerprints using fMRI time-series data (Dachuan Song, 2025) View paper
  - Metastability and Discrete State Representations (2 papers)
  - [3] A parsimonious description of global functional brain organization in three spatiotemporal patterns (Taylor Bolt, 2022) View paper
  - [49] Hierarchical Characterization of Brain Dynamics via State Space-based Vector Quantization (Yang Yanwu, 2025) View paper
- Multi-Subject and Cross-Modal Representation Learning
  - Shared Latent Manifolds Across Subjects (2 papers)
  - [9] Learning shared neural manifolds from multi-subject FMRI data (Jessie Huang, 2022) View paper
  - [10] Nonlinear latent representations of high-dimensional task-fMRI data: Unveiling cognitive and behavioral insights in heterogeneous spatial maps (M. Zabihi, 2024) View paper
  - Cross-Modal Synthesis and Alignment (3 papers)
  - [4] Catd: Unified representation learning for eeg-to-fmri cross-modal generation (Weiheng Yao, 2025) View paper
  - [16] Alignment of auditory artificial networks with massive individual fMRI brain data leads to generalisable improvements in brain encoding and downstream tasks (MaÃ«lle Freteault, 2025) View paper
  - [38] Intuition in Silico: Representational Alignment of Deep Neural Networks with Human Brain Dynamics in Intuitive Reasoning (Li, 2025) View paper
- Task-Based and Predictive Representation Learning
  - Task Activation Prediction from Resting-State (2 papers)
  - [33] TARDRL: Task-Aware Reconstruction for Dynamic Representation Learning of fMRI (Yunxi Zhao, 2024) View paper
  - [34] Predicting task-related brain activity from resting-state brain dynamics with fMRI Transformer (Junbeom Kwon, 2024) View paper
  - Phenotype and Clinical Outcome Prediction (2 papers)
  - [14] Deep representations for time-varying brain datasets (Sikun Lin, 2022) View paper
  - [30] 498. Individual Differences in Multimodal Neurodevelopment Along a Sensorimotor-To-Association Axis: Implications for Early Adolescent Psychopathology (Katherine L. Bottenhorn, 2024) View paper
- Specialized Decoding and Analysis Frameworks
  - Multivariate Pattern Analysis and Temporal Decoding (2 papers)
  - [17] Decoding dynamic brain patterns from evoked responses: a tutorial on multivariate pattern analysis applied to time series neuroimaging data (Tijl Grootswagers, 2017) View paper
  - [24] Characterizing the dynamics of mental representations: the temporal generalization method (J. King, 2014) View paper
  - Compression and Efficient Encoding (2 papers)
  - [13] A Compact Implicit Neural Representation for Efficient Storage of Massive 4D Functional Magnetic Resonance Imaging (Li Ruoran, 2025) View paper
  - [41] A compressed code for memory discrimination (Dale Zhou, 2025) View paper
  - Specialized Analytical and Generative Frameworks (3 papers)
  - [37] Deep Generative Analysis for Task-Based Functional MRI Experiments (Daniela de Albuquerque, 2021) View paper
  - [44] Learning diverse causally emergent representations from time series data (Christos Kaplanis, 2024) View paper
  - [48] Learning Image Derived PDE-Phenotypes from fMRI Data (Bica Ion, 2024) View paper
- Dimensionality Reduction and Latent Feature Models
  - Adaptive and Piecewise Linear Methods (1 papers)
  - [5] Adaptive latent feature sharing for piecewise linear dimensionality reduction (Farooq Adam, 2024) View paper
  - Independent Component Analysis and Real-Time Methods (1 papers)
  - [43] Real-time independent component analysis of fMRI time-series (F. Esposito, 2003) View paper
- Semantic and Cognitive Representation Models (2 papers)
  - [36] Neural representation of association strength and prediction error during novel symbol-speech sounds learning (Gorka Fraga GonzÃ¡lez, 2023) View paper

    - [46] Predicting neural activity patterns associated with sentences using a neurobiologically motivated model of semantic representation (Andrew James Anderson, 2017) View paper
- Survey and Review Literature (3 papers)
    - [27] Advances in fMRI-Based Brain Function Mapping: A Deep Learning Perspective (Zhao, 2025) View paper
    - [29] Representation Learning on Brain Data (Sikun, 2022) View paper
    - [42] fmri brain decoding and its applications in brainâ□□computer interface: A survey (Bing Du, 2022) View paper

Narrative

Core task: learning abstract representations of brain dynamics from fMRI time series. The field has organized itself around several complementary perspectives on how to extract meaningful structure from high-dimensional neuroimaging data. Graph-based spatiotemporal methods treat brain regions as nodes and model evolving connectivity patterns, often using dynamic graph neural networks or temporal graph learning frameworks such as BrainTGL[6] and Dynamic Graph Neuroimaging[2]. Foundation models and self-supervised pretraining approaches borrow ideas from large-scale vision and language modeling, applying masked autoencoding or contrastive objectives to learn general-purpose brain representations that transfer across tasks and subjects, as seen in works like BrainMAE[39] and Foundational fMRI Model[26]. Generative models focus on reconstructing external stimuli or internal cognitive states from brain activity, while temporal dynamics and state-space modeling emphasize capturing the sequential evolution of neural states. Multi-subject and cross-modal branches address alignment and transfer across individuals or modalities, and task-based methods optimize representations for specific predictive goals. Dimensionality reduction and semantic representation models round out the taxonomy by exploring low-dimensional manifolds and cognitive content.

Within the foundation model branch, a particularly active line of work centers on masked autoencoding and predictive architectures that learn to reconstruct or predict missing or future brain activity patterns. Brain Semantoks[0] sits squarely in this cluster, proposing a discrete tokenization scheme that quantizes spatiotemporal brain dynamics into a compact vocabulary before applying masked prediction. This approach contrasts with continuous embedding methods like Brain JEPA[20], which uses joint-embedding predictive architectures without explicit discretization, and with Foundational fMRI Model[26], which scales up pretraining across diverse datasets but retains continuous latent codes. The trade-off revolves around whether discrete tokens offer better interpretability and compositionality or whether continuous representations preserve richer temporal detail. Across these branches, open questions persist about how to balance model scale, inductive biases for spatiotemporal structure, and the degree of supervision needed to capture clinically or cognitively relevant brain dynamics.

Related Works in Same Category

The following **2 sibling papers** share the same taxonomy leaf node with the original paper:

1. Brain-jepa: Brain dynamics foundation model with gradient positioning and spatiotemporal masking

Authors: Dong Zijian, Li Ruilin, Zijian Dong, Wu Yilei, Ruilin Li, et al. (18 authors total) | **Year/Venue**: 2024 | **URL**: View paper

Abstract

We introduce Brain-JEPA, a brain dynamics foundation model with the Joint-Embedding Predictive Architecture (JEPA). This pioneering model achieves state-of-the-art performance in demographic prediction, disease diagnosis/prognosis, and trait prediction through fine-tuning. Furthermore, it excels in off-the-shelf evaluations (e.g., linear probing) and demonstrates superior generalizability across different ethnic groups, surpassing the previous large model for brain activity significantly. Brain-...

Relationship Analysis

Both papers belong to the Masked Autoencoding and Predictive Architectures category, employing self-supervised pretraining strategies to learn robust latent representations from fMRI time series. They overlap in their use of transformer-based architectures and focus on learning abstract representations suitable for downstream tasks without extensive fine-tuning. However, the original paper (Brain-Semantoks) uses a self-distillation objective across temporal views with a semantic tokenizer aggregating regional signals into functional network tokens, while the candidate paper (Brain-JEPA) employs a joint-embedding predictive architecture that predicts masked representations in latent space with brain gradient positioning and spatiotemporal masking strategies.

---

2. A Foundational fMRI Model for Representing Continuous Brain States

Authors: Li Yang, Lei Guo, Yixuan Yuan, Junwei Han, Xintao Hu, et al. (6 authors total) | **Year/Venue**: 2025 | **URL**: View paper

Abstract

Foundational models have significant potential to advance brain function research, particularly in understanding the dynamics of brain states. However, most existing models process brain signals within fixed time windows, restricting their ability to capture the full temporal complexity of brain activity. In this study, we propose BrainSN (Brain States Network), a novel fMRI foundational model designed to represent continuous brain state information and support diverse downstream tasks. First, l...

Relationship Analysis

Both papers belong to the masked autoencoding and predictive architectures category, using transformer-based foundation models for learning fMRI representations. They overlap in their use of self-supervised pretraining on large-scale unlabeled fMRI data and evaluation on downstream clinical/cognitive tasks. However, Brain-Semantoks focuses on self-distillation across temporal views with a semantic tokenizer aggregating ROIs into functional network tokens, while BrainSN emphasizes continuous brain state modeling through multi-scale reconstruction and future prediction objectives without the network-level tokenization approach.

Contributions Analysis

Overall novelty summary. The paper introduces Brain-Semantoks, a self-supervised framework combining a semantic tokenizer that aggregates regional fMRI signals into functional network tokens with a self-distillation objective for temporal stability. It resides in the 'Masked Autoencoding and Predictive Architectures' leaf under 'Foundation Models and Self-Supervised Pretraining', alongside two sibling papers. This leaf represents a moderately populated research direction within the broader foundation model branch, which itself contains six papers across two sub-categories. The taxonomy shows this is an active but not overcrowded area, with the paper positioned among methods that learn general-purpose brain representations through reconstruction or prediction objectives.

The taxonomy reveals several neighboring research directions that contextualize this work. The sibling 'Graph Contrastive and Modular Pretraining' category explores alternative self-supervised objectives using graph structure and contrastive learning rather than masked reconstruction. Adjacent branches include 'Temporal Dynamics and State-Space Modeling', which emphasizes explicit temporal evolution through recurrent or state-space formulations, and 'Metastability and Discrete State Representations', which also quantizes brain dynamics but focuses on metastable configurations rather than functional network tokens. The paper bridges foundation model pretraining with discrete tokenization approaches, connecting masked autoencoding traditions with state-based representations while maintaining focus on abstract, noise-robust features.

Among 21 candidates examined across three contributions, none were identified as clearly refuting the proposed methods. The self-distillation framework examined 10 candidates with no refutable overlap, the semantic tokenizer examined 10 candidates with similar results, and the training curriculum examined 1 candidate without refutation. This suggests that within the limited search scope, the specific combination of semantic tokenization, self-distillation, and curriculum learning appears relatively unexplored. However, the modest search scale means substantial prior work may exist beyond the top-K semantic matches examined. The semantic tokenizer contribution, despite examining 10 candidates, shows no direct precedent in the retrieved literature.

Based on the limited literature search of 21 candidates, the work appears to occupy a distinctive position combining discrete tokenization with self-supervised learning for fMRI. The taxonomy structure indicates this sits in an active but not saturated research area, with clear differentiation from continuous embedding methods in sibling papers. The absence of refuting candidates across all contributions suggests novelty within the examined scope, though the search scale leaves open the possibility of relevant work in adjacent communities or earlier literature not captured by semantic similarity.

This paper presents **3 main contributions**, each analyzed against relevant prior work:

Contribution 1: Self-distillation framework for learning abstract brain dynamics representations

Description: The authors propose a novel pretraining approach that shifts from reconstruction-based objectives to learning high-level, stable phenotypic signatures through self-distillation across temporal views. This framework explicitly trains models to capture abstract representations suitable for transfer learning rather than modeling low-level signal details.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. LLEDAâLifelong self-supervised domain adaptation

URL: View paper

Brief Assessment

LLEDA[68] focuses on lifelong domain adaptation with memory replay mechanisms for continual learning across domains, not on learning abstract representations of brain dynamics through self-distillation. The technical approaches and application domains are fundamentally different.

---

2. Self-supervised learning of brain dynamics from broad neuroimaging data

URL: View paper

Brief Assessment

Self-Supervised Brain Dynamics[63] focuses on self-supervised learning for neuroimaging using NLP-inspired sequence modeling techniques (autoencoding, causal/masked language modeling). The original paper's self-distillation approach with semantic tokenization and teacher-guided temporal regularization represents a distinct methodological contribution not present in the candidate.

---

3. Explainable Self-Supervised Dynamic Neuroimaging Using Time Reversal

URL: View paper

Brief Assessment

Time Reversal Neuroimaging[66] uses a time reversal pretraining task with LSTM models for schizophrenia classification, not a self-distillation framework for learning abstract representations across temporal views.

---

4. Self-supervised learning for electroencephalogram: A systematic survey

URL: View paper

Brief Assessment

Self-Supervised EEG Survey[62] focuses on EEG signal analysis and provides a systematic survey of SSL methods for electroencephalography, not fMRI brain dynamics or self-distillation frameworks for neuroimaging foundation models.

---

5. Population transformer: Learning population-level representations of neural activity

URL: View paper

Brief Assessment

Population Transformer[64] focuses on learning population-level codes for multi-channel intracranial neural recordings with spatially-sparse electrodes, not on self-distillation across temporal views for fMRI brain dynamics. The technical approaches and data modalities are fundamentally different.

---

6. Adaptive-Similarity-Based Brain Dynamic Functional Connectivity with Spatial-Temporal Attention and Domain Adaptation for Schizophrenia Diagnosis

URL: View paper

Brief Assessment

Adaptive Similarity Schizophrenia[71] focuses on dynamic functional connectivity construction using Kalman filters and domain adaptation for schizophrenia diagnosis, not on self-distillation frameworks for learning abstract representations suitable for transfer learning.

---

7. Longitudinal self-supervised learning

URL: View paper

Brief Assessment

Longitudinal Self-Supervised[70] focuses on disentangling temporal factors (e.g., brain age) across longitudinal MRI sequences using factor disentanglement and autoencoding, not on self-distillation across temporal views for learning stable phenotypic signatures from fMRI time series as in the original paper.

---

8. Self-supervised Learning for Encoding Between-Subject Information in Clinical EEG

URL: View paper

Brief Assessment

Between-Subject EEG[67] focuses on contrastive learning using subject identities for single-channel EEG pathology detection, not self-distillation across temporal views for abstract fMRI brain dynamics as in the original paper.

---

9. GMAEEG: A Self-Supervised Graph Masked Autoencoder for EEG Representation Learning

URL: View paper

Brief Assessment

GMAEEG[69] uses a reconstruction-based masked autoencoder approach for EEG signals, fundamentally different from the original paper's self-distillation framework that explicitly avoids reconstruction objectives in favor of learning abstract phenotypic signatures through temporal view consistency.

10. BENDR: Using transformers and a contrastive self-supervised learning task to learn from massive amounts of EEG data

URL: View paper

Brief Assessment

BENDR[65] focuses on contrastive self-supervised learning for EEG data using a wav2vec 2.0-inspired approach, not self-distillation across temporal views for fMRI as in the original paper. The modalities (EEG vs. fMRI), architectures (contrastive prediction vs. student-teacher distillation), and objectives differ fundamentally.

Contribution 2: Semantic tokenizer for functional brain networks

Description: The authors introduce a neuroscientifically-grounded tokenizer that aggregates information from brain regions within functional networks into single robust tokens. This creates a more compact, semantically meaningful input sequence compared to treating individual noisy ROI signals as tokens.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. BACE: Behavior-adaptive connectivity estimation for interpretable graphs of neural dynamics

URL: View paper

Brief Assessment

BACE[55] focuses on learning directed connectivity between brain regions from intracranial LFP recordings, using per-region temporal encoders (GRUs) to aggregate micro-contacts within anatomical regions. This differs from the original paper's semantic tokenizer, which aggregates ROI signals within functional networks (e.g., default mode network) for fMRI analysis using a novel multi-scale convolutional architecture.

2. TFAGL: A novel agent graph learning method using time-frequency EEG for major depressive disorder detection

URL: View paper

Brief Assessment

TFAGL[58] focuses on EEG-based depression detection using agent nodes for brain region interactions, not fMRI foundation models with semantic tokenization of functional networks.

3. The disturbed functional brain network in major depressive disorder identified by graph theory analysis

URL: View paper

Brief Assessment

Depression Network Graph[52] focuses on graph theory analysis of functional connectivity matrices in major depressive disorder, not on creating semantic tokens from brain network signals for foundation model training.

4. Reconfiguration of dynamic largeâ□□scale brain network functional connectivity in generalized tonicâ□□clonic seizures

URL: View paper

Brief Assessment

Seizure Network Reconfiguration[60] focuses on dynamic functional connectivity variability in epilepsy patients using traditional network analysis methods, not on developing tokenization approaches for transformer-based foundation models or representation learning frameworks.

5. Disorder-specific neurodynamic features in schizophrenia inferred by neurodynamic embedded contrastive variational autoencoder model

URL: View paper

Brief Assessment

The candidate paper focuses on neurodynamic modeling for schizophrenia using fMRI data, not on tokenizing functional brain networks for foundation models. The candidate uses LSTM networks to extract temporal states and parameters from brain regions, which is fundamentally different from the original paper's semantic tokenizer that aggregates regional signals into robust network-level tokens for transformer-based learning.

6. Joint learning of multi-level dynamic brain networks for autism spectrum disorder diagnosis

URL: View paper

Brief Assessment

Autism Multi-Level Networks[61] focuses on graph convolutional networks for multi-level brain network analysis in ASD diagnosis, not on tokenization strategies for transformer-based foundation models. The candidate addresses network-level feature learning through GCNs rather than creating semantic tokens for self-attention mechanisms.

7. EEG emotion classification based on graph convolutional network

URL: View paper

Brief Assessment

EEG Graph Emotion[56] focuses on regional-level EEG representations for emotion recognition using graph convolutional networks, not on creating semantic tokens from fMRI functional brain networks for foundation model pretraining.

8. Hierarchical Encoding and Fusion of Brain Functions for Depression Subtype Classification

URL: View paper

Brief Assessment

Depression Subtype Encoding[54] focuses on pre-training features from individual brain regions for depression subtype classification using graph neural networks, not on creating semantic tokens from functional networks for general fMRI foundation models.

9. RTGMFF: Enhanced fmri-based brain disorder diagnosis via roi-driven text generation and multimodal feature fusion

URL: View paper

Brief Assessment

RTGMFF[53] focuses on deterministic ROI-level text generation from activation statistics for clinical reporting, not on creating semantic tokens from functional brain networks for foundation model pretraining. The candidate's approach generates textual descriptions of brain regions, while the original work aggregates regional signals into robust network-level tokens for self-supervised learning.

10. Cognitmoe: A cognition-aware collaborative multi-expert network for bipolar disorder diagnosis

URL: View paper

Brief Assessment

CognitMoE[57] focuses on brain-region-level imaging features for bipolar disorder diagnosis, not on creating semantic tokens from functional brain networks for foundation model pretraining.

Contribution 3: Teacher-guided Temporal Regularizer training curriculum

Description: The authors develop a principled training curriculum that stabilizes self-distillation on low signal-to-noise fMRI data by initially guiding the model to learn time-averaged network representations before modeling complex temporal variations. This regularizer prevents convergence to poor solutions during early training.

This contribution was assessed against **1 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. A Survey on Joint Embedding Predictive Architectures and World Models

URL: View paper

Brief Assessment

Joint Embedding Architectures[51] is a survey paper discussing general architectural principles. The candidate mentions curriculum-based masking schedules and self-distillation in passing, but does not present a specific training curriculum for stabilizing self-distillation on low signal-to-noise fMRI data through temporal regularization as the original paper does.

Appendix: Text Similarity Detection

No high-similarity text segments were detected across any compared papers.

References

- [0] Brain-Semantoks: Learning Semantic Tokens of Brain Dynamics with a Self-Distilled Foundation Model View paper
- [1] Learning dynamic graph representation of brain connectome with spatio-temporal attention View paper
- [2] Dynamic graph representation learning for spatio-temporal neuroimaging analysis View paper
- [3] A parsimonious description of global functional brain organization in three spatiotemporal patterns View paper
- [4] Catd: Unified representation learning for eeg-to-fmri cross-modal generation View paper
- [5] Adaptive latent feature sharing for piecewise linear dimensionality reduction View paper
- [6] BrainTGL: A dynamic graph representation learning model for brain network analysis View paper
- [7] High-resolution image reconstruction with latent diffusion models from human brain activity View paper
- [8] Large-scale Graph Representation Learning of Dynamic Brain Connectome with Transformers View paper
- [9] Learning shared neural manifolds from multi-subject FMRI data View paper
- [10] Nonlinear latent representations of high-dimensional task-fMRI data: Unveiling cognitive and behavioral insights in heterogeneous spatial maps View paper
- [11] Neurospectrum: A Geometric and Topological Deep Learning Framework for Uncovering Spatiotemporal Signatures in Neural Activity View paper
- [12] Reconstructing seen image from brain activity by visually-guided cognitive representation and adversarial learning View paper
- [13] A Compact Implicit Neural Representation for Efficient Storage of Massive 4D Functional Magnetic Resonance Imaging View paper
- [14] Deep representations for time-varying brain datasets View paper
- [15] Optimized AI-based neural decoding from BOLD fMRI signal for analyzing visual and semantic ROIs in the human visual system View paper
- [16] Alignment of auditory artificial networks with massive individual fMRI brain data leads to generalisable improvements in brain encoding and downstream tasks View paper
- [17] Decoding dynamic brain patterns from evoked responses: a tutorial on multivariate pattern analysis applied to time series neuroimaging data View paper
- [18] Learning Dynamic Brain Connectome with Graph Transformers for Psychiatric Diagnosis Classification View paper
- [19] Decoding realistic images from brain activity with contrastive self-supervision and latent diffusion View paper
- [20] Brain-jepa: Brain dynamics foundation model with gradient positioning and spatiotemporal masking View paper
- [21] Reconstructing natural scenes from fmri patterns using bigbigan View paper
- [22] A Brain Graph Foundation Model: Pre-Training and Prompt-Tuning for Any Atlas and Disorder View paper
- [23] Multi-scale dynamic graph learning for brain disorder detection with functional MRI View paper
- [24] Characterizing the dynamics of mental representations: the temporal generalization method View paper
- [25] Characterization of regional differences in resting-state fMRI with a data-driven network model of brain dynamics View paper
- [26] A Foundational fMRI Model for Representing Continuous Brain States View paper
- [27] Advances in fMRI-Based Brain Function Mapping: A Deep Learning Perspective View paper
- [28] Leveraging brain modularity prior for interpretable representation learning of fMRI View paper
- [29] Representation Learning on Brain Data View paper
- [30] 498. Individual Differences in Multimodal Neurodevelopment Along a Sensorimotor-To-Association Axis: Implications for Early Adolescent Psychopathology View paper
- [31] State-space model with deep learning for functional dynamics estimation in resting-state fMRI View paper
- [32] Mindldm: Reconstruct visual stimuli from fmri using latent diffusion model View paper
- [33] TARDRL: Task-Aware Reconstruction for Dynamic Representation Learning of fMRI View paper
- [34] Predicting task-related brain activity from resting-state brain dynamics with fMRI Transformer View paper

- [35] Learning brain dynamics of evolving manifold functional MRI data using geometric-attention neural network View paper
- [36] Neural representation of association strength and prediction error during novel symbol-speech sounds learning View paper
- [37] Deep Generative Analysis for Task-Based Functional MRI Experiments View paper
- [38] Intuition in Silico: Representational Alignment of Deep Neural Networks with Human Brain Dynamics in Intuitive Reasoning View paper
- [39] BrainMAE: A Region-aware Self-supervised Learning Framework for Brain Signals View paper
- [40] â¦ , highly accurate brain decoding of subtly distinct brain states from functional MRI using intrinsic functional networks and long short-term memory recurrent neural â¦ View paper
- [41] A compressed code for memory discrimination View paper
- [42] fmri brain decoding and its applications in brainâcomputer interface: A survey View paper
- [43] Real-time independent component analysis of fMRI time-series View paper
- [44] Learning diverse causally emergent representations from time series data View paper
- [45] Reconstructing brain causal dynamics for subject and task fingerprints using fMRI time-series data View paper
- [46] Predicting neural activity patterns associated with sentences using a neurobiologically motivated model of semantic representation View paper
- [47] Learning deep temporal representations for fMRI brain decoding View paper
- [48] Learning Image Derived PDE-Phenotypes from fMRI Data View paper
- [49] Hierarchical Characterization of Brain Dynamics via State Space-based Vector Quantization View paper
- [50] Temporal Graph Representation Learning for Autism spectrum disorder Brain Networks View paper
- [51] A Survey on Joint Embedding Predictive Architectures and World Models View paper
- [52] The disturbed functional brain network in major depressive disorder identified by graph theory analysis View paper
- [53] RTGMFF: Enhanced fmri-based brain disorder diagnosis via roi-driven text generation and multimodal feature fusion View paper
- [54] Hierarchical Encoding and Fusion of Brain Functions for Depression Subtype Classification View paper
- [55] BACE: Behavior-adaptive connectivity estimation for interpretable graphs of neural dynamics View paper
- [56] EEG emotion classification based on graph convolutional network View paper
- [57] Cognitmoe: A cognition-aware collaborative multi-expert network for bipolar disorder diagnosis View paper
- [58] TFAGL: A novel agent graph learning method using time-frequency EEG for major depressive disorder detection View paper
- [59] Disorder-specific neurodynamic features in schizophrenia inferred by neurodynamic embedded contrastive variational autoencoder model View paper
- [60] Reconfiguration of dynamic largeâscale brain network functional connectivity in generalized tonicâclonic seizures View paper
- [61] Joint learning of multi-level dynamic brain networks for autism spectrum disorder diagnosis View paper
- [62] Self-supervised learning for electroencephalogram: A systematic survey View paper
- [63] Self-supervised learning of brain dynamics from broad neuroimaging data View paper
- [64] Population transformer: Learning population-level representations of neural activity View paper
- [65] BENDR: Using transformers and a contrastive self-supervised learning task to learn from massive amounts of EEG data View paper
- [66] Explainable Self-Supervised Dynamic Neuroimaging Using Time Reversal View paper
- [67] Self-supervised Learning for Encoding Between-Subject Information in Clinical EEG View paper
- [68] LLEDAâLifelong self-supervised domain adaptation View paper
- [69] GMAEEG: A Self-Supervised Graph Masked Autoencoder for EEG Representation Learning View paper
- [70] Longitudinal self-supervised learning View paper
- [71] Adaptive-Similarity-Based Brain Dynamic Functional Connectivity with Spatial-Temporal Attention and Domain Adaptation for Schizophrenia Diagnosis View paper