CARformer: Class-Aware Representation Learning for Small-Cohort, Imbalanced Psychiatric Disorder Classification in Brain MRI

Xingyue Fu[1], Ruoxi Deng[1], Euijoon Ahn[2,3], Mayuresh S. Korgaonkar[4,5], Jinman Kim[1(✉)]

[1] Biomedical Data Analysis and Visualisation (BDAV) Lab, School of Computer Science, The University of Sydney, Sydney, NSW 2006, Australia
[2] College of Science and Engineering, James Cook University, Cairns, QLD 4870, Australia
[3] Centre for AI and Data Science Innovation, James Cook University, Cairns, QLD 4870, Australia
[3] Brain Dynamics Centre, Westmead Institute for Medical Research, The University of Sydney, Westmead, NSW 2145, Australia
[4] Discipline of Psychiatry, Sydney Medical School, Westmead, NSW 2145, Australia
jinman.kim@sydney.edu.au

Abstract. Structural MRI (sMRI) has revealed morphological correlates associated with psychiatric disorders, motivating neuroimaging-based diagnostic support tools. However, multi-class classification from sMRI remains challenging in small-cohort psychiatric settings, where site-level datasets often contain only a few hundred subjects and exhibit class imbalance, and where subtle inter-class differences further complicate discrimination. We propose a class-aware representation learning framework (CARformer) that adapts domain-pretrained encoders for 3D sMRI multi-class psychiatric classification. The framework includes (1) a class-aware Bag-of-Queries (CA-BoQ) module that learns class-aligned query representations to aggregate class-specific dis-criminative evidence under inter-class differences, and (2) a Class-Aware Contrastive (CAC) learning strategy that strengthens class separation under imbalance. We evaluated the framework across two sites from a public brain MRI dataset to assess robustness under differences in diagnostic composition, class imbalance, and imaging acquisition protocols. Our results demonstrate consistent improvements in macro AUC and AUPRC over existing standard 3D Convolutional Neural Network (CNN) backbones and domain-pretrained fine-tuning frameworks across both sites, supporting the practical utility of class-aware adaptation of domain pretraining in small-cohort, imbalanced psychiatric imaging tasks. Source code is available at: https://github.com/xyfu33/CARformer.

Keywords: Multi-class Classification, Psychiatric Disease, Structural MRI.

1 Introduction

Mental health disorders place a significant burden on individuals and society. Severe psychiatric diseases such as bipolar disorder (BD) and schizophrenia (SZ) are leading

causes of disability worldwide and often require long-term, complex treatment and follow-up [1, 2]. In practice, psychiatric diagnosis relies largely on subjective clinical observation and self-reported symptoms [3, 4]. However, symptom overlap and comorbidities are common across disorders, particularly between major depressive disorder (MDD) and other severe psychiatric diseases, which can lead to diagnostic uncertainty and inconsistent decisions. These then expose patients to inappropriate treatment [1, 3]. Motivated by this, we study multi-disorder differential diagnosis involving MDD and severe disorders under realistic small-cohort settings.

With advances in neuroimaging, structural brain MRI (sMRI) studies have identified disorder-associated morphological abnormalities. Cross-disorder analyses report substantial shared variance in morphometric abnormality patterns across major psychiatric disorders, involving distributed cortical and subcortical regions, particularly fronto–limbic and striatal circuits [5, 6]. Disorder-specific deviations have also been highlighted, including stronger effects for MDD in the rostral anterior cingulate and medial orbitofrontal cortex, and for SZ in the superior temporal gyrus and thalamocortical regions [5, 7]. These findings motivate AI models that leverage sMRI for decision support while capturing subtle, disorder-specific differences needed for differential diagnosis among clinically overlapping psychiatric disorders.

Recent clinical AI studies in psychiatric neuroimaging predominantly formulate diagnosis as binary classification, typically distinguishing a single disorder from healthy controls. While effective for patient–control separation, this setting does not address the clinically important challenge of differentiating between disorders with overlapping phenotypes. Prior work has explored single-modality sMRI, particularly T1-weighted MRI, for binary schizophrenia classification [8], as well as multimodal graph neural networks integrating sMRI, functional MRI (fMRI), and clinical information for MDD diagnosis [9]. In contrast, multi-disorder classification remains less explored, partly due to the need for larger, well-curated cohorts to learn subtle inter-disease differences under class imbalance and heterogeneity. For example, a multi-class sMRI framework employing a patch-based hierarchical network required a curated dataset of 2,490 subjects for robust training and generalization [10]. More recently, domain pretraining and transfer learning have been introduced to mitigate limited-label regimes. Structural MRI encoders pretrained via functional connectivity network supervision have demonstrated improved downstream classification performance [11]. However, even on the full SRPBS cohort, performance in clinically imbalanced multi-class classification remains moderate [11], suggesting that effectively adapting pretrained representations for subtle multi-disorder discrimination remains an open challenge.

A key practical barrier is data availability and cross-site heterogeneity in acquisition protocols and scanners. The sMRI are not routinely acquired for all psychiatric patients during routine visits, which limits clinical data collection and availability [12]. Single-site datasets often contain only a few hundred subjects and exhibit pronounced class imbalance [12]. Multi-site collections further introduce heterogeneity, which is a well-known confounder AI models [13]. These factors make it difficult to train robust classifiers that reliably distinguish between closely related psychiatric disorders.

With the development of domain-pretrained models and transfer learning frameworks, valuable but limited clinical datasets can be exploited more effectively, enabling

improved performance under small-cohort and imbalanced settings. In this study, we propose a class-aware representation learning framework for multi-class psychiatric disorder classification from brain sMRI under small-cohort, imbalanced data distributions. Leveraging a domain-pretrained 3D Uniformer encoder, we introduce CARformer, which couples (i) a Class-Aware Bag-of-Queries (CA-BoQ) module that aggregates class-specific discriminative representations for classification and (ii) a Class-Aware Contrastive (CAC) learning strategy for transfer learning. We validated CARformer on two cohorts from distinct clinical sites with different disease compositions, acquisition protocols, and class distributions, and show improved performance compared with widely used baselines for psychiatric MRI classification. Overall, these findings indicate the potential of CARformer for multi-disorder psychiatric classification in small-cohort, imbalanced, single-modality settings, improving the practical representation learning from limited clinical sMRI data.

2 Method

2.1 Overview

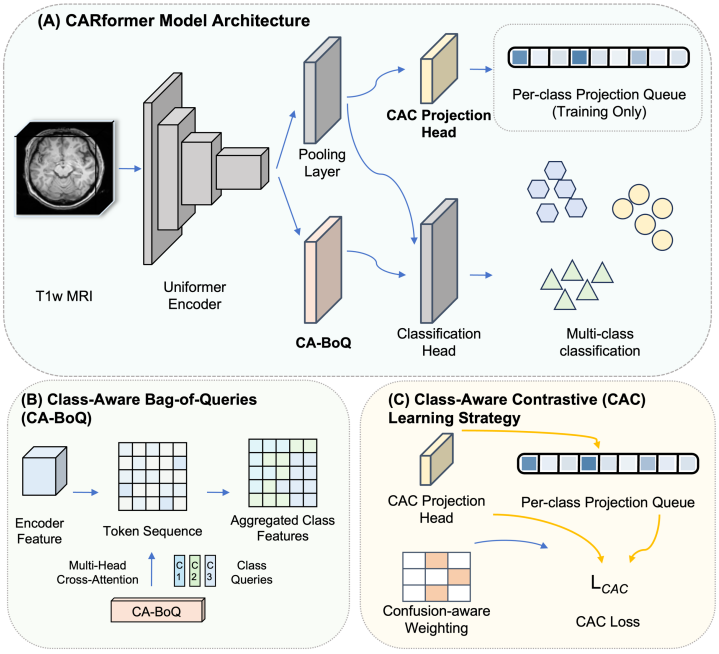

Fig. 1. CARformer framework with Class-Aware Bag-of-Queries and Class-Aware Contrastive learning strategy for sMRI 3-class classification.

The proposed CARformer framework integrates two complementary class-aware components to address multi-disorder discrimination under small-cohort, imbalanced psychiatric sMRI settings, specifically using T1w MRI: (1) CA-BoQ module that learns class-aligned query aggregation to extract class-specific discriminative evidence, and (2) a CAC learning strategy which improves inter-class separability through class-aware supervised contrastive supervision with confusion-informed negative reweighting. As illustrated in Fig. 1, the Uniformer encoder provides shared feature representations, while global pooling and CA-BoQ operate in parallel to produce complementary classification signals, and CAC is applied on globally pooled features to regularize representation learning.

2.2 CARformer

CARformer builds upon a domain-pretrained 3D Uniformer encoder trained on large-scale multiparametric brain MRI (mpMRI) [14]. The encoder extracts volumetric feature maps, from which global average pooling (GAP) produces a global logit Z_{GAP}, while the proposed CA-BoQ produces class-specific logits Z_{BoQ}using class-aligned queries and a diagonal per-query head. The final logits are fused by weighted additive blending with fixed $\alpha = 0.3$, enabling complementary class-aware and global feature learning while preserving the stability of the pretrained backbone.

2.3 Class-Aware Bag-of-Queries (CA-BoQ)

Inspired by the Bag-of-Queries mechanism [15], we introduce a class-aligned query aggregation module to capture class-specific discriminative features under small-cohort and imbalanced conditions. Let B denote the batch size, C the feature dimension, and $D, H, W$the spatial dimensions of the encoder feature map $F \in \mathbb{R}^{B\times C\times D\times H\times W}$, spatial locations are reshaped into tokens $X \in \mathbb{R}^{B\times T\times C}$, where $T = DHW$. We define K learnable query vectors $Q \in \mathbb{R}^{K\times C}$, where K denotes the number of classes. The queries attend to spatial tokens via multi-head cross-attention:

$$Z = \mathrm{MHA}(Q, X, X), Z \in \mathbb{R}^{B\times K\times C}, \qquad (1)$$

where MHA computes query–token attention weights and returns a query-specific weighted aggregation of token features. Layer normalization is applied to queries and tokens before attention and to the resulting query embeddings to ensure stable feature conditioning. Each query embedding is mapped independently through a shared diagonal slot-wise projection head:

$$\hat{y}_k = f(z_k), \qquad (2)$$

where $z_k \in \mathbb{R}^C$ is the embedding of the k-th query slot, and $f(\cdot)$is a shared linear projection applied to each slot to produce the corresponding class logit. A query-diversity penalty ($\lambda_{div} = 10^{-3}$) to reduce query collapse. The resulting CA-BoQ logits are combined with GAP-based classifier logits, enabling complementary global and class-aware feature learning.

2.4 Class-Aware Contrastive (CAC) learning

A class-aware supervised contrastive learning strategy applied to globally pooled encoder features. The CAC framework maintains a class-partitioned memory queue that stores recent feature embeddings for each class, enabling stable positive and negative sampling even when minibatches are imbalanced. To mitigate representation bias toward majority classes, anchor contributions are weighted inversely proportional to class frequency, increasing the influence of minority-class samples in the contrastive objective.

CAC introduces a confusion-aware negative weighting mechanism using a class-pair weighting matrix $\mathbf{W} \in \mathbb{R}^{K\times K}$, updated from validation predictions. Before the first validation pass, all entries are 1; after validation, W is row-normalized, clipped to [1,2], has diagonal entries reset to 1, and is exponentially smoothed with decay 0.5. During contrastive learning, W reweights only negative pairs between frequently confused classes, while positives remain weight 1. The overall training objective combines classification loss and CAC supervision:

$$L = L_{CE}^{adj} + \lambda L_{CAC}, \tag{3}$$

where L_{CE}^{adj} is the class-adjusted cross-entropy loss, $\mathcal{L}_{CAC}$ is the confusion-aware supervised contrastive loss, and λ is a balancing coefficient controlling the contribution of contrastive supervision.

3 Experiments and Result

3.1 Dataset Description

The Japanese Strategic Research Program for the Promotion of Brain Science (SRPBS) dataset is a publicly available, large-scale, multi-site neuroimaging cohort comprising resting-state fMRI and structural MRI across multiple psychiatric disorders [12]. In this study, we used two SRPBS sites, University of Tokyo Hospital (UTO) and Kyoto University TimTrio (KUT), because these are the only sites that include two of our target disorders together with healthy controls, enabling a three-class multi-class classification setting within each site. Details of disease composition, participant demographics, and imaging protocols are provided in Table 1. For each site, dataset splits undertaken stratified sampling to preserve the diagnostic distribution and to balance sex and age across splits.

Table 1. Participant demographics and sMRI acquisition details for the two sites from SRPBS dataset. Data are shown as means (standard deviation). M, men; W, women; HC, healthy controls; MDD, major depressive disorder; BD, bipolar disorder, SZ, schizophrenia spectrum disorders; UTO, University of Tokyo Hospital; KUT, Kyoto University TimTrio; FoV: Field of view; TR, repetition time; TE, echo time; TI, inversion time.

Demographic characteristics	Imaging protocols

		HC	MDD	BD	SZ	MRI scanner	FoV, mm	Matrix	Voxel size, mm3	TR, TE, TI (ms)	Flip angle, deg
UTO	No., M:W Age, y	96, 33:63 47.0 (15.5)	62, 36:26 38.7 (11.6)	41, 26:15 34.2 (9.1)		GE MR750w	240	256 × 256	1 × 1 ×1.2	7.7, 3.1, 400	11
KUT	No., M:W Age, y	159, 93:66 36.5 (13.6)	16, 10:6 42.6 (12.5)		45, 21:24 41.4 (10.8)	Siemens TimTrio	225 × 240	240 × 256	0.9375× 0.9375× 1.0	2000, 3.4, 990	8

3.2 Implementation Details

Our models were trained with PyTorch on NVIDIA RTX A6000 GPUs. Pre-trained BrainMVP Uniformer variants were used as initial weights [14]. Training employed the Adam [16] optimizer with a cosine-annealing learning-rate schedule (initial rate 3×10^{-4}) and decaying to $< 1 \times 10^{-6}$. With gradient accumulation, the effective batch size was 16, and all networks were trained for 200 epochs. Data augmentations included center cropping, random mirror flipping, random intensity offset, and random intensity scaling. We report macro-averaged and per-class, threshold-free one-vs-rest under the receiver operating characteristic curve (AUC) and area under the precision–recall curve (AUPRC).

3.3 Experiment Result

We evaluated CARformer against widely used 3D CNN baselines trained from scratch and domain-pretrained fine-tuning frameworks on the UTO (Table 2) and KUT (Table 3) sites. CARformer achieves the best macro-averaged AUC and AUPRC across both sites. On UTO, CARformer improves macro AUC and macro AUPRC over BrainMVP fine-tuning (0.771/0.708 vs. 0.675/0.565), with clear gains for the minority class (e.g., BD AUPRC 0.709 vs. 0.356). On KUT, CARformer improves macro AUC and macro AUPRC over PCRLv2 fine-tuning (0.895/0.836 vs. 0.726/0.547) and yields higher AUPRC for MDD (0.700) than all baselines. Overall, the results indicate consistent improvements under small-cohort, imbalanced multi-class settings. Qualitative attention maps from the CA-BoQ module (Fig. 2) further illustrate distinct class-specific spatial patterns across both sites.

Table 2. Comparison of macro-averaged and per-class AUC and AUPRC on the SRPBS-UTO site. The best and second-best results are bolded and underlined

Model	Training	Macro average		HC		MDD		BD	
		AUC	AUPRC	AUC	AUPRC	AUC	AUPRC	AUC	AUPRC
VGG 11 [8]	From scratch	0.546	0.471	0.620	0.596	0.595	**0.599**	0.422	0.218
ResNet18		0.572	0.496	0.620	0.596	0.643	0.492	0.453	<u>0.401</u>
PCRLv2 [17]	Fine tune	0.612	0.480	0.660	0.684	**0.738**	0.528	0.438	0.229
BrainMVP [14]		<u>0.675</u>	<u>0.565</u>	<u>0.710</u>	<u>0.774</u>	<u>0.690</u>	0.566	<u>0.625</u>	0.356

CARformer (Ours)	**0.771**	**0.708**	**0.810**	**0.822**	**0.738**	0.594	**0.765**	**0.709**

Table 3. Comparison of macro-averaged and per-class AUC and AUPRC on the SRPBS-KUT site. The best and second-best results are bolded and underlined

Model	Training	Macro average		HC		MDD		SZ	
		AUC	AUPRC	AUC	AUPRC	AUC	AUPRC	AUC	AUPRC
VGG 11 [8]	From scratch	0.701	0.498	0.750	0.886	0.643	0.163	0.711	0.445
ResNet18		0.887	0.725	**0.946**	**0.977**	0.714	0.196	**1.00**	**1.00**
PCRLv2 [17]	Fine tune	0.726	0.547	0.741	0.877	0.714	0.327	0.722	0.436
BrainMVP [14]		0.613	0.484	0.630	0.612	0.631	0.472	0.578	0.368
CARformer (Ours)		**0.895**	**0.836**	0.902	0.952	**0.929**	**0.700**	0.856	0.856

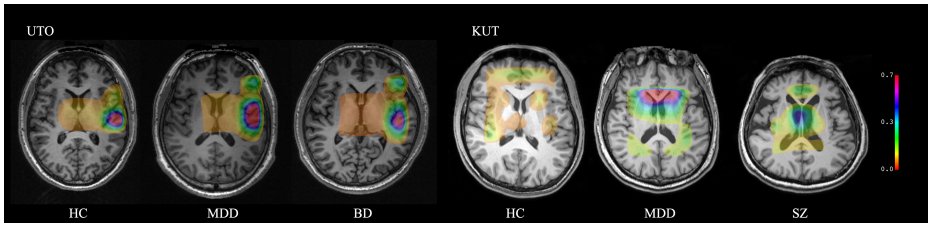

Fig. 2. Query-specific spatial attention maps from the CA-BoQ module within the UTO and KUT differential-diagnosis tasks, visualizing task-specific class attention.

3.4 Ablation Study

We conducted an ablation study to evaluate the contributions of the proposed class-aware components (CA-BoQ and CAC) and domain pretraining (Table 4). Both CA-BoQ and CAC improve performance under small-cohort, imbalanced multi-class classification, confirming their effectiveness in enhancing class-specific feature discrimination and inter-class separability. Notably, these improvements remain evident even without domain pretraining (i.e., with ImageNet-1K initialization), indicating that the class-aware modules provide substantial benefits beyond initialization. Domain pretraining further improves performance when combined with CA-BoQ and CAC, demonstrating complementary gains. Overall, the full model achieves the best and most consistent performance across both sites.

Table 4. Ablation analysis of proposed components of classification performance on two sites. *(Domain-pretrain ✓: brain MRI pretraining; Domain-pretrain ✗: ImageNet-1K initialization.)*

			UTO		KUT	
			Macro average		Macro average	
CA-BoQ	CAC	Domain-pretrain	AUC	AUPRC	AUC	AUPRC
✓	✓	✓	0.771	0.708	0.895	0.836

✗	✓	✓	0.726	0.642	0.825	0.678
✗	✗	✓	0.704	0.611	0.730	0.553
✓	✓	✗	0.761	0.624	0.786	0.748

4 Discussion

In this study, we evaluated CARformer on two clinical sites, UTO and KUT, from the SRPBS dataset. The two sites differ in disease composition, acquisition protocols, and class distributions. This setting reflects practical constraints in psychiatric MRI, where differential diagnosis is often performed under small-cohort and imbalanced conditions and where protocol heterogeneity can further challenge generalization.

On the UTO site, differential diagnosis is performed among HC, MDD, and BD with an imbalanced class distribution of 96/62/41. This task requires separating clinically overlapping disorders with subtle structural differences on sMRI. In this setting, CARformer substantially improved minority-class detection, with BD AUPRC increasing to 0.709 compared with 0.356 for BrainMVP and 0.401 for ResNet18, representing an approximately twofold improvement over prior pretrained and from-scratch baselines. This indicates improved robustness to class imbalance in multi-class differential diagnosis. Qualitative analysis of query-specific spatial attention maps from the CA-BoQ module (Fig. 2) shows that class-aware queries consistently attend to fronto–limbic and subcortical regions, including the insula, anterior cingulate, and basal ganglia, supporting the ability of CA-BoQ to capture biologically meaningful and discriminative structural patterns consistent with prior neuroimaging findings [5–7].

To assess generalizability under a different imaging protocol and more severe imbalance, we further evaluated CARformer on the KUT site, which distinguishes HC, MDD, and SZ with a highly skewed class distribution of 159/16/45. CARformer achieved the best macro AUC and macro AUPRC and yielded the highest AUPRC for the minority MDD class, indicating improved minority-class discrimination under severe imbalance. Qualitative CA-BoQ attention maps (Fig. 2) further show distinct class-specific spatial patterns for MDD and SZ despite the limited MDD samples, supporting robust minority-class representation under severe imbalance. In contrast, several CNN baselines showed substantially lower minority-class AUPRC, consistent with majority-dominated predictions in highly imbalance settings. CARformer also shows higher overall performance on KUT than on UTO. This is plausibly related to stronger morphometric differences for schizophrenia reported in large-scale studies, including larger residual effects in the superior temporal gyrus [5], and suggests that class-aware aggregation and confusion-aware contrastive supervision are beneficial under combined domain shift and imbalance.

Overall, results across two sites indicate that combining domain-pretrained representations with class-aware aggregation and imbalance-aware supervision can improve multi-disorder psychiatric MRI classification under low-data constraints. However, the study is limited to two sites from a single public dataset, which limits cross-dataset and cross-ethnicity generalizability, and our future work will evaluate generalization on

clinically realistic datasets and further investigate the spatial interpretability and robustness of the framework.

5 Conclusion

We presented CARformer, a class-aware representation learning framework for multi-class psychiatric disorder classification from structural MRI under small-cohort and imbalanced settings. Across two clinical sites with different label sets and acquisition protocols, CARformer achieves the best macro AUC and macro AUPRC among evaluated baselines and shows consistent improvements in minority-class performance. These findings support the potential of class-aware aggregation and confusion-aware contrastive transfer for clinically realistic psychiatric MRI classification.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Jacob, K.S., Patel, V.: Classification of mental disorders: A global mental health perspective. The Lancet. 383, 1433–1435 (2014). https://doi.org/10.1016/S0140-6736(13)62382-X.
2. Kieling, C., Buchweitz, C., Caye, A., Silvani, J., Ameis, S.H., Brunoni, A.R., Cost, K.T., Courtney, D.B., Georgiades, K., Merikangas, K.R., Henderson, J.L., Polanczyk, G.V., Rohde, L.A., Salum, G.A., Szatmari, P.: Worldwide prevalence and disability from mental disorders across childhood and adolescence: Evidence from the global burden of disease study. JAMA Psychiatry. 81, 347 (2024). https://doi.org/10.1001/jamapsychiatry.2023.5051.
3. First, M.B., Gaebel, W., Maj, M., Stein, D.J., Kogan, C.S., Saunders, J.B., Poznyak, V.B., Gureje, O., Lewis-Fernández, R., Maercker, A., Brewin, C.R., Cloitre, M., Claudino, A., Pike, K.M., Baird, G., Skuse, D., Krueger, R.B., Briken, P., Burke, J.D., Lochman, J.E., Evans, S.C., Woods, D.W., Reed, G.M.: An organization- and category-level comparison of diagnostic requirements for mental disorders in ICD -11 and DSM -5. World Psychiatry. 20, 34–51 (2021). https://doi.org/10.1002/wps.20825.
4. Abi-Dargham, A., Moeller, S.J., Ali, F., DeLorenzo, C., Domschke, K., Horga, G., Jutla, A., Kotov, R., Paulus, M.P., Rubio, J.M., Sanacora, G., Veenstra-VanderWeele, J., Krystal, J.H.: Candidate biomarkers in psychiatric disorders: State of the field. World Psychiatry. 22, 236–262 (2023). https://doi.org/10.1002/wps.21078.
5. Opel, N., Goltermann, J., Hermesdorf, M., Berger, K., Baune, B.T., Dannlowski, U.: Cross-Disorder Analysis of Brain Structural Abnormalities in Six Major Psychiatric Disorders: A Secondary Analysis of Mega- and Meta-analytical Findings From the ENIGMA Consortium. Biological Psychiatry. 88, 678–686 (2020). https://doi.org/10.1016/j.biopsych.2020.04.027.
6. Redlich, R., Almeida, J.J.R., Grotegerd, D., Opel, N., Kugel, H., Heindel, W., Arolt, V., Phillips, M.L., Dannlowski, U.: Brain Morphometric Biomarkers Distinguishing Unipolar and Bipolar Depression: A Voxel-Based Morphometry–Pattern Classification Approach. JAMA Psychiatry. 71, 1222 (2014). https://doi.org/10.1001/jamapsychiatry.2014.1100.

7. for the ENIGMA-Major Depressive Disorder Working Group, Schmaal, L., Veltman, D.J., Van Erp, T.G.M., Sämann, P.G., Frodl, T., Jahanshad, N., Loehrer, E., Tiemeier, H., Hofman, A., Niessen, W.J., Vernooij, M.W., Ikram, M.A., Wittfeld, K., Grabe, H.J., Block, A., Hegenscheid, K., Völzke, H., Hoehn, D., Czisch, M., Lagopoulos, J., Hatton, S.N., Hickie, I.B., Goya-Maldonado, R., Krämer, B., Gruber, O., Couvy-Duchesne, B., Rentería, M.E., Strike, L.T., Mills, N.T., De Zubicaray, G.I., McMahon, K.L., Medland, S.E., Martin, N.G., Gillespie, N.A., Wright, M.J., Hall, G.B., MacQueen, G.M., Frey, E.M., Carballedo, A., Van Velzen, L.S., Van Tol, M.J., Van Der Wee, N.J., Veer, I.M., Walter, H., Schnell, K., Schramm, E., Normann, C., Schoepf, D., Konrad, C., Zurowski, B., Nickson, T., McIntosh, A.M., Papmeyer, M., Whalley, H.C., Sussmann, J.E., Godlewska, B.R., Cowen, P.J., Fischer, F.H., Rose, M., Penninx, B.W.J.H., Thompson, P.M., Hibar, D.P.: Subcortical brain alterations in major depressive disorder: findings from the ENIGMA Major Depressive Disorder working group. Mol Psychiatry. 21, 806–812 (2016). https://doi.org/10.1038/mp.2015.69.
8. Zhang, J., Rao, V.M., Tian, Y., Yang, Y., Acosta, N., Wan, Z., Lee, P.-Y., Zhang, C., Kegeles, L.S., Small, S.A., Guo, J.: Detecting schizophrenia with 3D structural brain MRI using deep learning. Sci Rep. 13, 14433 (2023). https://doi.org/10.1038/s41598-023-41359-z.
9. Liu, S., Zhou, J., Zhu, X., Zhang, Y., Zhou, X., Zhang, S., Yang, Z., Wang, Z., Wang, R., Yuan, Y., Fang, X., Chen, X., Wang, Y., Zhang, L., Wang, G., Jin, C.: An objective quantitative diagnosis of depression using a local-to-global multimodal fusion graph neural network. Patterns. 5, 101081 (2024). https://doi.org/10.1016/j.patter.2024.101081.
10. Li, Q., Wang, W., Guo, Q., Jiang, L., Qiao, K., Hu, Y., Zhang, X., Wang, Z., Peng, D., Fan, Q., Zhao, M., Fang, Y., Wang, J., Qiu, H., Wang, J., Li, G., Sheng, J., Tang, Y., Jin, C., Shen, D., Niznikiewicz, M.A., Li, C., Yang, Z.: Toward neuroimaging-based diagnostic support: a deep learning approach with a closed-loop system for psychiatric classification. iScience. 28, 114100 (2025). https://doi.org/10.1016/j.isci.2025.114100.
11. Hu, X., Wang, W., Xiao, L.: Learning 3D Medical Image Models From Brain Functional Connectivity Network Supervision For Mental Disorder Diagnosis, http://arxiv.org/abs/2503.04205, (2025). https://doi.org/10.48550/arXiv.2503.04205.
12. Tanaka, S.C., Yamashita, A., Yahata, N., Itahashi, T., Lisi, G., Yamada, T., Ichikawa, N., Takamura, M., Yoshihara, Y., Kunimatsu, A., Okada, N., Hashimoto, R., Okada, G., Sakai, Y., Morimoto, J., Narumoto, J., Shimada, Y., Mano, H., Yoshida, W., Seymour, B., Shimizu, T., Hosomi, K., Saitoh, Y., Kasai, K., Kato, N., Takahashi, H., Okamoto, Y., Yamashita, O., Kawato, M., Imamizu, H.: A multi-site, multi-disorder resting-state magnetic resonance image database. Sci. Data. 8, 227 (2021). https://doi.org/10.1038/s41597-021-01004-8.
13. Cackowski, S., Barbier, E.L., Dojat, M., Christen, T.: ImUnity: A generalizable VAE-GAN solution for multicenter MR image harmonization. Medical Image Analysis. 88, 102799 (2023). https://doi.org/10.1016/j.media.2023.102799.
14. Rui, S., Chen, L., Tang, Z., Wang, L., Liu, M., Zhang, S., Wang, X.: Multi-modal vision pre-training for medical image analysis. (2024). https://doi.org/10.48550/ARXIV.2410.10604.
15. Ali-Bey, A., Chaib-draa, B., Giguère, P.: BoQ: A Place is Worth a Bag of Learnable Queries, https://arxiv.org/abs/2405.07364, (2024). https://doi.org/10.48550/ARXIV.2405.07364.
16. Kingma, D.P., Ba, J.: Adam: A Method for Stochastic Optimization, https://arxiv.org/abs/1412.6980, (2014). https://doi.org/10.48550/ARXIV.1412.6980.
17. Zhou, H.-Y., Lu, C., Chen, C., Yang, S., Yu, Y.: PCRLv2: A Unified Visual Information Preservation Framework for Self-supervised Pre-training in Medical Image Analysis, https://arxiv.org/abs/2301.00772, (2023). https://doi.org/10.48550/ARXIV.2301.00772.