BCPL: Convolutional Prototype Learning for Brain Networks for Depression Diagnosis

Dongmei Zhi[1,2], Xiaohong Ma[3,4], Chuanyue Wang[5,6], Luxian Lv[7,8], Xianbin Li[5,6], Vince D. Calhoun[9], Weizheng Yan[1,2], Dongren Yao[1,2], Shile Qi[9], Rongtao Jiang[1,2], Jianlong Zhao[1], Xiao Yang[3,4], Yongfeng Yang[7,8], Miao Pan[7,8], Jing Sui[1,2,9,10*]

[1]*Brainnetome Center and National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, China*
[2]*University of Chinese Academy of Sciences, Beijing, China*
[3]*Psychiatric Laboratory and Mental Health Center, the State Key Laboratory of Biotherapy, West China Hospital of Sichuan University, Chengdu, Sichuan, China*
[4]*Huaxi Brain Research Center, West China Hospital of Sichuan University, Chengdu, Sichuan, China*
[5]*The National Clinical Research Center for Mental Disorders & Beijing Key Laboratory of Mental Disorders, Beijing Anding Hospital, Capital Medical University, Beijing, China*
[6]*Advanced Innovation Center for Human Brain Protection, Capital Medical University, Beijing, China*
[7]*Henan Key Lab of Biological Psychiatry, Xinxiang Medical University, Xinxiang Henan, China*
[8]*Department of Psychiatry, Henan Mental Hospital, The Second Affiliated Hospital of Xinxiang Medical University, Xinxiang, Henan, China*
[9]*Tri-Institutional Center for Translational Research in Neuroimaging and Data Science (TReNDS) [Georgia State University, Georgia Institute of Technology, Emory University], Atlanta, GA, USA*
[10]*CAS Centre for Excellence in Brain Science and Intelligence Technology, Institute of Automation, Chinese Academy of Sciences, Beijing, China*

INTRODUCTION: Depression is a disabling disorder, which has been reported to be associated with alterations in brain structure and function (Kennis, et al., 2019), and current diagnosis of depression is based on diagnosis criteria drawn from self-reported clinical symptoms without any objective biomarkers. Although numerous methods have been proposed to find reliable and robust biomarkers for pre-diagnosis of depression onset, most of them are purely discriminative models and essentially learn a partition of the whole feature space (Gao, et al., 2018). However, prototype learning, which can be viewed as a generative model based on the Gaussian assumption, can learn a robust representation and is appropriate for rejection and open set problems (Liu and Nakagawa, 2001).

METHODS: Motivated by the ability of the BrainnetCNN which can leverage the topological locality of functional brain networks by edge-to-edge, edge-to-node, node-to-graph convolutional filters and fully-connected layers (Kawahara, et al., 2017), and prototype learning, we propose a convolutional prototype learning for brain networks (BCPL), which can extract discriminative features and learn multiple prototypes to represent different classes. The loss function (Figure 2A) is composed of distance based cross-entropy loss between samples and prototypes, and prototype loss between the

samples and its closest prototype to learn inter-class separable and intra-class compact representations, in which λ is used to control the weight of prototype loss. As the flowchart is displayed in Figure 1, fMRI data were collected from 208 depression patients and 210 healthy controls from three sites, and were preprocessed based on SPM12. The functional connectivity matrix was estimated using Pearson correlation based on Brainnetome atlas for each subject (Figure 1A). Then functional connectivity matrix was input into the BCPL model to learn a prototype for each class, and the classification was implemented by finding the nearest prototype in the feature space. The nested k-fold cross-validation and leave-one-site-out strategies were used to evaluate the performance of the BCPL model and the traditional classifiers, including Linear support vector machine, Adaboost, Random forest, and Neural network. Furthermore, to increase interpretability, the Vanilla backpropagation algorithm was performed to identify the most contributing functional connectivities (FCs) (Simonyan, et al., 2013), which were then compared between depression and healthy controls using two-sample t-test.

RESULTS: As summarized in Figure 2, the learned prototypes within the same class became more and more compact with the increasing of the weight λ on prototype loss (Figure 2A), demonstrating that our BCPL model can learn more robust and discriminative representations. Accuracies of 70% and 65% were obtained respectively for the multi-sites pooling and leave-one-site transfer classification by BCPL, achieving 4% higher accuracy than traditional classifiers. Among the top 30 contributing features, increased FCs were found between middle frontal gyrus and amygdala, and hippocampus ($P<0.005$, FDR corrected), and decreased FCs were found between frontal lobe and superior parietal lobe, thalamus, and basal ganglia ($p<0.05$, FDR corrected).

CONCLUSIONS: This is the first attempt to propose a novel deep learning framework, BCPL, making feature extractor and the prototypes being learned jointly from the functional connectivity matrix. In addition, the BCPL model was applied to improve depression classification, demonstrating that the most discriminative FCs were located between frontal lobe and subcortical nuclei, which shows its potential for detecting biomarkers in mental disorders. More importantly, results suggested that the proposed model can learn intra-class compact and inter-class separable representations, which can be used for rejection and incremental category learning tasks based on the distance to the prototypes, deserving further investigation.

REFERENCES

Gao, S., Calhoun, V.D., Sui, J. (2018) Machine learning in major depression: From classification to treatment outcome prediction. CNS neuroscience & therapeutics, 24:1037-1052.

Kawahara, J., Brown, C.J., Miller, S.P., Booth, B.G., Chau, V., Grunau, R.E., Zwicker, J.G., Hamarneh, G. (2017) BrainNetCNN: convolutional neural networks for brain networks; towards predicting neurodevelopment. NeuroImage, 146:1038-

1049.
Kennis, M., Gerritsen, L., van Dalen, M., Williams, A., Cuijpers, P., Bockting, C. (2019) Prospective biomarkers of major depressive disorder: a systematic review and meta-analysis. Molecular psychiatry:1-18.
Liu, C.-L., Nakagawa, M. (2001) Evaluation of prototype learning algorithms for nearest-neighbor classifier in application to handwritten character recognition. Pattern Recognition, 34:601-615.
Simonyan, K., Vedaldi, A., Zisserman, A. (2013) Deep inside convolutional networks: Visualising image classification models and saliency maps. arXiv preprint arXiv:1312.6034.

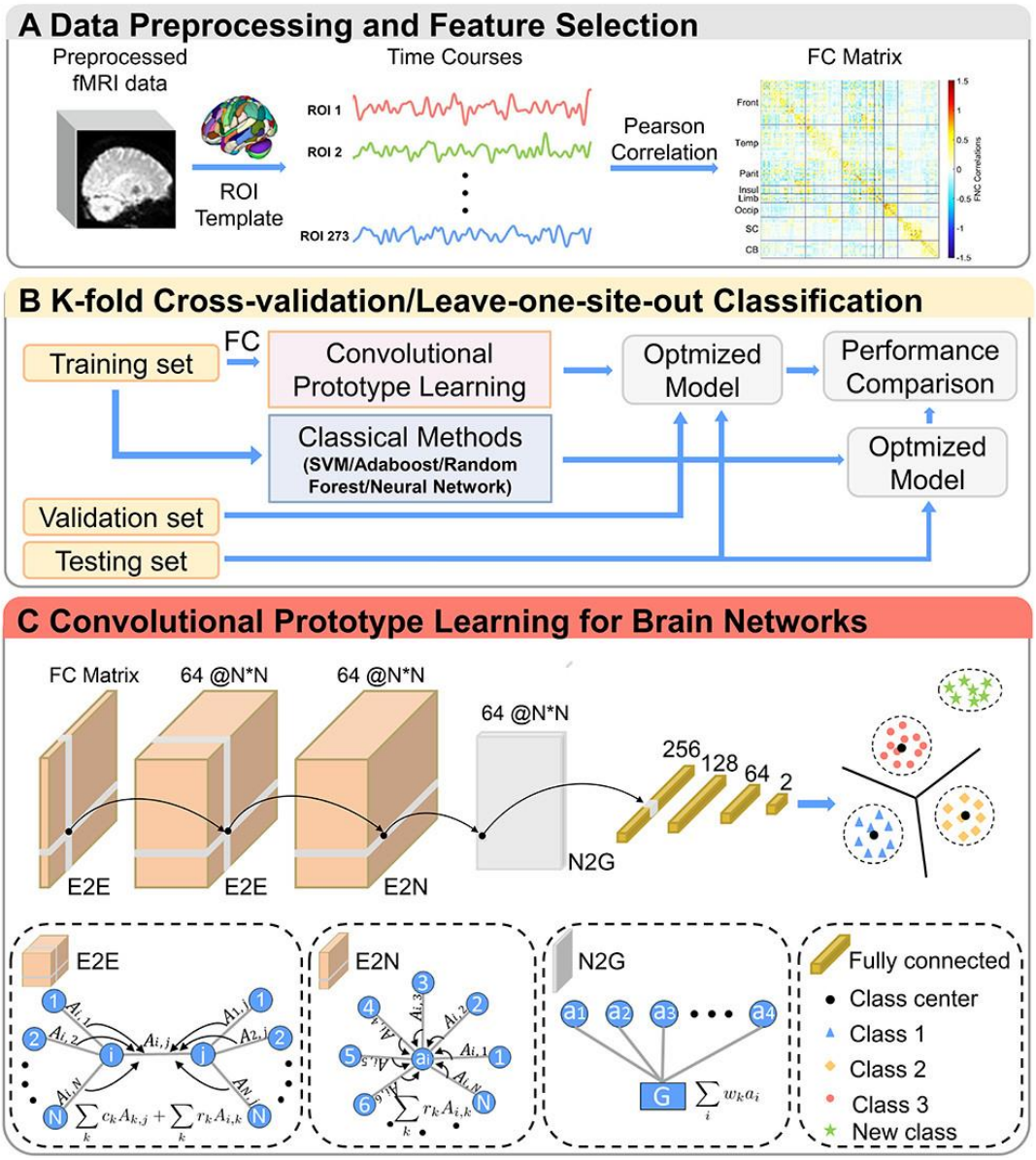

Figure 1. The framework of the convolutional prototype learning for brain networks in distinguishing depression from healthy controls. (A) Data preprocessing and feature selection. Functional connectivity matrix was computed using Pearson correlation

based on Brainnetome Atlas. (B) The functional connectivity data were randomly split into training, validation, and testing sets. In multi-site pooling classification, all three datasets were pooled together, and then k-fold cross-validation strategies were used for evaluating classification performance. In leave-one-site-out transfer prediction, the samples of a given imaging site were left for testing, and the sample of other sites were used for training. (C) Details of the convolutional prototype learning for brain networks (BCPL) model. In the bottom of the BCPL, convolutional neural networks for brain networks are used for extracting discriminative features from functional connectivity matrix, and in the top of the BCPL, prototype learning was used to build multiple prototypes to represent different classes.

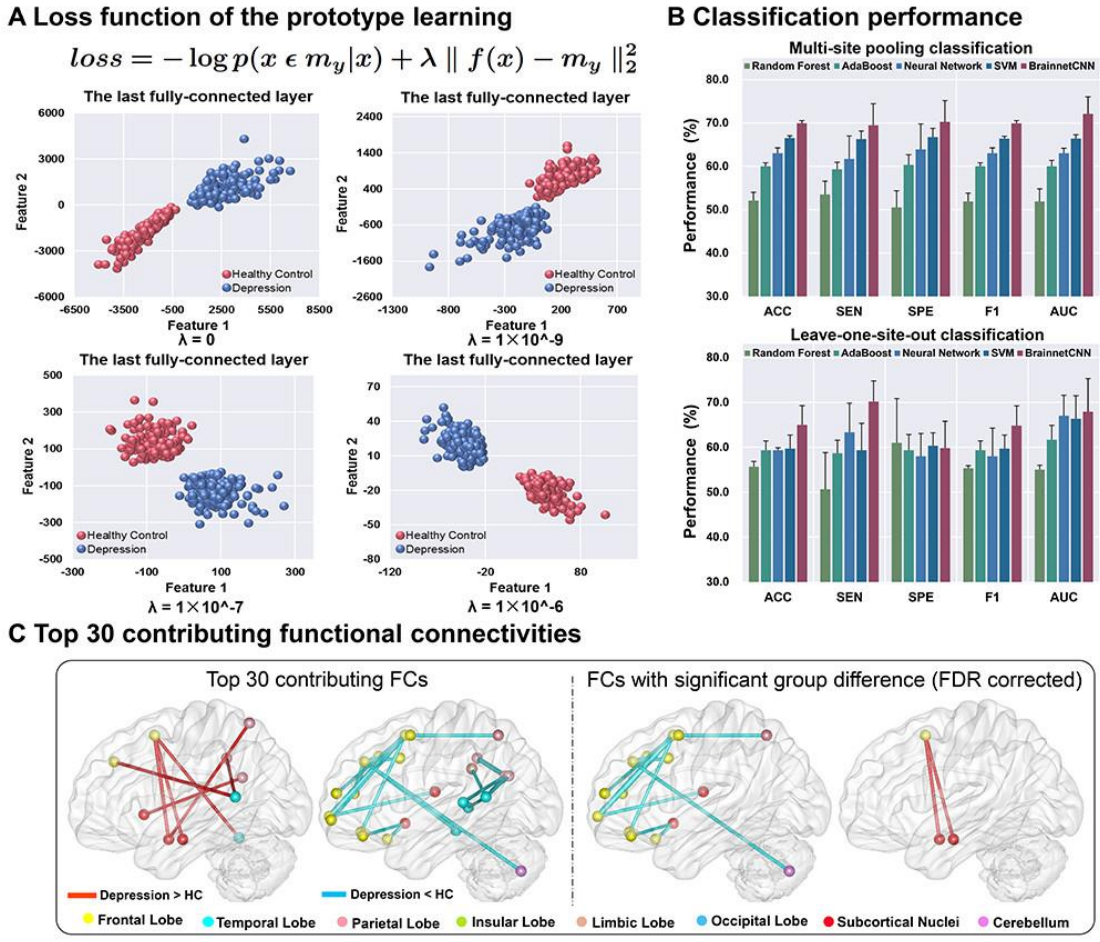

Figure 2. Prototype loss function and classification results of multi-site pooling and leave-one-site-out transfer classification. (A) The loss function is composed of similarity between samples and prototypes and prototype loss between the samples and its closest prototype, in which λ is used to control the weight of prototype loss. (B) 10-fold cross-validation and leave-one-site-out classification performance. (C) Top 30 discriminative functional connectivities discovered using the Vanilla backpropagation algorithm.