Brain-Environment Cross-Attention (BECA) Meta-Matching: A New Perspective of Brain Connectome Zero-Shot Learning

Ziquan Wei[1 [0000−0001−6553−4482]], Tingting Dan[1 [0000−0001−6936−2649]], and Guorong Wu[1 [0000−0002−0550−6145]]

University of North Carolina at Chapel Hill, Chapel Hill NC 27599, USA
{ziquanw,tingting_dan,grwu}@med.unc.edu

Abstract. Zero-shot learning (ZSL) is critical for deep learning models being deployed in unseen downstream applications. Given that fMRI studies of the human connectome with respect to cognitive disorders are boutique and lack sufficient labeled samples, a reliable and interpretable ZSL technology is necessary to empower the brain foundation model for clinical applications. Although self-supervised learning and transfer learning on data reconstruction and semantic information, respectively, have achieved success in ZSL performance for language and vision, little attention has been paid to the recognition of brain disordering. In contrast to stereotypical language or vision data, the human brain is a dynamically wired system where distributed regions communicate through functional connectivity and spontaneously respond to stimuli from environmental exposures. Thus, functional neuroimages are often associated with phenotypic traits underlying brain-environment interactions (BEIs), such as cognitive states and clinical outcomes. By capitalizing on large-scale functional neuroimages as well as a rich collection of BEI data, we break the frame of self-supervised and transfer learning by using logical regression as the pre-training objective for brain connectome. We formulate ZSL on unseen classes by identifying a reliable matching across environmental variables, which is derived from a decoder-only model for BEI prediction from functional connectivity. Together, we present a novel learning schema of brain-environment cross-attention (BECA) meta-matching, which is a new horizon of ZSL for brain connectome. In experiments, all fMRI data in HCP-young adult and HCP-aging datasets are utilized for pre-training, and BECA is evaluated on disease early diagnosis of Autism, Parkinson's disease, and Schizophrenia, where promising results indicate the great potential to facilitate current neuroimaging applications in clinical routines.

Keywords: Zero-shot learning · Brain foundation model · fMRI · Functional connectivity · Autism · Parkinson's · Schizophrenia.

1 Introduction

Zero-shot and few-shot learning (ZSL and FSL) have demonstrated remarkable success in various domains, including natural language processing and computer

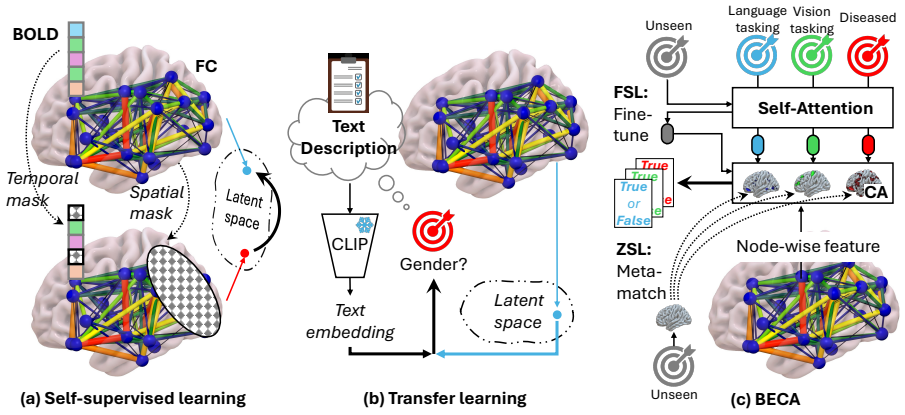

Fig. 1. Previous solutions and ours for zero-shot learning (ZSL) and few-shot learning (FSL) on brain connectome. (a) Self-supervised learning inherits the idea of which in large foundation models for better latent feature extraction. (b) Transfer learning brings the power of a pre-trained CLIP model to brain embedding, where ZSL is achieved by mapping unseen tasks to the continuous text embedding space. (c) Brain-Environment Cross-Attention (BECA) learns from the brain logical state existing in large-scale datasets, where finetuning and meta-matching are implemented by injecting new tokens and t-test, respectively. 'CA' stands for cross-attention.

vision. Their capability of knowledge transfer, composition, and generative prediction is impressive and widely applied in real world applications. To that end, tremendous efforts have been made to pre-train large models on extensive unlabeled fMRI data using scalable self-supervised and transfer learning for brain foundation models. This is under exploration since BrainLM [16] applied masked autoencoder, then followed by various applications of general masking strategies [26, 9, 28] and transfer learning [12] for phenotypic prediction, e.g. gender and age. Nonetheless, clinical outcome is rarely the focus in these foundation models. Meanwhile, in the recognition of brain disorders, where labeled samples are often limited and expensive to obtain, a robust ZSL framework is vital for deploying brain foundation models.

ZSL requires the model to generalize to unseen data and tasks without direct supervision. As shown in Fig. 1, previous work can be categorized into two groups [19], (a) self-supervised learning and (b) transfer learning. Self-supervised learning is derived from self-regressive methodology that has shown the impressive performance of generative models in computer vision and natural language processing. Although brain-dedicated masking [16, 26] and advanced encoding embeddings [9] are proposed to better extract latent features, reconstruction supervision is suboptimal for downstream tasks. The reason is that these models are purposed the same as the image/language to reconstruct the raw signal from its masked versions, leaving the complex clinical outcomes not well utilized. Furthermore, the prediction of unseen classes is commonly done by an average of

the prediction of seen classes [28]. On the other hand, [12] introduced text-based neuroscience knowledge of a pre-trained Biomed CLIP [29] to brain connectome transfer learning that shifts latent features to text embedding space. This keeps the benefit of the CLIP model that bridges the gap between seen and unseen classes through semantic information [13, 19], so that brain connectome ZSL is reliable for arbitrary phenotypic prediction. However, the relationship between seen and unseen tasks is bounded by the language model and cannot be learned directly from large-scale neuroimaging data.

Unlike language or vision data that require manual annotation, brain fMRI is often associated with nonimaging phenotypes regarding brain-environment interactions (BEIs), e.g. cognitive state and mental health, as shown by colored aims in Fig. 1 (c), which are defined as variable brain status and are commonly recorded during fMRI acquisition. According to the nature of fMRI, which always has such brain logical state, a new perspective of the brain foundation model is necessary to learn from logical regression instead of self-regression or transfer learning. Thus, FSL and ZSL on unseen classes from boutique studies can be formulated as a meta-matching, the ZSL manner, or a transformation, the FSL manner, of the rich environmental variables existing in large-scale datasets, even subjects of which are all healthy.

To this end, as shown in Fig. 1 (c), we propose Brain-Environment Cross-Attention (BECA) as a framework for brain connectome ZSL and FSL from brain logical state under various environments in large-scale datasets. *Firstly*, the functional connectivity (FC) as input feature is directly multiplied with tokenized environmental variables denoted by colored capsules, which are produced by a self-attention between BEI tokens. *Then*, the cross-attention (CA) denoted by the brain surface as the brain activation map combines the FC of activated nodes as output with the supervision of the truth of BEI. *Lastly*, FSL is achieved by injecting a new token for the unseen task, that is, a BEI, to join the self-attention with seen BEIs. In parallel, ZSL meta-matches the brain activation map of the unseen to all seen data via t-test so that the suspiciousness can be represented by p value.

Our contributions are threefold: (1) A decoder-only brain foundation model for brain disordering recognition with released model weights pre-trained by two HCP datasets. (2) A statistical meta-matching for brain connectome zero-shot learning based on the proposed foundation model. (3) Promising performance in the recognition of Autism, Parkinson's disease, and Schizophrenia using fine-tuning data from 0% to 100%.

2 Motivations

The dynamic signal of brain functional MRI (fMRI) is a blood-oxygen-dependent level (BOLD). BOLD signal, which is influenced by a mixture of factors and distorted by non-neuronal fluctuations, has a relatively low signal-to-noise ratio (SNR) [3]. Brain connectomes, on the other hand, increase the SNR in raw signals by representing brain activity via the Pearson correlation coefficient (PCC),

which is also called functional connectivity (FC). Various works have shown superior performance using FC compared to the raw BOLD signal for downstream applications. For example, benchmark papers [6, 20, 7] evaluated the performance by using the BOLD or the correlation as the input, and the BOLD signal has consistently demonstrated lower accuracy. In addition, both static FC and dynamic FC (using the sliding window technique) outperform the BOLD [24]. Therefore, we focus on the brain connectome ZSL for disease early diagnosis.

The idea of meta-matching between unseen data and seen data for brain connectome analysis is first proposed in [11], where a diverse range of environmental variables have been found to be inter-correlated with each other given the static FC. Even without fine-tuning, a kernel ridge regression (KRR) can outperform the trained version after a basic meta-matching between phenotypic labels. This motivates us to further develop the idea of meta-matching between BEIs via cross-attention. Simultaneously, a recent work [18] replaces the fully connected multilayer perceptron (MLP) with the cross-attention decoder between image and logical prediction to gain better interpretability. This demonstrates the feasibility of learning the brain cross-attention with respect to the logical BEI for downstream ZSL given the high inter-correlation between BEIs [11].

3 Methods

To ensure the robustness of BECA map and so does the ZSL, we pre-train a large brain foundation model coined as large connectome model (LCM) on multiple datasets including healthy and diseased subjects with various brain status. Afterward, ZSL and FSL are implemented by the t-test between BECA maps of testing and pre-training data and injecting new tokens, respectively.

3.1 Model Pre-training: A Large Connectome Model

As shown in Fig. 2 left part, we build the LCM by stacking multiple layers of the pair of self- and cross-attention. Assume that FC is denoted by $\mathbf{M} \in \mathbb{R}^{N\times N}$ with N the number of nodes, and BEI is denoted by $\mathbf{V} \in \mathbb{R}^{P\times E}$ with P the number of BEI and E the dimensionality of randomly initialized token vector. Then, l^{th} layer updates BEI token firstly via self-attention as follows

$$\mathbf{V} = \texttt{Softmax}\left((\mathbf{V}\bar{\boldsymbol{\alpha}}_h)(\mathbf{V}\bar{\boldsymbol{\beta}}_h)^T/\sqrt{D}\right)(\mathbf{V}\bar{\boldsymbol{\gamma}}_h)\,, \tag{1}$$

where $\bar{\boldsymbol{\alpha}}_h, \bar{\boldsymbol{\beta}}_h, \bar{\boldsymbol{\gamma}}_h \in \mathbb{R}^{E\times D}$ are learnable parameters of linear layers in the self-attention block, h is the head index, and D is the hidden channel. BEI tokens are then transformed from FC tokens according to the BECA map $\boldsymbol{\mathcal{B}} \in \mathbb{R}^{P\times N}$ as follows

$$\mathbf{V} = \boldsymbol{\mathcal{B}}\,(\mathbf{M}\hat{\boldsymbol{\gamma}}_h)\,,\ \boldsymbol{\mathcal{B}} = \texttt{Softmax}\left((\mathbf{M}\hat{\boldsymbol{\alpha}}_h)(\mathbf{V}\hat{\boldsymbol{\beta}}_h)^T/\sqrt{D}\right)^T \tag{2}$$

where $\hat{\boldsymbol{\alpha}}_h, \hat{\boldsymbol{\gamma}}_h \in \mathbb{R}^{N\times D}, \hat{\boldsymbol{\beta}}_h \in \mathbb{R}^{E\times D}$ are learnable parameters of linear layers in the cross-attention block. Lastly, the BEI prediction readout is performed by an

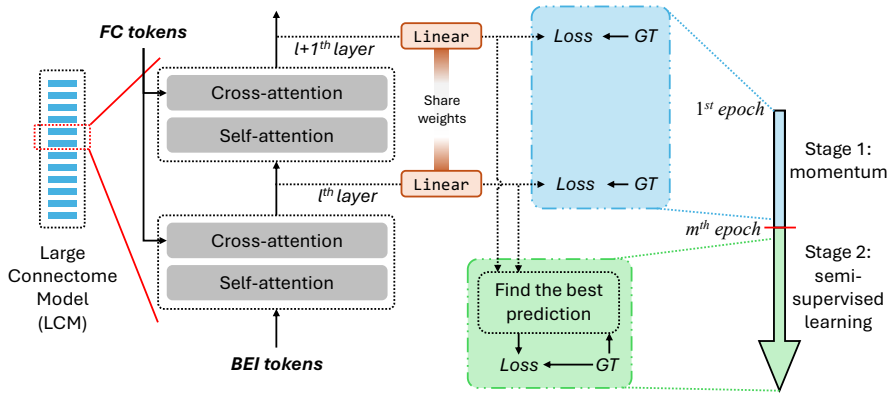

Fig. 2. Model pre-training for LCM has two stages. Stage one, gain momentum for LCM in the first m epochs via supervising on all layers. Stage two, a semi-supervision updates model weights of partial layers to prevent overfitting on a single BEI. Loss is implemented by cross-entropy.

additional linear layer $\rho : \mathbb{R}^{P\times D} \to \mathbb{R}^{P\times 1}$ marked by the orange block in Fig. 2. Note that the bias in linear layers is omitted in this section for clarity.

Then, as shown in Fig. 2 right, LCM is pre-trained with two stages to prevent overfitting on a single BEI. BEIs differ in the complexity of feature representation, e.g. Parkinson's disease at different stages of treatment and function-specified tasking state. In this case, we propose that LCM predicts each BEI at different layers of the model via a semi-supervision. Obviously, the initialization of semi-supervised learning is important to have a proper starting point and a correct learning direction. Thus, training the LCM has two stages, (1) utilize the average prediction from all layers to update the LCM parameters in the first m epochs and (2) examine only the best prediction in the rest of the epochs. Namely, stage 1 produces a 'momentum' that can push the training of LCM to the correct direction, and hence LCM can achieve a diverse and correct feature representation for different BEIs in the following stage 2. During testing, the layer with the highest prediction is used for the final output of the LCM.

3.2 Zero-Shot Learning: Cross-Attention Meta-Matching

As shown in Fig. 3, ZSL is implemented via meta-matching (MM) unseen tasks with seen tasks via an independent t-test determining if correlations of the BECA map between diverse populations are significantly different. MM is a t-test between two groups of PCC between $\boldsymbol{\mathcal{B}}_i$ of the unseen subject and the positive and negative seen subjects, where i is the index of BEI and p value by the t-test reflects the suspicion that the brain is diseased. The probability of prediction s

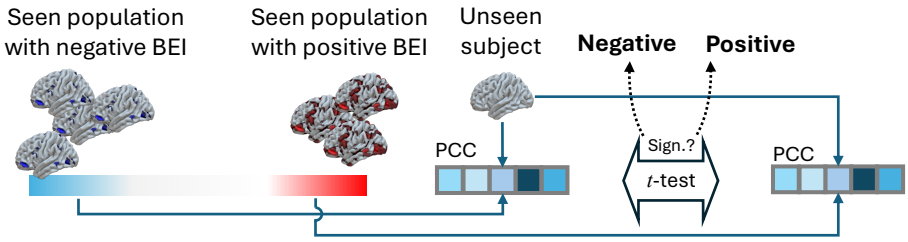

Fig. 3. Meta-Matching (MM) between BECA maps for ZSL on a subject with unseen BEI, where the pial surface indicates a BECA map, 'PCC' stands for Pearson correlation coefficient, and 'Sign.' stands for significant.

is then calculated with the regularized p value as follows

$$s = \begin{cases} 0.5 + \frac{p - min(p)}{2(max(p) - min(p))}, & \text{if } p > \tau \\ \frac{p - min(p)}{2(max(p) - min(p))}, & \text{otherwise} \end{cases} \tag{3}$$

where τ is the threshold for p value. By default $\tau = 0.05$ and the average p is used with $i = 1, \cdots, P$. Since our ZSL considered the prediction as binary classification, Eq. 3 can restrict s into ranges $[0, 0.5]$ and $[0.5, 1.0]$, respectively, stratifying the probability of healthy and disordered brains.

Given the design of our ZSL and model architecture, we propose two types of FSL. First, the design of LCM allows it to be easily fine-tuned with unseen datasets that have Q new phenotypes by concatenating new tokens as $[\mathbf{V}, \hat{\mathbf{V}}]$ with randomly initialized $\hat{\mathbf{V}} \in \mathbb{R}^{Q \times E}$. We denote the first type as BECA (Sup.). The second is to search for the hyperparameters i and τ of MM as described in Section 3.2 with few training data. This type is denoted by BECA (MM).

4 Experiments

We partition brain regions using the AAL atlas [22] through all experiments. The pre-training data is preprocessed with FMRIprep [10]. Performance is evaluated with 10-fold cross-validation (CV).

Datasets and implementations: HCP-young adult (HCPYA) [23], HCP-aging (HCPA) [2], and Alzheimer's Disease Neuroimaging Initiative (ADNI) [25] are used to pre-train LCM. HCPs are instrumental in task recognition research, offering a comprehensive view of young adults and the aging process, respectively. HCPYA has over 1100 healthy young adults, and each of which is associated with seven human behaviors, e.g. language and working memory. HCPA includes data from 717 subjects, encompassing fMRI records ($n = 4,863$) with human behaviors associated with memory, sensory-motor, and the resting state. ADNI provides BEI on brain disorders with a collection of pre-processed fMRI ($n = 138$) including clinical diagnostic labels of Alzheimer's Disease (AD). These pre-training datasets result in 13 BEI tokens (7 tasks in HCPYA, 3 tasks

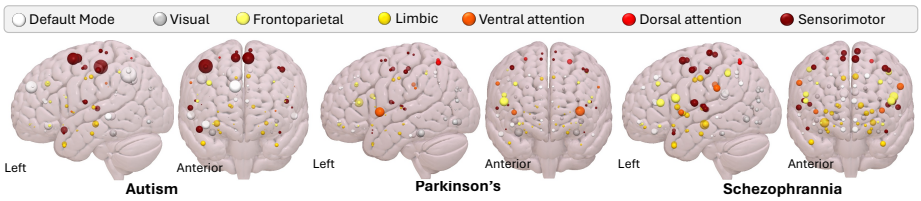

Fig. 4. The average BECA map of all test data at the readout layer. The node size indicates the relative attention weight.

in HCPA, 1 resting state, 1 healthy state, and 1 AD state) implemented in our model. Autism Brain Imaging Data Exchange (ABIDE) [27], Parkinson's Progression Markers Initiative (PPMI) [27], and Schizophrenia (SZ) are pre-processed and used for ZSL and FSL evaluation, where ABIDE contains 537/488 ($n = 1,025$) Control Normal (CN)/Autism subjects, PPMI contains 15/113 ($n = 128$) CN/Parkinson's, and SZ is in-house data of this work with 159/30 ($n = 189$) high-risk CN (sibling of diseased subject)/SZ. LCM is implemented with 8 heads, 32 layers, and 2048 hidden channels containing 1.2B learnable parameters. The pre-training took 14.4 hours using a RTX 6000 Ada GPU. Codes and model weights can be found here[1]. We compare with one state-of-the-art (SOTA) brain connectome foundation model, BrainMass (30M) [28], along with some SOTA brain-dedicated models without pre-training, BolT [1] (1.6M) and NeuroPath [24] (0.7M). BrainMass is pre-trained with our data since their released model uses a different atlas, hidden channels are all set as the same for a fair comparison.

4.1 Classification Results

Evaluation metrics for the classification experiments are the average weighted F1 score and the Area Under ROC Curve (AUC). Since datasets are boutique with a class imbalance issue, F1 score indicating the precision and recall sometimes cannot reflect performance differences for PPMI and SZ. In contrast, AUC can quantify the performance of the predicted probability. We test models with 0%, i.e., the ZSL, 1%, 10%, 50%, and 100% training data in each CV fold, where each percentage has at least one sample for each class. Note that it is one-shot learning in the case 1% for PPMI and SZ due to the limited sample size.

The performance of ZSL and FSL is listed in Table 1. Compared to the SOTA models, the proposed BECA always has the best F1 score for each section on all datasets except for 1% SZ. BECA shows a minor shortage of 0.31% lower F1 but 8.26% higher AUC than BolT on 1% SZ. In general, BECA demonstrated a better ZSL performance than BrainMass using an average guess, and it performed a new SOTA performance when using all training data.

[1] https://github.com/Chrisa142857/brain_network_decoder/tree/zero-shot-learning

Table 1. Zero-shot and few-shot learning performance, where 'FT %' stands for the percentage of finetuning data size, and **bold** indicates the best score in a section.

FT %		ABIDE		PPMI		SZ	
		F1 score	AUC	F1 score	AUC	F1 score	AUC
0%	BrainMass	$33.92_{\pm 4.91}$	$50.51_{\pm 6.35}$	$6.19_{\pm 5.82}$	$27.02_{\pm 11.26}$	$\mathbf{78.60}_{\pm 8.74}$	$41.67_{\pm 20.44}$
	BECA (MM)	$\mathbf{39.70}_{\pm 5.56}$	$\mathbf{51.50}_{\pm 11.49}$	$\mathbf{16.80}_{\pm 14.67}$	$\mathbf{38.10}_{\pm 18.16}$	$\mathbf{78.60}_{\pm 9.77}$	$\mathbf{52.15}_{\pm 25.35}$
1%	BolT	$47.73_{\pm 6.47}$	$\mathbf{58.42}_{\pm 3.61}$	$71.40_{\pm 16.46}$	$40.55_{\pm 28.81}$	$\mathbf{72.55}_{\pm 17.68}$	$73.69_{\pm 16.03}$
	NeuroPath	$45.73_{\pm 12.62}$	$51.65_{\pm 3.83}$	$73.69_{\pm 14.17}$	$51.19_{\pm 23.90}$	$71.04_{\pm 19.49}$	$69.65_{\pm 12.54}$
	BrainMass	$43.08_{\pm 2.58}$	$49.88_{\pm 4.46}$	$72.32^{*}_{\pm 14.67}$	$50.00^{*}_{\pm 0.00}$	$4.30^{*}_{\pm 2.75}$	$50.00^{*}_{\pm 0.00}$
	BECA (MM)	$44.91_{\pm 5.96}$	$50.00_{\pm 2.99}$	$52.03_{\pm 25.69}$	$40.49_{\pm 41.04}$	$50.32_{\pm 34.96}$	$57.61_{\pm 20.66}$
	BECA (Sup.)	$\mathbf{54.06}_{\pm 3.30}$	$55.63_{\pm 4.91}$	$\mathbf{74.36}_{\pm 16.74}$	$\mathbf{66.00}_{\pm 25.46}$	$72.24_{\pm 13.33}$	$\mathbf{81.95}_{\pm 16.27}$
10%	BolT	$58.97_{\pm 4.89}$	$61.82_{\pm 5.71}$	$69.47_{\pm 18.41}$	$62.69_{\pm 25.21}$	$79.70_{\pm 9.76}$	$\mathbf{84.56}_{\pm 16.68}$
	NeuroPath	$54.12_{\pm 6.73}$	$55.80_{\pm 4.60}$	$75.12_{\pm 15.62}$	$59.21_{\pm 32.71}$	$78.60_{\pm 9.77}$	$77.12_{\pm 15.31}$
	BrainMass	$46.75_{\pm 7.62}$	$43.23_{\pm 9.78}$	$75.72_{\pm 14.22}$	$38.74_{\pm 24.38}$	$78.60_{\pm 8.74}$	$36.69_{\pm 27.68}$
	BECA (MM)	$50.14_{\pm 8.23}$	$50.40_{\pm 8.57}$	$43.18_{\pm 23.77}$	$24.58_{\pm 29.47}$	$78.84_{\pm 9.75}$	$44.73_{\pm 25.25}$
	BECA (Sup.)	$\mathbf{65.43}_{\pm 3.64}$	$\mathbf{65.83}_{\pm 3.46}$	$\mathbf{76.57}_{\pm 16.47}$	$\mathbf{71.55}_{\pm 36.50}$	$\mathbf{81.63}_{\pm 8.58}$	$74.50_{\pm 10.93}$
50%	BolT	$63.12_{\pm 2.81}$	$63.25_{\pm 3.60}$	$\mathbf{77.41}_{\pm 14.10}$	$63.95_{\pm 20.44}$	$80.16_{\pm 12.68}$	$76.91_{\pm 24.10}$
	NeuroPath	$60.23_{\pm 2.77}$	$60.13_{\pm 3.59}$	$75.72_{\pm 15.90}$	$62.60_{\pm 25.59}$	$80.16_{\pm 12.68}$	$\mathbf{82.47}_{\pm 12.31}$
	BrainMass	$60.12_{\pm 4.93}$	$66.01_{\pm 3.40}$	$75.72_{\pm 14.22}$	$40.55_{\pm 21.55}$	$78.60_{\pm 8.74}$	$32.09_{\pm 13.65}$
	BECA (MM)	$53.36_{\pm 3.11}$	$52.24_{\pm 4.49}$	$61.99_{\pm 31.43}$	$44.70_{\pm 32.66}$	$77.52_{\pm 9.58}$	$56.76_{\pm 16.83}$
	BECA (Sup.)	$\mathbf{66.24}_{\pm 4.41}$	$\mathbf{66.41}_{\pm 4.74}$	$75.72_{\pm 15.90}$	$\mathbf{89.98}_{\pm 7.94}$	$\mathbf{81.73}_{\pm 11.48}$	$70.52_{\pm 20.90}$
100%	BolT	$64.41_{\pm 1.97}$	$64.05_{\pm 3.86}$	$74.92_{\pm 15.18}$	$68.02_{\pm 26.84}$	$81.14_{\pm 9.75}$	$\mathbf{83.42}_{\pm 8.48}$
	NeuroPath	$65.04_{\pm 5.13}$	$64.53_{\pm 6.28}$	$76.03_{\pm 16.06}$	$64.98_{\pm 16.60}$	$81.74_{\pm 8.90}$	$78.98_{\pm 14.01}$
	BrainMass	$63.80_{\pm 5.69}$	$69.65_{\pm 4.39}$	$75.72_{\pm 14.22}$	$72.10_{\pm 20.73}$	$78.60_{\pm 8.74}$	$24.94_{\pm 19.34}$
	BECA (MM)	$48.31_{\pm 3.05}$	$48.38_{\pm 5.29}$	$73.23_{\pm 16.21}$	$31.26_{\pm 33.01}$	$72.81_{\pm 7.75}$	$53.78_{\pm 24.43}$
	BECA (Sup.)	$\mathbf{74.40}_{\pm 14.36}$	$\mathbf{77.98}_{\pm 12.65}$	$\mathbf{86.08}_{\pm 11.28}$	$\mathbf{89.81}_{\pm 16.04}$	$\mathbf{83.61}_{\pm 6.01}$	$74.35_{\pm 15.32}$

*Note: Predictions by BrainMass are identical with one-shot learning.

4.2 BECA Map Visualization

The visualizations of the average BECA map of test data with the same label in ABIDE, PPMI, and SZ are shown in Fig. 4. We can observe BECA map showing activated in default mode (DMN) and sensorimotor networks (SMN) for ABIDE, which aligns with current neuroscience knowledge [17, 30]. For PPMI, BECA is attentive to ventral attention (VAN) [21] and frontoparietal (FPN) [4]. For SZ, BECA also agrees with [5, 14, 15, 8] to be attentive to DMN, FPN, SMN, VAN, and limbic.

5 Conclusion

In conclusion, we propose a novel ZSL framework coined as brain-environment cross-attention (BECA) for brain disease early diagnosis. Unlike language or

vision data that require manual annotation, brain fMRI has nonimaging phenotypes regarding brain-environment interactions (BEIs), e.g. cognitive state and mental health. Thus, we break the frame of self-supervised and transfer learning for brain connectome by formulating FSL and ZSL on unseen classes as a meta-matching and a transformation, respectively, of the rich environmental variables existing in large-scale fMRI datasets, even subjects of which are all healthy. We pre-train an LCM on multiple large-scale datasets, and test BECA on ABIDE, PPMI, and SZ, where promising performance using data from 0% to 100% illustrates the great potential to facilitate current neuroimaging applications in clinical routines. We released the LCM weights pre-trained on HCPs and ADNI.

Disclosure of Interests

The authors have no competing interests in the paper as required by the publisher.

References

1. Bedel, H.A., Sivgin, I., Dalmaz, O., Dar, S.U., Çukur, T.: Bolt: Fused window transformers for fmri time series analysis. Medical Image Analysis **88**, 102841 (2023)
2. Bookheimer, S.Y., Salat, D.H., Terpstra, M., Ances, B.M., Barch, D.M., Buckner, R.L., Burgess, G.C., Curtiss, S.W., Diaz-Santos, M., Elam, J.S., et al.: The lifespan human connectome project in aging: an overview. Neuroimage **185**, 335–348 (2019)
3. Caballero-Gaudes, C., Reynolds, R.C.: Methods for cleaning the bold fmri signal. Neuroimage **154**, 128–149 (2017)
4. Cascone, A.D., Langella, S., Sklerov, M., Dayan, E.: Frontoparietal network resilience is associated with protection against cognitive decline in parkinson's disease. Communications biology **4**(1), 1021 (2021)
5. Chen, J., Müller, V.I., Dukart, J., Hoffstaedter, F., Baker, J.T., Holmes, A.J., Vatansever, D., Nickl-Jockschat, T., Liu, X., Derntl, B., et al.: Intrinsic connectivity patterns of task-defined brain networks allow individual prediction of cognitive symptom dimension of schizophrenia and are linked to molecular architecture. Biological psychiatry **89**(3), 308–319 (2021)
6. Cui, H., Dai, W., Zhu, Y., Kan, X., Gu, A.A.C., Lukemire, J., Zhan, L., He, L., Guo, Y., Yang, C.: Braingb: a benchmark for brain network analysis with graph neural networks. IEEE transactions on medical imaging **42**(2), 493–506 (2022)
7. Ding, J., Dan, T., Wei, Z., Cho, H., Laurienti, P.J., Kim, W.H., Wu, G.: Machine learning on dynamic functional connectivity: Promise, pitfalls, and interpretations. arXiv preprint arXiv:2409.11377 (2024)
8. Dong, D., Wang, Y., Chang, X., Luo, C., Yao, D.: Dysfunction of large-scale brain networks in schizophrenia: a meta-analysis of resting-state functional connectivity. Schizophrenia bulletin **44**(1), 168–181 (2018)
9. Dong, Z., Li, R., Wu, Y., Nguyen, T.T., Chong, J.S.X., Ji, F., Tong, N.R.J., Chen, C.L.H., Zhou, J.H.: Brain-jepa: Brain dynamics foundation model with gradient positioning and spatiotemporal masking. arXiv preprint arXiv:2409.19407 (2024)
10. Esteban, O., Markiewicz, C.J., Blair, R.W., Moodie, C.A., Isik, A.I., Erramuzpe, A., Kent, J.D., Goncalves, M., DuPre, E., Snyder, M., et al.: fmriprep: a robust preprocessing pipeline for functional mri. Nature methods **16**(1), 111–116 (2019)

11. He, T., An, L., Chen, P., Chen, J., Feng, J., Bzdok, D., Holmes, A.J., Eickhoff, S.B., Yeo, B.T.: Meta-matching as a simple framework to translate phenotypic predictive models from big to small data. Nature neuroscience **25**(6), 795–804 (2022)
12. He, Z., Li, W., Liu, Y., Liu, X., Han, J., Zhang, T., Yuan, Y.: Fm-app: Foundation model for any phenotype prediction via fmri to smri knowledge transfer. IEEE Transactions on Medical Imaging (2024)
13. Liu, X., Zhang, F., Hou, Z., Mian, L., Wang, Z., Zhang, J., Tang, J.: Self-supervised learning: Generative or contrastive. IEEE transactions on knowledge and data engineering **35**(1), 857–876 (2021)
14. Luvsannyam, E., Jain, M.S., Pormento, M.K.L., Siddiqui, H., Balagtas, A.R.A., Emuze, B.O., Poprawski, T.: Neurobiology of schizophrenia: a comprehensive review. Cureus **14**(4) (2022)
15. Mastrandrea, R., Piras, F., Gabrielli, A., Banaj, N., Caldarelli, G., Spalletta, G., Gili, T.: The unbalanced reorganization of weaker functional connections induces the altered brain network topology in schizophrenia. Scientific Reports **11**(1), 15400 (2021)
16. Ortega Caro, J., Oliveira Fonseca, A.H., Averill, C., Rizvi, S.A., Rosati, M., Cross, J.L., Mittal, P., Zappala, E., Levine, D., Dhodapkar, R.M., et al.: Brainlm: A foundation model for brain activity recordings. bioRxiv pp. 2023–09 (2023)
17. Padmanabhan, A., Lynch, C.J., Schaer, M., Menon, V.: The default mode network in autism. Biological Psychiatry: Cognitive Neuroscience and Neuroimaging **2**(6), 476–486 (2017)
18. Paul, D., Chowdhury, A., Xiong, X., Chang, F.J., Carlyn, D.E., Stevens, S., Provost, K.L., Karpatne, A., Carstens, B., Rubenstein, D., Stewart, C., Berger-Wolf, T., Su, Y., Chao, W.L.: A simple interpretable transformer for fine-grained image classification and analysis. In: The Twelfth International Conference on Learning Representations (2024), https://openreview.net/forum?id=bkdWThqE6q
19. Pourpanah, F., Abdar, M., Luo, Y., Zhou, X., Wang, R., Lim, C.P., Wang, X.Z., Wu, Q.J.: A review of generalized zero-shot learning methods. IEEE transactions on pattern analysis and machine intelligence **45**(4), 4051–4070 (2022)
20. Said, A., Bayrak, R., Derr, T., Shabbir, M., Moyer, D., Chang, C., Koutsoukos, X.: Neurograph: Benchmarks for graph machine learning in brain connectomics. Advances in Neural Information Processing Systems **36**, 6509–6531 (2023)
21. Tang, S., Wang, Y., Liu, Y., Chau, S.W., Chan, J.W., Chu, W.C., Abrigo, J.M., Mok, V.C., Wing, Y.K.: Large-scale network dysfunction in α-synucleinopathy: A meta-analysis of resting-state functional connectivity. EBioMedicine **77** (2022)
22. Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M.: Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain. Neuroimage **15**(1), 273–289 (2002)
23. Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E., Yacoub, E., Ugurbil, K., Consortium, W.M.H., et al.: The wu-minn human connectome project: an overview. Neuroimage **80**, 62–79 (2013)
24. Wei, Z., Dan, T., Ding, J., Wu, G.: Neuropath: A neural pathway transformer for joining the dots of human connectomes. In: The Thirty-eighth Annual Conference on Neural Information Processing Systems (2024)
25. Weiner, M.W., Veitch, D.P., Aisen, P.S., Beckett, L.A., Cairns, N.J., Cedarbaum, J., Donohue, M.C., Green, R.C., Harvey, D., Jack Jr, C.R., et al.: Impact of the alzheimer's disease neuroimaging initiative, 2004 to 2014. Alzheimer's & Dementia **11**(7), 865–884 (2015)

26. Wen, G., Cao, P., Liu, L., Yang, J., Zhang, X., Wang, F., Zaiane, O.R.: Graph self-supervised learning with application to brain networks analysis. IEEE Journal of Biomedical and Health Informatics **27**(8), 4154–4165 (2023)
27. Xu, J., Yang, Y., Huang, D., Gururajapathy, S.S., Ke, Y., Qiao, M., Wang, A., Kumar, H., McGeown, J., Kwon, E.: Data-driven network neuroscience: On data collection and benchmark. Advances in Neural Information Processing Systems **36**, 21841–21856 (2023)
28. Yang, Y., Ye, C., Su, G., Zhang, Z., Chang, Z., Chen, H., Chan, P., Yu, Y., Ma, T.: Brainmass: Advancing brain network analysis for diagnosis with large-scale self-supervised learning. IEEE Transactions on Medical Imaging (2024)
29. Zhang, S., Xu, Y., Usuyama, N., Xu, H., Bagga, J., Tinn, R., Preston, S., Rao, R., Wei, M., Valluri, N., et al.: Biomedclip: a multimodal biomedical foundation model pretrained from fifteen million scientific image-text pairs. arXiv preprint arXiv:2303.00915 (2023)
30. Zhang, Z., Chan, M.Y., Han, L., Carreno, C.A., Winter-Nelson, E., Wig, G.S., (ADNI, A.D.N.I., et al.: Dissociable effects of alzheimer's disease-related cognitive dysfunction and aging on functional brain network segregation. Journal of Neuroscience **43**(46), 7879–7892 (2023)