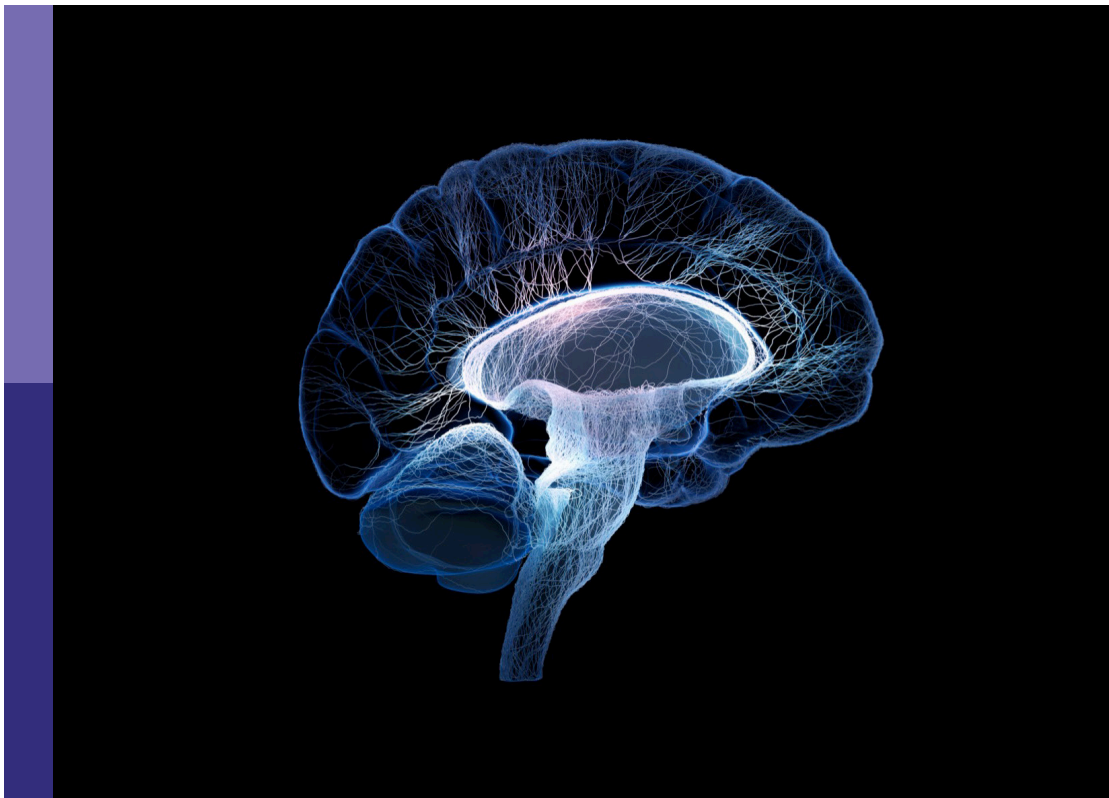
frontiers | Research Topics

Innovative applications with artificial intelligence methods in neuroimaging data analysis

Edited by
Yao Wu, Feng Liu, Li Zhao and Yuan-Chiao Lu

Published in
Frontiers in Neuroscience
Frontiers in Human Neuroscience
Frontiers in Neuroimaging



ISSN 1664-8714
ISBN 978-2-83251-189-3
DOI 10.3389/978-2-83251-189-3

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Innovative applications with artificial intelligence methods in neuroimaging data analysis

Topic editors

Yao Wu — Children's National Hospital, United States
Feng Liu — Tianjin Medical University General Hospital, China
Li Zhao — Zhejiang University, China
Yuan-Chiao Lu — Children's National Hospital, United States

Citation

Wu, Y., Liu, F., Zhao, L., Lu, Y.-C., eds. (2023). *Innovative applications with artificial intelligence methods in neuroimaging data analysis*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-83251-189-3

Table of contents

frontiers | Frontiers in Human Neuroscience

TYPE Editorial
PUBLISHED 16 December 2022
DOI 10.3389/fnhum.2022.1108253

OPEN ACCESS

EDITED AND REVIEWED BY
Mingzhou Ding,
University of Florida, United States

*CORRESPONDENCE
Yao Wu
ywu@childrensnational.org

SPECIALTY SECTION
This article was submitted to Brain Imaging and Stimulation, a section of the journal Frontiers in Human Neuroscience

RECEIVED 25 November 2022
ACCEPTED 05 December 2022
PUBLISHED 16 December 2022

CITATION
Liu F, Zhao L, Lu Y-C and Wu Y (2022) Editorial: Innovative applications with artificial intelligence methods in neuroimaging data analysis. *Front. Hum. Neurosci.* 16:1108253. doi: 10.3389/fnhum.2022.1108253

Editorial: Innovative applications with artificial intelligence methods in neuroimaging data analysis

Feng Liu[1], Li Zhao[2], Yuan-Chiao Lu[3] and Yao Wu[4]*

[1]Department of Radiology and Tianjin Key Laboratory of Functional Imaging, Tianjin Medical University General Hospital, Tianjin, China, [2]Key Laboratory for Biomedical Engineering of Ministry of Education, College of Biomedical Engineering and Instrument Science, Zhejiang University, Hangzhou, China, [3]Center for Neuroscience and Regenerative Medicine, Henry M. Jackson Foundation for the Advancement of Military Medicine, Bethesda, MD, United States, [4]Developing Brain Institute, Children's National Hospital, Washington, DC, United States

KEYWORDS

neuroimaging, medical image processing, artificial intelligence, computational modeling, brain disease diagnosis

Editorial on the Research Topic
Innovative applications with artificial intelligence methods in neuroimaging data analysis

Developing advanced analytic techniques to process neuroimaging data is crucial in advancing our understanding of the human brain structure and function. Most traditional image processing methods are unable to meet the accuracy and efficiency requirements of clinical practice and neuroscience research. Alternatively, advanced data processing methods such as artificial intelligence have yielded promising results in medical image analysis (Shen et al., 2017; England and Cheng, 2019), such as tumor detection (Saba et al., 2020; Sharif et al., 2020), brain registration (Wu et al., 2015; Fu et al., 2020; Wei et al., 2021), tissue segmentation (Wu et al., 2014; Ronneberger et al., 2015; Zhao et al., 2022), image reconstruction (Kainz et al., 2015; Cerrolaza et al., 2018), and neuropsychiatric disease diagnosis (Liu et al., 2015, 2017). Advanced artificial intelligence methods have shown improved accuracy and efficiency of neuroimaging data processing. As a result, this will advance the understanding of the human brain, which may assist in early diagnosis and developing intervention and/or surgery in patients with brain disorders. Currently, artificial intelligence is still in its infancy regarding its application to medical field and has the potential to be extensively used in clinical settings.

This Research Topic focuses on developing and applying artificial intelligence methods in medical image analysis, especially in the human brain, as well as using

novel data processing methods and tools to address neuroimaging-related clinical and neuroscience questions. A total of 14 articles were exclusively selected and published in this topic.

Brain structure segmentation

Two studies applied deep learning methods in brain structure segmentation using magnetic resonance imaging (MRI) scans. Theaud et al. proposed a DenseUNet-based deep learning segmentation algorithm for 10 tissues (i.e., white matter, gray matter, cerebrospinal fluid, ventricles, putamen, pallidum, hippocampus, caudate, amygdala, and thalamus) in diffusion weighted MR images. This method was trained and validated on 1,000 individuals from 22 to 90 years old from 5 public databases. Segmentation accuracy was superior to Freesurfer and FSL-FAST and the impacts on tractography were evaluated. Chai et al. proposed a contrast attention U-Net for deep gray matter nuclei segmentation in concatenated T1-weighted and quantitative susceptibility mapping sequences. This method was evaluated on two datasets acquired using different parameters from different MRI devices. Their results also suggested that sufficient data augmentation, deep supervision, and non-uniform patch sampling contributed to improving the segmentation accuracy.

Brain functional images

Ren et al. employed Pearson's correlation and nearest neighbor to identify individuals in different conditions including right-handed tapping, left-handed tapping, foot tapping and resting state on functional near-infrared spectroscopy. Wang et al. proposed a functional MRI encoding model to study the hierarchy of neural auditory processing in the human brain through an unsupervised deep convolutional auto-encoder model. Their findings showed that the neural representation of hierarchical auditory features is not limited to the superior temporal gyrus, but is also related to the bilateral insula, ventral visual cortex, and thalamus. Yu et al. developed a computational framework that incorporates both spatial and temporal characteristics of the brain to investigate brain states and high-level semantic features from naturalistic functional MRI. The framework is shown to be effective in classifying audio categories and identifying semantically meaningful high-level features. Su et al. utilized support vector machine to identify brain activity changes in response to short-term real-world visual experience in a group of radiologists, which may provide novel insights into the neural mechanism of visual experts. Banerjee et al. proposed a seizure onset zone localization algorithm, namely "EPIK," based on independent components derived from resting-state functional MRI in children with drug resistant epilepsy. EPIK outperforms support vector machine and convolutional neural network and shows consistent performance across different demographic subgroups.

Feature extraction

An et al. proposed a semiautomatic prediction model for the rupture risk estimation of aneurysms, which consisted of multidimensional feature fusion, feature selection, and the construction of classification methods. Features included morphological features, radiomics features, clinical features, and deep learning features. Three dimensional EfficientNet-B0 was used to extract and analyze the classification capabilities of three sets of deep learning features (no-sigmoid features, sigmoid features, and binarization features). Five classification models were compared, and the *k*-nearest neighbor produced the best results. This study suggests that the full use of multidimensional feature fusion can improve the performance of aneurysm rupture risk assessment. Jiang et al. proposed a multi-scale feature extraction by the neural network with multi-task learning in continuous blood pressure estimation. Specifically, segmentation, denoising, and normalization were used to preprocess the target (electrocardiograph and photoplethysmography) and label signals (arterial blood pressure), and then a neural network with multi-task learning was designed to extract multi-scale features related to blood pressure from preprocessed target signals. Three blood pressure values (systolic blood pressure, diastolic blood pressure, and mean arterial pressure) were estimated simultaneously through multi-task learning, thus improving the accuracy of blood pressure estimation.

Brain disease diagnosis

Liu et al. developed a method using decomposition-based correlation learning to capture the relationship between structural and functional MRI data. This method was evaluated in the classification of multiple neuropsychiatric disorders including schizophrenia, bipolar disorder, and attention deficit hyperactivity disorder. Chen et al. used MR spectroscopy to measure biochemical metabolites in prefrontal white matter and hippocampus in bipolar disorder patients with and without suicidal ideation, and combined brain biochemical metabolites with support vector machine algorithm to predict the severity of suicide risk in patients with bipolar disorder. Zhao, Han et al. proposed a hierarchical sub-network strategy to construct functional connectivity network from resting-state functional MRI based on matrix variate normal distribution theory. This method showed promising results in the classification of patients with autism spectrum disorder and normal controls. Zhao, Pan et al. proposed a scheme for constructing a high-order brain functional network from electroencephalography data based on sliding window, correlation, and clustering. Results demonstrate

the efficiency of the high-order brain functional network in the identification of major depressive disorder. Hou et al. successfully employed linear support vector machine to classify patients with obstructive sleep apnea from healthy controls based on whole-brain resting-state functional connectivity, indicating these features can serve as neuroimaging biomarkers for this disorder.

The articles in this Research Topic proposed and applied advanced processing techniques in medical image analysis, mainly focusing on the human brain. This topic may benefit researchers and clinicians who are interested in artificial intelligence methods and neuroimaging data analysis.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

This study was supported by the Brain and Behavior Research Foundation-NARSAD (28218), National Natural Science Foundation of China (82072001), Tianjin Key Medical Discipline (Specialty) Construction Project (TJYXZDXK-001A), the Alzheimer's Association through AARF-18-566347, the MOE Frontier Science Center for Brain Science and Brain-Machine Integration, Zhejiang University, Zhejiang Provincial Natural Science Foundation of China under Grant Nos. LGJ22H180004, 2020R01003, and 2022C03057, and Alibaba Cloud.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Cerrolaza, J. J., Sinclair, M., Li, Y., Gomez, A., Ferrante, E., Matthew, J., et al. (2018). "Deep learning with ultrasound physics for fetal skull segmentation," in *Proceedings—International Symposium on Biomedical Imaging* 2018-April(Isbi), 564–567. doi: 10.1109/ISBI.2018.8363639

England, J. R., and Cheng, P. M. (2019). Artificial intelligence for medical image analysis: a guide for authors and reviewers. *Am. J. Roentgenol.* 212, 513–519. doi: 10.2214/AJR.18.20490

Fu, Y., Lei, Y., Wang, T., Curran, W. J., Liu, T., and Yang, X. (2020). Deep learning in medical image registration: a review. *Phys. Med. Biol.* 65, 20TR01. doi: 10.1088/1361-6560/ab843e

Kainz, B., Steinberger, M., Wein, W., Kuklisova-Murgasova, M., Malamateniou, C., Keraudren, K., et al. (2015). Fast volume reconstruction from motion corrupted stacks of 2D slices. *IEEE Trans. Med. Imaging* 34, 1901–1913. doi: 10.1109/TMI.2015.2415453

Liu, F., Guo, W., Fouche, J.-P., Wang, Y., Wang, W., Ding, J., et al. (2015). Multivariate classification of social anxiety disorder using whole brain functional connectivity. *Brain Struct. Funct.* 220, 101–115. doi: 10.1007/s00429-013-0641-4

Liu, F., Wang, Y., Li, M., Wang, W., Li, R., Zhang, Z., et al. (2017). Dynamic functional network connectivity in idiopathic generalized epilepsy with generalized tonic–clonic seizure. *Hum. Brain Mapp.* 38, 957–973. doi: 10.1002/hbm.23430

Ronneberger, O., Fischer, P., and Brox, T. (2015). "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 234–241. doi: 10.1007/978-3-319-24574-4_28

Saba, T., Mohamed, A. S., El-Affendi, M., Amin, J., and Sharif, M. (2020). Brain tumor detection using fusion of hand crafted and deep learning features. *Cogn. Syst. Res.* 59, 221–230. doi: 10.1016/j.cogsys.2019.09.007

Sharif, M. I., Li, J. P., Naz, J., and Rashid, I. (2020). A comprehensive review on multi-organs tumor detection based on machine learning. *Pattern Recognit. Lett.* 131, 30–37. doi: 10.1016/j.patrec.2019.12.006

Shen, D., Wu, G., and Suk, H.-I. (2017). Deep learning in medical image analysis. *Annu. Rev. Biomed. Eng.* 19, 221. doi: 10.1146/annurev-bioeng-071516-044442

Wei, D., Ahmad, S., Guo, Y., Chen, L., Huang, Y., Ma, L., et al. (2021). Recurrent tissue-aware network for deformable registration of infant brain MR images. *IEEE Trans. Med. Imaging* 41, 1219–1229. doi: 10.1109/TMI.2021.3137280

Wu, Y., Liu, G., Huang, M., Guo, J., Jiang, J., Yang, W., et al. (2014). Prostate segmentation based on variant scale patch and local independent projection. *IEEE Trans. Med. Imaging* 33, 1290–1303. doi: 10.1109/TMI.2014.2308901

Wu, Y., Wu, G., Wang, L., Munsell, B. C., Wang, Q., Lin, W., et al. (2015). Hierarchical and symmetric infant image registration by robust longitudinal-example-guided correspondence detection. *Med. Phys.* 42, 4174–4189. doi: 10.1118/1.4922393

Zhao, L., Asis-Cruz, J. D., Feng, X., Wu, Y., Kapse, K., Largent, A., et al. (2022). Automated 3D fetal brain segmentation using an optimized deep learning approach. *Am. J. Neuroradiol.* 43, 448–454. doi: 10.3174/ajnr.A7419

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 10 February 2022
doi: 10.3389/fnins.2021.810431

Hierarchical Synchronization Estimation of Low- and High-Order Functional Connectivity Based on Sub-Network Division for the Diagnosis of Autism Spectrum Disorder

*Feng Zhao[1], Zhongwei Han[1], Dapeng Cheng[1], Ning Mao[2], Xiaobo Chen[1], Yuan Li[3], Deming Fan[4] and Peiqiang Liu[1]**

[1] School of Computer Science and Technology, Shandong Technology and Business University, Yantai, China, [2] Department of Radiology, Yantai Yuhuangding Hospital, Yantai, China, [3] School of Management Science and Engineering, Shandong Technology and Business University, Yantai, China, [4] School of Information Science and Technology, Qingdao University of Science and Technology, Qingdao, China

OPEN ACCESS

Edited by:
Feng Liu,
Tianjin Medical University General Hospital, China

Reviewed by:
Qiang Guo,
Shandong University of Finance and Economics, China
Biao Jie,
Anhui Normal University, China
Qiguang Miao,
Xidian University, China
Heng Chen,
Guizhou University, China

***Correspondence:**
Peiqiang Liu
liupq@126.com

Specialty section:
This article was submitted to Brain Imaging Methods, a section of the journal Frontiers in Neuroscience

Received: *09 November 2021*
Accepted: *27 December 2021*
Published: *10 February 2022*

Citation:
Zhao F, Han Z, Cheng D, Mao N, Chen X, Li Y, Fan D and Liu P (2022) Hierarchical Synchronization Estimation of Low- and High-Order Functional Connectivity Based on Sub-Network Division for the Diagnosis of Autism Spectrum Disorder. Front. Neurosci. 15:810431. doi: 10.3389/fnins.2021.810431

Functional connectivity network (FCN) calculated by resting-state functional magnetic resonance imaging (rs-fMRI) plays an increasingly important role in the exploration of neurologic and mental diseases. Among the presented researches, the method of constructing FCN based on Matrix Variate Normal Distribution (MVND) theory provides a novel perspective, which can capture both low- and high-order correlations simultaneously with a clear mathematical interpretability. However, when fitting MVND model, the dimension of the parameters (i.e., population mean and population covariance) to be estimated is too high, but the number of samples is relatively quite small, which is insufficient to achieve accurate fitting. To address the issue, we divide the brain network into several sub-networks, and then the MVND based FCN construction algorithm is implemented in each sub-network, thus the spatial dimension of MVND is reduced and more accurate estimates of low- and high-order FCNs is obtained. Furthermore, for making up the functional connectivity which is lost because of the sub-network division, the rs-fMRI mean series of all sub-networks are calculated, and the low- and high-order FCN across sub-networks are estimated with the MVND based FCN construction method. In order to prove the superiority and effectiveness of this method, we design and conduct classification experiments on ASD patients and normal controls. The experimental results show that the classification accuracy of "hierarchical sub-network method" is greatly improved, and the sub-network found most related to ASD in our experiment is consistent with other related medical researches.

Keywords: functional connectivity network, resting-state functional magnetic resonance imaging, matrix variate normal distribution, autism spectrum disorder, hierarchical sub-network method

INTRODUCTION

Functional connectivity networks (FCN), usually calculated from resting-state functional magnetic resonance imaging (rs-fMRI), using blood oxygenation level dependent (BOLD) signals as neurophysiological indicators, are playing an increasingly important role in exploring the working mechanism of the brain and investigating the brain's functional variations of some mental disorders, such as autism spectrum disorder (ASD) (Felouat and Oukid-Khouas, 2020; Sun et al., 2021), major depressive disorder (Mousavian et al., 2020), Alzheimer's disease (Jones et al., 2012; Wang et al., 2017), and its early stage, i.e., mild cognitive impairment (Chen et al., 2016; Zhang et al., 2020), et al.

FCN is a weighted network based on the graph theory, which takes the regions of interest (ROIs) in the brain as the nodes, the correlation of the rs-fMRI time series between different ROIs as the functional connectivity (FC) and the FC strength as the weight of the edge (Smith et al., 2013). Among all the methods for FC estimation, the most classic and popular example is Pearson's Correlation (PC) (Chen et al., 2016; Zhao et al., 2018; Sun et al., 2021). So far, it has been commonly known that the brain network structures and edge weights of the patients are different from those of the normal population due to the occurrence of pathological changes (Greicius et al., 2003).

At present, researchers have proposed many FCN models for disease diagnosis, which can be roughly divided into two categories. The first class is the so-called "low-order FCN" (Zhou et al., 2018b) that can only reflect FC characteristics between any two ROIs. For example, the conventional FCN assumes that all the rs-fMRI time series are static during the whole scanning period. Under such assumption, FC is quantified with the correlation (e.g., Pearson's correlation) between a pair of rs-fMRI time series derived from two ROIs (Achard, 2006). The dynamic FCN overcomes the drawback that the conventional FCN cannot reflect the dynamic information of brain activity. Based on the sliding window strategy, the rs-fMRI time series are divided into a set of short time series fragments, and the conventional FCN is constructed on each fragment. This can capture dynamic FC changes over time to a certain extent (Kudela et al., 2017). Notice that the low-order FCNs only calculate the pairwise correlation between two brain ROIs while fail to reflect deeper linkage mechanism involving multiple ROIs inside the brain. And the functional connectivity involving multiple ROIs may contain complementary information to low-order FC. The second class of FCN model is the so-called "high-order FCN" (Song et al., 2020) that can capture deeper brain information by designing FC model of multiple ROIs. For example, on the basis of dynamic FCN, Chen et al. (2016) and Zhao et al. (2018) took each FC time series as the network node and the correlation coefficient of FC time series of each ROI pair as the edge weight to construct a high-order FCN, which fills the interaction between paired ROI and other ROI pairs. Zhang et al. (2016) proposed a novel method to capture second-level relationship between two brain regions using inter-regional resemblance of the FC topographical profiles, which complements the discovery of more biologically meaningful inter-group differences. Furthermore, Zhao et al. (2020) combined inter-regional resemblance of the FC topographical profiles with dynamic network and central moment to explore dynamic and high-order relationships between two brain regions, which mines the dynamic FC relationship of multiple ROIs from multiple perspectives. Of note, the above methods all share the "correlation's correlation" strategy. In addition, in the literatures, many authors (e.g., Zhang et al., 2016, 2017) have presented the importance of high-order FC and explained potential biological meanings of high-order FC networks in dedicated studies. Since this paper mainly focuses on the applications of high-order FCNs for diagnosis, detailed discussion about general biological meanings of high-order FC networks can be found in these published works.

Zhou et al. (2018a) proposed a novel FC estimation method based on Matrix Variate Normal Distribution (MVND) theory. Compared with other higher-order models, MVND-based FCN can simultaneously obtain both low- and high-order FCNs with a clear mathematical explanation, and has demonstrated superior performance in identifying MCI patients from NCs. Specifically, the FCN sequence is constructed with the sliding window strategy, and then the so-constructed FCNs are taken as the samples to estimate the final low-order and high-order FCNs. In other words, each FCN is regarded as a random variable matrix (RVM) which obeys MVND, and all the FCNs in the sequence are taken together as the sample population to fit an MVND model. Like the other models mentioned above, Zhou's work is an FCN construction method based on fully brain network (FBN). So, we use the term "fully network FCN method" to refer to the method presented by Zhou et al. (2018a).

However, the "fully network FCN method" has the problem of "high dimension but small sample," which makes it actually impossible to fit an MVND model accurately. Theoretically, when fitting any distribution, the more samples there are, the more accurate the distribution will be. Besides, the higher spatial dimension where the distribution is located, the more samples will be needed in a fitting task. However, there exist the following facts in the "fully network FCN method": (1) each FCN is represented as a 116×116 matrix. (2) Each rs-fMRI time series contains only 137 volumes at most leading that no more than 137 FCNs can be generated even through the sliding window strategy. In fact, it is almost impossible to fit such a high-dimensional distribution with such a small number of samples.

In general, for fitting a more accurate MVND, either reducing the dimension of the RVM or increasing the number of RVM will be helpful. In other words, the fitting accuracy of MVND can be improved by reducing the ratio between the dimension of RVM and the number of RVM. However, as mentioned above, it is impossible to generate more than 137 FCNs through sliding windows even in extreme cases, then increasing the number of RVM is not feasible. Therefore, we put forward the "hierarchical sub-network method" to improve the "fully network FCN method" from the perspective of reducing the dimension of RVM in this paper. Specifically, the brain network is divided into several sub-networks, and each sub-network contains only part of rs-fMRI time series. Firstly, the MVND based FCN construction algorithm is implemented in each sub-network, so as to reduce the spatial dimension of MVND and obtain more accurate estimates of intra-sub-network low- and

high-order FCNs. Furthermore, the rs-fMRI mean series of all sub-networks are obtained, and the low- and high-order FCN across sub-networks are estimated according to the same strategy to compensate for the loss of FC information caused by sub-network division.

We propose the "hierarchical sub-network method" based on the following two motivations. On one hand, the ratio of the dimension of RVM to the number of RVM can be effectively decreased, so as to improve the fitting effect of MVND through the sub-network strategy. In fact, in this paper, the brain is divided into six relatively independent sub-networks according to the BrainNet Viewer software (Xia et al., 2013) like: the default mode network (DMN), the execution and attention network (EAN), etc. In other word, each sub-network is a relatively independent functional area and just contains a little part of ROIs. We take the largest sub-network as an example to illustrate the effectiveness of this method in improving MVND fitting. The largest sub-network only contains 26 ROIs, so in the MVND-based FCN construction method, RVM is expressed as a 26×26 matrix. Each ROI measured 170 signal elements, the ratio of the dimension of RVM to the number of RVM is 3.97 in our method, while in the "fully network FCN method," as analyzed earlier, the RVM is represented as a 116×116 matrix and the dimension quantity ratio of RVM is 79.15. Therefore, our method can reduce the difficulty of MVND fitting from the perspective of spatial dimension.

On the other hand, although the sub-network strategy can achieve more accurate fitting of MVND and more accurate extraction of FC information in sub-networks, we have to point out the fact that merely building FCN in sub-networks inevitably loses FC information of ROIs across different sub-networks, which can be understood more clearly by comparing FC information captured in fully network (**Figure 1A**) and sub-network (**Figure 1B**) of brain. **Figure 1A** represents the eight FCs among the six ROIs before network division. **Figure 1B** reflects that the attention to FC information of ROI within the sub-network ignores the two FCs belonging to ROI of different networks. In this paper, we defuse this problem skillfully through the correlation of any pair of sub-networks. Specifically, we first average all the rs-fMRI time series in each sub-network to get 6 (the number of sub-networks) mean time series, and then take all sub-networks as nodes to construct the low- and high-order FCN with MVND based FCN construction method. Corresponding to intra-sub-network features, these features are called inter-sub-network features. Finally, both intra-sub-network features and inter-sub-network features are used as the basis for autism classification experiments.

In summary, the advantages of the "hierarchical sub-network method" are as follows: first, combining the MVND-based FCN construction method with functional sub-networks can reduce the spatial dimension of MVND and achieve more accurate fitting of MVND; Second, capturing intra-sub-network features and inter-sub-network features from macro and micro perspectives to achieve the full expression of FC information in brain networks. In order to verify that the "hierarchical sub-network method" is superior to the "fully network FCN method," we apply both methods to the Autism Brain Imaging Data Exchange (ABIDE) database for individual based classification between ASD patients and NCs.

MATERIALS AND METHODS

Data Acquisition and Preprocessing

In this study, 92 rs-fMRI images of subjects with ages ranging from 7 to 15 years old from the publicly available Autism Brain Image Data Exchange Database (ABIDE) (Di Martino et al., 2014) are used, including 45 ASD patients and 47 NCs In order to avoid the influences of the heterogeneity of multi-site data on the results due to the difference in medical device, collection protocol, etc., we chose 45 ASD patients (36 males and 9 females) and 47 NC subjects (36 males and 11 females) with ages ranging from 7 to 15 years old. The mean frame-wise displacement was computed to describe head motion for each individual. The individuals were excluded if their mean FD is larger than 1 mm (Lin et al., 2015; Ray et al., 2015). All these considered subjects had no excessive head motion with a displacement of < 1.5 mm or an angular rotation of < 1.5 in any of three directions. The detailed

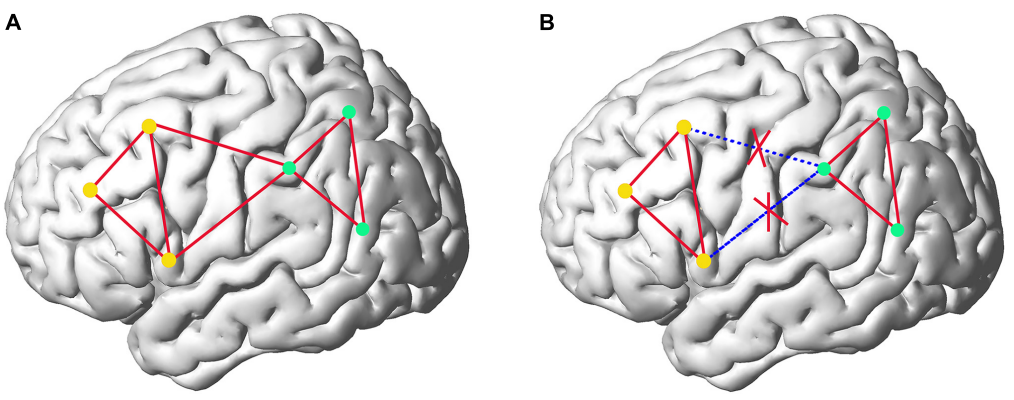


FIGURE 1 | The comparison of the FC information captured in fully network **(A)** and sub-network **(B)** of brain. The dots represent ROIs and the red lines represent FCs. Yellow dots and green dots depict two different sub-networks.

demographic information of these subjects is summarized in **Table 1**. As shown in **Table 1**, there are no significant differences ($p > 0.05$) in gender, age, and FIQ between two groups.

The observed rs-fMRI images are scanned at New York University (NYU) Langone Medical Center using a 3 -T Siemens Allegra scanner with the following parameters: flip angle = 90, 33 slices, TR/TE = 2,000/15 ms, 180 volumes, and voxel thickness = 4 mm. More details on the data collection, exclusion criteria, and scan parameters can be obtained from the ABIDE website.[1]

The acquired rs-fMRI data is preprocessed by the Statistical Parametric Mapping (SPM8) software.[2] Then, the brain is parcellated into 116 ROIs using the Automated Anatomical Marker (AAL) atlas (Tzourio-Mazoyer et al., 2002), and the average rs-fMRI time series for each ROI are calculated and expressed as a data matrix $X \in R^{170 \times 116}$, where 170 denotes the total number of temporal image volumes and 116 denotes the total number of brain ROIs.

The Pipeline of the "Hierarchical Sub-Network Method"

The pipeline of our proposed "hierarchical sub-network method" is shown in **Figure 2**, which mainly includes the following four steps: (1) Sub-network division. The division labels of the sub-network are obtained according to the BrainNet Viewer software (Xia et al., 2013), and the rs-fMRI time series of each subject are divided into 6 groups according to the division labels. (2) Intra-sub-network feature extraction. In each sub-network, the FCN sequence is constructed with sliding window, and the MVND is fitted with the FCN sequence being the RVM sample to obtain the intra-sub-network features. (3) Inter-sub-network feature extraction. The mean time series of each subnetwork is calculated, and then the low-order and high-order FCNs of the fully network are estimated synchronously with the MVND-based FCN construction method. (4) Feature normalization, feature selection and feature fusion. The features obtained in steps (2)—(3) are normalized. Then we use *T*-test and LASSO algorithms to select the most relevant features for the classification task. (5) ASD classification. We use SVM with linear kernel for ASD classification.

In the following subsections, we describe the above steps in detail. The meanings of the mathematical symbols are that bold uppercase letters represent matrices (i.e., $\boldsymbol{M}$), regular uppercase letters represent total values (i.e.,M), bold lowercase letters represent vectors (i.e., $\boldsymbol{m}$), and regular lowercase letters represent scalars (i.e., m).

Dividing the Brain Into Sub-Networks

For each subject, we define $\boldsymbol{x}_i = (x_{i1}, x_{i2}, \cdots, x_{iM})(i = 1, 2, \cdots, N)$ as the average rs-fMRI time series across all voxels belonging to the i-th ROI, where M denotes the total number of temporal image volumes, and N denotes the total number of ROIs. According to the experimental data mentioned above, here $M = 170$ and $N = 116$. Divide all ROIs into U different sub-networks $\{\Omega_1, \Omega_2, \cdots, \Omega_u, \cdots, \Omega_U\}$, where U consists of index i if x_i is included in the u-th sub-network. In the current study, the 116 ROIs in the Automated Anatomical Labeling (AAL) template were divided into six common functional networks according to the BrainNet Viewer software (Xia et al., 2013): the default mode network (DMN), the execution and attention network (EAN), the sensorimotor network (SMN), the visual network (Visual), the subcortical nuclei (SBN) regions and the cerebellum (Cerebel), so here $U = 6$. Of note, we choose this division method for the following two reasons. On one hand, the generated six sub-networks based on the BrainNet Viewer software have clear biological explanation, which makes this study have a broader medical reference value. On the other hand, dividing six sub-networks is enough to satisfy the dimensionality reduction needs of this study. Since the number of ROIs varies in each sub-network, we can use N_u to denote the total number of ROIs in the u-th sub-network. **Figure 3** gives an intuitive view of the division of the sub-network. In section "Discussion," we also discuss the sub-network division method based on similarity.

Constructing the Functional Connectivity Network Time Series With Sliding-Window Strategy

In **Figure 4**, **step 1** illustrate the construction of the FCN time series with sliding-window strategy vividly. Let the correlation between the i-th and the j-th ROIs be:

$$c_{ij} = corr(\boldsymbol{x_i}, \boldsymbol{x_j}) \tag{1}$$

Then, an FCN can be established using the classical method by taking $\boldsymbol{x_i}$ as nodes and c_{ij} as weights of edges. Here, c_{ij} is the weight of the edge connecting the i-th ROI and the j-th ROI. In the u-th sub-network, $i, j \in \Omega_u$, the total number of nodes of FCN is N_u, thus FCN can be expressed as a symmetric matrix, defined as follows:

$$\boldsymbol{W_u} = (c_{ij}) \; i, j \in \Omega_u \tag{2}$$

where $\boldsymbol{W}_u \in R^{N_u \times N_u}$ represents the FCN in the u-th sub-network. Next, the sliding window strategy is introduced. The

[1] http://fcon_1000.projects.nitrc.org/indi/abide/abide_I.html
[2] http://www.fil.ion.ucl.ac.uk/spm/software/spm8/

TABLE 1 | Demographic information of the subjects.

Characteristic	NC	ASD	p-value
Gender (M/F)	36/9	36/11	0.2135[a]
Age (mean ± SD)	11.1 ± 2.3	11.0 ± 2.3	0.773[b]
FIQ (mean ± SD)	106.8 ± 17.4	113.3 ± 14.1	0.0510[b]
ADI-R (mean ± SD)	32.2 ± 14.3[c]	-	-
ADOS (mean ± SD)	13.7 ± 5.0	-	-
FD (mm)(mean ± SD)	0.14 ± 0.05	0.15 ± 0.07	0.36[b]

ASD, autism spectrum disorders; NC, normal control; M, male; F, female; FIQ, Full Intelligence Quotient; ADI-R, Autism Diagnostic Interview-Revised; ADOS, autism diagnostic observation schedule.
[a] Thep-value was obtained by χ^2-test.
[b] The p-value was obtained by two-sample two-tailed t-test.
[c] Two patients do not have the ADI-R score.

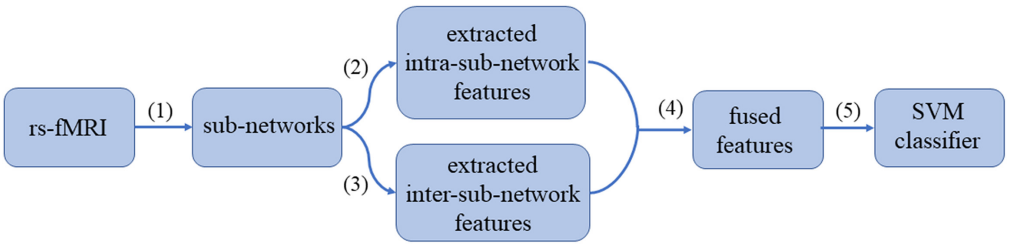


FIGURE 2 | The pipeline of the "hierarchical sub-network method."

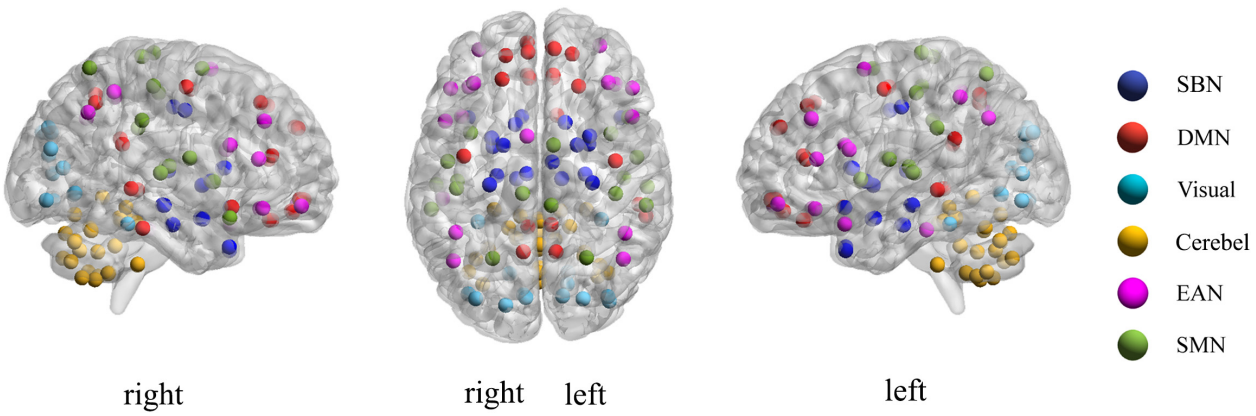


FIGURE 3 | Visualization of the location of each sub-network in brain.

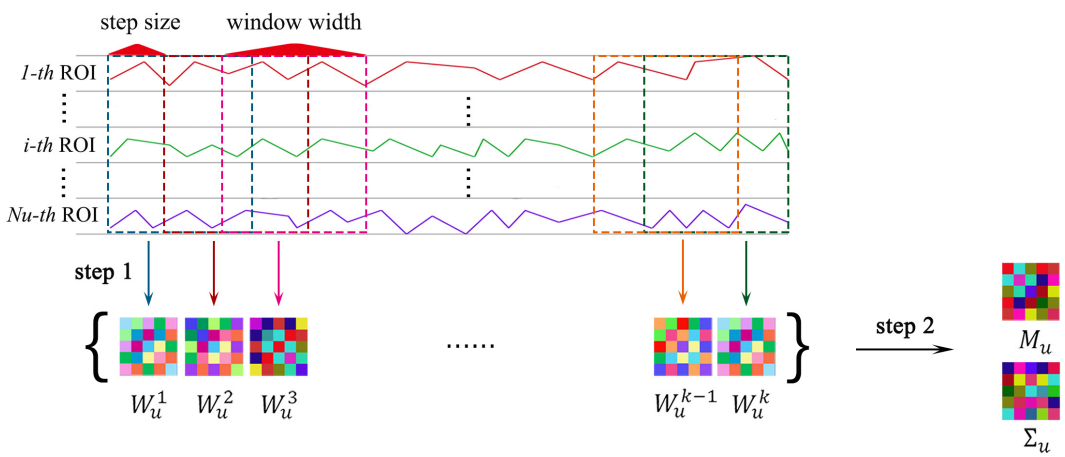


FIGURE 4 | The MVND based FCN construction method. Where, **step 1** shows the sliding-window strategy, **step 2** shows the MVND based feature extraction.

entire rs-fMRI time series of all ROIs is divided into K segments by window sliding, and corresponding FCNs are established on each rs-fMRI time series segment, thus forming a sequence containing K FCNs where K is determined by the window width l_w and step size l_s of the sliding window.

Specifically, taking the u-th sub-network as an example, the total number of ROIs is N_u, the total number of timing image voxels is M, and a sequence containing K FCNs, denoted by $\{\boldsymbol{W}_u^1, \boldsymbol{W}_u^2, \cdots, \boldsymbol{W}_u^K\}$, will be obtained through the sliding window strategy, where $K = \left[\frac{M-l_w}{l_s}\right] + 1, M = 170$, l_w and l_s are variable parameters.

Extracting the Intra-Sub-Network Features

In Figure 4, step 2 displays the pipeline of the extraction of intra-sub-network features. In each sub-network, we regard the obtained FCN sequence as a sample population obeying a

multivariate Gaussian distribution, then each FCN is regarded as a random variable matrix sample, then

$$W_u^k \sim N(M_u, \Sigma_u)\, 1 \le k \le K \tag{3}$$

where, $M_u \in R^{N_u \times N_u}$ is the population mean or mathematical expectation, and $\Sigma_u \in R^{N_u^2 \times N_u^2}$ is the population covariance of W_u. As mentioned in the introduction, M_u and Σ_u correspond to the low-order FC Features and high-order FC Features of brain networks, respectively. Since the dimension of Σ_u is too high, in order to avoid overfitting in the classification experiment, and consistent with the method in Zhou et al. (2018a) we replace the population variance with the form of Kronecker product decomposition (Gupta and Nagar, 2000), i.e., $\Sigma = C_u^1 \otimes C_u^2$, where C_u^1, $C_u^2 \in R^{N_u \times N_u}$ are positive semi-definite, representing the column and row covariance matrices of W_u, respectively. Since W_u is a symmetric matrix, $C_u^1 = C_u^2$, we can use $C_u = C_u^1 = C_u^2$ to replace Σ_u with the advantage of not losing information, so as to achieve the dimension reduction of Σ_u. Specifically, according to maximum likelihood estimation (MLE) theory of MVND, in each sub-network, the MLE of M_u is

$$M_u = \frac{1}{K} \sum_{k=1}^{K} W_u^k \tag{4}$$

The MLE of C_u can be achieved by the following iteration formula:

$$C_u = \frac{1}{KN_u} \sum_{k=1}^{K} (W_u^k - M_u) C_u^{-1} (W_u^k - M_u)^T \tag{5}$$

where, $1 \le k \le K, 1 \le u \le U$.

Extracting the Inter-Sub-Network Features

As mentioned in the introduction, after the sub-network division, we must consider both intra-sub-network and inter-sub-network features. The overview of the extraction of inter-sub-network features is vividly illustrated in **Figure 5** and the extraction of inter-sub-network features is divided into two steps: (1) Calculating the mean correlation time series for each sub-network (see **Figure 5A**). (2) Estimating low- and high-order FCNs simultaneously with the MVND based FCN construction method from rs-fMRI mean time series (see **Figure 5B**). The estimated low- and high-order FCNs are the inter-low-order features and the inter-high-order features, respectively.

The inter-sub-network feature extraction method is equivalent to the construction of FCN in the whole brain scale and the FCN construction method is the same as that in the sub-network scale. Both are constructed by MVND based FCN construction method, which can be referred to section "Constructing the FCN Time Series With Sliding-Window Strategy" and section "Extracting the Intra-Sub-Network Features." Here we describe in detail the generation of mean time series of each sub-network by taking the u-th subnetwork as an example.

The mean correlation time series y_u of the u-th sub-network can be calculated by averaging those rs-fMRI time series assigned to this sub-network. Specifically, each element in y_u is defined as:

$$y_u^m = \frac{\sum i \in \Omega_u\, y_i^m}{\|\Omega_u\|},\ 1 \le m \le M \tag{6}$$

Where, m represents the subscript of the element in y_u, and $\|_u\|$ represents the total number of rs-fMRI time series contained in the u-th sub-network.

Feature Normalization, Selection, Fusion, and Classification

All the features we have obtained include the intra-sub-network features and the inter-sub-network features, each of which consists of both high-order features and low-order features. Let's call them f_1, f_2, f_3, and f_4. These four feature vectors are acquired in different ways, so there are inevitably scale differences

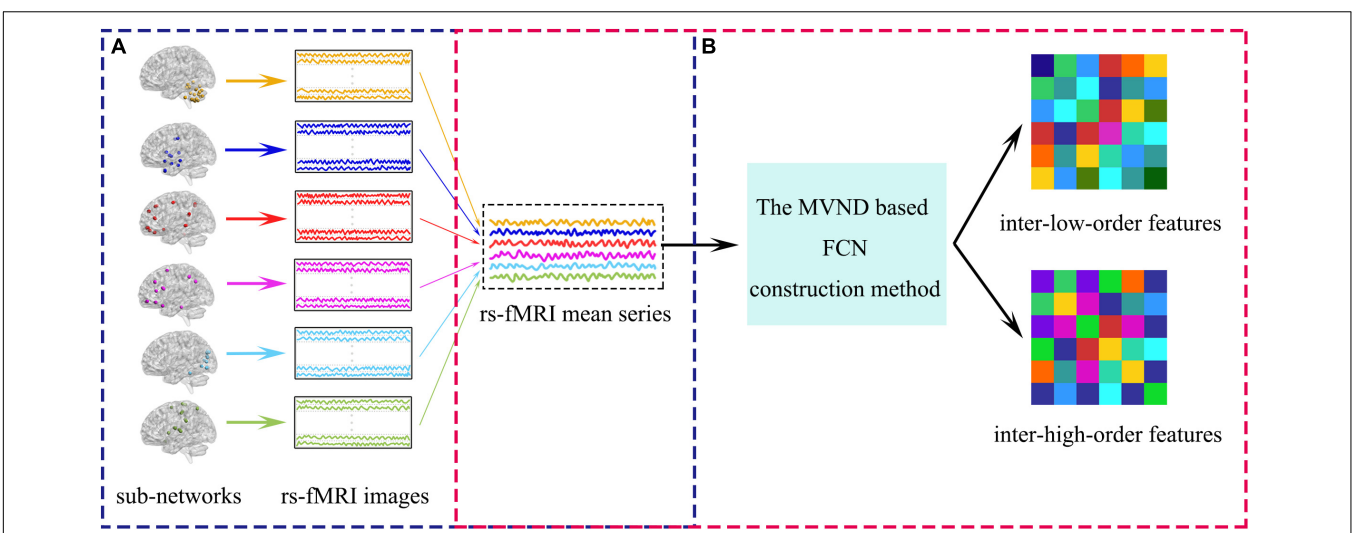


FIGURE 5 | The overview of the extraction of inter-sub-network features. **(A)** Shows the calculation of rs-fMRI mean series. **(B)** Shows the pipeline of the extraction of the inter-sub-network features.

among different features. In order to treat each feature equally, we normalize them in the same way. Here, $\boldsymbol{f}_1, \boldsymbol{f}_2, \boldsymbol{f}_3$, and $\boldsymbol{f}_4$ are normalized by the "min-max normalization" method, respectively. Take $\boldsymbol{f}_1$ as an example:

$$f'_{1i} = \frac{\boldsymbol{f}_{1i} - \min(\boldsymbol{f}_1)}{\max(\boldsymbol{f}_1) - \min(\boldsymbol{f}_1)} \quad (7)$$

Where, $\boldsymbol{f}_1$ represents the vector of intra-sub-network high-order features. $\boldsymbol{f}_{1i}$ represents the i-th element in $\boldsymbol{f}_1$, $min(\boldsymbol{f}_1)$ represents the minimum value in $\boldsymbol{f}_1$, and $max(\boldsymbol{f}_1)$ represents the maximum value in $\boldsymbol{f}_1$. The four types of features obtained by the "sub-network FCN method" just reflect the functional connectivity relationship between or among ROIs from four perspectives and they are complementary and homogeneous. Therefore, our fusion method is to simply combine them as a whole. In other words, the normalized feature data of four feature vectors are concatenated and expressed as a long vector f, that is $\mathrm{f} = [f'_1, f'_2, f'_3, f'_4]$.

However, the intra-sub-network low-order features and the inter-sub-network low-order features expressed by $\boldsymbol{f}_2$ and $\boldsymbol{f}_4$ exists as the form of FCN. FCN is a symmetric matrix, and the repeated feature leads to redundancy. So, we vectorize their lower off-diagonal-triangular parts to redefine the feature vectors. In this way, the original feature represented by $\boldsymbol{f}$ is replaced by a new one denoted by $\boldsymbol{f}_a$. Obviously, $\boldsymbol{f}_a$ may still contain features unrelated to ASD disease. In order to reduce the interference of irrelevant features and improve the generalization performance, we use the two-stage feature selection strategy to select a small set of most discriminative features for ASD diagnosis.

The first step is to perform a two-sample t-test between NCs and ASD subjects for each feature in the $\boldsymbol{f}_a$. Those features whose p-value is smaller than a certain threshold are preserved. At this point, we label the newly obtained feature set as $\boldsymbol{f}_b$. In the second step, we apply the L_1-norm regularized least squares regression, known as *LASSO* (Tibshirani, 1996), to further select the discriminative features from $\boldsymbol{f}_b$. Specifically, we used $\boldsymbol{f}_b^l$ to denote the features of the l-th subject and I^l to represent the label of the l-th subject. If the l-th subject is a patient with ASD, $I^l = 1$; otherwise, $I^l = -1$. Let $\boldsymbol{w}$ represents the weight vector for the feature selection task. The *LASSO* model is expressed by mathematical formula as:

$$\min \frac{1}{2} \sum_{l=1}^{L} \left\| I^l - (\boldsymbol{f}_b^l)^T \boldsymbol{w} \right\|_2^2 \lambda \|\boldsymbol{w}\|_1 \quad (8)$$

Where, L represents the total number of subjects, and $L = 92$ in this experiment. λ is a parameter, controlling the model's sparsity based on the L_1-norm regularization. The larger the value of λ, the sparser the model is. Different from the t-test, which selects feature separately, *LASSO* investigates all features synchronously. The t-test method and the LASSO method select features from different perspectives. As a binary classification problem, the t-test method can effectively select the features with high significance in ASD subjects and NC subjects. However, t-test method treats each feature independently without taking into account their inherent correlation, thus possibly resulting in many redundant features. Therefore, we further use LASSO method for the second selection which is able to consider the relationship between features. Therefore, we combine the two methods and design a two-stage feature selection strategy. We use $\boldsymbol{F}$ to represent the final feature for classification. In the classification phase, we use SVM (Chang and Lin, 2011) with a simple linear kernel for ASD identification. SVM seeks a maximum margin hyper-plane to separate the two kinds of samples. By adjusting the hyperparameter γ, the empirical risk of the training data and the complexity of the model can be balanced, so as to obtain good generalization performance on unlabeled test data.

Evaluation Methodology

We use nested fivefold cross-validation strategy which consists of two nested loops to evaluate classification performance in this experiment. In outer loop, 92 subjects are divided into 5 subsets of the roughly same size, where one subset is selected as the test-set, and the other 4 subsets are used as the training-set. In inner loop, the data of the training-set are combined and redivided into five subsets of similar size, four of which are used for tuning the hyperparameters and one for model evaluation. The performance of our method is mainly affected by three hyperparameters, they are p and λ in feature selection and γ in SVM model. The optimal hyperparameters can be determined when the average classification accuracy reaches its optimum. we determine the optimal values for the parameters in the following range: $\mathrm{p} \in [0.01 : 0.01 : 0.1]$, $\lambda \in [0.1 : 0.1 : 0.9]$, and $\gamma \in \left[2^{-4}, \cdots, 2^4\right]$. When the optimal hyperparameters are selected in inner loop, they are returned to the outer loop where the model will be trained based on the training dataset and evaluated on the testing dataset. Besides classification accuracy (ACC), we use sensitivity or true positive rate (TPR), specificity or true negative rate (TNR), positive predictive value (PPV), and negative predictive value (NPV)[3] to comprehensively evaluate the classification performance of the two methods.

RESULTS

Autism Spectrum Disorder Classification Performance

In this work, we compare the performance of the "hierarchical sub-network method" and the "fully network FCN method" in the ASD classification experiment. Specifically, we use the fusion of all the features extracted by each method to perform classification experiments. The experimental results are shown in **Table 2** and can be found with *Sub-Fusion* and *Fully Fusion* as pointers. Furthermore, in order to analyze the influence of different types of features in two compared methods on the experimental results, we carry out separate experiments on intra-sub-network high-order features, intra-sub-network low-order features, inter-sub-network high-order features, inter-sub-network low-order features, fully network high-order features and fully network low-order features. In **Table 2**, they are

[3]https://en.wikipedia.org/wiki/Sensitivity_and_specificity

TABLE 2 | ASD classification performance using different features.

Feature type	ACC (%)	TPR (%)	TNR (%)	PPV (%)	NPV (%)
Sub-Intra-Low	74 ± 0.21	73 ± 0.33	75 ± 0.39	75 ± 0.38	73 ± 0.49
Sub-Intra-High	77 ± 0.30	73 ± 0.45	81 ± 0.14	79 ± 0.47	76 ± 0.12
Sub-Intra-Fusion	79 ± 0.49	76 ± 0.45	82 ± 0.23	81 ± 0.13	77 ± 0.25
Sub-Inter-Low	66 ± 0.30	62 ± 0.22	70 ± 0.21	67 ± 0.45	66 ± 0.00
Sub-Inter-High	72 ± 0.38	69 ± 0.22	74 ± 0.40	72 ± 0.09	71 ± 0.43
Sub-Inter-Fusion	73 ± 0.45	71 ± 0.45	72 ± 0.34	71 ± 0.11	72 ± 0.34
Sub-Fusion	**81 ± 0.44**	**78 ± 0.30**	**83 ± 0.11**	**81 ± 0.45**	**80 ± 0.44**
Fully Low	74 ± 0.30	78 ± 0.29	70 ± 0.30	71 ± 0.45	77 ± 0.37
Fully High	71 ± 0.45	65 ± 0.37	77 ± 0.45	72 ± 0.50	69 ± 0.25
Fully Fusion	75 ± 0.18	72 ± 0.34	74 ± 0.48	74 ± 0.22	73 ± 0.14
Con-static	74 ± 0.04	72 ± 0.23	76 ± 0.01	74 ± 0.05	73 ± 0.07
Con-dynamic	75 ± 0.12	73 ± 0.14	76 ± 0.29	74 ± 0.23	75 ± 0.08

Values highlighted in bold mean the best results.

abbreviated as *Sub-Intra-High*, *Sub-Intra-Low*, *Sub-Inter-High*, *Sub-Inter-Low*, *Fully High,* and *Fully Low*. In addition, we conduct experiments on the fusion of intra-sub-network features and inter-sub-network features in the "hierarchical sub-network method" and the results can be found with the pointer *Sub-Intra-Fusion* and *Sub-Inter-Fusion* in **Table 2**. Finally, we experimented with two traditional methods under the same data, and reported the experimental results in **Table 2**. Traditional static FCN method and low-order dynamic FCN method are abbreviated as *con-static* and *con-dynamic,* respectively, in **Table 2**.

Table 2 shows the mean classification performance for each compared feature type. From the experimental results shown in **Table 2**, we can make the following judgments: (1) The classification accuracy of the intra-sub-network low- and high-order features of the "hierarchical sub-network method" (i.e., *Sub-Intra-Low*, *Sub-Intra-High*) is better than the corresponding features extracted by the "fully-network FCN method" (i.e., *Fully-Low*, *Fully-High*). (2) Both in the "hierarchical sub-network method" and the "fully-network FCN method", the classification performance of fusion features is significantly better than those of each type of features alone. (3) The performance of the fusion features (i.e., *Sub-Fusion*) of the "hierarchical sub-network method" is significantly higher than those (i.e., *Fully-Fusion*) of the "fully-network FCN method". (4) The classification performance of the fusion of the Intra-sub-network features extracted by "hierarchical sub-network method" (i.e., *Sub-Intra-Fusion*) is significantly better than the fusion of features extracted by "fully-network FCN method"(i.e., *Fully-Fusion*) . (5) Both the "fully-network FCN method" and the "hierarchical sub-network FCN method" perform better than the two traditional FCN methods, and the "hierarchical sub-network FCN method" has the most obvious advantages.

Influence of Parameters on Accuracy

In the "hierarchical sub-network method", we use the sliding window strategy to generate FCN sequences. There are two key parameters of the sliding window strategy that have a crucial impact on feature extraction and further affect the final recognition accuracy. They are the window width **(W)** and the step size **(S)** of the sliding window. In order to evaluate the influence of these two parameters on the experimental results, we conducted an ASD classification experiment under different parameter combinations. The window width is set as [30:10:120] and the step size is set as [1:1:12]. **Figure 6** shows the average accuracy of ASD classification under different parameter combinations. Referring to **Figure 6**, we can draw the following conclusions: (1) Sliding window parameters have great influence on classification performance. In the "hierarchical sub-network method," the maximum recognition accuracy is obtained when the window width is 60 and the step size is 4; The best performance of the "fully network FCN method" is achieved when the window width is 50 and the step size is 7. (2) In the performance comparison between the "hierarchical sub-network method" and the "fully network FCN method" under the same sliding window parameters, the "hierarchical sub-network method" is superior to the "fully network FCN method" in the majority of cases. (3) In each method, the best classification performance is achieved on average when the window width is between 50 and 90.

In addition, in the stage of feature fusion and classifier training, three hyperparameters have great impact on the results, that is, p-values in t-test, λ in lasso and γ in SVM. In this experiment, we explored the effects of different combinations of λ and γ on the results. Before that, the window width and step size are fixed as 60 and 4, respectively, which are also the parameter when the "hierarchical sub-network method" reaches the maximum. **Figure 7** shows the classification accuracy under different combinations of λ and γ in the two methods when the hyperparameter $p = 1$ of the t-test. From **Figure 7**, we can see that hyperparameter λ and γ have significant influence on the experimental results, and the effects are different in the two experiments.

The Most Discriminative Sub-Networks and Features for Autism Spectrum Disorder Diagnosis

According to the feature selection method mentioned in ***Feature normalization, selection, and classification,****t*-test and *LASSO*

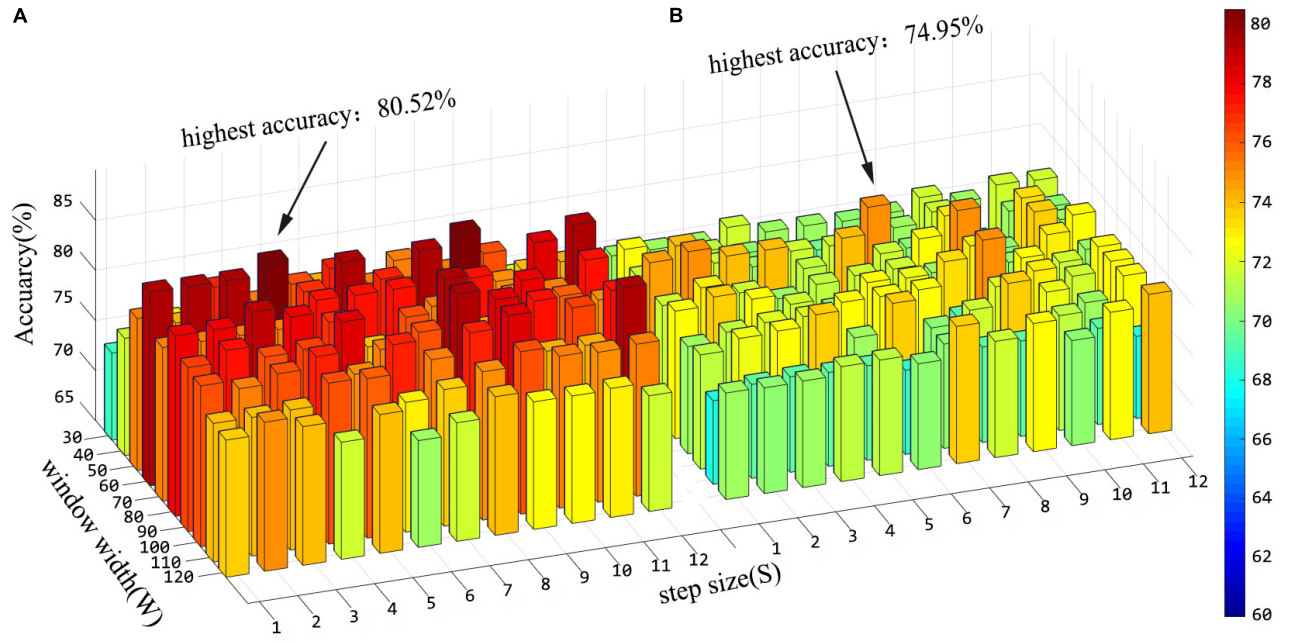


FIGURE 6 | Histogram of classification accuracy of the "hierarchical sub-network method" **(A)** and the "fully-network FCN method" **(B)** under various sliding window parameters.

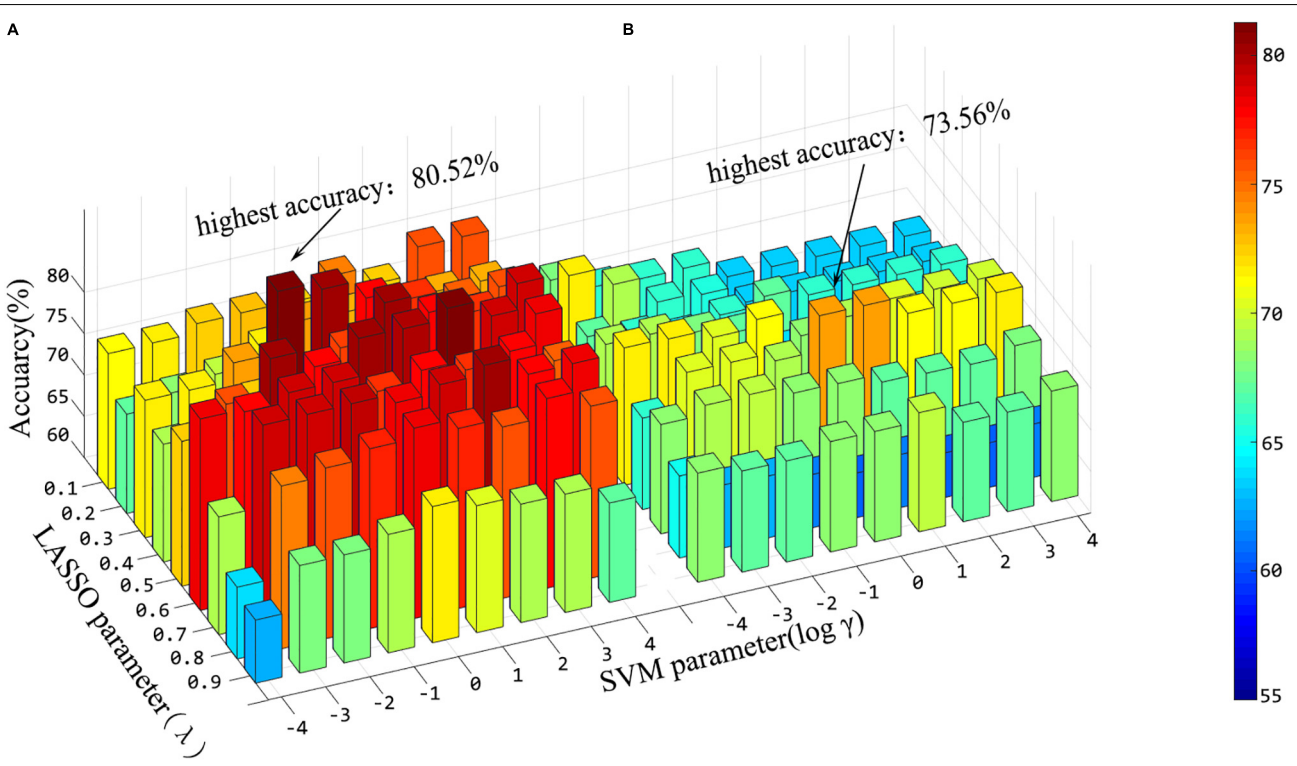


FIGURE 7 | Histogram of classification accuracy of the "hierarchical sub-network method" **(A)** and the "full-network FCN method" **(B)** under various hyper-parameter combinations.

are used to extract the most discriminative features from the original features in two steps for the ASD classification experiment. From all fivefold validation experiments, we take out and analyze the features used for training classifier each time. We trace each feature to each sub-network and count how often each sub-network is tracked. According to the frequency, the contribution of each sub-network to ASD recognition is calculated. The higher the frequency, the greater the contribution of the sub-network. **Figure 8** shows the contribution and distribution of different sub-networks. In order to have a more vivid and deep impression, the distribution of sub-network contribution is displayed on a surface rendering of the brain using the BrainNet viewer software (see **Figure 9**). The larger the volume of the ball, the greater the contribution rate of the sub-network to ASD recognition. Each sphere represents an ROI, we only use the set of spheres with the same color to represent the sub-network to observe the relationship between each sub-network and ASD.

Combined with **Figures 8**, **9**, we can see that only four sub-networks provide discriminative features for classification experiments, and DMN is the sub-network that contributes the most. This suggests that sub-network DMN is closely related to the diagnosis of ASD. Sub-network SMN and Visual provide zero contribution in this study, and precise judgments need further research. In addition, we believe that tracing the FC features that contribute most is also a convincing perspective to compare the differences between the two methods. The intra-sub-network low-order FCN and the fully network low-order FCN are used in the classification experiment. Then *t*-test and *LASSO* regression are used to select the features twice to get the final features for training. This part of the feature is considered the most discriminating. Each of these features represents an FC between a pair of ROIs. The features extracted in 10 repeated experiments are counted, and the top 10 features with the highest frequency are selected and shown in **Figure 10**. The name of the ROIs and brain anatomic areas shown in **Figure 10** are referred to the file ("Node_AAL116.node") provided by BrainNet Viewer software.

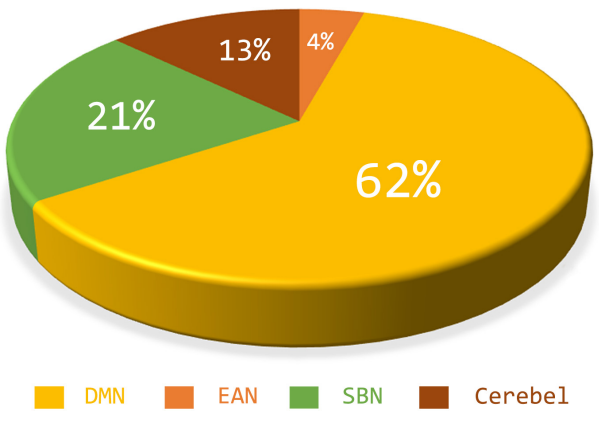


FIGURE 8 | The feature contribution rate of different sub-networks to classifier training.

Although certain genes have been found to be involved in ASD, the affected brain regions and the mechanisms behind specific defects are still poorly understood. According to **Figure 10**, except that the functional connections of REC in the left FRO region and PreCG in the right FRO region are selected by both methods, other features are different. The functional connection features selected by the "hierarchical sub-network method" are mostly concentrated in the FRO region. In fact, current studies have confirmed the relationship between FRO lesions and ASD disease (Scott-Van Zeeland et al., 2010; Solso et al., 2016).

DISCUSSION

We proposed "hierarchical sub-network method" based on MVND theory. This method not only inherits the advantages of MVND based FCN construction, being able to simultaneously obtain high-order features reflecting FC information among multiple ROIs and low-order features reflecting FC information between any two ROIs, but also improves the fitting effect of MVND with the help of sub-network division, so as to capture the functional connections of the brain more accurately and provide more discriminative features. We believe that compared with the "fully network FCN method," the "hierarchical sub-network method" can fully mine the disease-disturbed FCN variation information and has a better performance in ASD classification experiments. We will give a more detailed discussion on the comparison of the two methods.

In order to have an intuitive understanding of the dimension of features extracted by the "hierarchical sub-network method" and the "fully network FCN method," we select the dimension of intra-sub-network low-order features and fully network low-order features as the representative to display. In detail, we set the window width parameter as 60 and the step size parameter as 1, which is the combination of sliding window parameters when the intra-sub-network-low-order features have the best performance in ASD classification experiment.

We can take the area of the feature image as a reference to perceive the dimension of the feature extracted by the two methods. The larger the area, the higher the dimension. In this experiment, the number of ROI in each sub-network is as follows: 18 in SMN, 14 in Visual, 17 in EAN, 22 in DMN, 19 in SBC, and 26 in Cerebel. Each sub-network is shown in order from top to bottom in **Figure 11**. We can see that the intra-sub-network low-order feature dimensions extracted by different sub-networks are, respectively, about 2.41, 1.45, 2.15, 3.60, 2.68, 5.02% of the fully network low-order feature dimensions. Therefore, the "high dimensional but small-sample of RVM" problem is greatly improved in the calculation of fitting multivariate Gaussian distribution in each sub-network. Since the MVND based FCN construction method generates a high-order FCN and a low-order FCN with the same dimensions at one time, the dimension comparison of the high-order FCNs in the two methods is similar to that of the low-order FCNs. Overall, intra-sub-network features are more concise and discriminative than the fully network features.

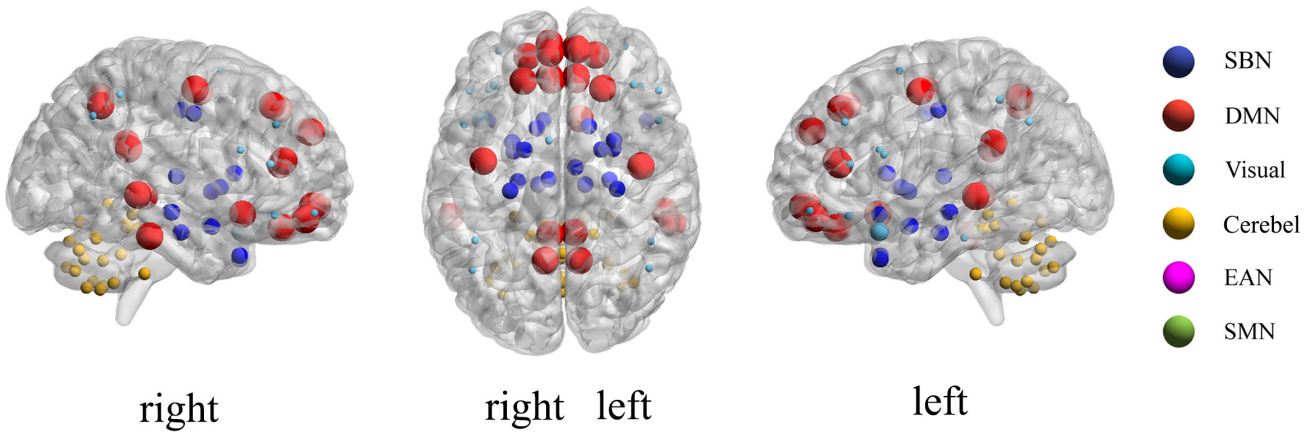


FIGURE 9 | The sub-network associated with ASD and the strength of its contribution to ASD.

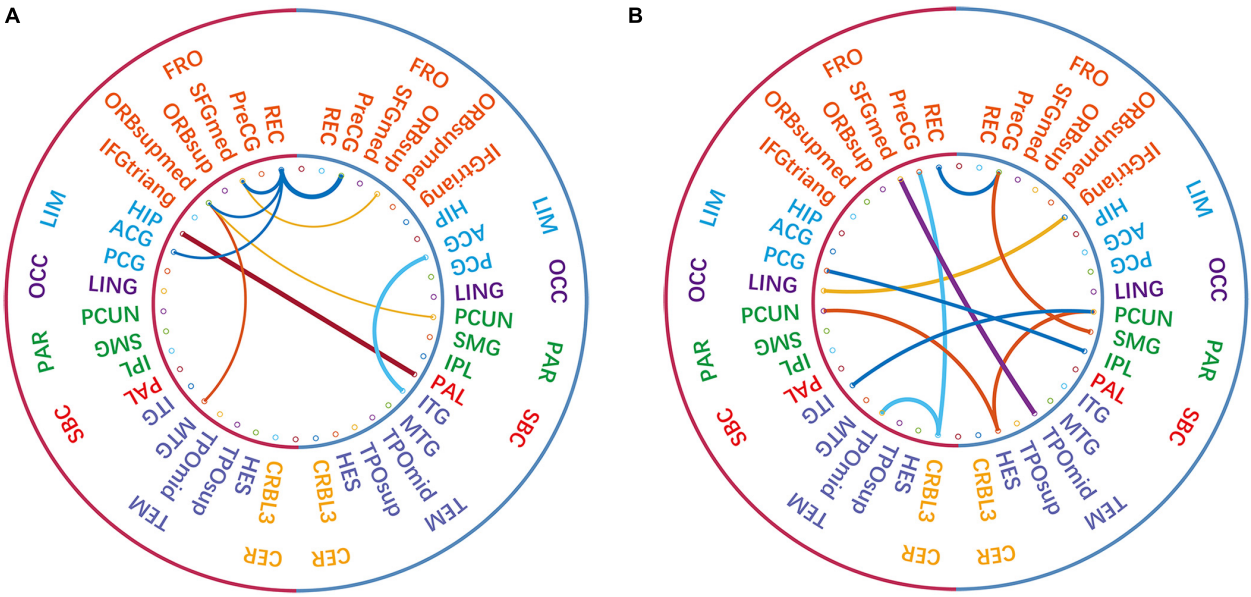


FIGURE 10 | The comparison of the most discriminative features in the "hierarchical sub-network method" **(A)** and the "fully-network FCN method" **(B)**. The left-side of each diagram represents the left hemisphere of the brain, and the right-side represents the right hemisphere of the brain. In the inner circle, each line connects two ROIs, and the thickness of the line represents the strength of its identification ability.

The results of ASD classification experiments show that: (1) when the high-order and low-order features extracted by the "hierarchical sub-network method" are trained separately, the classification accuracy is higher than that of the "fully network FCN method," and the classification performance of the "hierarchical sub-network method" is better than that of the "fully network FCN method." This means that the "hierarchical sub-network method" can capture the FC changes more accurately. There are two factors that play a role together: first, in technology, the dimensionality of the FCN is reduced due to the division of the sub-network, i.e., the dimension decline of RVM, which makes the MVND fitting more accurately. Another factor is about the biological mechanism. The sub-network division of the brain integrates the ROIs which are closely related in function, and focuses on observing the functional connection relationship among the ROIs belonging to the same sub-network, so that the information reflected by the RVM is more concise and effective, and the extracted low-order features and high-order features have a more discriminative power in ASD classification experiment. (2) The fusion of features in each method leads to better classification performance, respectively. To be sure, high-order features and low-order features can provide complementary information, which is the reason why the fusion features have better classification performance in ASD classification experiments.

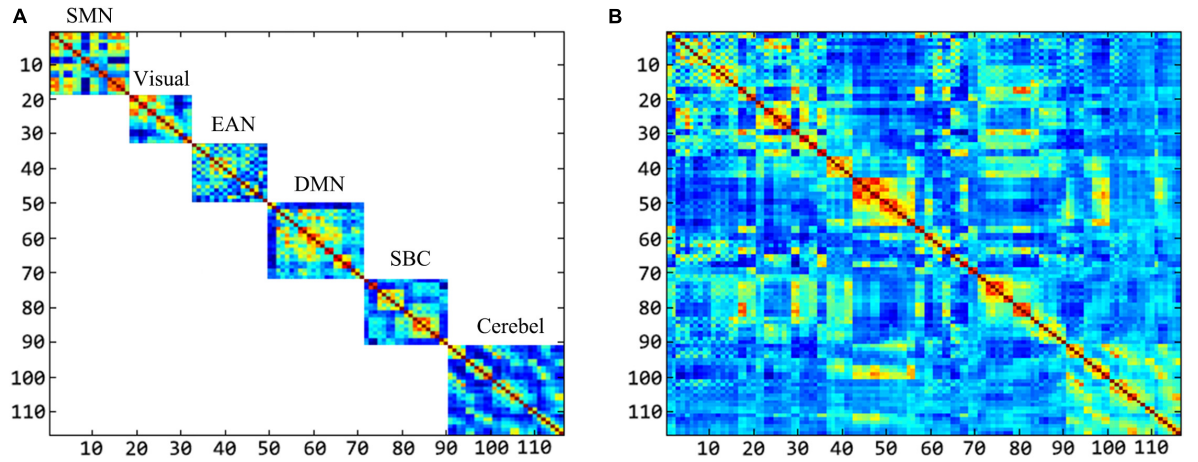


FIGURE 11 | The comparison of the feature dimension in both methods. **(A)** Shows the feature dimension of each sub-network. **(B)** Shows the feature dimension of the fully-network.

In the experiment to explore the influence of sliding window parameters on the classification accuracy, we found that the classification results of the two methods were changed with the combination of window width and step size parameters, especially the window width. Short window width and long window width have their own advantages and disadvantages. Short window width can provide rich short-term dynamic change information, but it is not stable due to the lack of low-frequency cycle (Sakoğlu et al., 2010). Long window width can make FC estimation more robust (Wang et al., 2018). From **Figure 6**, we can see the phenomenon that the maximum classification accuracy can be achieved in the middle length of the window width. In addition, under the same parameter combination, the "hierarchical sub-network method" is almost always more accurate than that of the "fully network FCN method." Since the dimension of the RVM is lower in each sub-network, more accurate fitting can be obtained. This proves in practice that the superiority of the "hierarchical sub-network method" is not a special result under special conditions.

To further demonstrate the validity of the "hierarchical sub-network method" in the diagnosis of ASD disease, we trace the sub-network to which the features of the classifier used for training belong. The experimental results showed that features in DMN was selected most frequently, suggesting that DMN was closely related to the pathogenesis of ASD. We found that this conclusion was basically consistent with the conclusion of other studies (Rutter et al., 2009; Simon and Engstrom, 2015; Churchill et al., 2018), and abnormalities of the DMN were commonly regarded as prominent ASD neurobiological features (Padmanabhan et al., 2017). From a biomedical perspective, DMN plays a crucial role in socially related stimuli because it is involved in the mental state of self-reflective thinking and considering the perspective of others, which is consistent with the fact that ASD is characterized by difficulties in social communication and interaction (Padmanabhan et al., 2017). Some studies have reported that the widely decreased of the FC in DMN in ASD not only contributes to the core defect of ASD, but also has a significant impact on the symptom severity (Assaf et al., 2010; Weng et al., 2010; Kiselev, 2014). For example, Assaf et al. (2010) pointed out that the decrease of functional connectivity in DMN of ASD patients was negatively correlated with the severity of social and communication disorders.

TABLE 3 | Experimental results of the application of similarity based sub-network division.

Feature type	ACC (%)	TPR (%)	TNR (%)	PPV (%)	NPV (%)
Intra-cluster	73 ± 0.91	77 ± 0.78	70 ± 0.21	71 ± 0.43	76 ± 0.74
Inter-cluster	61 ± 0.96	55 ± 0.56	68 ± 0.09	62 ± 0.50	61 ± 0.54
Cluster-Fusion	78 ± 0.26	77 ± 0.78	78 ± 0.72	77 ± 0.78	75 ± 0.72

In this study, we choose the medical template as the framework of sub-network division because of its advantage of biological interpretation. Of note, this is not the only scheme for sub-network partition. For example, ROI grouping based on the similarity of rs-fMRI time series can also be used as a sub-network division method. Specifically, k-means algorithm is used to cluster rs-fMRI time series, and the number of sub-networks is determined by specifying the number of clusters. By fixing the number of clusters from 6 to 11, we try to apply the similarity based sub-network division as an alternative to the "hierarchical sub network," and verify it in the classification experiments of ASD and NC. When the number of clusters is fixed at 8, the classification accuracy gets maximum, and the experimental results are shown in **Table 3**. In order to distinguish from the existing features, we use "*intra-cluster*" and "*inter-cluster*" to represent the intra-sub-network and inter-sub-network features in this method, respectively, and "*Cluster-Fusion*" to represent the fusion features. From the results, the random division of sub-networks according to the similarity of rs-fMRI time series does not perform better than the existing methods, and as far as we know, this method has two obvious shortcomings: first, although it shows better performance, it cannot make a biological

explanation for the results. Second, the number of clusters is not easy to determine which is greatly affected by subjects.

CONCLUSION

This paper proposes a new strategy for mental illness diagnosis based on FCN. The proposed method is based on the following two considerations: Technically, the FCN based on MVND is not well constructed in the fully network domain, and there exists the problem of "high dimension but small sample of RVM." From the biological point of view, many mental diseases reflect the sub-network property of brain function, and the aggregation of functional linkage makes the diagnosis of diseases more targeted. The results of ASD classification experiments show that the "hierarchical sub-network method" is comparable to the "fully network FCN method," and the biomedical findings obtained are consistent with other studies.

Besides Pearson's correlation, we can also utilize other candidates, such as Flexible Least Squares (FLS) method provided by the DynamicBC toolbox, to construct low-order FC network. In comparison with Pearson's correlation, FLS method has the advantage that more dynamic FC networks can be calculated by avoiding the sliding-window approach. The influence of different low-order FC networks to the performance of high-order FC network will be one of our directions for further study. One limitation of this work is that the ROIs corresponding to the higher-order features cannot be traced in the ASD classification experiment and this makes higher-order features useless for the discovery of ASD lesions. Further exploration of physiological markers of ASD and effective algorithms is our future work. Another limitation is that the "hierarchical sub-network method" could only explore the network-wise inter-network FCs, but would miss the ROI-wise inter-network FCs. How to compensate for the lost FCs with ROI-wise inter-network FCs needs to be further explored.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

ETHICS STATEMENT

The study was reviewed and approved by the Ethics Committee of Shandong Technology and Business University. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

AUTHOR CONTRIBUTIONS

FZ: conceptualization, methodology, and writing—review and editing. ZH: conceptualization, software, writing—original draft, methodology, formal analysis, investigation, and validation. DC, NM, and YL: writing—review and editing. XC and DF: conceptualization and writing—review and editing. PL: writing—review and editing, investigation, and supervision. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported in part by the National Natural Science Foundation of China (61773244, 61976125, 61272319, 61873117, 61972235, 61976124, and 62176140) and the Shandong Provincial Key Research and Development Program of China (2019GGX101069).

REFERENCES

Achard, S. (2006). A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. *J. Neurosci.* 26, 63–72. doi: 10.1523/JNEUROSCI.3874-05.2006

Assaf, M., Jagannathan, K., Calhoun, V. D., Miller, L., Stevens, M. C., Sahl, R., et al. (2010). Abnormal functional connectivity of default mode sub-networks in autism spectrum disorder patients. *NeuroImage* 53, 247–256. doi: 10.1016/j.neuroimage.2010.05.067

Chang, C. C., and Lin, C. J. (2011). LIBSVM: a library for support vector machines. *Assoc. Comput. Mach.* 2, 2157–6904. doi: 10.1145/1961189.1961199

Chen, X., Zhang, H., Gao, Y., Wee, C.-Y., Li, G., and Shen, D. (2016). High-order resting-state functional connectivity network for MCI classification. *Hum. Brain Mapp.* 37, 3282–3296. doi: 10.1002/hbm.23240

Churchill, N. W., Hutchison, M. G., Graham, S. J., and Schweizer, T. A. (2018). Connectomic markers of symptom severity in sport-related concussion: whole-brain analysis of resting-state fMRI. *NeuroImage Clin.* 18, 518–526. doi: 10.1016/j.nicl.2018.02.011

Di Martino, A., Yan, C.-G., Li, Q., Denio, E., Castellanos, F. X., Alaerts, K., et al. (2014). The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Mol. Psychiatry* 19, 659—-667. doi: 10.1038/mp.2013.78

Felouat, H., and Oukid-Khouas, S. (2020). "Graph convolutional networks and functional connectivity for identification of autism spectrum disorder," in *Proceedings of the 2020 2nd International Conference on Embedded & Distributed Systems (EDiS)*, Vol. 2020, Oran, 27–32. doi: 10.1109/EDiS49545.2020.9296476

Greicius, M. D., Krasnow, B., Reiss, A. L., and Menon, V. (2003). Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proc. Natl. Acad. Sci. U.S.A.* 100, 253–258. doi: 10.1073/pnas.0135058100

Gupta, A. K., and Nagar, D. K. (2000). *Matrix Variate Distributions*. Boca Raton, FL: Chapman & Hall/CRC.

Jones, D. T., Vemuri, P., Murphy, M. C., Gunter, J. L., Senjem, M. L., Machulda, M. M., et al. (2012). Non-stationarity in the "Resting Brain's" modular architecture. *PLoS One* 7:e39731. doi: 10.1371/journal.pone.0039731

Kiselev, S. (2014). Deficit of executive abilities as a risk factor for weakness in grammar understanding in Russian-speaking children — longitudinal research. *Int. J. Psychophysiol.* 94:212. doi: 10.1016/j.ijpsycho.2014.08.850

Kudela, M., Harezlak, J., and Lindquist, M. A. (2017). Assessing uncertainty in dynamic functional connectivity. *NeuroImage* 149, 165–177. doi: 10.1016/j.neuroimage.2017.01.056

Lin, H. Y., Tseng, W. Y. I., Lai, M. C., Matsuo, K., and Gau, S. S. F. (2015). Altered resting-state frontoparietal control network in children with

attention-deficit/hyperactivity disorder. *J. Int. Neuropsychol. Soc.* 21, 271–284. doi: 10.1017/s135561771500020x

Mousavian, M., Chen, J., and Greening, S. (2020). "Depression detection using atlas from fMRI images," in *Proceedings of the 2020 19th IEEE International Conference on Machine Learning and Applications (ICMLA)*, Miami, FL, 1348–1353. doi: 10.1109/ICMLA51294.2020.00210

Padmanabhan, A., Lynch, C. J., Schaer, M., and Menon, V. (2017). The default mode network in autism. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* 2, 476–486. doi: 10.1016/j.bpsc.2017.04.004

Ray, S., Gohel, S., and Biswal, B. B. (2015). Altered functional connectivity strength in abstinent chronic cocaine smokers compared to healthy controls. *Brain Connect.* 5, 476–486. doi: 10.1089/brain.2014.0240

Rutter, L., Carver, F. W., Holroyd, T., Nadar, S. R., Mitchell-Francis, J., Apud, J., et al. (2009). Magnetoencephalographic gamma power reduction in patients with schizophrenia during resting condition. *Hum. Brain Mapp.* 30, 3254–3264. doi: 10.1002/hbm.20746

Sakoğlu, Ü, Pearlson, G. D., Kiehl, K. A., Wang, Y. M., Michael, A. M., and Calhoun, V. D. (2010). A method for evaluating dynamic functional network connectivity and task-modulation: application to schizophrenia. *Magn. Reson. Mater. Phys. Biol. Med.* 23, 351–366. doi: 10.1007/s10334-010-0197-8

Scott-Van Zeeland, A. A., Abrahams, B. S., Alvarez-Retuerto, A. I., Sonnenblick, L. I., Rudie, J. D., Ghahremani, D., et al. (2010). Altered functional connectivity in frontal lobe circuits is associated with variation in the autism risk gene CNTNAP2. *Sci. Transl. Med.* 2:56ra80. doi: 10.1126/scitranslmed.3001344

Simon, R., and Engstrom, M. (2015). The default mode network as a biomarker for monitoring the therapeutic effects of meditation. *Front. Psychol.* 06:776. doi: 10.3389/fpsyg.2015.00776

Smith, S. M., Vidaurre, D., Beckmann, C. F., Glasser, M. F., Jenkinson, M., Miller, K. L., et al. (2013). Functional connectomics from resting-state fMRI. *Trends Cogn. Sci.* 17, 666–682. doi: 10.1016/j.tics.2013.09.016

Solso, S., Xu, R., Proudfoot, J., Hagler, D. J., Campbell, K., Venkatraman, V., et al. (2016). Diffusion tensor imaging provides evidence of possible axonal overconnectivity in frontal lobes in autism spectrum disorder toddlers. *Biol. Psychiatry* 79, 676–684. doi: 10.1016/j.biopsych.2015.06.029

Song, X., Elazab, A., and Zhang, Y. (2020). Classification of mild cognitive impairment based on a combined high-order network and graph convolutional network. *IEEE Access* 8, 42816–42827. doi: 10.1109/ACCESS.2020.2974997

Sun, L., Xue, Y., Zhang, Y., Qiao, L., Zhang, L., and Liu, M. (2021). Estimating sparse functional connectivity networks via hyperparameter-free learning model. *Artif. Intell. Med.* 111:102004. doi: 10.1016/j.artmed.2020.102004

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B* 58, 267–288. doi: 10.1111/j.2517-6161.1996.tb02080.x

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage* 15, 273–289. doi: 10.1006/nimg.2001.0978

Wang, Y., Zhu, L., Zou, Q., Cui, Q., Liao, W., Duan, X., et al. (2018). Frequency dependent hub role of the dorsal and ventral right anterior insula. *NeuroImage* 165, 112–117. doi: 10.1016/j.neuroimage.2017.10.004

Wang, Z., Zhu, X., Adeli, E., Zhu, Y., Nie, F., Munsell, B., et al. (2017). Multi-modal classification of neurodegenerative disease by progressive graph-based transductive learning. *Med. Image Anal.* 39, 218–230. doi: 10.1016/j.media.2017.05.003

Weng, S. J., Wiggins, J. L., Peltier, S. J., Carrasco, M., Risi, S., Lord, C., et al. (2010). Alterations of resting state functional connectivity in the default network in adolescents with autism spectrum disorders. *Brain Res.* 1313, 202–214. doi: 10.1016/j.brainres.2009.11.057

Xia, M., Wang, J., and He, Y. (2013). BrainNet viewer: a network visualization tool for human brain connectomics. *PLoS One* 8:e68910. doi: 10.1371/journal.pone.0068910

Zhang, H., Chen, X., Shi, F., Li, G., Kim, M., Giannakopoulos, P., et al. (2016). Topographical information-based high-order functional connectivity and its application in abnormality detection for mild cognitive impairment. *J. Alzheimers Dis.* 54, 1095–1112. doi: 10.3233/JAD-160092

Zhang, Y., Zhang, H., Adeli, E., Chen, X., and Shen, D. (2020). Multiview feature learning with multiatlas-based functional connectivity networks for MCI diagnosis. *IEEE Trans. Cybern.* 1–12. doi: 10.1109/tcyb.2020.3016953

Zhang, Y., Zhang, H., Chen, X., and Shen, D. (2017). "Constructing multi-frequency high-order functional connectivity network for diagnosis of mild cognitive impairment," in *Connectomics in NeuroImaging. CNI 2017. Lecture Notes in Computer Science*, Vol. 10511, eds G. Wu, P. Laurienti, L. Bonilha, and B. Munsell (Cham: Springer), 9–16. doi: 10.1007/978-3-319-67159-8_2

Zhao, F., Chen, Z., Rekik, I., Lee, S.-W., and Shen, D. (2020). Diagnosis of autism spectrum disorder using central-moment features from low- and high-order dynamic resting-state functional connectivity networks. *Front. Neurosci.* 14:258. doi: 10.3389/fnins.2020.00258

Zhao, F., Zhang, H., Rekik, I., An, Z., and Shen, D. (2018). Diagnosis of autism spectrum disorders using multi-level high-order functional networks derived from resting-state functional MRI. *Front. Hum. Neurosci.* 12:184. doi: 10.3389/fnhum.2018.00184

Zhou, Y., Zhang, L., Teng, S., Qiao, L., and Shen, D. (2018b). Improving sparsity and modularity of high-order functional connectivity networks for MCI and ASD identification. *Front. Neurosci.* 12:959. doi: 10.3389/fnins.2018.00959

Zhou, Y., Qiao, L., Li, W., Zhang, L., and Shen, D. (2018a). Simultaneous estimation of low- and high-order functional connectivity for identifying mild cognitive impairment. *Front. Neuroinform.* 12:3. doi: 10.3389/fninf.2018.00003

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 11 February 2022
doi: 10.3389/fnins.2022.813293

Identifying Individuals by fNIRS-Based Brain Functional Network Fingerprints

Haonan Ren[1], *Shufeng Zhou*[1], *Limei Zhang*[1], *Feng Zhao*[2] *and Lishan Qiao*[1]*

[1] School of Mathematics Science, Liaocheng University, Liaocheng, China, [2] School of Computer Science and Technology, Shandong Technology and Business University, Yantai, China

Individual identification based on brain functional network (BFN) has attracted a lot of research interest in recent years, since it provides a novel biometric for identity authentication, as well as a feasible way of exploring the brain at an individual level. Previous studies have shown that an individual can be identified by its BFN fingerprint estimated from functional magnetic resonance imaging, electroencephalogram, or magnetoencephalography data. Functional near-infrared spectroscopy (fNIRS) is an emerging imaging technique that, by measuring the changes in blood oxygen concentration, can respond to cerebral activities; in this paper, we investigate whether fNIRS-based BFN could be used as a "fingerprint" to identify individuals. In particular, Pearson's correlation is first used to calculate BFN based on the preprocessed fNIRS signals, and then the nearest neighbor scheme is used to match the estimated BFNs between different individuals. Through the experiments on an open-access fNIRS dataset, we have two main findings: (1) under the cases of cross-task (i.e., resting, right-handed, left-handed finger tapping, and foot tapping), the BFN fingerprints generally work well for the individual identification, and, more interestingly, (2) the accuracy under cross-task is well above the accuracy under cross-view (i.e., oxyhemoglobin and de-oxyhemoglobin). These findings indicate that fNIRS-based BFN fingerprint is a potential biometric for identifying individual.

Keywords: functional near-infrared spectroscopy, brain functional network (BFN), cross-task, individual identification, cross-view

OPEN ACCESS

Edited by:
Feng Liu,
Tianjin Medical University General Hospital, China

Reviewed by:
Tao Zhang,
University of Electronic Science and Technology of China, China
Qi Han Zhang,
Tianjin Normal University, China
Yanbei Liu,
Tianjin Polytechnic University, China

***Correspondence:**
Lishan Qiao
qlishan@163.com

Specialty section:
This article was submitted to Brain Imaging Methods, a section of the journal Frontiers in Neuroscience

Received: *11 November 2021*
Accepted: *03 January 2022*
Published: *11 February 2022*

Citation:
Ren H, Zhou S, Zhang L, Zhao F and Qiao L (2022) Identifying Individuals by fNIRS-Based Brain Functional Network Fingerprints. Front. Neurosci. 16:813293. doi: 10.3389/fnins.2022.813293

1. INTRODUCTION

Identifying individuals from a group is a significant task that has largely related to social security system and health care system (Ginther et al., 1992; Schmidt et al., 2005; Gershon et al., 2009). The mainstream identification characteristics, including face, fingerprint, and so on, are easily counterfeited, unstable in time, and involve privacy implications (Prabhakar et al., 2003; Jain et al., 2004, 2006). Advanced studies indicate that brain functional network (BFN) estimated by the temporal correlation between pairs of brain regions has the advantages of anti-imitation and stability in timing (Hilger et al., 2017, 2020; Wang et al., 2019b; Sastry et al., 2021). More importantly, the BFN-based "fingerprint" provides a potential way of exploring the brain at the individual level.

Up to now, several modalities of data have been utilized for constructing BFN fingerprints (Finn et al., 2015; Wang et al., 2019a, 2020; da Silva Castanheira et al., 2021; Sareen et al., 2021), including functional magnetic resonance (fMRI), electroencephalography (EEG), and magnetoencephalography (MEG). Among these modalities, fMRI was first used to estimate and identify BFN "fingerprint" by Finn et al. (2015). Their experimental result indicated that the fMRI-based BFN fingerprint can lead to a high identification accuracy, and individual functional connectivity is intrinsic and reliable. After Finn's work, Wang et al. (2019a) represented EEG signals as BFNs and used the deep intrinsic features of BFNs captured by the graph neural network for subject identification, showing that BFNs demonstrated more robust biometric traits than univariate features such as power spectral density functions and coefficients of auto-regressive stochastic models. Furthermore, da Silva Castanheira et al. (2021) generated functional connectivity fingerprints from MEG that measures the resting-state brain activity, and achieved a similar recognition rates to fMRI in the individual identification task.

As a complementary functional neuroimaging technique to fMRI and MEG, the emerging functional near-infrared spectroscopy (fNIRS) fNIRS has successfully explored the functional activation of shallow cerebral cortex during human behavior (Quaresima and Ferrari, 2019). The fNIRS simultaneously provides the concentration changes in deoxyhemoglobin (Deoxy-Hb) and oxyhemoglobin (Oxy-Hb), and the latter delivers additional information with respect to the fMRI signal (Irani et al., 2007; Duan et al., 2012). Also, the insensitivity of fNIRS to movements and the portability of the device make it possible for long-term monitoring and repeated measurements of cortical activities possible in various scenarios, such as outdoor activity or resting state. More importantly, the relatively low-cost and non-invasive technology makes the fNIRS applicable among larger groups, including infants and children (Strangman et al., 2002). Based on these advantages, fNIRS is naturally suitable for the study of individual identification under motion stimulating condition.

In this study, we mainly investigate whether the bio-specific BFNs extracted from fNIRS data are discriminative enough to identify individuals. More specifically, we use an open-access fNIRS dataset (Bak et al., 2019) from 30 subjects with multiple tasks, including resting state (REST), right-handed tapping (RHT), left-handed tapping (LHT), and foot tapping (FT) in this study. Note that we regard the resting state as a special task. The BFN fingerprints corresponding to each task are first calculated by Pearson's correlation (PC). Then, based on the nearest neighbor scheme, we demonstrate that an individual-specific BFN fingerprint extracted from one task can be used to match those from another. The results show that BFN fingerprints estimated from different tasks are strongly intrinsically linked and that they are stable and reliable biometric features for individual identification. Additionally, since the BFNs from different views (i.e., Oxy-Hb and Deoxy- Hb) are involved, we can naturally design cross-view experiment to explore the possibility of individual identification. Furthermore, we believe that BFN fingerprinting has a potential in the brain exploration and patient identification for medical systems, which also presents a viable thinking for decoding the brain functional states at the individual level.

The rest of this paper is organized as follows. In Section 2, we introduce the fNIRS data preparation, BFN fingerprints estimation, and their identification. In Section 3, we report the identification accuracy across different tasks. In Section 4, we analyze the experimental results and point out some limitations of the involved scheme. Finally, we summarize this paper in Section 5.

2. MATERIALS AND METHODS

In this section, we describe the data preparation (including acquisition and preprocessing), BFN estimation, and BFN-based fingerprint identification.

2.1. Data Preparation

2.1.1. Data Acquisition

In this paper, an open-access dataset of fNIRS with three kinds of tasks (including RHT, LHT, and FT) is used to conduct individual identification experiments. In particular, 30 subjects (23.4 ± 2.5 years old) participated in the experiment. All of them declared that they have no psychiatric and neurological disorder that could affect the experimental results. The data are freely downloaded from https://figshare.com, and more details about the dataset can be found in Bak et al. (2019).

Following the literature (Bak et al., 2019), the equipment used in the experiment was a three-wavelength continuous-time multi-channel fNIRS system (LIGHTNIRS, Shimadzu, Kyoto, Japan) consisting of eight light sources and eight detectors, which formed 20 channels to record changes in blood oxygen concentration. As shown in **Figure 1**, the light source and detector were located 3 cm apart and evenly distributed around C3 and C4 that represent the motor cortex (Georgopoulos, 1988; Gratton et al., 2006).

During the data collection, all subjects were informed to seat in front of a 27-inch monitor and executed random commands that appeared on the screen. The detection pipeline for a individual consisted of three sessions. Each session contained 25 trials, and each trial lasted an average of 30 s as shown in **Figure 2**. The 30-s trial consisted of three phases: the first 2 s were the introduction period, during which instructions appeared randomly on the screen; the next 10 s were the experiment period, during which the subjects need to make corresponding actions; and the last 17–19 s were the rest period, during which the brain returned to the resting state.

2.1.2. Data Preprocessing

The original optical density information is collected using the multi-channel fNIRS system and need to be converted into the changes in blood oxygen concentration by modified Lambert–Beer law (Khan et al., 2020). Due to different motion artifacts (like heartbeat, respiration, and Mayer wave) and the interference from the instrument, the signals need to be further processed for the subsequent identification experiment (Delpy et al., 1988). First, a band-pass filtering is implemented by the third-order

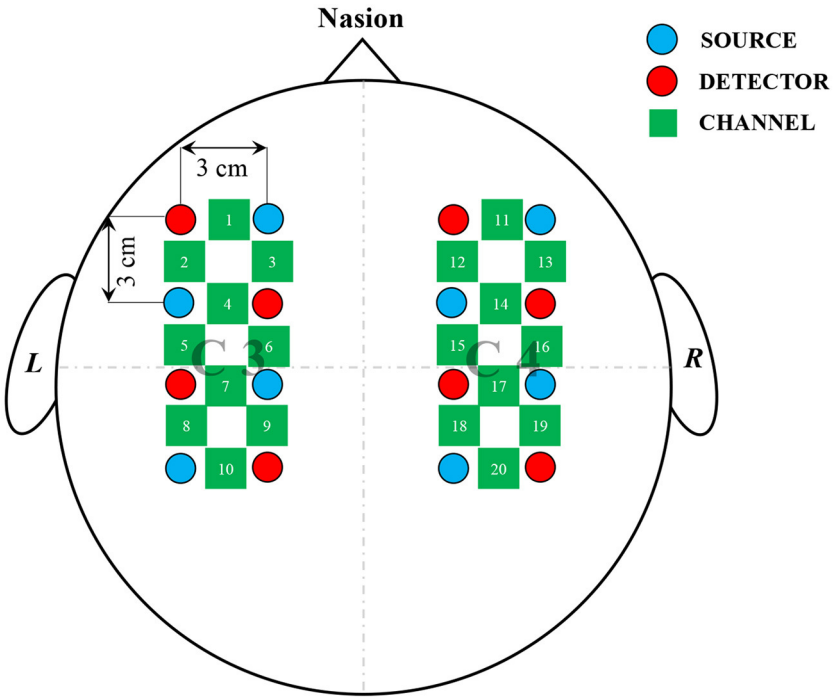


FIGURE 1 | Arrangement of channels. Light sources and detectors were placed around C3, C4; red circles, blue circles, and green squares represented sources, detectors, and channels, respectively; total of eight sources and eight detectors, having separation of 3 cm between each source-detector pair, formed 20 channels to record cerebral activities.

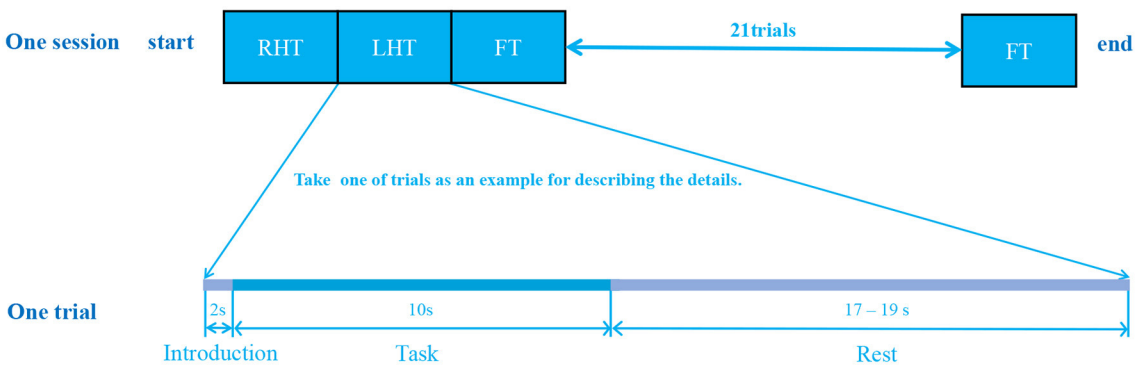


FIGURE 2 | Paradigm for data acquisition. The data acquisition experiment was separated into three sessions, each session contained 25 trials, and each trial lasted about 30 seconds. As a result, for each subject, 75=25*3 trials were recorded with a complete experiment duration of about 2250 seconds = 75*30 seconds. Three tasks were involved, and each trial randomly selected one task to perform. A single trial consisted of three stages: 2-second introduction phase followed by 10-second task period and again followed by 17–19 seconds rest stage.

Butterworth filter with a cut-off frequency of 0.01–0.1 Hz to eliminate the physiological noise (approximately 0.1 Hz for Mayer wave, 0.25 Hz for respiration, and 1 Hz for a heartbeat). Then, the baseline correction is used to subtract the global signal (i.e., the average signals of all channels) from all signals (Nguyen et al., 2018; Zhang et al., 2021). All the preprocessing mentioned above is done through the BBCI toolkit (Blankertz et al., 2016).

Note that, in order to maintain data integrity, we do not abandon any channel signal (even with low signal-to-noise ratio) since they may contain some individual variability.

2.2. BFN Fingerprints Estimation and Identification

2.2.1. BFN Estimation

After data preprocessing, we convert the signals into BFNs according to the pipeline shown in **Figure 3**. Since the involved subjects are treated in a similar way, we only take one subject

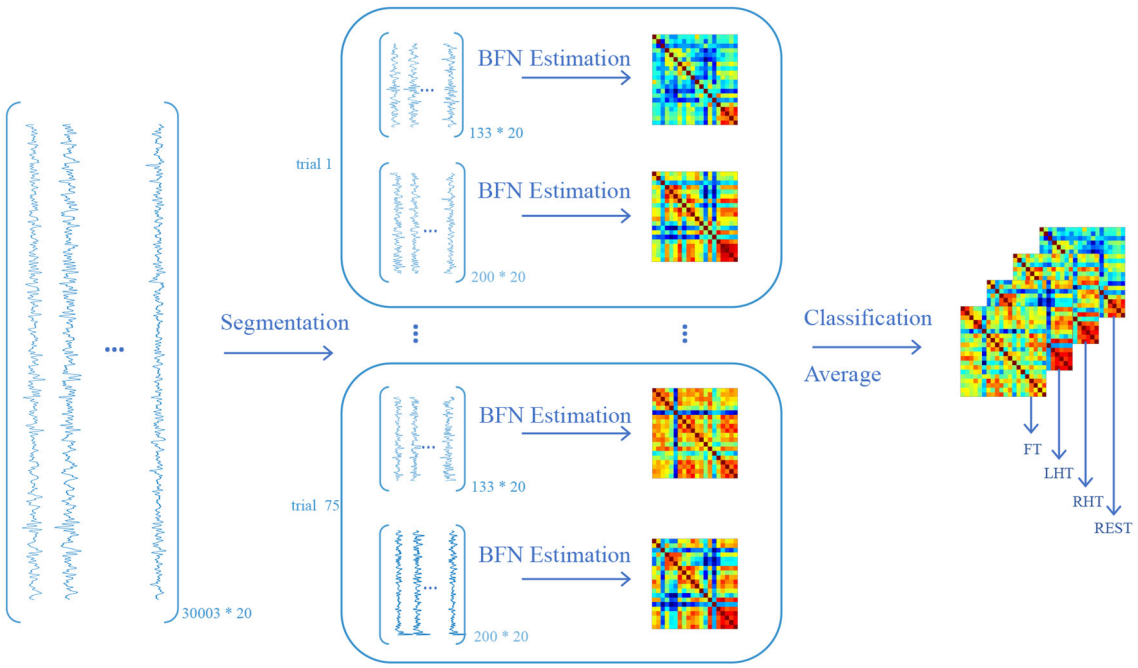


FIGURE 3 | The pipeline of converting functional near-infrared spectroscopy (fNIRS) signals to brain functional networks (BFNs) (only take the first subject in the dataset as an example). The size of the original Oxy-Hb data is 30003*20, where the 30003 is the number of time points, and the 20 is the number of channels. By segmenting the original data according to the phase of task and resting state for each trial, we obtain 150 data matrices, half of which have a size of 133*20 under task state, and the remaining have a size of 200*20 under resting state. Note that, when conducting the segmentation operation, we truncate the time points that lie at task phase and the intermediate of the rest phase, respectively, for ensuring signal purity and eliminating the signal noise caused by task-resting switching. The data matrix is then used to estimate BFN fingerprint by PC. As a result, the subject obtains 150 BFNs, each of which has a size of 20*20. All BFNs are classified into four classes according to four tasks (i.e., RHT, LHT, FT, and REST). Each task class contains 25 BFNs and the resting class contains 75 BFNs. Since the size of each BFN is same, the subject would obtain a new BFN by averaging the element of the corresponding position of all BFNs within one class. Eventually each subject obtains 4 BFNs corresponding to 4 tasks.

under Oxy-Hb view in the dataset as an example to illustrate the conversion process. The first step is signal segmentation according to the trials. In consideration of the time point of one trial consisting of task and resting states, we segment task and resting data. As a result, we acquire 150 signal matrices, half of which correspond to task state and the remaining correspond to resting state. The second step is BFN estimation based on these signal matrices. In particular, the nodes of the BFN correspond to 20 channels, and the edge or edge weights are estimated as the PC between the time series (the columns of the signal matrix) associated with the channels. Finally, these estimated BFNs are divided into four groups according to the type of tasks (including RHT, LHT, FT, and REST), and the BFNs in each group are then averaged to generate a representative BFN for each task. Consequently, we acquire four BFNs for each subject, corresponding to four different tasks.

2.2.2. BFN Identification

Individual identification is performed across tasks based on the "source set" and "target set," as shown in **Figure 4A**. In particular, given a target BFN x_t^*, we calculate its similarity to each BFN $x_s^{(i)}$ in the source set, denoted by $Sim(x_t^*, x_s^{(i)}), i = 1, 2, \cdots, 30$, where the similarity is defined as PC between two BFNs. Then, we use the nearest neighbor principle to predict the label ID^* of the target BFN as follows:

$$ID^* = arg \max_{i \in \{1,2,\cdots,30\}} Sim(x_t^*, x_s^{(i)}) \quad (1)$$

If the predicted label is equal to the actual label, the prediction score is counted as 1, and 0, otherwise.

After matching all pairs of BFNs in the target-source set, we can get the recognition accuracy as

$$ACC = \frac{the\ sum\ of\ prediction\ scores}{the\ total\ number\ of\ subjects} \quad (2)$$

Since different tasks are involved in the dataset, we conduct cross-task individual identification experiment by setting all pairs of target-source mode, as shown in **Figure 4B**.

3. RESULTS

We eventually get the prediction accuracies within all pairs of target-source modes and the experimental results between the same views are shown in **Figure 5**. The accuracy varies from 20/30 (63%) to 30/30 (100%) in

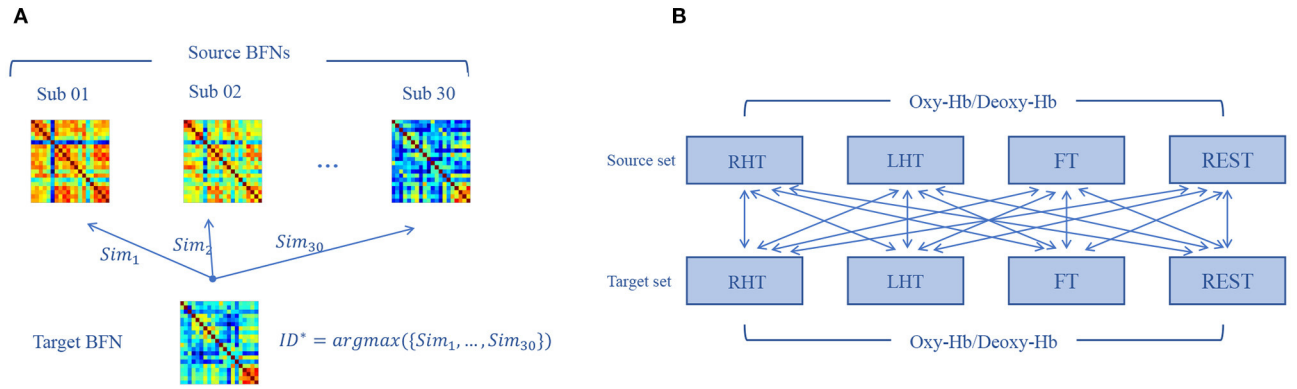


FIGURE 4 | Individual identification procedure. **(A)** We calculate the similarity between a given target brain functional network (BFN) and all the source BFNs, respectively, and the label of the source BFN corresponding to the maximum similarity is the prediction label of the target matrix. **(B)** After data preprocessing, each subject obtains four BFNs corresponding to four tasks. By setting all BFNs from one state as source set and another state as target set for identification, we can get 16 possible combinations of the target-source separation.

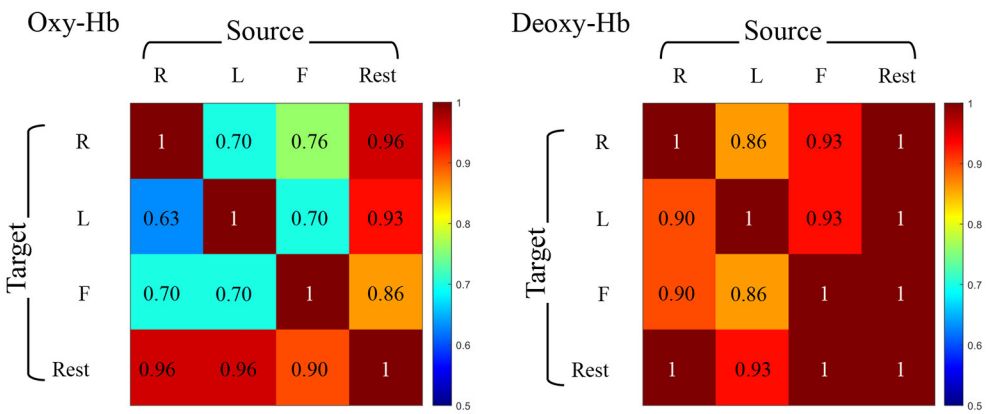


FIGURE 5 | Accuracy of identification under cross-task condition. Individual identification accuracy for the source set from one task and the target set from another task. Each row represents the same target set and each column means the same source set. Oxy-Hb is on the left and Deoxy-Hb is on the right.

different modes (REST vs. RHT, REST vs. LHT, REST vs. FT, RHT vs. LHT, RHT vs. FT, LHT vs. FT for Oxy-Hb and Deoxy-Hb, respectively).

Since the BFNs from different views are involved, we can naturally design cross-view experiment to place BFNs from Oxy-Hb and Deoxy-Hb into source/target sets, respectively. However, compared to the performance associated with the same views, the overall recognition accuracy under cross-view condition shown in **Figure 6** changes significantly, ranging from 9/30 (23%) to 23/30 (76%). Meanwhile, we acquire the highest accuracy in REST–REST mode and the lowest accuracy in RHT–FT mode.

4. DISCUSSION

4.1. Results Comparison

On the basis of experimental results, we have the following findings: (1) The identification accuracy under cross-task is 95.69% ± 0.29 for Deoxy-Hb and 86.00% ± 1.87 for Oxy-Hb, meaning that the identification performance based on Deoxy-Hb is better and more stable than Oxy-Hb in this experiment. Meanwhile, the identification accuracy under cross-view of using Deoxy-Hb as the source set to identify Oxy-Hb as the target set is higher than the opposite view setting. This finding encourages us to explore more potential on cross-view identification based on fNIRS data. (2) The accuracy under cross-task is well above the accuracy under cross-view. To explain this finding, we retrace the experimental process, and discover that the similarity between BFNs estimated under different tasks by one subject is much higher than the similarity between BFNs estimated under the same task state by different subjects. That is, the variability of BFNs is closely related to individual behavioral differences, but is more dependent on the inherent structure and function of the brain itself. (3) In most cases, the resting-state BFNs can be uniquely identified by a given BFN obtained from another task state. This phenomenon illustrates that the BFN in the rest phase

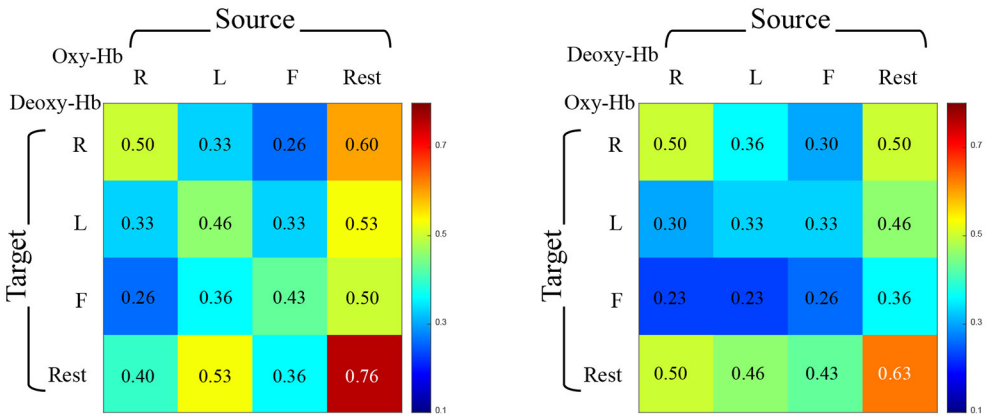


FIGURE 6 | Accuracy of identification under cross-view condition. Individual identification accuracy between Oxy-Hb and Deoxy-Hb consists of four tasks. On the left, each row represents the target set of Deoxy-Hb, and each column represents the source set of Oxy-Hb. On the right side, it is the opposite.

is more suitable for individual identification than the BFNs in the task phase.

4.2. Limitations

However, there are some limitations that need to be considered. At the data level, in addition to the unavoidable small sample, all channels are located in motion cortex, which prevents us to estimate more representative BFN fingerprints based on the whole-brain changes in blood oxygen concentration. At the method level, we combine PC and the nearest neighbor scheme to conduct the identification experiment. In particular, PC is the simplest and wildly used method for estimating BFNs, but it captures the full correlation between pairs of channels and does not remove the confounding effect of other channels. According to previous studies (Hiwa et al., 2016; Guo et al., 2021; Sun et al., 2021; Xue et al., 2022), we can estimate more discriminative BFNs to identify individual. This is a direction for future research.

4.3. Additional Considerations

Note that, since the subjects' behaviors are constantly changing, BFNs are also vary dynamically and significantly within short periods of time (Hutchison et al., 2013). Hence, longer measurement time is one of the prerequisites for obtaining the discriminative and stable BFNs, and future work should focus on the correlation between the length of the signal and the discriminative nature of the BFN. In addition, we find that all signals of one subject in cross-task experiment are collected on the same day, and it is unclear to what extent does the interval between sessions affect the discriminative BFN fingerprints. Future work should focus on the stability or variability of BFN fingerprints over several months or years rather than days.

5. CONCLUSION

In this paper, we conduct individual identification experiments on fNIRS data under cross-task and cross-view conditions, respectively. The identification process includes the BFNs estimation and identification. In particular, we calculate PC between BFNs as similarity, and then evaluate the feasibility of subject recognition. The experimental results show that fNIRS-based BFN fingerprints have good bio-specificity and the properties of difficulty to imitation, which have the potential to serve as an alternative biometric feature for identifying individuals. However, this method, in this paper, only considers the similarity of BFNs estimated between different states or views, without mining the association between BFNs. Therefore, we plan to explore the consistency of BFNs based on fNIRS from the perspective of multi-task, even multi-view in the future.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

HR and LQ designed the study. HR downloaded and analyzed the data, performed experiments, and drafted the manuscript. HR, SZ, LZ, FZ, and LQ revised the manuscript. All the authors read and approved the final manuscript.

FUNDING

This work was partly supported by the National Natural Science Foundation of China (nos. 61976110, 62176112, and 11931008), Natural Science Foundation of Shandong Province (no. ZR202102270451), and the Open Project of Liaocheng University Animal Husbandry Discipline (no. 319312101-01).

REFERENCES

Bak, S., Park, J., Shin, J., and Jeong, J. (2019). Open-access fnirs dataset for classification of unilateral finger-and foot-tapping. *Electronics* 8, 1486. doi: 10.3390/electronics8121486

Blankertz, B., Acqualagna, L., Dähne, S., Haufe, S., Schultze-Kraft, M., Sturm, I., et al. (2016). The berlin brain-computer interface: progress beyond communication and control. *Front. Neurosci.* 10:530. doi: 10.3389/fnins.2016.00530

da Silva Castanheira, J., Orozco, H. D., Misic, B., and Baillet, S. (2021). Meg, myself, and i: individual identification from neurophysiological brain activity. *bioRxiv*. doi: 10.1101/2021.02.18.431803

Delpy, D. T., Cope, M., van der Zee, P., Arridge, S., Wray, S., and Wyatt, J. (1988). Estimation of optical pathlength through tissue from direct time of flight measurement. *Phys. Med. Biol.* 33,1433. doi: 10.1088/0031-9155/33/12/008

Duan, L., Zhang, Y.-J., and Zhu, C.-Z. (2012). Quantitative comparison of resting-state functional connectivity derived from fnirs and fmri: a simultaneous recording study. *Neuroimage* 60, 2008–2018. doi: 10.1016/j.neuroimage.2012.02.014

Finn, E. S., Shen, X., Scheinost, D., Rosenberg, M. D., Huang, J., Chun, M. M., et al. (2015). Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nat. Neurosci.* 18, 1664–1671. doi: 10.1038/nn.4135

Georgopoulos, A. P. (1988). Neural integration of movement: role of motor cortex in reaching. *FASEB J.* 2, 2849–2857. doi: 10.1096/fasebj.2.13.3139485

Gershon, A., Wang, C., Guan, J., Vasilevska-Ristovska, J., Cicutto, L., and To, T. (2009). Identifying individuals with physcian diagnosed copd in health administrative databases. *COPD* 6, 388–394. doi: 10.1080/15412550903140865

Ginther, C., Issel-Tarver, L., and King, M.-C. (1992). Identifying individuals by sequencing mitochondrial dna from teeth. *Nat. Genet.* 2, 135–138. doi: 10.1038/ng1092-135

Gratton, G., Brumback, C. R., Gordon, B. A., Pearson, M. A., Low, K. A., and Fabiani, M. (2006). Effects of measurement method, wavelength, and source-detector distance on the fast optical signal. *Neuroimage* 32, 1576–1590. doi: 10.1016/j.neuroimage.2006.05.030

Guo, T., Zhang, Y., Xue, Y., Qiao, L., and Shen, D. (2021). Brain function network: higher order vs. more discrimination. *Front. Neurosci.* 1033:696639. doi: 10.3389/fnins.2021.696639

Hilger, K., Ekman, M., Fiebach, C. J., and Basten, U. (2017). Intelligence is associated with the modular structure of intrinsic brain networks. *Sci. Rep.* 7, 1–12. doi: 10.1038/s41598-017-15795-7

Hilger, K., Fukushima, M., Sporns, O., and Fiebach, C. J. (2020). Temporal stability of functional brain modules associated with human intelligence. *Hum. Brain Mapp.* 41, 362–372. doi: 10.1002/hbm.24807

Hiwa, S., Hanawa, K., Tamura, R., Hachisuka, K., and Hiroyasu, T. (2016). Analyzing brain functions by subject classification of functional near-infrared spectroscopy data using convolutional neural networks analysis. *Comput. Intell. Neurosci.* 2016:1841945. doi: 10.1155/2016/1841945

Hutchison, R. M., Womelsdorf, T., Allen, E. A., Bandettini, P. A., Calhoun, V. D., Corbetta, M., et al. (2013). Dynamic functional connectivity: promise, issues, and interpretations. *Neuroimage* 80, 360–378. doi: 10.1016/j.neuroimage.2013.05.079

Irani, F., Platek, S. M., Bunce, S., Ruocco, A. C., and Chute, D. (2007). Functional near infrared spectroscopy (fnirs): an emerging neuroimaging technology with important applications for the study of brain disorders. *Clin. Neuropsychol.* 21, 9–37. doi: 10.1080/13854040600910018

Jain, A. K., Pankanti, S., Prabhakar, S., Hong, L., and Ross, A. (2004). "Biometrics: a grand challenge," in *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004., Vol. 2* (Cambridge, UK: IEEE), 935–942.

Jain, A. K., Ross, A., and Pankanti, S. (2006). Biometrics: a tool for information security. *IEEE Trans. Inf. Forensics Security* 1, 125–143. doi: 10.1109/TIFS.2006.873653

Khan, R. A., Naseer, N., Saleem, S., Qureshi, N. K., Noori, F. M., and Khan, M. J. (2020). Cortical tasks-based optimal filter selection: an fnirs study. *J. Healthc. Eng.* 2020:9152369. doi: 10.1155/2020/9152369

Nguyen, H.-D., Yoo, S.-H., Bhutta, M. R., and Hong, K.-S. (2018). Adaptive filtering of physiological noises in fnirs data. *Biomed. Eng. Online* 17, 1–23. doi: 10.1186/s12938-018-0613-2

Prabhakar, S., Pankanti, S., and Jain, A. K. (2003). Biometric recognition: Security and privacy concerns. *IEEE Security Privacy* 1, 33–42. doi: 10.1109/MSECP.2003.1193209

Quaresima, V., and Ferrari, M. (2019). Functional near-infrared spectroscopy (fnirs) for assessing cerebral cortex function during human behavior in natural/social situations: a concise review. *Organ. Res. Methods* 22, 46–68. doi: 10.1177/1094428116658959

Sareen, E., Zahar, S., Van De Ville, D., Gupta, A., Griffa, A., and Amico, E. (2021). Exploring meg brain fingerprints: evaluation, pitfalls and interpretations. *bioRxiv*. doi: 10.1016/j.neuroimage.2021.118331

Sastry, N. C., Roy, D., and Banerjee, A. (2021). Unsupervised characterization of dynamic functional connectivity reveals age-associated differences in temporal stability and connectivity states during rest and task. *bioRxiv*. doi: 10.1101/2021.07.08.451590

Schmidt, M. I., Duncan, B. B., Bang, H., Pankow, J. S., Ballantyne, C. M., Golden, S. H., et al. (2005). Identifying individuals at high risk for diabetes: The atherosclerosis risk in communities study. *Diabetes Care* 28, 2013–2018. doi: 10.2337/diacare.28.8.2013

Strangman, G., Boas, D. A., and Sutton, J. P. (2002). Non-invasive neuroimaging using near-infrared light. *Biol. Psychiatry* 52, 679–693. doi: 10.1016/S0006-3223(02)01550-0

Sun, L., Xue, Y., Zhang, Y., Qiao, L., Zhang, L., and Liu, M. (2021). Estimating sparse functional connectivity networks via hyperparameter-free learning model. *Artif. Intell. Med.* 111:102004. doi: 10.1016/j.artmed.2020.102004

Wang, M., El-Fiqi, H., Hu, J., and Abbass, H. A. (2019a). Convolutional neural networks using dynamic functional connectivity for eeg-based person identification in diverse human states. *IEEE Trans. Inf. Forensics Security* 14, 3259–3272. doi: 10.1109/TIFS.2019.2916403

Wang, M., Hu, J., and Abbass, H. A. (2020). Brainprint: Eeg biometric identification based on analyzing brain connectivity graphs. *Pattern Recognit.* 105:107381. doi: 10.1016/j.patcog.2020.107381

Wang, M., Yuan, Z., and Niu, H. (2019b). Reliability evaluation on weighted graph metrics of fnirs brain networks. *Quant. Imaging Med. Surg.* 9, 832. doi: 10.21037/qims.2019.05.08

Xue, Y., Zhang, Y., Zhang, L., Lee, S.-W., Qiao, L., and Shen, D. (2022). Learning brain functional networks with latent temporal dependency for mci identification. *IEEE Trans. Biomed. Eng.* 69, 590–601. doi: 10.1109/TBME.2021.3102015

Zhang, F., Cheong, D., Khan, A. F., Chen, Y., Ding, L., and Yuan, H. (2021). Correcting physiological noise in whole-head functional

near-infrared spectroscopy. *J. Neurosci. Methods* 360:109262. doi: 10.1016/j.jneumeth.2021.109262

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflictof interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 17 February 2022
doi: 10.3389/fnins.2022.813056

Intracranial Aneurysm Rupture Risk Estimation With Multidimensional Feature Fusion

*Xingwei An[1,2], Jiaqian He[1], Yang Di[1], Miao Wang[1], Bin Luo[1,3], Ying Huang[3] and Dong Ming[1,2,4]**

[1] Academy of Medical Engineering and Translational Medicine, Tianjin University, Tianjin, China, [2] Tianjin Center for Brain Science, Tianjin, China, [3] Department of Neurosurgery, Huanhu Hospital of Tianjin University, Tianjin, China, [4] Department of Biomedical Engineering, School of Precision Instruments and Optoelectronics Engineering, Tianjin University, Tianjin, China

The rupture of aneurysms is the main cause of spontaneous subarachnoid hemorrhage (SAH), which is a serious life-threatening disease with high mortality and permanent disability rates. Therefore, it is highly desirable to evaluate the rupture risk of aneurysms. In this study, we proposed a novel semiautomatic prediction model for the rupture risk estimation of aneurysms based on the CADA dataset, including 108 datasets with 125 annotated aneurysms. The model consisted of multidimensional feature fusion, feature selection, and the construction of classification methods. For the multidimensional feature fusion, we extracted four kinds of features and combined them into the feature set, including morphological features, radiomics features, clinical features, and deep learning features. Specifically, we applied the feature extractor 3D EfficientNet-B0 to extract and analyze the classification capabilities of three different deep learning features, namely, no-sigmoid features, sigmoid features, and binarization features. In the experiment, we constructed five distinct classification models, among which the *k*-nearest neighbor classifier showed the best performance for aneurysm rupture risk estimation, reaching an F2-score of 0.789. Our results suggest that the full use of multidimensional feature fusion can improve the performance of aneurysm rupture risk assessment. Compared with other methods, our method achieves the state-of-the-art performance for aneurysm rupture risk assessment methods based on CADA 2020.

Keywords: intracranial aneurysm, risk estimation, feature fusion, machine learning, radiomics

OPEN ACCESS

Edited by:
Feng Liu,
Tianjin Medical University General Hospital, China

Reviewed by:
Yang Yang,
Fourth Military Medical University, China
Cheng Luo,
University of Electronic Science and Technology of China, China
Dong Wen,
University of Science and Technology Beijing, China

***Correspondence:**
Dong Ming
richardming@tju.edu.cn

Specialty section:
This article was submitted to Brain Imaging Methods, a section of the journal Frontiers in Neuroscience

Received: *11 November 2021*
Accepted: *05 January 2022*
Published: *17 February 2022*

Citation:
An X, He J, Di Y, Wang M, Luo B, Huang Y and Ming D (2022) Intracranial Aneurysm Rupture Risk Estimation With Multidimensional Feature Fusion. Front. Neurosci. 16:813056. doi: 10.3389/fnins.2022.813056

INTRODUCTION

An intracranial aneurysm is an abnormal local dilatation of the cerebral artery due to the weakness of the vessel wall. It occurs in approximately 2–5% of the population and is the main cause of non-traumatic subarachnoid hemorrhage (SAH) (Xu et al., 2019). SAH caused by aneurysm rupture is a serious neurological disease with high mortality and morbidity. Despite treatment technology advances and imaging technology improvements currently, the death rate of SAH is approximately 40–50% and leaves approximately half of survivors with permanent neurological deficits (Boulouis et al., 2017; Roked and Reddy, 2020). Therefore, early detection of aneurysms and rupture risk assessment of unruptured aneurysms are clinically significant for the treatment and prognosis of patients.

Aneurysmal morphology such as shape and size, patient-specific clinical factors such as hypertension, smoking, a history of SAH, sex, and population, as well as hemodynamics of aneurysms are known to be risk factors associated with intracranial aneurysm rupture (Abboud et al., 2017; Boulouis et al., 2017). At present, digital subtraction angiography (DSA), computed tomography angiography (CTA), and magnetic resonance angiography (MRA) are primary imaging techniques clinically for rupture risk assessment of aneurysm. Doctors comprehensively assess the rupture risk of aneurysm mainly based on the high-resolution angiographic images and patient-specific clinical factors. However, due to the variations in the level of experience and proficiency among physicians, the evaluation is highly subjective and lacks consistency among experts. Therefore, it is necessary to develop a computer-aided diagnosis system for assessing the rupture risk of the aneurysm to assist doctors in diagnosis and decision-making to avoid overtreatment and risks associated with surgery.

As an important branch of artificial intelligence, machine learning (ML) enables to identify and process complex relationships between features in big data sets and can be rapidly applied to unknown data for prediction (Senders et al., 2018). Some recent studies have shown that ML plays an important role in the rupture risk assessment of aneurysm. Silva et al. (2019) demonstrated the ability of ML to distinguish ruptured and unruptured aneurysms based on conventional radiographic characteristics of aneurysms and patient-specific clinical features. Tanioka et al. (2020) constructed ML classification models for the identification of ruptured aneurysms by applying manually measured morphological variables and hemodynamic parameters. However, for the assessment of the rupture state of aneurysms, incorporating abundant variables into the classification model is the key to affecting the assessment performance.

Radiomics refers to the technology of analyzing and mining high volumes of quantitative features extracted from medical images and then developing a robust model based on the key information that works to support the clinical decision ultimately (Limkin et al., 2017). It has shown considerable potential in many medical challenges, such as auxiliary diagnosis, classification, and grading of diseases (Yun et al., 2019; Peeken et al., 2021). Recently, the application of radiomics combined with ML in the rupture assessment of intracranial aneurysms has shown initial results. A preliminary study (Ou et al., 2021) employed conventional morphological features and radiomics features to construct an ML classification model, which proves the potential use of radiomics signatures in predicting aneurysm rupture. Alwalid et al. (2021) developed a radiomics classification model on CTA images to identify patients with ruptured aneurysms. However, the ability of radiomics features characterizing regions of interest is subject to low-level properties to some extent (Hua et al., 2020). In recent years, deep learning methods, especially convolutional neural networks (CNNs), have achieved excellent results in dealing with the tasks of classification, segmentation, and detection in medical imaging (Zeng et al., 2020; Yang et al., 2021). The convolution and pooling operations in the network automatically learn and capture the local details as well as more complex information and structure features of images, so as to obtain the abstract representation of the image at various scales.

Thus, we deemed that the complementary advantages of deep learning and radiomics technologies could enrich feature representations of medical images and further improve the prediction performance for the rupture risk of the aneurysm. In this study, we explored multidimensional features derived from both high-resolution angiographic images and high-quality three-dimensional aneurysm modeling data to build a semiautomatic prediction model for rupture risk estimation of the aneurysm.

MATERIALS AND METHODS

Dataset

The challenge for aneurysm rupture risk estimation is task 3 in cerebral aneurysm detection and analysis (CADA) challenge. The challenge organizers provided 110 datasets with 128 annotated aneurysms. The image data of patients were acquired utilizing the digital subtraction AXIOM Artis C-arm system by a rotational acquisition time of 5 s with 126 frames. Postprocessing was performed using LEONARDO InSpace 3D (Siemens, Forchheim, Germany). All segmentation masks provided by a skilled annotator were checked by an experienced neurosurgeon later. **Figure 1** shows the example of the two types of segmentation masks (stereolithography files and image files) for the same aneurysm. In addition, the rupture state and rupture information of each aneurysm are provided. After removing 3 cases with missing information, the remaining 125 cases are included for model training and validation.

Feature Extraction

In this study, we extracted multidimensional features derived from both angiographic images and three-dimensional aneurysm modeling data, consisting of radiomics features, morphological features, deep learning features, and clinical information. Details are described as follows.

Radiomics Features

Before radiomics feature extraction, image preprocessing with intensity normalization to the grayscale range of [0, 100] and isotropic resampling to a uniform pixel dimension of $0.5 \times 0.5 \times 0.5$ mm^3 was performed. We extracted radiomics features including aneurysm intensity, shape-based, and texture

FIGURE 1 | An example case for angiographic image of the aneurysm and corresponding two types of segmentation masks.

features from regions of interest defined by the angiography images and segmentation masks (image files) using the open-source PyRadiomics package (version 3.0.1) (van Griethuysen et al., 2017). Texture features are visual features that reflect the uniformity of the image and the slow or periodic changes on the surface of the object. Specifically, these extracted characteristics were divided into the following seven groups, including first-order statistics (18 features), shape 3D-based (14 features), gray-level co-occurrence matrix (24 features), gray-level run-length matrix (16 features), gray-level size zone matrix (16 features), neighboring gray-tone difference matrix (5 features), and gray-level dependence matrix (14 features). Most features listed above were in accordance with the recommendations of the Imaging Biomarker Standardization Initiative (IBSI) (Zwanenburg et al., 2020).

Morphological Features Based on Stereolithography Files

Currently, shape-based features have shown to be beneficial in assessing the rupture risk of the aneurysm (Abboud et al., 2017; Silva et al., 2019; Tanioka et al., 2020). Therefore, we extracted morphological features of the aneurysm based on the three-dimensional modeling data for a more reliable estimation, including the length, width, height, surface area, and volume. In addition, we considered that curvature features provided additional representations for describing the morphology of aneurysms. The extracted curvature features of aneurysms included the principal curvature, Gaussian curvature, and mean curvature. The maximum, minimum, mean and standard deviation of curvature were calculated, respectively. In this study, 25 morphological features were extracted for each case.

Deep Learning Features

To acquire high-level image features, we employed the convolutional neural network method to mine the abstract features. In this study, we selected and trained a 3D EfficientNet-B0 as the feature extractor (Tan and Le, 2019), which balanced the depth, width, and resolution of the model with a highly effective compound coefficient, thereby achieving satisfactory accuracy. **Figure 2** shows the network architecture, and its main building block is MBConv (Sandler et al., 2018) with squeeze-and-excitation optimization (Hu et al., 2020), as shown in **Figure 3**. We, respectively, took the image and mask as the input of the convolutional neural network and explored three various deep learning features from convolutional neural network outputs.

1) No-sigmoid features: outputs of the feature maps only through the final fully connected layer.
2) Sigmoid features: outputs of the feature maps through the final fully connected layer and sigmoid function successively.
3) Binarization features: outputs of the feature maps through the final fully connected layer, sigmoid function, and binarization operation successively. Binarization operation can be calculated as follows:

$$f(x) = \begin{cases} 1, & x \geq 0.5 \\ 0, & x < 0.5 \end{cases} \quad (1)$$

In summary, we combined the multidimensional features above to enrich the feature representation of rupture risk. All the feature sets are shown in **Figure 4**.

We obtained three types of feature sets as original multidimensional feature sets, namely, no-sigmoid type, sigmoid type, and binarization type. Each original feature set contained four parts of features, which were morphological features (25 features), radiomics features (107 features), the corresponding type of deep learning features (2 features, different feature sets have different deep learning features), and patient-specific clinical factors (2 features, sex, and age). That is, each of the three types of original multidimensional feature sets contained 136 features.

Nested Cross-Validation

Cross-validation can evaluate the generalization ability of ML algorithms to data sets independent of training data and prevent over-fitting effectively (Arlot and Celisse, 2010). Stratified sampling was used in this study to perform cross-validation to ensure that the proportion of samples in each target class in the training set and validation set is the same as the full set. Considering that this traditional cross-validation method cannot solve the problem of optimal model selection and model parameter tuning well, we used nested cross-validation (Varoquaux et al., 2017) in order to search for hyperparameters by estimating the generalization error of the basic model to obtain the best parameters of the model. The process of 8-fold cross-validation is shown in **Figure 5**, which contains a two-loop nested cross-validation scheme. Hyperparameters were optimized using grid search as part of the inner loop. The optimal hyperparameters were then used for testing on the outer loop.

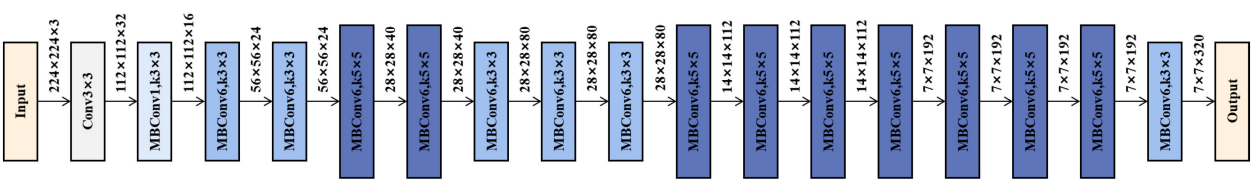


FIGURE 2 | EfficientNet-B0 structure.

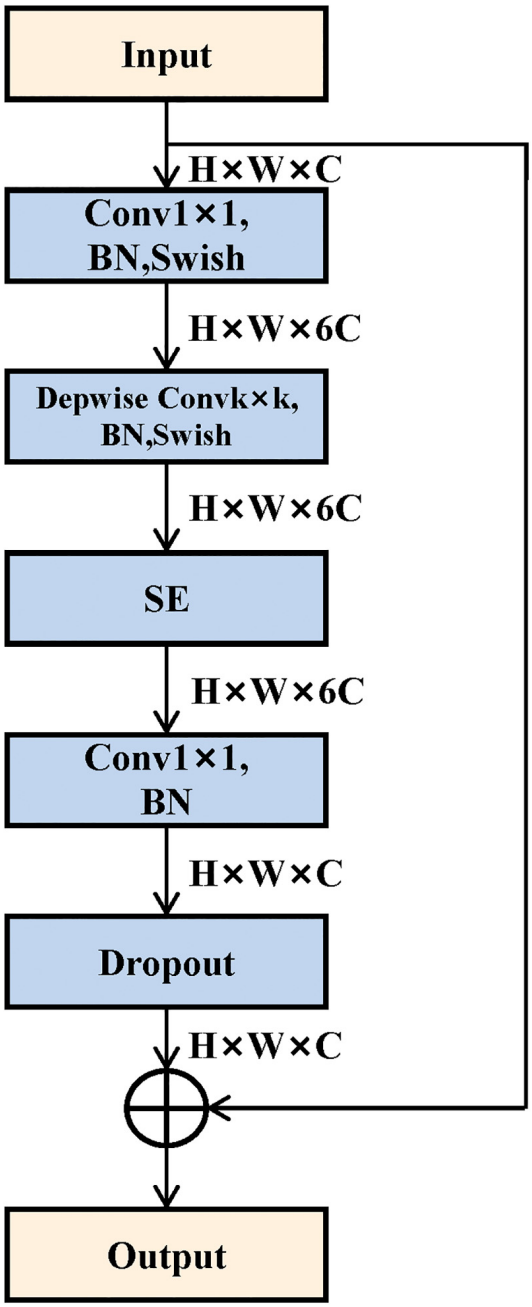


FIGURE 3 | Mobile inverted bottleneck convolution (MBConv) module structure. It mainly consists of depthwise convolution and squeeze-and-excitation block.

Feature Selection

We uniformly standardized all the features by removing the mean and scaling to unit variance. To improve the stability and generalization performance of the model, it is critical to select discriminative features. We chose the random forest (RF) (Genuer et al., 2010) and XGboost (Chen and Guestrin, 2016) methods for feature selection. Both methods could generate a ranking of the feature importance after training and further select some profitable features by setting the threshold. In this study, the following two feature selection steps were applied.

Step 1: We took the features selected simultaneously by RF and XGboost based on the training set of the original multidimensional feature set of each cross-validation fold and merged the features retained by all 8 folds to get the feature set M.

Step 2: We accumulated the important scores of the features in each fold and counted the top 1/2 features of the RF and Xgboost methods in the M set to obtain feature set M_r and M_x, respectively. M_r represented feature selection by RF. M_x represented feature selection by Xgboost. The feature subset N was obtained by $N = M_r \cap M_x$.

Therefore, the corresponding feature subsets were obtained from the three original feature sets, among which 22 features were retained for the no-sigmoid type feature subset, 24 features for the sigmoid type feature subset, and 24 for the binarization type feature subset. The selected features in the three types of feature subsets are shown in **Supplementary Table 1**.

Classification Model

To find an optimal classifier for the classification task of ruptured and unruptured aneurysms, five distinct ML models were used to build the classification model using the selected features, respectively, including support vector machine (SVM) (Cortes and Vapnik, 1995), RF (Breiman, 2001), *k*-nearest neighbor (KNN) (Fix and Hodges, 1989), logistic regression (LR) (Berkson, 1946), and XGBoost (Chen and Guestrin, 2016) classifiers. To ensure the robustness of the experimental results, we adopted 8-fold cross-validation and then took the average of classification metrics as the final result.

Model Evaluation

Considering that the identification of aneurysms at risk is more important than the avoidance of false-positive risk classification, F2-score was selected as the final score metric in the rupture risk estimation challenge. The F2-score integrates two indicators of recall and precision, and it is considered that recall is twice as important as precision, as shown in Eq. 2. In addition, we also calculated other metrics including accuracy (ACC), the area under the curve (AUC), recall, and precision.

$$F2 = \frac{5 * Precision * Recall}{4 * Precision + Recall} \quad (2)$$

RESULTS

Implementation Details

For the EfficientNet-B0 feature extractor, we trained this model on an NVIDIA GeForce RTX 3090 GPU with 24 GB memory. All the images were employed spacing normalization to a common spacing of 0.5 × 0.5 × 0.5 mm^3 and intensity normalization to the grayscale range of [0, 1]. We resized all the images to 128 × 128 × 128 and set total epochs to 100 for each fold of cross-validation, with the learning rate 3e-4 and batch size 4. The AdamW algorithm was adopted to optimize the feature

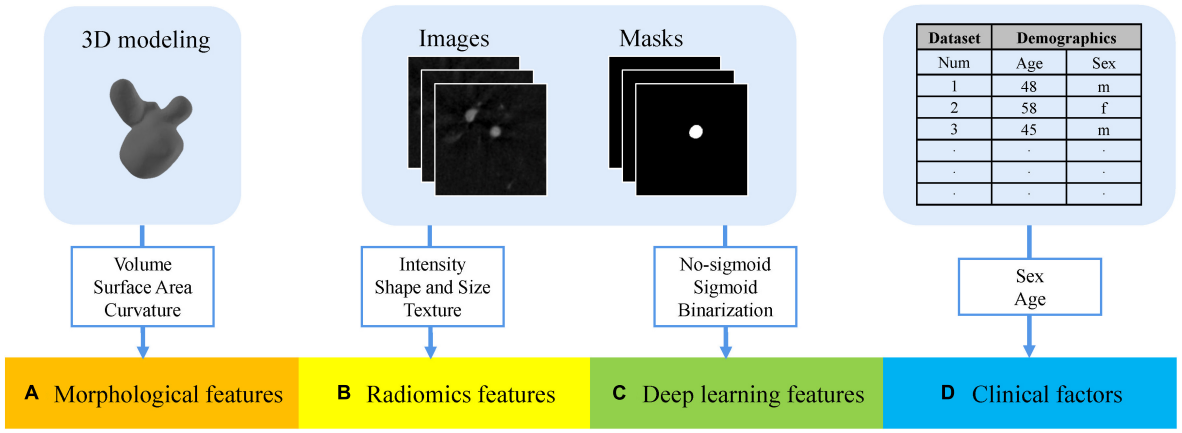


FIGURE 4 | Multidimensional feature set consists of four groups. **(A)** Morphological features extracted from stereolithography files. **(B)** Radiomics features extracted from the angiography images and segmentation masks. **(C)** Deep learning features extracted from the angiography images and segmentation masks. **(D)** Patient-specific clinical factors.

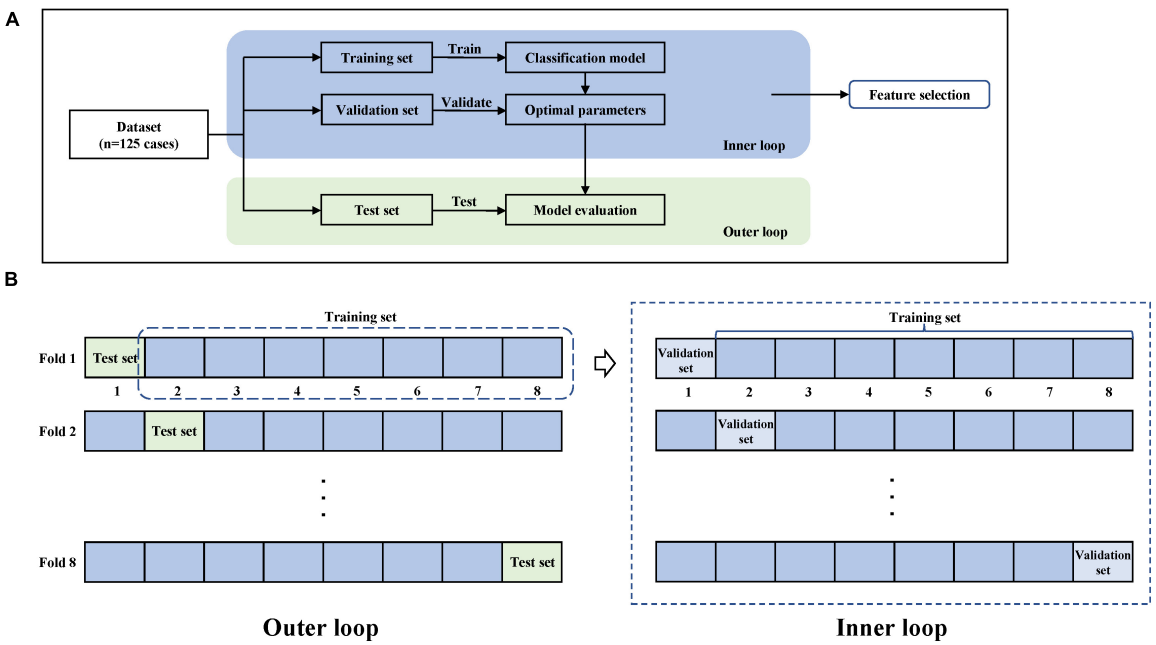


FIGURE 5 | Model training and nested cross-validation. **(A)** General overview. **(B)** 8-fold cross-validation.

extraction network. We also used weight decay with 1e-8. Our other experiments were implemented on an AMD Ryzen 5 5600H CPU @3.30 GHz with 16 GB RAM.

Rupture Risk Estimation Results

After feature selection, we constructed five different classification models with the three feature subsets, and the final results are shown in **Table 1**. It can be seen that the KNN model based on the sigmoid type feature set achieved the best mean performance on the F2-score. Thus, we chose the sigmoid feature subset as the final feature subset. A heat map was constructed to show the association between selected features and aneurysm rupture status based on the feature subset, as shown in **Supplementary Figure 1**.

Table 2 shows the comparison of results among different classifiers constructed with this feature set. Based on the F2-score, the KNN model shows the best result with a mean F2-score of 0.789 on the test set. Thus, the KNN algorithm was chosen as the final model. For the presented five classification models, the KNN classifier shows the best performance on most metrics for aneurysm rupture risk estimation. Evaluating the performance of accuracy, the KNN model shows the best result with a mean of 0.791.

TABLE 1 | The mean F2-score for different feature sets and classification methods on the test set.

Classifier	Binarization	Sigmoid	No-sigmoid
SVM	**0.730**	0.724	0.609
LR	**0.747**	0.731	0.644
RF	0.695	0.675	**0.707**
XGBoost	0.708	**0.715**	0.698
KNN	0.752	**0.789**	0.580

The best results for each specified classifier are highlighted in bold red.

TABLE 2 | Comparison of the results of different classifiers based on the sigmoid feature set.

Classifier	F2-score	ACC	AUC	Precision	Recall
SVM	0.724	0.775	0.820	**0.779**	0.732
LR	0.731	0.776	**0.834**	0.761	0.732
RF	0.675	0.751	0.810	0.771	0.660
XGBoost	0.715	0.767	0.803	0.773	0.714
KNN	**0.789**	**0.791**	0.811	0.755	**0.803**

The best results for each of these metrics are highlighted in bold red.

In addition, the corresponding mean receiver operating characteristic (ROC) curve over all outer folds based on the optimal model is shown in **Figure 6**. The KNN model shows a good performance in the classification of aneurysm rupture with a mean AUC of 0.811 on the test set.

Ivantsits et al. (2021) and Liu et al. (2021), respectively, proposed two excellent semiautomatic aneurysm rupture risk estimation methods on the CADA dataset. **Table 3** shows the comparison on the metrics of our approach with two related works. It is observed that our approach achieved better performance than both related works on the CADA dataset. Under the condition of using the same classifier, our methods performed better, which proves that the features we extracted were more suitable and effective for this task.

DISCUSSION

Intracranial aneurysm rupture is a catastrophic medical event with high mortality and permanent disability risk. A timely and accurate rupture risk estimation of aneurysms is necessary for clinical treatment. At present, the widespread availability of vascular neuroimaging has allowed more unruptured aneurysms to be discovered incidentally, but the treatment decision-making for aneurysms is still a challenge that the clinic needs to face because doctors are required to weigh the risk of SAH along with the risks of surgical or endovascular treatments and subsequent complications with discretion (Boulouis et al., 2017).

The morphology of the intracranial aneurysm is considered to be associated with the rupture state of the aneurysm (Abboud et al., 2017; Boulouis et al., 2017). Most of the previous

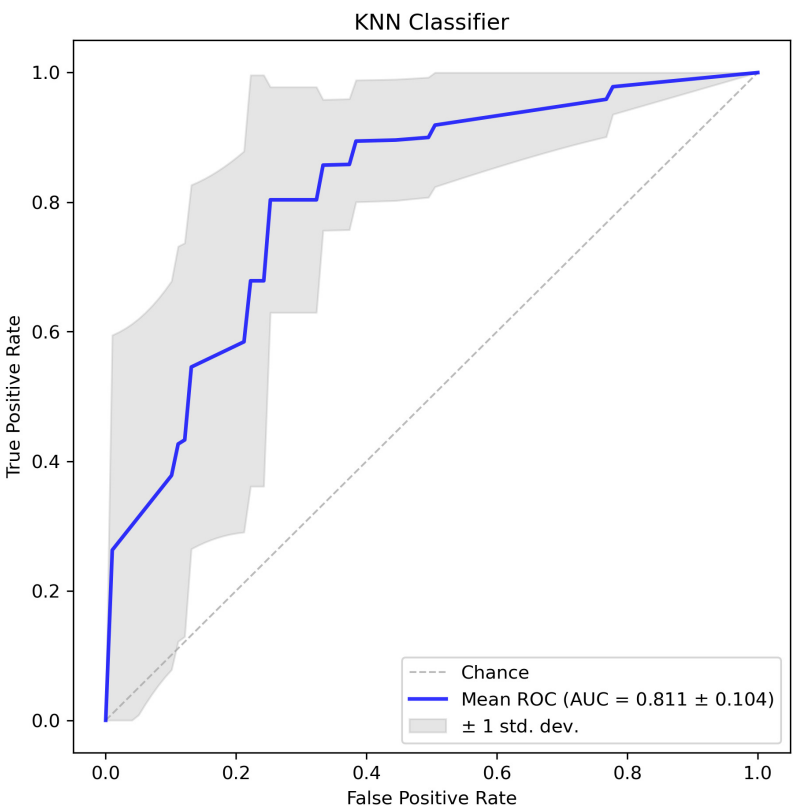


FIGURE 6 | The mean receiver operating characteristic (ROC) curve of the *k*-nearest neighbor (KNN) classifier based on 8-fold cross-validation.

TABLE 3 | Aneurysm rupture risk estimation performance of our method and two related works based on CADA dataset.

Classifier	Method	F2-score	ACC	AUC	Precision	Recall
XGBoost	Liu et al., 2021	0.673	0.652	n/a	0.583	0.700
	Ours	**0.715**	**0.767**	**0.803**	**0.773**	**0.714**
KNN	Ivantsits et al., 2021	0.690	0.660	n/a	n/a	n/a
	Ours	**0.789**	**0.791**	**0.811**	**0.755**	**0.803**
RF	Ivantsits et al., 2021	n/a	0.690	n/a	n/a	n/a
	Ours	**0.675**	**0.751**	**0.810**	0.771	**0.660**
SVM	Ours	0.724	0.775	0.820	**0.779**	0.732
LR	Ours	0.731	0.776	**0.834**	0.761	0.732

Better results for each specified classifier are highlighted in bold black. The best results for each of these metrics are highlighted in bold red.

reports employed manually measured morphological indicators to identify the rupture risk (Liu et al., 2018; Silva et al., 2019; Tanioka et al., 2020), which did not fully explore the rich information of angiography images.

In this study, we proposed a classification method based on diverse types of risk factors for the assessment of aneurysm rupture state, so as to promote timely management of patients and provide some guidance for follow-up treatment decisions. In the pipeline for assessing the rupture risk of aneurysm, our proposed method consisted of multidimensional feature fusion, feature selection to capture the discriminative variables, and followed by the construction of classification models. We combined multidimensional feature representations related to rupture risk factors of aneurysms including radiomics features, morphological features, deep learning features, and patient-specific clinical factors. Considering the powerful feature extraction capability of deep learning for images, we took the deep learning network as a feature extractor to extract and analyze the classification capability of three different deep learning features. The results indicate a great potential of the sigmoid type feature subset as a risk factor for intracranial aneurysm rupture estimation.

The sigmoid type feature subset included deep learning descriptors, shape descriptors, first-order histogram descriptors, and texture descriptors. As high-level semantic features, the deep learning features proposed in this study could learn complex information patterns and structural features in image data, which are invisible to human eyes. Curvature features represented as additional morphological features may reflect changes related to the aneurysm rupture state. Radiomics features are calculated in a pixel-by-pixel manner, which can quantitatively describe the morphology of the 3D lesion. In this study, nine radiomics features were finally retained. It proves the potential of radiomics features for the classification of aneurysm rupture, which is consistent with previous studies (Alwalid et al., 2021; Ou et al., 2021). Texture patterns within the aneurysm region especially the aneurysmal lumen may be caused by the uneven distribution of contrast agents, which were thought to be related to turbulence flow (Ou et al., 2021). It is generally considered that turbulent flow could activate inflammatory mechanisms and could be associated with higher-risk lesions (George et al., 2016). This further explains why texture features could be used as the risk factor for assessing aneurysm rupture. For clinical variables, both sex and age were not included in the final feature subset. It could be due to its complicated mechanism on aneurysm rupture and the experiment being based on a small data set, and further studies are needed to prove the relationship between clinical variables and rupture outcome.

Our study has some limitations that are worth noting. One is that it takes some computational cost to extract deep learning features due to the large size of angiographic images. The other is that due to the limited amount of data provided, further verification is required on external data. In the future, we plan to collect clinical data to verify the robustness of our approach and take steps to further optimize the performance of our model to achieve a more efficient automatic aneurysm rupture risk assessment. Recent studies have shown that computer-aided diagnosis algorithms for aneurysm detection have the potential to shorten reading times and enhance the performances of radiologists (Shi et al., 2020). A further idea is considered to effectively integrate this work with aneurysm detection to build a complete automatic aneurysm diagnosis system, which may improve efficiency in the radiology department (Alwalid et al., 2021) and promote timely management for patients.

CONCLUSION

In this study, we assumed that multidimensional feature fusion and feature selection strategies are necessary to enhance the level of aneurysm rupture risk assessment. Based on the inspiration, we combined morphological features, radiomics features, clinical features, and deep learning features with the feature extractor 3D EfficientNet-B0 to propose a novel semiautomatic ML algorithm for aneurysm rupture risk assessment. Our results demonstrate that the multidimensional risk factors we proposed could improve the ability to identify the ruptured state of the aneurysm. Compared with other methods, our method outperforms the state-of-the-art aneurysm rupture risk assessment method based on CADA 2020, which shows the good prospect of application in decision support systems for patients with aneurysms.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. It can be found at CADA Rupture Risk Estimation Challenge: https://cada-rre.grand-challenge.org.

AUTHOR CONTRIBUTIONS

DM and YH contributed to the conception and design of the study. JH, YD, and MW performed data analysis. JH and XA

drafted the manuscript. XA edited the manuscript and supervised the entire study. BL interpreted the annotated data. All authors contributed to the study and approved the submitted version.

FUNDING

This study was supported in part by the National Natural Science Foundation of China (Grant Nos. 81925020 and 81630051) and the Tianjin Science and Technology Program (Grant No. 20JCZDJC00620).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2022.813056/full#supplementary-material

REFERENCES

Abboud, T., Rustom, J., Bester, M., Czorlich, P., Vittorazzi, E., Pinnschmidt, H. O., et al. (2017). Morphology of Ruptured and Unruptured Intracranial Aneurysms. *World Neurosurg.* 99, 610–617. doi: 10.1016/j.wneu.2016.12.053

Alwalid, O., Long, X., Xie, M., Yang, J., Cen, C., Liu, H., et al. (2021). CT Angiography-Based Radiomics for Classification of Intracranial Aneurysm Rupture. *Front. Neurol.* 12:619864. doi: 10.3389/fneur.2021.619864

Arlot, S., and Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statist. Surv.* 4, 40–79. doi: 10.1214/09-SS054

Berkson, J. (1946). Limitations of the Application of Fourfold Table Analysis to Hospital Data. *Biometrics Bull.* 2:47. doi: 10.2307/3002000

Boulouis, G., Rodriguez-Régent, C., Rasolonjatovo, E. C., Ben Hassen, W., Trystram, D., Edjlali-Goujon, M., et al. (2017). Unruptured intracranial aneurysms: an updated review of current concepts for risk factors, detection and management. *Rev. Neurol.* 173, 542–551. doi: 10.1016/j.neurol.2017.05.004

Breiman, L. (2001). Random forests. *Mach. Learn.* 45, 5–32. doi: 10.1023/A:1010933404324

Chen, T., and Guestrin, C. (2016). "XGBoost: A Scalable Tree Boosting System," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. (San Francisco: ACM), 785–794. doi: 10.1145/2939672.2939785

Cortes, C., and Vapnik, V. (1995). Support-vector networks. *Mach. Learn.* 20, 273–297. doi: 10.1007/BF00994018

Fix, E., and Hodges, J. L. (1989). Discriminatory Analysis. Nonparametric Discrimination: consistency Properties. *Int. Stat. Rev.* 57:238. doi: 10.2307/1403797

Genuer, R., Poggi, J. M., and Tuleau-Malot, C. (2010). Variable selection using random forests. *Pattern Recognit. Lett.* 31, 2225–2236. doi: 10.1016/j.patrec.2010.03.014

George, E., Giannopoulos, A. A., Aghayev, A., Rohatgi, S., Imanzadeh, A., Antoniadis, A. P., et al. (2016). Contrast inhomogeneity in CT angiography of the abdominal aortic aneurysm. *J. Cardiovasc. Comput. Tomogr.* 10, 179–183. doi: 10.1016/j.jcct.2015.11.006

Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., and Chen, L.-C. (2018). MobileNetV2: Inverted Residuals and Linear Bottlenecks. in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (Salt Lake, UT: IEEE)*, 4510–4520. doi: 10.1109/CVPR.2018.00474

Hu, J., Shen, L., Albanie, S., Sun, G., and Wu, E. (2020). Squeeze-and-Excitation Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* 42, 2011–2023. doi: 10.1109/TPAMI.2019.2913372

Hua, W., Xiao, T., Jiang, X., Liu, Z., Wang, M., Zheng, H., et al. (2020). Lymph-vascular space invasion prediction in cervical cancer: exploring radiomics and deep learning multilevel features of tumor and peritumor tissue on multiparametric MRI. *Biomed. Signal Process Control* 58:101869. doi: 10.1016/j.bspc.2020.101869

Ivantsits, M., Huellebrand, M., Kelle, S., Kuehne, T., and Hennemuth, A. (2021). ""Intracranial Aneurysm Rupture Risk Estimation Utilizing Vessel-Graphs and Machine Learning"," in *Cerebral Aneurysm Detection Lecture Notes in Computer Science*, eds A. Hennemuth, L. Goubergrits, M. Ivantsits, and J.-M. Kuhnigk (Cham: Springer International Publishing), 93–103. doi: 10.1007/978-3-030-72862-5_10

Limkin, E. J., Sun, R., Dercle, L., Zacharaki, E. I., Robert, C., Reuzé, S., et al. (2017). Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology. *Annals. Oncol.* 28, 1191–1206. doi: 10.1093/annonc/mdx034

Liu, J., Chen, Y., Lan, L., Lin, B., Chen, W., Wang, M., et al. (2018). Prediction of rupture risk in anterior communicating artery aneurysms with a feed-forward artificial neural network. *Eur. Radiol.* 28, 3268–3275. doi: 10.1007/s00330-017-5300-3

Liu, Y., Yang, Y., Lin, Y., Li, Y., Wei, D., Ma, K., et al. (2021). ""Cerebral Aneurysm Rupture Risk Estimation Using XGBoost and Fully Connected Neural Network"," in *Cerebral Aneurysm Detection Lecture Notes in Computer Science*, eds A. Hennemuth, L. Goubergrits, M. Ivantsits, and J. M. Kuhnigk (Cham: Springer International Publishing), 87–92. doi: 10.1007/978-3-030-72862-5_9

Ou, C., Chong, W., Duan, C.-Z., Zhang, X., Morgan, M., and Qian, Y. (2021). A preliminary investigation of radiomics differences between ruptured and unruptured intracranial aneurysms. *Eur. Radiol.* 31, 2716–2725. doi: 10.1007/s00330-020-07325-3

Peeken, J. C., Asadpour, R., Specht, K., Chen, E. Y., Klymenko, O., Akinkuoroye, V., et al. (2021). MRI-based delta-radiomics predicts pathologic complete response in high-grade soft-tissue sarcoma patients treated with neoadjuvant therapy. *Radiot. Oncol.* 164, 73–82. doi: 10.1016/j.radonc.2021.08.023

Roked, F., and Reddy, U. (2020). Management of subarachnoid haemorrhage. *Anaesth. Intensive Care Med.* 21, 305–311. doi: 10.1016/j.mpaic.2020.03.013

Senders, J. T., Staples, P. C., Karhade, A. V., Zaki, M. M., Gormley, W. B., Broekman, M. L. D., et al. (2018). Machine Learning and Neurosurgical Outcome Prediction: a Systematic Review. *World Neurosurg.* 109, 476.e–486.e. doi: 10.1016/j.wneu.2017.09.149

Shi, Z., Hu, B., Schoepf, U. J., Savage, R. H., Dargis, D. M., Pan, C. W., et al. (2020). Artificial Intelligence in the Management of Intracranial Aneurysms: current Status and Future Perspectives. *AJNR Am. J. Neuroradiol.* 41, 373–379. doi: 10.3174/ajnr.A6468

Silva, M. A., Patel, J., Kavouridis, V., Gallerani, T., Beers, A., Chang, K., et al. (2019). Machine Learning Models can Detect Aneurysm Rupture and Identify Clinical Features Associated with Rupture. *World Neurosurg.* 131, e46–e51. doi: 10.1016/j.wneu.2019.06.231

Tan, M., and Le, Q. (2019). "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," in *Proceedings of the 36th International Conference on Machine Learning (PMLR), 6105–6114*. Available Online at: https://proceedings.mlr.press/v97/tan19a.html. [accessed on Nov 17, 2021]

Tanioka, S., Ishida, F., Yamamoto, A., Shimizu, S., Sakaida, H., Toyoda, M., et al. (2020). Machine Learning Classification of Cerebral Aneurysm Rupture Status with Morphologic Variables and Hemodynamic Parameters. *Radiol. Artif. Intell.* 2:e190077. doi: 10.1148/ryai.2019190077

van Griethuysen, J. J. M., Fedorov, A., Parmar, C., Hosny, A., Aucoin, N., Narayan, V., et al. (2017). Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res.* 77, e104–e107. doi: 10.1158/0008-5472.CAN-17-0339

Varoquaux, G., Raamana, P. R., Engemann, D. A., Hoyos-Idrobo, A., Schwartz, Y., and Thirion, B. (2017). Assessing and tuning brain decoders: cross-validation, caveats, and guidelines. *NeuroImage* 145, 166–179. doi: 10.1016/j.neuroimage.2016.10.038

Xu, Z., Rui, Y.-N., Hagan, J. P., and Kim, D. H. (2019). Intracranial Aneurysms: pathology, Genetics, and Molecular Mechanisms. *Neuromol Med.* 21, 325–343. doi: 10.1007/s12017-019-08537-7

Yang, J., Xie, M., Hu, C., Alwalid, O., Xu, Y., Liu, J., et al. (2021). Deep Learning for Detecting Cerebral Aneurysms with CT Angiography. *Radiology* 298, 155–163. doi: 10.1148/radiol.2020192154

Yun, J., Park, J. E., Lee, H., Ham, S., Kim, N., and Kim, H. S. (2019). Radiomic features and multilayer perceptron network classifier: a robust MRI classification strategy for distinguishing glioblastoma from primary central nervous system lymphoma. *Sci. Rep.* 9:5746. doi: 10.1038/s41598-019-42276-w

Zeng, Y., Liu, X., Xiao, N., Li, Y., Jiang, Y., Feng, J., et al. (2020). Automatic Diagnosis Based on Spatial Information Fusion Feature for Intracranial Aneurysm. *IEEE Trans. Med. Imag.* 39, 1448–1458. doi: 10.1109/TMI.2019.2951439

Zwanenburg, A., Leger, S., Vallières, M., and Löck, S. (2020). Image biomarker standardisation initiative. *Radiology* 295, 328–338. doi: 10.1148/radiol.2020191145

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 24 March 2022
doi: 10.3389/fnins.2022.843988

Exploring Hierarchical Auditory Representation *via* a Neural Encoding Model

*Liting Wang[1], Huan Liu[1], Xin Zhang[2], Shijie Zhao[1], Lei Guo[1], Junwei Han[1] and Xintao Hu[1]**

[1] School of Automation, Northwestern Polytechnical University, Xi'an, China, [2] Institute of Medical Research, Northwestern Polytechnical University, Xi'an, China

By integrating hierarchical feature modeling of auditory information using deep neural networks (DNNs), recent functional magnetic resonance imaging (fMRI) encoding studies have revealed the hierarchical neural auditory representation in the superior temporal gyrus (STG). Most of these studies adopted supervised DNNs (e.g., for audio classification) to derive the hierarchical feature representation of external auditory stimuli. One possible limitation is that the extracted features could be biased toward discriminative features while ignoring general attributes shared by auditory information in multiple categories. Consequently, the hierarchy of neural acoustic processing revealed by the encoding model might be biased toward classification. In this study, we explored the hierarchical neural auditory representation *via* an fMRI encoding framework in which an unsupervised deep convolutional auto-encoder (DCAE) model was adopted to derive the hierarchical feature representations of the stimuli (naturalistic auditory excerpts in different categories) in fMRI acquisition. The experimental results showed that the neural representation of hierarchical auditory features is not limited to previously reported STG, but also involves the bilateral insula, ventral visual cortex, and thalamus. The current study may provide complementary evidence to understand the hierarchical auditory processing in the human brain.

Keywords: hierarchical auditory representation, deep convolutional auto-encoder, naturalistic experience, neural encoding, fMRI

OPEN ACCESS

Edited by:
Feng Liu,
Tianjin Medical University General Hospital, China

Reviewed by:
Dahua Yu,
Inner Mongolia University of Science and Technology, China
Yizhang Jiang,
Jiangnan University, China

****Correspondence:***
Xintao Hu
xhu@nwpu.edu.cn

Specialty section:
This article was submitted to Brain Imaging Methods, a section of the journal Frontiers in Neuroscience

Received: *27 December 2021*
Accepted: *16 February 2022*
Published: *24 March 2022*

Citation:
Wang L, Liu H, Zhang X, Zhao S, Guo L, Han J and Hu X (2022) Exploring Hierarchical Auditory Representation via a Neural Encoding Model. Front. Neurosci. 16:843988. doi: 10.3389/fnins.2022.843988

INTRODUCTION

There are growing evidences supporting the hierarchy of auditory representations during auditory processing in the human brain (Chevillet et al., 2011; Sharpee et al., 2011; Durschmid et al., 2016; De Heer et al., 2017; Kell et al., 2018). For example, the neural processing of narrative speech involves hierarchical representations starting from the primary auditory areas and laterally to the temporal lobe (De Heer et al., 2017). In addition, the localization and identification of relevant auditory objects are accomplished *via* parallel "where" and "what" pathways (Ahveninen et al., 2006; Lomber and Malhotra, 2008; Bizley and Cohen, 2013). The hierarchy of neural auditory representation is important to understand what sensory information is processed as one traverses the sensory pathways from the primary sensory areas to higher-order areas.

In light of their hierarchical feature representation ability, recently advanced deep neural networks (DNNs) have gained increasing interest in exploring the hierarchy of neural auditory representation. These studies offer promising prospects to understand the fundamental mechanisms of brain functions responding to external stimuli. Specifically, brain encoding models (Naselaris et al., 2011; Han et al., 2014; Mesgarani et al., 2014; Du et al., 2019) have been used to establish the relationship between acoustic features represented in different layers of DNNs and brain activities. Brain regions of interest that selectively respond to extracted features in different layers were then inferred according to encoding performance. Using such a neural encoding framework, researchers have revealed a representational gradient in the superior temporal gyrus (STG) during auditory information processing (Evans and Davis, 2015; Kell et al., 2018; O'Sullivan et al., 2019; Kiremitçi et al., 2021). For example, Kell et al. (2018) found that latent features in intermediate network layers best predicted neural responses in the primary auditory cortex, while features in deeper layers can better explain brain activities in anterior, lateral and posterior directions of the non-primary areas.

In the majority of existing studies, the hierarchical features of external acoustic stimuli were derived using supervised DNNs that are designed for specific tasks, such as audio genre classification (Güçlü et al., 2016) or speech recognition (Kell et al., 2018). One possible limitation is that the supervised hierarchical representations could be biased toward discriminative features while ignoring the common ones shared by auditory excerpts in different categories. Consequently, the hierarchical organization of neural auditory processing revealed by the encoding model may be confined to classification or recognition domain. However, the neural processing of auditory information during naturalistic experience is not restricted to classification or recognition (Hasson and Honey, 2012; Fasano et al., 2020). Unlike supervised DNNs that use predefined labels as targets for model optimization, unsupervised DNNs such as deep convolutional auto-encoder (DCAE) adopts data reconstruction errors as objective functions and hence learn intrinsic and hierarchical features of input data directly (Masci et al., 2011). Thus, unsupervised DNNs may serve as possible tools to comprehensively map the hierarchy of neural auditory processing.

In this manuscript, we proposed an fMRI encoding framework to explore the hierarchy of neural auditory processing in the human brain. In brief, an unsupervised DCAE model (Masci et al., 2011), instead of supervised DNNs used in existing studies (Güçlü et al., 2016; Kell et al., 2018), was trained to derive unbiased hierarchical feature representations of naturalistic auditory excerpts in three semantic categories (pop music, classic music, and speech). An encoding model based on LASSO algorithm (Tibshirani, 2011) was learned to predict fMRI brain activities using acoustic features represented in each layer of the DCAE model. Brain regions that selectively response to the hierarchical features were inferred according to the encoding performance subsequently.

MATERIALS AND METHODS

Overview

As illustrated in Figure, we acquired fMRI data when the participants were freely listening to naturalistic auditory excerpts (**Figure 1A**). Then the hierarchical feature representations of each audio excerpt were derived *via* an unsupervised DCAE model (Masci et al., 2011; **Figure 1B**). Afterward, the hierarchical acoustic features were correlated to fMRI brain activities using an encoding model based on LASSO algorithm (Tibshirani, 2011; **Figure 1C** and Section "Encoding Model and Group-Wise Analysis"). In brief, the hierarchical feature representation was used to predict fMRI brain activities with a sparsity regularization, and the prediction accuracies was used to measure how well the acoustic features and brain activities were correlated. After that, a group-wise analysis was performed to identify brain regions whose activities were predicted with accuracies significantly above chance to infer hierarchical auditory representation in the brain.

Functional Magnetic Resonance Imaging Acquisition and Preprocessing

Auditory excerpts in three semantic categories (classical music, pop music, and speech) were used as naturalistic stimuli in fMRI data acquisition. Each category was composed of seven excerpts and each excerpt was around 90 s. All excerpts were taken from legal copies of compressed MP3 audio files. These audio excerpts were aggregated in a random order to avoid the influence of the internal structure of audio data on human brain's perception. FMRI data were acquired using a GE 3T Signa MRI system (GE Healthcare, Milwaukee, WI, United States) with an 8-channel head coil at the Bio-Imaging Research Center of the University of Georgia (UGA) under UGA Institutional Review Board (IRB) approval. Six healthy university students voluntarily participated in the study. The audio stimuli were delivered to the participants

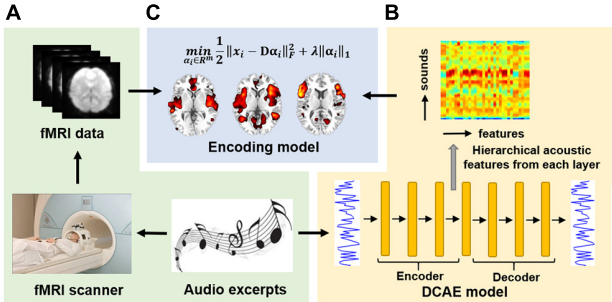


FIGURE 1 | The schematic illustration of the study. **(A)** fMRI acquisition using naturalistic auditory excerpts as stimuli. **(B)** Hierarchical feature representation of the naturalistic auditory stimuli *via* an unsupervised DCAE model. **(C)** Hierarchical acoustic features were correlated to fMRI brain activities using an encoding model based on LASSO to infer hierarchical auditory representation in the brain.

using an MRI-compatible audio headphone (Nordic NeuroLab, Bergen, Norway).

The detailed fMRI acquisition parameters were as follows: TR = 1.5 s, TE = 25 ms, 64 × 64 matrix, 30 axis slices, 4 mm slice thickness, 220 mm Field of View (FOV). FMRI data were pre-processed using FSL FEAT (FMRI Expert Analysis Tool) (Smith et al., 2004). The preprocessing included brain skull removal, slice timing and motion correction, spatial smoothing with 5 mm full-width at half-maximum (FWHM) Gaussian kernel, high pass temporal filtering, and linear registration to the standard Montreal Neurological Institute (MNI) brain template. After preprocessing, the time course of each voxel was normalized to have zero mean and unit standard deviation.

Hierarchical Feature Representation Based on Deep Convolutional Auto-Encoder

Deep Convolutional Auto-Encoder Model

The DCAE model used in this study is composed of an encoding block and a decoding block, as shown in **Figure 2**. The encoder transforms the input data into a detailed feature representation (feature maps), and the decoder performs data reconstruction (Masci et al., 2011). The objective of the DCAE model is to minimize the reconstruction errors between the input auditory signals and reconstructed ones.

Each block in the encoder consists of a convolutional layer and a max-pooling layer. A convolutional layer acts as feature extractor and the max-pooling layer reduces computational cost in the upper convolutional layer and gains translation/scale-invariance (Peterson et al., 2018; Song et al., 2018). Each block in the decoder consists of a deconvolution layer and an un-pooling layer. It is notable that the max-pooling operation is not invertible. To address this problem, we adopted a switch-based un-pooling approach (Zeiler and Fergus, 2014). The "switches" record the exact location of the max value in each pooling region during max-pooling, and then these "switches" are placed to its original position with corresponding max values (Huang et al., 2017). A linear activation function was applied in the first convolutional layer in the encoder and the last deconvolution layer in the decoder. The Rectified Linear Unit (ReLU) (Dahl et al., 2013) was used as activation function elsewhere. The objective function of the DCAE model consists of two terms. The first term represents the reconstruction error. The second term is an L2 regularization applied on weights to prevent overfitting and make the learned features more interpretable (Bilgic et al., 2014).

The number of layers in the DCAE model here was empirically set to balance the effectiveness of hierarchical feature learning and the interpretability of the subsequent inference of the hierarchical neural auditory processing. Intuitively, a larger number of layers would result in a finer featural representation of the input auditory excerpts. However, this would bring difficulties in interpreting the cortical hierarchy of acoustic feature processing in the human brain. In contrast, a smaller number of layers may not sufficient to learn the hierarchical feature representations of the input acoustic excerpts and consequently interrupt the encoding inference.

Deep Convolutional Auto-Encoder Parameter Settings and Model Training

During model training, the length of an input training sample was the same as the TR (1.5 s) in fMRI acquisition. The naturalistic auditory stimuli used in fMRI acquisition contribute 1,260 samples, which are not sufficient to train the DCAE model. To address this problem, we constructed additional 36,000 samples from the MagnaTagATune Dataset (Law et al., 2009) and the LibriSpeech Corpus (Panayotov et al., 2015) to pre-train the model (Data 1). The pre-trained model was then fine-tuned using the samples from the fMRI stimuli (Data 2). We implemented the DCAE model using Keras (Chollet, 2015) with CUDA and cuDNN. Based on our prior experiences (Huang et al., 2017), hyper-parameters in the DCAE including the number

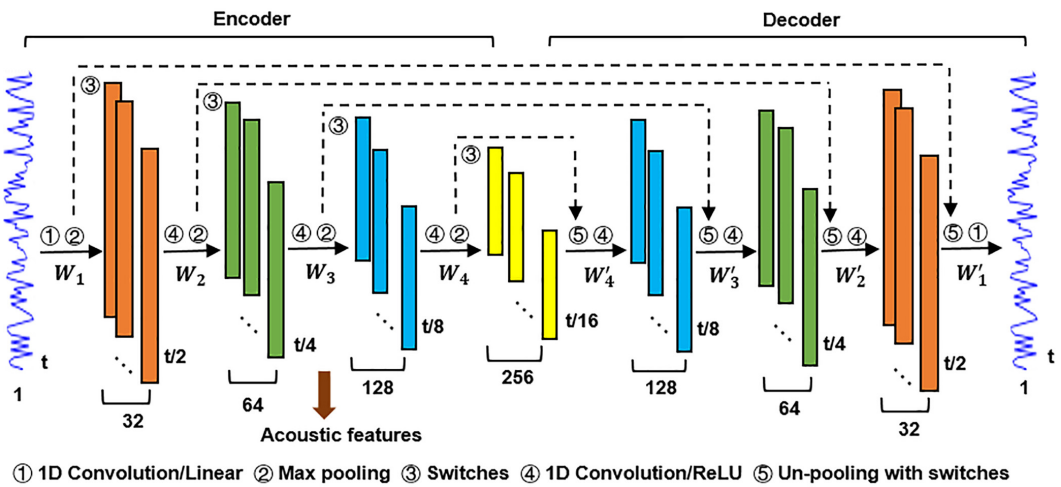


FIGURE 2 | The DCAE model used for the hierarchical feature modeling.

and the length of the filters were detailed in **Table 1**. The regularization parameter κ is experimentally set as 0.001. We used the Adam optimizer with default parameters $\beta_1 = 0.5$, $\beta_2 = 0.999$, epsilon = $1e^{-8}$ and a mini-batch size of 32 to train the model. We manually tuned the learning rate $\alpha = 0.0002$ and weight decay = 0.001 to iteratively minimize the mean square error (MSE) loss function. The DCAE model converged after about 5,000 epochs.

Hierarchical Acoustic Feature Representation

Similar to a previous study (Kell et al., 2018), the acoustic features encoded in each of the four max-pooling layers in the encoder were regarded as a single level of the hierarchical feature representation of an input auditory sample. For each input sample (1.5 s*16k/s = 24k*1), its hierarchical feature maps on the four max-pooling layers are in the dimension of t_i*c_i, where t_i is the length of sample in the output of i-th max-pooling layer (24k, 12k, 6k, and 3k for $i = 1, \ldots, 4$, respectively). c_i is the number of filters (channels) in the i-th convolutional layer. Following the feature dimensionality reduction strategy used in Güçlü et al. (2016), the high dimensional feature map on each max-pooling layer was temporally averaged, resulted in a c_i-dimensional feature vector. For a given auditory excerpt consisting of 60 samples that was used as stimulus in fMRI acquisition, its hierarchical feature representation is in the dimension of 60*c_i. Subsequently, each column of these hierarchical acoustic features was convolved with the canonical double-gamma hemodynamic response function (HRF).

Encoding Model and Group-Wise Analysis

Linear encoding models are preferred in fMRI encoding studies due to their good interpretability (Naselaris et al., 2011). Compared to other linear regression models such as ridge regression and support vector regression (SVR) with a linear kernel, LASSO enforces a sparse encoding model that is able to identify a more compact set of variables of interest. Thus, an encoding model based on LASSO algorithm (Tibshirani, 2011) was trained to predict fMRI responses using the hierarchical feature representation described above. In the encoding model, we treated each 60-s auditory excerpt in fMRI acquisition and the corresponding individual excerpt-specific fMRI data as a single sample, resulting in a collection of 126 (3 auditory categories × 7 excerpts in each category × 6 participants) samples. The encoding model can be formulated as a matrix factorization with a sparsity penalty:

$$min_{\alpha i} \in R^m \frac{1}{2}||x_i - D\alpha_i||_2^2 + \lambda||\alpha_i||_1 \quad (1)$$

TABLE 1 | The number and length of filters in the DCAE model.

Filter number/filter length	Layer 1	Layer 2	Layer 3	Layer 4
Encoder	32/64	64/32	128/16	256/8
Decoder	256/8	128/16	64/32	32/64

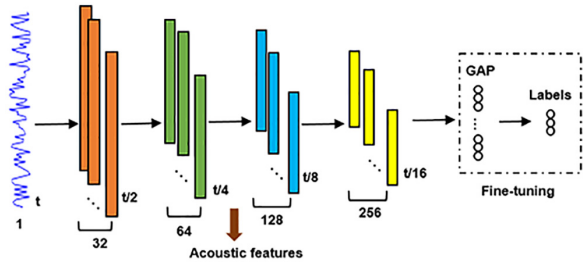


FIGURE 3 | The architecture of the supervised DNNs for audio classification. GAP, global average pooling.

where x_i is the fMRI signal of each voxel in an individual participant, **D** is the corresponding hierarchical feature representation in each layer, a_i is the encoding coefficients, and λ is a sparsity controlling parameter. The encoding model was trained for each voxel independently. The encoding performance for each voxel was calculated as the Pearson correlation coefficient (PCC) between the predicted fMRI activities and the recorded ones. Repeating encoding model training and performance evaluating for each voxel resulted in an encoding performance map for each sample. The parameter λ balances the regression residual and sparsity level. The encoding model with a smaller λ better predicts x_i using a larger subset of **D** at the risk of over-fitting, while a larger λ decreases the prediction accuracy using a more compact subset of features. In our study, λ was varied from 0.05 to 0.15 with interval of 0.05 and was optimized *via* a leave-one-out cross-validation strategy to maximize the average encoding performance in the testing set.

A group-wise analysis was then performed to infer the corresponding brain regions that selectively encoded each level of the hierarchical feature representations in the DCAE model. In brief, for a given level of the hierarchical feature representations, the encoding performance map for each sample was independently normalized and aggregated to perform one-sample t-test to infer the corresponding brain regions that have encoding accuracy significantly above chance ($p < 0.01$, $Z \geq 2.3$).

A Comparison Study

A comparison study was performed to compare the neural encoding of unsupervised hierarchical feature representations with that of a supervised classification model described as follows. A global average pooling (GAP) layer (Yu et al., 2017) followed by a fully connected soft-max classification layer were connected to the fourth max-pooling layer of the unsupervised DCAE model (**Figure 3**). Adopting cross-entropy loss function, Adam optimizer, early stopping strategy and batch size of 32, it was pre-trained using Data 1 and followed by fine-tuning using Data 2. Supervised hierarchical feature representations of input auditory excerpts were derived from the converged classification model. The neural encoding of supervised hierarchical features was probed using the same encoding framework described in Section "Encoding Model and Group-Wise Analysis" and was compared with that of the unsupervised DCAE model.

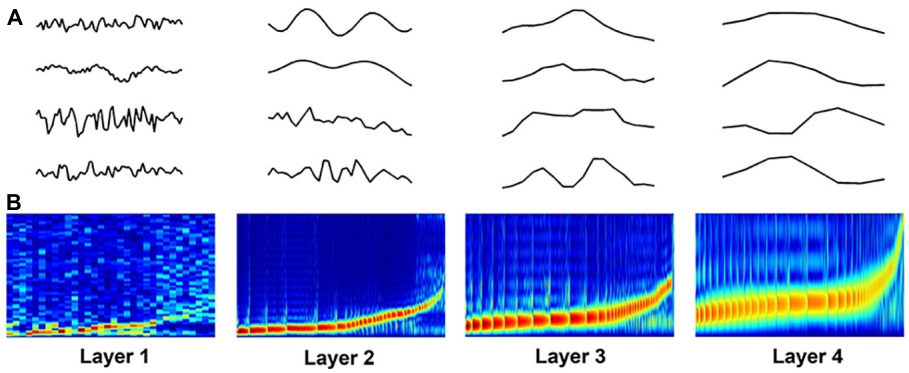


FIGURE 4 | Visualization of learned filters in the DCAE model. **(A)** Examples of the learned filters in each layer. **(B)** The power-spectrum patterns of learned filters. The *x*-axis represents the index of filters, the *y*-axis represents the frequency ranging from 0 to 8000 Hz.

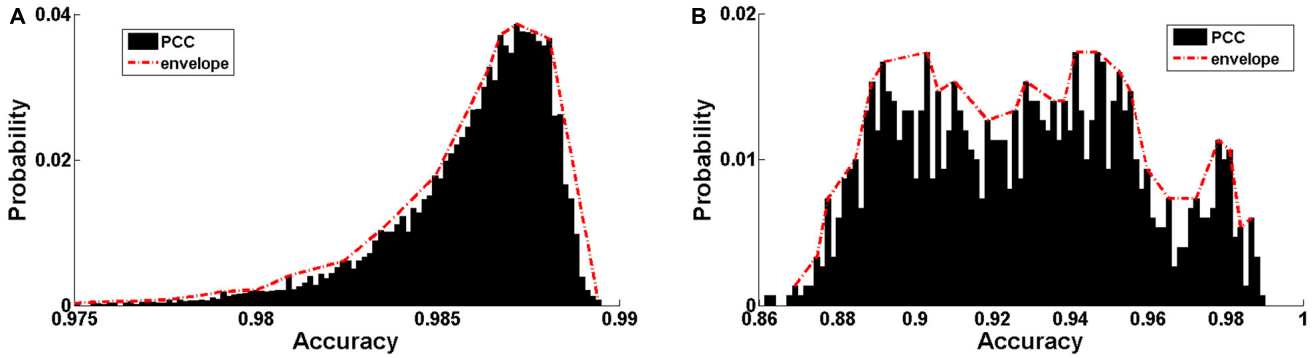


FIGURE 5 | Encoding performance of the trained DCAE model. **(A)** The distribution of Pearson correlation coefficients (PCC) between the input audio signals and reconstructed ones in the MagnaTagATune dataset and LibriSpeech Corpus. **(B)** The distribution of PCC in the auditory samples from the fMRI stimuli.

RESULTS

Evaluation of Hierarchical Feature Learning

Figure 4A shows some examples of the learned filters in the DCAE model. The power-spectrum patterns of the learned filters are depicted in **Figure 4B**, where the filters in each layer are sorted according to the frequency (low to high) at which its magnitude reaches the maximum (Lee et al., 2018). In the first layer, the frequency of the filters increases approximately linearly in low frequency filter banks whereas filters that are selective for higher frequency are more spread out. As the layer goes deeper, the trend of frequency becomes non-linearly steeper in high frequency filter banks. These spectrum patterns are consistent with those in frame-level end-to-end learning for music classification (Dieleman and Schrauwen, 2014; Lee et al., 2018), suggesting the effectiveness of hierarchical feature learning in the DCAE model.

The distribution of Pearson correlation coefficients (PCCs) between the input audio signals and reconstructed ones is shown in **Figure 5**. The PCC is relatively high in both Data 1 (0.9859 ± 0.0024, **Figure 5A**) and Data 2 (0.9274 ± 0.0297, **Figure 5B**). The discriminative ability of the hierarchical features learned by the DCAE model was then examined using a classification task based on support vector machine (SVM) with an RBF kernel. The classification performance in 5-fold cross-validations is summarized in **Table 2** for each layer. The classification accuracy slowly increases as the layer goes deeper. Both the high data reconstruction performance and high classification accuracy indicate that the trained DCAE model could well capture the intrinsic features of the input samples. Similar classification results are observed in the supervised model (**Table 3**).

Encoding Performance

The optimal sparsity controlling parameter $\lambda = 0.1$ maximized the overall encoding performance depicted in **Figure 6** for the unsupervised DCAE (**Figure 6A**) and supervised classification model (**Figure 6B**). Each subgraph shows the PCC between the original fMRI signals and the ones predicted by the hierarchical feature representations in each layer. In general, the distribution of brain regions in each layer is similar in the unsupervised DCAE and supervised classification model. The primary auditory cortex is selective to acoustic features learned in the first layer, the non-primary auditory cortex in the superior temporal gyrus (STG) is more sensitive to intermediate-layer acoustic

TABLE 2 | The classification accuracies in different layers of the DCAE model (mean ± std).

	Layer 1	Layer 2	Layer 3	Layer 4
Data 1	0.7787 ± 0.0326	0.9079 ± 0.0060	0.9168 ± 0.0096	0.9198 ± 0.0077
Data 2	0.7528 ± 0.0221	0.9044 ± 0.0104	0.9084 ± 0.0169	0.9181 ± 0.0220

TABLE 3 | The classification accuracies in different layers of the supervised model.

	Layer 1	Layer 2	Layer 3	Layer 4
Data 1	0.9093 ± 0.0160	0.9489 ± 0.0110	0.9558 ± 0.0086	0.9679 ± 0.0025
Data 2	0.8545 ± 0.1661	0.9309 ± 0.0139	0.9531 ± 0.0056	0.9552 ± 0.0033

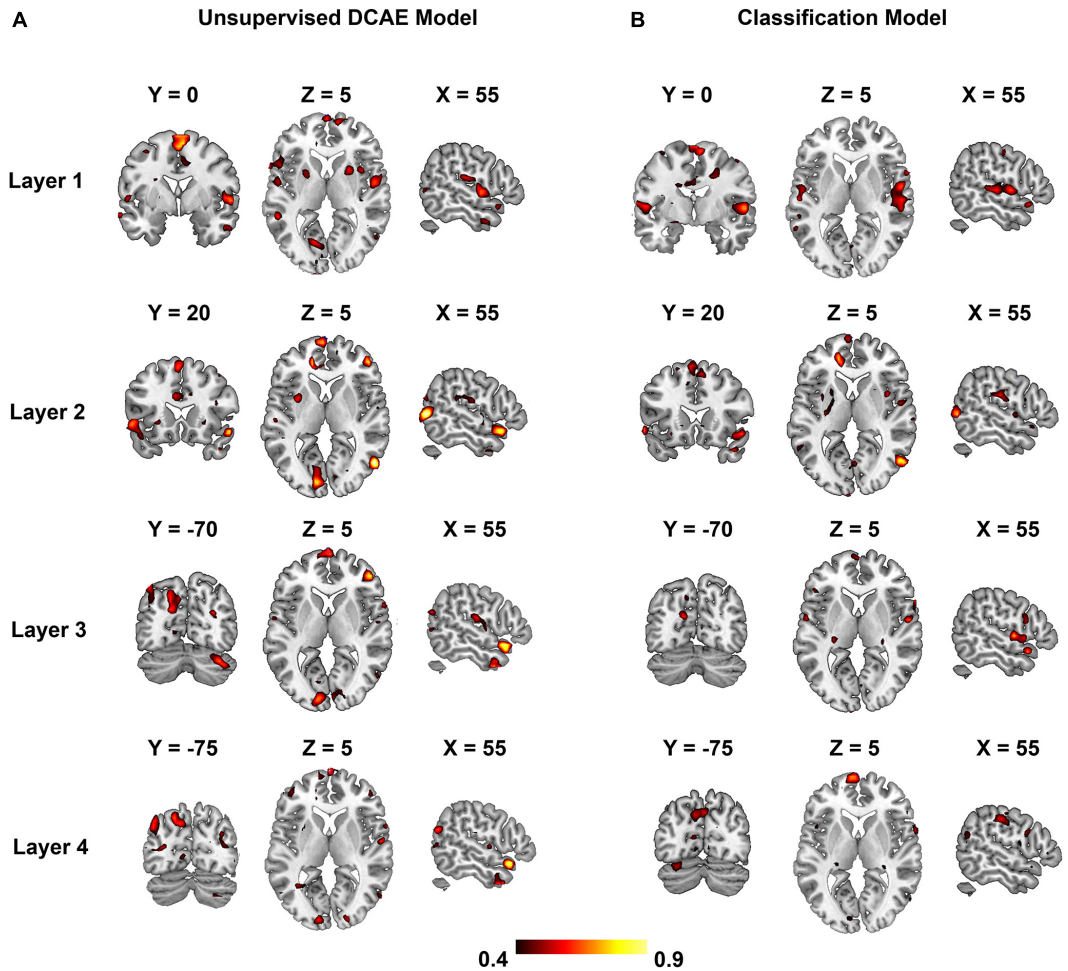


FIGURE 6 | The encoding performance for each layer in the unsupervised DCAE model **(A)** and supervised classification model **(B)**.

features, while the prefrontal cortex, visual cortex, and precuneus are involved in the processing of higher-level features learned in the last layer.

We further adopted a paired-sample *t*-test to compare the encoding performance between the unsupervised DCAE and supervised classification models. It is observed that the encoding performance in some brain regions in the unsupervised DCAE model is significantly higher ($p \leq 0.01$, $Z \geq 2.3$) than those in the supervised classification model, including the primary auditory cortex (A1) in the first layer, part of middle temporal gyrus (MTG) and visual cortex in the second layer, anterior STG, posterior STG, part of prefrontal cortex (PFC), cuneus and precuneus in the third and last layer (**Figure 7**). No obvious brain regions showed significantly higher encoding performance in supervised classification model compared to the unsupervised one. These results suggested that the hierarchical features learned in the unsupervised DCAE model can achieve better encoding performance.

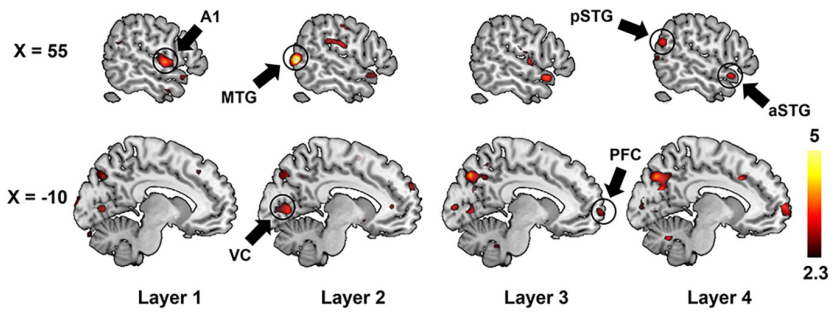


FIGURE 7 | The comparison of encoding performance between the unsupervised DCAE model and supervised classification model in each layer. A1, primary auditory cortex; aSTG, anterior superior temporal gyrus; pSTG, posterior superior temporal gyrus; MTG, middle temporal gyrus; VC, visual cortex; PFC, prefrontal cortex.

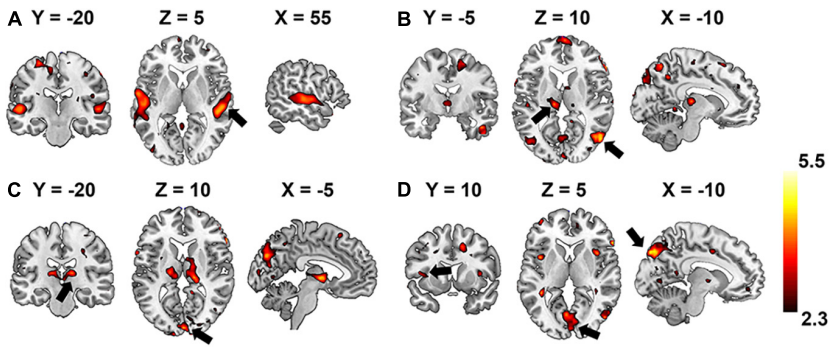


FIGURE 8 | Brain regions that are selectively activated by the hierarchical acoustic features represented in each encoder layer of the unsupervised DCAE model. Panels **(A–D)** represent the first four layers in the unsupervised DCAE model.

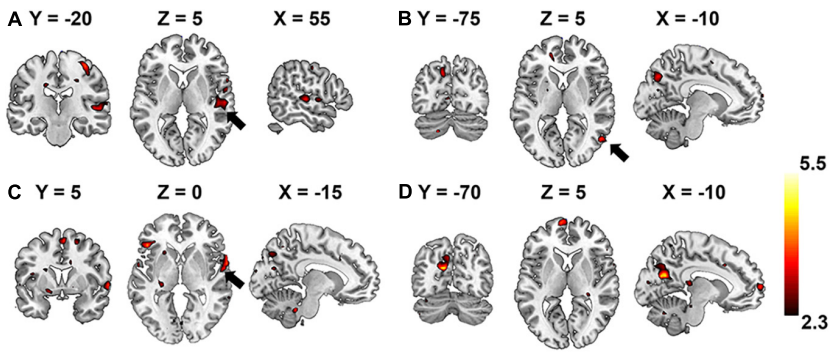


FIGURE 9 | Brain regions that are selectively activated by the hierarchical acoustic features represented in each layer of the supervised classification model. Panels **(A–D)** represent the first four layers in the supervised classification model.

Hierarchical Neural Auditory Representation

Group-wise analysis was used to evaluate whether the encoding performance was significantly above chance ($Z \geq 2.3$) for each voxel independently. Brain regions of interest that were selective to each level of the hierarchical acoustic feature representation were inferred accordingly to probe the hierarchy of neural auditory processing. **Figure 8** shows the Z-maps of encoding performance for each encoder layer in the unsupervised DCAE model. In the first layer (**Figure 8A**), brain activities in the primary and association auditory cortices were with significantly ($Z \geq 2.3$) high encoding accuracy, indicating that the features learned in the first layer may represent basic acoustic features. Part of the middle temporal gyrus (MTG) was activated in the second and third layer (**Figures 8B,C**). In the fourth layer, bilateral insula and ventral visual cortex were with significantly high encoding accuracy (**Figure 8D**). In addition, we observed that the thalamus was activated by the features represented in

the second and third layers. The statistical details of these brain regions are listed in **Supplementary Table 1**.

In comparison, the hierarchy of neural auditory processing revealed by the encoding model using supervised feature learning model is partly in line with the one in the unsupervised model, as shown in **Figure 9**. For example, the primary auditory cortex and visual cortex were selectively activated by the features represented in the first and fourth layer of the supervised model, respectively. However, those selective brain regions were much sparser and scattered distributed compared to the ones in the unsupervised model. In addition, the bilateral insula and thalamus were not activated in the supervised classification model.

DISCUSSION

In this study, we investigated the hierarchy of neural acoustic processing in the human brain *via* an fMRI encoding model. Compared to existing studies that used supervised feature learning models that are designed for classification or recognition to achieve a hierarchical feature representation of input acoustic information (Kell et al., 2018; O'Sullivan et al., 2019), the novelty of the current study is adopting an unsupervised DCAE feature learning model to derive intrinsic and unbiased hierarchical feature representation of naturalistic auditory stimuli in fMRI acquisition. Our experimental results showed that the neural representation of hierarchical auditory features is not limited to previously reported STG (Kell et al., 2018; O'Sullivan et al., 2019), but also involves the bilateral insula, ventral visual cortex and thalamus.

In the current study, our experimental results showed that the visual cortex and insula are related to the encoding of high-level acoustic features represented in the deepest layers of the DCAE model. It may indicate that these high-level features carry higher-order attributes such as emotion (Gu et al., 2013) and visual imagery (Vetter et al., 2014) elicited by auditory excerpts. For example, an fMRI study that uses auditory stimulation to examine the activity in the early visual cortex suggested that the auditory input enables the visual system to predict incoming information and could confer a survival advantage (Vetter et al., 2014). It also has been reported that the higher-level abstract or categorical information of acoustic stimulation is fed down to early visual cortex (Cate et al., 2009; Vetter et al., 2014). In addition, we observed that the thalamus may encode middle-level features. It has been reported that the thalamus plays an important role in auditory processing (Schonwiesner et al., 2006), especially for sound source localization (Proctor and Konishi, 1997), and tones modulated by attention (Frith and Friston, 1996). Our findings, in conjunction with previous results on the visual and auditory cortical representations (King and Nelken, 2009; Khalighrazavi and Kriegeskorte, 2014; Cichy et al., 2016), suggest that the existence of multiple representational gradients that processes increasingly complex conceptual information as we have experienced the sensory hierarchy of the human brain.

In the comparison study, the supervised model achieved better classification performance compared to the unsupervised DCAE model (**Tables 2**, **3**). However, the unsupervised DCAE model outperformed the supervised model in terms of encoding performance (**Figures 8**, **9**). More importantly, the cortical hierarchy pattern inferred by the supervised model was much sparser and scattered distributed compared to the ones in the unsupervised model (**Figures 6**, **7**). These observations indicate that the intrinsic and unbiased hierarchical features learned in the DCAE model may provide additional evidence to understand the cortical hierarchy in neural auditory processing compared to the features learned in the supervised model that were biased toward discriminative ones while ignoring general attributes shared by auditory information in multiple categories.

In summary, the findings in this study may provide complementary evidences to understand the hierarchical auditory processing in the human brain. The current study can be improved in several ways in the future. It is expected to validate the findings using larger-scale fMRI datasets that recruit more participants. In the current study, we adopted an unsupervised DCAE model to derive the hierarchical feature representations of the acoustic stimuli in fMRI acquisition. The architecture and some of the hyperparameters (e.g., the number of layers, the number and length of the filters) of the DCAE model were empirically set. Although this DCAE model was able to effectively learn the hierarchical feature representation of the input acoustic excerpts as indicated by the SVM-based classification tasks in our experiments, it could be optimized by automated machine learning technique such as neural architecture search neural architecture search (NAS) (Elsken et al., 2019). In addition, the recently advanced self-supervised learning models (Sermanet et al., 2018; Li et al., 2020) may serve as more efficient and ecological approaches to unsupervised acoustic feature learning, and thus could enrich our understanding of the cortical hierarchy of neural auditory processing in future studies.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Bio-Imaging Research Center of the University of Georgia. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

LW and XH designed the study, analyzed the data, and wrote the manuscript. HL analyzed the data. XZ, SZ, LG, and JH participated in the revision, reading, and approval of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was partially supported by the National Natural Science Foundation of China (NSFC) under grants 62076205, 61836006, and 61936007.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2022.843988/full#supplementary-material

REFERENCES

Ahveninen, J., Jaaskelainen, I. P., Raij, T., Bonmassar, G., Devore, S., Hamalainen, M., et al. (2006). Task-modulated "what" and "where" pathways in human auditory cortex. *Proc. Natl. Acad. Sci. U.S.A.* 103, 14608–14613. doi: 10.1073/pnas.0510480103

Bilgic, B., Chatnuntawech, I., Fan, A. P., Setsompop, K., Cauley, S. F., Wald, L. L., et al. (2014). Fast image reconstruction with L2-regularization. *J. Magn. Reson. Imaging* 40, 181–191. doi: 10.1002/jmri.24365

Bizley, J. K., and Cohen, Y. E. (2013). The what, where and how of auditory-object perception. *Nat. Rev. Neurosci.* 14, 693–707. doi: 10.1038/nrn3565

Cate, A. D., Herron, T. J., Yund, E. W., Stecker, G. C., Rinne, T., Kang, X., et al. (2009). Auditory attention activates peripheral visual cortex. *PLoS One* 4:e4645. doi: 10.1371/journal.pone.0004645

Chevillet, M. A., Riesenhuber, M., and Rauschecker, J. P. (2011). Functional correlates of the anterolateral processing hierarchy in human auditory cortex. *J. Neurosci.* 31, 9345–9352. doi: 10.1523/JNEUROSCI.1448-11.2011

Chollet, F. (2015). *Keras*. Available online at: https://github.com/fchollet/keras (accessed May, 2021).

Cichy, R. M., Khosla, A., Pantazis, D., Torralba, A., and Oliva, A. (2016). Comparison of deep neural networks to spatio-temporal cortical dynamics of human visual object recognition reveals hierarchical correspondence. *Sci. Rep.* 6:27755. doi: 10.1038/srep27755

Dahl, G. E., Sainath, T. N., and Hinton, G. E. (2013). "Improving deep neural networks for LVCSR using rectified linear units and dropout," in *Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, Vancouver, 8609–8613. doi: 10.1109/ICASSP.2013.6639346

De Heer, W. A., Huth, A. G., Griffiths, T. L., Gallant, J. L., and Theunissen, F. E. (2017). The hierarchical cortical organization of human speech processing. *J. Neurosci.* 37, 6539–6557. doi: 10.1523/JNEUROSCI.3267-16.2017

Dieleman, S., and Schrauwen, B. (2014). "End-to-end learning for music audio," in *Proceedings of the 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Florence, 6964–6968. doi: 10.1109/ICASSP.2014.6854950

Du, C., Du, C., Huang, L., and He, H. (2019). Reconstructing perceived images from human brain activities with bayesian deep multiview learning. *IEEE Trans. Neural Netw.* 30, 2310–2323. doi: 10.1109/TNNLS.2018.2882456

Durschmid, S., Edwards, E., Reichert, C., Dewar, C., Hinrichs, H., Heinze, H., et al. (2016). Hierarchy of prediction errors for auditory events in human temporal and frontal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 113, 6755–6760. doi: 10.1073/pnas.1525030113

Elsken, T., Metzen, J. H., and Hutter, F. (2019). Neural architecture search: a survey. *J. Mach. Learn. Res.* 20, 1997–2017.

Evans, S., and Davis, M. H. (2015). Hierarchical organization of auditory and motor representations in speech perception: evidence from searchlight similarity analysis. *Cereb. Cortex* 25, 4772–4788. doi: 10.1093/cercor/bhv136

Fasano, M. C., Glerean, E., Gold, B. P., Sheng, D., Sams, M., Vuust, P., et al. (2020). Inter-subject similarity of brain activity in expert musicians after multimodal learning: a behavioral and neuroimaging study on learning to play a piano sonata. *Neuroscience* 441, 102–116. doi: 10.1016/j.neuroscience.2020.06.015

Frith, C. D., and Friston, K. J. (1996). The role of the thalamus in "Top Down" modulation of attention to sound. *Neuroimage* 4, 210–215. doi: 10.1006/nimg.1996.0072

Gu, X., Hof, P. R., Friston, K. J., and Fan, J. (2013). Anterior insular cortex and emotional awareness. *J. Comp. Neurol.* 521, 3371–3388. doi: 10.1002/cne.23368

Güçlü, U., Thielen, J., Hanke, M., and Van Gerven, M. (2016). Brains on beats. *Adv. Neural Inf. Process. Syst.* 29, 2101–2109.

Han, J., Zhao, S., Hu, X., Guo, L., and Liu, T. (2014). Encoding brain network response to free viewing of videos. *Cogn. Neurodyn.* 8, 389–397. doi: 10.1007/s11571-014-9291-3

Hasson, U., and Honey, C. J. (2012). Future trends in Neuroimaging: neural processes as expressed within real-life contexts. *Neuroimage* 62, 1272–1278. doi: 10.1016/j.neuroimage.2012.02.004

Huang, H., Hu, X., Zhao, Y., Makkie, M., Dong, Q., Zhao, S., et al. (2017). Modeling task fMRI data via deep convolutional autoencoder. *IEEE Trans. Med. Imaging* 37, 1551–1561. doi: 10.1109/TMI.2017.2715285

Kell, A., Yamins, D., Shook, E. N., Normanhaignere, S., and Mcdermott, J. H. (2018). A task-optimized neural network replicates human auditory behavior, predicts brain responses, and reveals a cortical processing hierarchy. *Neuron* 98, 630–644. doi: 10.1016/j.neuron.2018.03.044

Khalighrazavi, S., and Kriegeskorte, N. (2014). Deep supervised, but not unsupervised, models may explain IT cortical representation. *PLoS Comput. Biol.* 10:e1003915. doi: 10.1371/journal.pcbi.1003915

King, A. J., and Nelken, I. (2009). Unraveling the principles of auditory cortical processing: Can we learn from the visual system? *Nat. Neurosci.* 12, 698–701. doi: 10.1038/nn.2308

Kiremitçi, I., Yilmaz, Ö., Çelik, E., Shahdloo, M., Huth, A. G., and Çukur, T. (2021). Attentional modulation of hierarchical speech representations in a multitalker environment. *Cereb. Cortex* 31, 4986–5005. doi: 10.1093/cercor/bhab136

Law, E., West, K., Mandel, M., Bay, M., and Downie, J. S. (2009). "Evaluation of algorithms using games: the case of music tagging," in *Proceedings of the 2009 10th International Society for Music Information Retrieval Conference (ISMIR)*, Kobe, 387–392.

Lee, J., Park, J., Kim, K. L., and Nam, J. (2018). SampleCNN: end-to-end deep convolutional neural networks using very small filters for music classification. *Appl. Sci.* 8:150. doi: 10.3390/app8010150

Li, D., Du, C., and He, H. (2020). Semi-supervised cross-modal image generation with generative adversarial networks. *Pattern Recogn.* 100:107085. doi: 10.1016/j.patcog.2019.107085

Lomber, S. G., and Malhotra, S. (2008). Double dissociation of 'what' and 'where' processing in auditory cortex. *Nat. Neurosci.* 11, 609–616. doi: 10.1038/nn.2108

Masci, J., Meier, U., Cireşan, D., and Schmidhuber, J. (2011). "Stacked convolutional auto-encoders for hierarchical feature extraction," in *Proceedings of the 2011 International Conference on Artificial Neural Networks*, Berlin, 52–59. doi: 10.1007/978-3-642-21735-7_7

Mesgarani, N., Cheung, C., Johnson, K., and Chang, E. F. (2014). Phonetic feature encoding in human superior temporal Gyrus. *Science* 343, 1006–1010. doi: 10.1126/science.1245994

Naselaris, T., Kay, K. N., Nishimoto, S., and Gallant, J. L. (2011). Encoding and decoding in fMRI. *Neuroimage* 56, 400–410. doi: 10.1016/j.neuroimage.2010.07.073

O'Sullivan, J., Herrero, J., Smith, E., Schevon, C., McKhann, G. M., Sheth, S. A., et al. (2019). Hierarchical encoding of attended auditory objects in multi-talker speech perception. *Neuron* 104, 1195–1209. doi: 10.1016/j.neuron.2019.09.007

Panayotov, V., Chen, G., Povey, D., and Khudanpur, S. (2015). "Librispeech: an ASR corpus based on public domain audio books," in *Proceedings of the 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, South Brisbane, 5206–5210. doi: 10.1109/ICASSP.2015.7178964

Peterson, J. C., Abbott, J. T., and Griffiths, T. L. (2018). Evaluating (and improving) the correspondence between deep neural networks and human representations. *Cogn. Sci.* 42, 2648–2669. doi: 10.1111/cogs.12670

Proctor, L., and Konishi, M. (1997). Representation of sound localization cues in the auditory thalamus of the barn owl. *Proc. Natl. Acad. Sci. U.S.A.* 94, 10421–10425. doi: 10.1073/pnas.94.19.10421

Schonwiesner, M., Krumbholz, K., Rubsamen, R., Fink, G. R., and Von Cramon, D. Y. (2006). Hemispheric asymmetry for auditory processing in the human

auditory brain stem, thalamus, and cortex. *Cereb. Cortex* 17, 492–499. doi: 10.1093/cercor/bhj165

Sermanet, P., Lynch, C., Chebotar, Y., Hsu, J., Jang, E., Schaal, S., et al. (2018). "Time-contrastive networks: self-supervised learning from video," in *Proceedings of the 2018 IEEE International Conference on Robotics and Automation (ICRA)*, Brisbane, 1134–1141. doi: 10.1109/ICRA.2018.8462891

Sharpee, T. O., Atencio, C. A., and Schreiner, C. E. (2011). Hierarchical representations in the auditory cortex. *Curr. Opin. Neurobiol.* 21, 761–767. doi: 10.1016/j.conb.2011.05.027

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23, S208–S219. doi: 10.1016/j.neuroimage.2004.07.051

Song, Z., Liu, Y., Song, R., Chen, Z., Yang, J., Zhang, C., et al. (2018). A sparsity-based stochastic pooling mechanism for deep convolutional neural networks. *Neural Netw.* 105, 340–345. doi: 10.1016/j.neunet.2018.05.015

Tibshirani, R. (2011). Regression shrinkage and selection via the lasso: a retrospective. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 73, 273–282. doi: 10.1111/j.1467-9868.2011.00771.x

Vetter, P., Smith, F. W., and Muckli, L. (2014). Decoding sound and imagery content in early visual cortex. *Curr. Biol.* 24, 1256–1262. doi: 10.1016/j.cub.2014.04.020

Yu, S., Jia, S., and Xu, C. (2017). Convolutional neural networks for hyperspectral image classification. *Neurocomputing* 219, 88–98. doi: 10.1016/j.neucom.2016.09.010

Zeiler, M. D., and Fergus, R. (2014). "Visualizing and understanding convolutional networks," in *Proceedings of the 2014 European Conference on Computer Vision*, Cham, 818–833. doi: 10.1007/978-3-319-10590-1_53

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Neuroscience

ORIGINAL RESEARCH
published: 06 May 2022
doi: 10.3389/fnins.2022.883693

Continuous Blood Pressure Estimation Based on Multi-Scale Feature Extraction by the Neural Network With Multi-Task Learning

Hengbing Jiang[1,2,3], Lili Zou[2,3], Dequn Huang[2,3] and Qianjin Feng[1]**

[1] School of Biomedical Engineering, Southern Medical University, Guangzhou, China, [2] Institute of Biological and Medical Engineering, Guangdong Academy of Sciences & National Engineering Research Center for Healthcare Devices, Guangzhou, China, [3] Guangdong Engineering Technology Research Center for Diagnosis and Rehabilitation of Dementia, Guangzhou, China

In this article, a novel method for continuous blood pressure (BP) estimation based on multi-scale feature extraction by the neural network with multi-task learning (MST-net) has been proposed and evaluated. First, we preprocess the target (Electrocardiograph; Photoplethysmography) and label signals (arterial blood pressure), especially using peak-to-peak time limits of signals to eliminate the interference of the false peak. Then, we design a MST-net to extract multi-scale features related to BP, fully excavate and learn the relationship between multi-scale features and BP, and then estimate three BP values simultaneously. Finally, the performance of the developed neural network is verified by using a public multi-parameter intelligent monitoring waveform database. The results show that the mean absolute error ± standard deviation for systolic blood pressure (SBP), diastolic blood pressure (DBP), and mean arterial pressure (MAP) with the proposed method against reference are 4.04 ± 5.81, 2.29 ± 3.55, and 2.46 ± 3.58 mmHg, respectively; the correlation coefficients of SBP, DBP, and MAP are 0.96, 0.92, and 0.94, respectively, which meet the Association for the Advancement of Medical Instrumentation standard and reach A level of the British Hypertension Society standard. This study provides insights into the improvement of accuracy and efficiency of a continuous BP estimation method with a simple structure and without calibration. The proposed algorithm for BP estimation could potentially enable continuous BP monitoring by mobile health devices.

Keywords: continuous blood pressure estimation, multi-scale features, neural networks, multi-task learning, photoplethysmography and electrocardiograph

OPEN ACCESS

Edited by:
Yao Wu,
Children's National Hospital, United States

Reviewed by:
Wonsang You,
Sun Moon University, South Korea
Ahmadreza Argha,
University of New South Wales, Australia

***Correspondence:**
Lili Zou
lilizou@mail.ustc.edu.cn
Qianjin Feng
fengqj99@fimmu.com

Specialty section:
This article was submitted to Brain Imaging Methods, a section of the journal Frontiers in Neuroscience

Received: *25 February 2022*
Accepted: *15 April 2022*
Published: *06 May 2022*

Citation:
Jiang H, Zou L, Huang D and Feng Q (2022) Continuous Blood Pressure Estimation Based on Multi-Scale Feature Extraction by the Neural Network With Multi-Task Learning. Front. Neurosci. 16:883693. doi: 10.3389/fnins.2022.883693

INTRODUCTION

The World Health Organization estimated that nearly 17.9 million people worldwide died of cardiovascular diseases in 2016 (World Health Organization, 2020), posing a serious threat to human health (El-Hajj and Kyriacou, 2020). Blood pressure (BP) monitoring plays an important part in the prevention, diagnosis, and prognosis of cardiovascular disease. The mercury sphygmomanometer is the most common method of measuring BP, but its measured value is instantaneous, random, and might be easily affected by human and environmental factors (O'brien et al., 2003). Therefore, efficient methods are needed to monitor BP continuously and accurately.

Intra-arterial monitoring is the gold standard method for continuous and accurate BP monitoring, but it can result in trauma to the human body and is not suitable for home monitoring. Compared with invasive intra-arterial continuous BP monitoring, non-invasive continuous BP monitoring is more secure and can be measured over a long time. At present, cuff-less arterial tonometry (Pressman and Newgard, 1963) and the volume-compensation method (Penaz, 1973) are mainly used to non-invasive monitor BP continuously. However, arterial tonometry is difficult to operate, which requires professional operation and may be greatly affected by human factors; the volume-compensation method has low precision and needs large equipment. In a word, monitoring theories of these methods limit their wide application in clinical and home use. Therefore, it is necessary to develop an easy-to-use and accurate method for continuous BP monitoring (Lázaro et al., 2019; Yang et al., 2021; Yen and Liao, 2022).

Exploring the relationship between the characteristic parameters of pulse waves and BP is a promising easy-to-use method for continuous and accurate BP monitoring. Recently, many studies have assessed the relationship between pulse wave transit time (PTT) and BP based on traditional methods to estimate BP, but its accuracy is low and these methods need to be calibrated (Chung et al., 2013; Mukkamala et al., 2015; Ding et al., 2016; Huynh et al., 2018). The combination of multiple features (e.g., PTT and pulse wave waveforms features) from photoplethysmography (PPG) and electrocardiograph (ECG) can improve the accuracy of BP estimation (Kachuee et al., 2017; Yoon et al., 2018; Thambiraj et al., 2020). However, these multiple features related to BP from pulse waves are mainly extracted through the feature engineering method, which has been identified as a heavy workload at work, and is difficult to find all of the features from PPG and ECG accurately. Besides, due to the many factors affecting BP, these traditional methods such as support vector machine, random forest, and adaptive boosting, are difficult to accurately fit the relationship between features and BP, which have limited accuracy.

With the development of artificial intelligence especially the deep learning (Ravì et al., 2017; Miotto et al., 2018; Argha et al., 2022), it is possible to extract multiple features related to BP from PPG and ECG and assess their relationship with BP accurately *via* the deep convolutional neural network (Radha et al., 2019; Yan et al., 2019; Li et al., 2020; Song et al., 2020; Paviglianiti et al., 2021). Some researchers have first extracted features manually and then used deep convolutional neural networks to estimate BP (Xu et al., 2017; Yan et al., 2019; Song et al., 2020; Paviglianiti et al., 2021). For example, Xu et al. (2017) have first manually extracted 15 features related to BP from PPG and ECG and then assessed their relationship with BP accurately by using artificial neural networks regression; Maqsood et al. (2021) have first manually extracted time-domain features, statistical features, and frequency domain features and then regressed BP values by using Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) regression. However, these methods also have the disadvantages of inaccuracy and time-consuming of manual feature extraction. Subsequently, the deep neural network methods based on end-to-end learning are used to automatically extract features related to BP and evaluate their relationship with BP, which achieves good performance (Eom et al., 2020). However, to improve the accuracy of BP estimation, the network structure (e.g., the number of network layers) used in these methods are complicated, which would increase the difficulty of model calculation and device deployment. In addition, most of them can only complete one task at a time using their model or complete multi-task using trained multiple models for BP estimation, which greatly reduces the efficiency of BP estimation (Gaurav et al., 2016; Rong and Li, 2021). Therefore, it is necessary to provide a simplified network with high accuracy and efficiency to monitor BP continuously.

In this study, in order to continuously and accurately estimate continuous BP without calibration from ECG and PPG signals, a new method for continuous BP estimation based on multi-scale feature extraction by the neural network with multi-task learning (MST-net) has been proposed and evaluated. Firstly, target signals and label signals arterial blood pressure (ABP) are preprocessed *via* segmenting, extracting labels, denoising, and normalizing. In particular, the interference of abnormal values and the false peak of wave signals are eliminated by peak amplitude and peak-to-peak timing limit. Subsequently, the multi-scale features related to BP are extracted from preprocessed target signals, and the relationship between multi-scale features and BP is trained and learned by the neural network with multi-task learning. Finally, the performance of the neural network is verified and compared with the Association for the Advancement of Medical Instrumentation (AAMI) standards, the British Hypertension Society (BHS) standards, and previous works. This model can not only estimate systolic blood pressure (SBP), diastolic blood pressure (DBP), and mean arterial pressure (MAP) simultaneously but also extract more scale features.

MATERIALS AND METHODS

The core concepts of continuous BP estimation based on multi-scale feature extraction by the neural network with multi-task learning proposed in this study are as follows: we adopt segmentation, denoising, and normalization to preprocess the target and label signals, especially using peak-to-peak timing limits of signals to eliminate the interference of the false peak of wave signal; we design a neural network with multi-task learning to extract multi-scale features related to BP from preprocessed target signals, fully excavate and learn the relationship between the multi-scale features and BP, and then estimate three BP values simultaneously through multi-task learning, thus improving the accuracy of BP estimation (**Figure 1**).

Problem Definition

To estimate BP continuously and accurately, target signals (PPG and ECG) and label signals (ABP) are synchronously divided into short segments with the same length, which are used as model inputs x_i and reference label BP values (SBP, DBP, and MAP) y_i^{SBP}, y_i^{DBP}, and y_i^{MAP}, respectively. Then all the x_i are used as the input

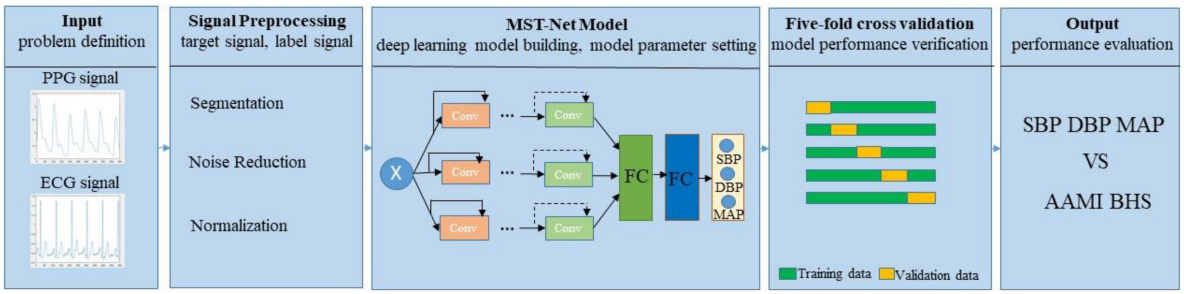


FIGURE 1 | Block diagram of the proposed continuous BP estimation method.

of the neural network, which is used to simultaneously estimate three BP values z_i^{SBP}, z_i^{DBP}, and z_i^{MAP}, and defined as follows:

$$z_i^{SBP} = F(x_i; \theta^{SBP}) \quad (1)$$

$$z_i^{DBP} = F(x_i; \theta^{DBP}) \quad (2)$$

$$z_i^{MAP} = F(x_i; \theta^{MAP}) \quad (3)$$

Where x_i represents input signals, $F(\cdot)$ represents the input-output function of the neural network with multi-task learning; SBP is the force exerted by blood on arterial walls during ventricular contraction, DBP is the force exerted by the artery walls during ventricular relaxation, and the MAP is the average pressure throughout the cardiac cycle; θ^{SBP}, θ^{DBP}, and θ^{MAP} represent the specific task parameters of the function; z_i^{SBP}, z_i^{DBP}, and z_i^{MAP} are the estimated SBP, DBP, and MAP output values of the network, and these three values are produced at the same time through an output layer (3 neurons) followed by the last fully connected (FC) layer.

The convergence of the neural network is evaluated by loss function MSE, and the MSE depends on the difference between the reference label BP values and estimated BP values and is defined as follows:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}[(y_i^{SBP} - z_i^{SBP})^2 + (y_i^{DBP} - z_i^{DBP})^2 + (y_i^{MAP} - z_i^{MAP})^2] \quad (4)$$

Where y_i ranges from 60 to 180 mmHg ($60 \leq y_i^{SBP}, y_i^{DBP}, y_i^{MAP} \leq 180$), and n is the number of signal segments. The smaller the MSE, the better performance of the model.

Preprocessing of Signals

Segmentation and label extraction. Our raw data comes from the University of California, Irvine (UCI) Machine Learning Repository dataset (Goldberger et al., 2000; Kachuee et al., 2015), which is derived from the public Multi-parameter Intelligent Monitoring in Intensive Care (MIMIC-II) database. This database contains multiple physiological signals collected simultaneously from intensive care unit patients. In this research, we extract simultaneous recordings of ECG, PPG, and ABP signals of 3,000 subjects from the database which was available at a 125 Hz sampling rate, and select signals with appropriate time (more than 8 min) as the data source. All selected signals were segmented into short segments of 8 s. Subsequently, the peak amplitude limit (80 mmHg ≤ SBP ≤ 180 mmHg, 60 mmHg ≤ DBP ≤ 130 mmHg; Kachuee et al., 2017) and the peak-to-peak time limit (greater than 0.6 s) were set for ABP signal to exclude abnormal value and false peaks. Then, the peaks and troughs were extracted from the processed ABP as reference values of SBP and DBP, respectively. The reference MAP value was calculated as the following formula:

$$MAP = \frac{(SBP + 2DBP)}{3} \quad (5)$$

Noise Reduction

The segmented ECG and PPG signals are first preprocessed using the discrete wavelet decomposition (DWT) filter with Daubechies 8 wavelet (db8) to remove high-frequency noise, baseline drift, and other noise (Singh and Tiwari, 2006). Specifically, combined with the DWT filter and Nyquist sampling theorem (Unser, 2000), all the ECG and PPG signals are sampled at 125 Hz, and then decomposed layers to extract the approximate coefficients (CAs) and detail coefficients (CDs), respectively. For ECG signals, the number of the decomposed layers is seven, and CDs of the first layer and CAs of the seventh are set to zero to remove the baseline drift (0∼0.5 Hz) and high-frequency noises (31.125∼62.25 Hz); For PPG signals, the number of the decomposed layers is eight, and CDs of the first layer and CAs of the eighth are set to zero to remove the baseline drift (0∼0.25 Hz) and high-frequency noises (31.125∼62.25 Hz). Subsequently, the rest of the CAs and CDs are denoised *via* a soft threshold and then reconstructed to obtain the target PPG and ECG signals.

Layer Normalization

Layer normalization is the normalization of a single training data to all neurons in a layer. Through layer normalization, the amplitude of the preprocessed target PPG and ECG signal is distributed within the range of [-1, 1], so that the input signals distribution of the model is similar, and the MST-net model has better converged. The normalization formula is defined as:

$$Normalized = 2 \times \frac{x - x_{min}}{x_{max} - x_{min}} - 1 \quad (6)$$

Where x is the amplitude of the target PPG and ECG signals, x_{max} and x_{min} are the maximum amplitude and minimum

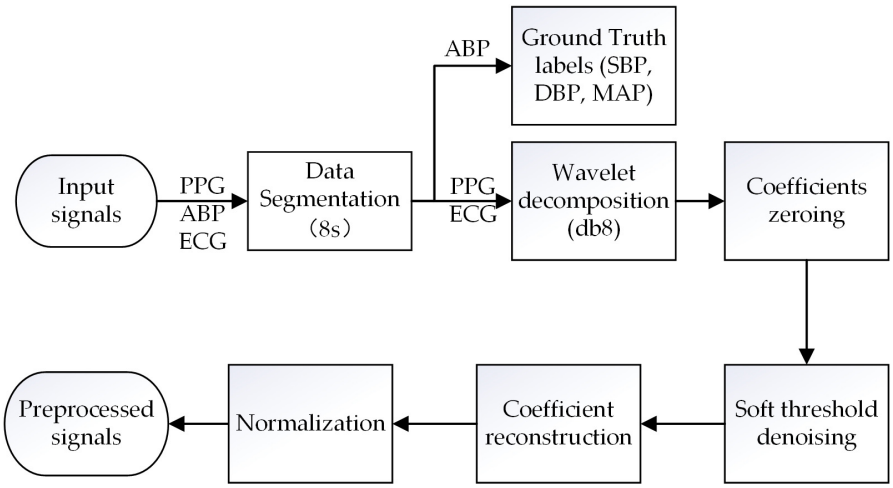


FIGURE 2 | Raw signals preprocessing pipeline.

amplitude, respectively, in the target signals. The pre-processing of signals is shown in **Figure 2**.

Model Architecture

The core concept of the neural network with multi-task learning we design is as follows (**Table 1**): we input the ECG and PPG signals (two one-dimensional ECG and PPG; input size: 2 × 1,000) at the same time and then process them using a one-dimensional (1D) convolution layer (Conv; with a convolution kernel of 15) to keep the information of the original signals as much as possible; we utilize the maximum pooling layer to remove redundant information and retain the main signal features; we adopt three network channels whose sizes of convolution kernels are 5, 7, and 9, respectively, to capture multi-scale features related to BP from target signals by using different receptive fields on each channel; we set four modules in each channel and set two convolution layers in each module to extract features; we set 64, 128, 256, and 512 filters on the four modules of each channel, respectively, to learn 1,536 features of target PPG and ECG signals; we set up two FC layers (1,536 and 256 neurons) for regression estimation of BP values, and the output layer consists of three neurons. In addition, to estimate SBP, DBP, and MAP simultaneously, we design a multi-task learning module in the designed neural network to reduce the over-fitting of specific tasks and improve the adaptability and efficiency of different tasks (Ruder, 2017).

Setting of Model Parameter

Batch normalization layer and activation function. Batch normalization is the normalization of individual neurons between different training data. The batch normalization layer can accelerate the convergence rate. Non-linear function rectifier linear unit (RELU) is introduced as the activation function after the normalization layer to avoid the gradient disappearance problem during the training process of the designed network and make the network train faster (Han and Moraga, 1995; Nair and Hinton, 2010; Chung et al., 2015). The RELU formula is defined as follows:

$$\mathrm{Relu}(x) = \begin{cases} 0, x < 0 \\ x, x \geq 0 \end{cases} \quad (7)$$

Adam

Adam can combine the advantages of AdaGrad (adjusting the learning rate (LR) of each different parameter) and Rmsprop

TABLE 1 | The network architecture of the MST-net model.

	MST (5)	MST (7)	MST (9)
		Input (2 × 1,000)	
	Stream 1	Stream 2	Stream 3
Layer 1		Conv (15)	
Layer 2		Max-pooling (3)	
Layer 3	Conv (5)-64 **	Conv (7)-64 **	Conv (9)-64 **
Layer 4	Conv (5)-64 **	Conv (7)-64 **	Conv (9)-64 **
Layer 5	Conv (5)-128 **	Conv (7)-128 **	Conv (9)-128 **
Layer 6	Conv (5)-128 **	Conv (7)-128 **	Conv (9)-128 **
Layer 7	Conv (5)-256 **	Conv (7)-256 **	Conv (9)-256 **
Layer 8	Conv (5)-256 **	Conv (7)-256 **	Conv (9)-256 **
Layer 9	Conv (5)-512 **	Conv (7)-512 **	Conv (9)-512 **
Layer 10	Conv (5)-512 **	Conv (7)-512 **	Conv (9)-512 **
Layer 11	AvgPool1d (1)	AvgPool1d (1)	AvgPool1d (1)
Layer 12	FC-512	FC-512	FC-512
Layer 13		FC-256	
		Output-3	

***Represents "Batch Normalization layer + non-linear function rectifier linear unit."*

(overcome the problem that the gradient of AdaGrad decreases sharply) optimization algorithms to update the parameter of the designed network to find the appropriate parameters and better convergence (Kingma and Ba, 2014).

Hyperparameters

In the training process, to train the designed network better and obtain the expected learning effect, the data input batch size is set to 100. The epoch is set to 150. The initial LR of the network is set to 0.01, and the fixed LR is decayed once every 5 epochs, which is defined as follows:

$$\mathrm{LR} = lr_{\mathrm{base}} \times \mathrm{gamma}^{\frac{\mathrm{step}}{5}} \tag{8}$$

Where lr_{base} is the original LR, gamma is the decay rate, and step represents the running number of epochs.

L2 Regularization

L2 regularization is added in the training process to improve the fitting effects and the generalization performance when the training set is small and the model is complex in the process of designed network training, that is, a constraint term is added to the MSE loss function $L(\theta)$ (Eq. 3) to generate a new loss function which is defined as follows:

$$L = L(\theta) + \lambda \sum_{i=1}^{k} w_i^2 \tag{9}$$

Where i is the layer number of the network, w is the weight of the network. λ is the coefficient of the L2 regularization which weighs the weight between the constraint term $\sum_{i=1}^{k} w_i^2$ and $L(\theta)$. Through the L2 regularization term, w can be reduced, and the smaller w, the lower the complexity and better the fitting of the network.

Model Performance Verification

The neural network with multi-task learning proposed in this study runs under the Pytorch1.8.1 framework, using Windows Server 2019 as the operating system. The server is equipped with an RTX 2080ti GPU with 11 G memory and an Intel Xeon Gold 5218 CPU with 32 cores and 64 GB memory. Based on existing methods for creating training/test data sets in BP estimation studies (Yan et al., 2019; Li et al., 2020; Miao et al., 2020a; Panwar et al., 2020), we set different training and test data. Due to the limited number of target signals data sets, five-fold cross-validation is used to evaluate the performance of the model. Our data is randomly divided into five equal-sized subsets, four of which are trained as training data, and the other one is tested as test data in turn, and the average of the five results is used as an estimate of the accuracy of the algorithm.

Model Performance Evaluation

To evaluate the BP estimation accuracy of the designed neural network, Pearson correlation coefficient (r), mean absolute error (MAE), mean error (ME), and standard deviation (SD) are examined. r represents the consistency of the estimated BP value and the reference BP value, ME represents the error between the estimated BP value and the reference BP value, MAE represents the absolute error between the estimated BP value and the reference BP value, and SD represents the degree of dispersion between the estimated BP value and the reference BP value. r, MAE, ME, and SD equations are defined as follows:

$$r = \frac{\sum_{i=1}^{n}(z_i - \overline{z})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(y_i - \overline{y})^2}\sqrt{\sum_{i=1}^{n}(z_i - \overline{z})^2}} \tag{10}$$

$$\mathrm{ME} = \frac{\sum_{i=1}^{n}(y_i - z_i)}{n} \tag{11}$$

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n}|z_i - y_i|}{n} \tag{12}$$

$$\mathrm{SD} = \sqrt{\frac{\sum_{i=1}^{n}(z_i - y_i - \mathrm{ME})^2}{n-1}} \tag{13}$$

Where y_i is the reference BP label value obtained from ABP, $\overline{y}$ is the average of y_i, z_i is the estimated BP value of the MST-net model, $\overline{z}$ is the average value of z_i, and n is the total number of target PPG and ECG signals segments in the test data set. Finally, BP estimation accuracy of the MST-net model is compared with the AAMI standards (Association for the Advancement of Medical Instrumentation, 2002) and the BHS standards (O'Brien et al., 1990) which are widely used as criteria for evaluating BP devices.

RESULTS AND DISCUSSION

Data Source

To accurately extract the true BP value (reference BP value) from ABP, the amplitude and peak-to-peak time of the ABP signal were restricted to exclude the interference of abnormal signals and false peaks. We can notice that multiple dicrotic wave peaks (false peaks) existed in the ABP signal (**Figure 3A**), and this might cause the number of detecting peak values to be more than the true number. We can extract the peak value more accurately after our restriction processing (**Figure 3A**; red dots). When limited the signal amplitude (<180 mmHg) and time constraints (>8 min) during ABP signals processing, the total number of subjects was reduced from 3,000 to 514. Then we obtained a total of 21,334 segments of BP data after the data segmentation process, that is, we use 213,340 beats in our model. The peaks and troughs were extracted from the processed ABP and used as the reference values of SBP and DBP, respectively, and calculated MAP based on SBP and DBP. The results showed that the DBP was mainly distributed in the range of 60 to 130 mmHg, the SBP was mainly distributed in the range of 80 to 180 mmHg, and the MAP was calculated based on DBP and SBP was mainly distributed between 70 and 135 mmHg (**Figure 3B**). This result was the same as the distribution of BP values obtained by Miao et al from ABP before (Kachuee et al., 2017; Miao et al., 2020b),

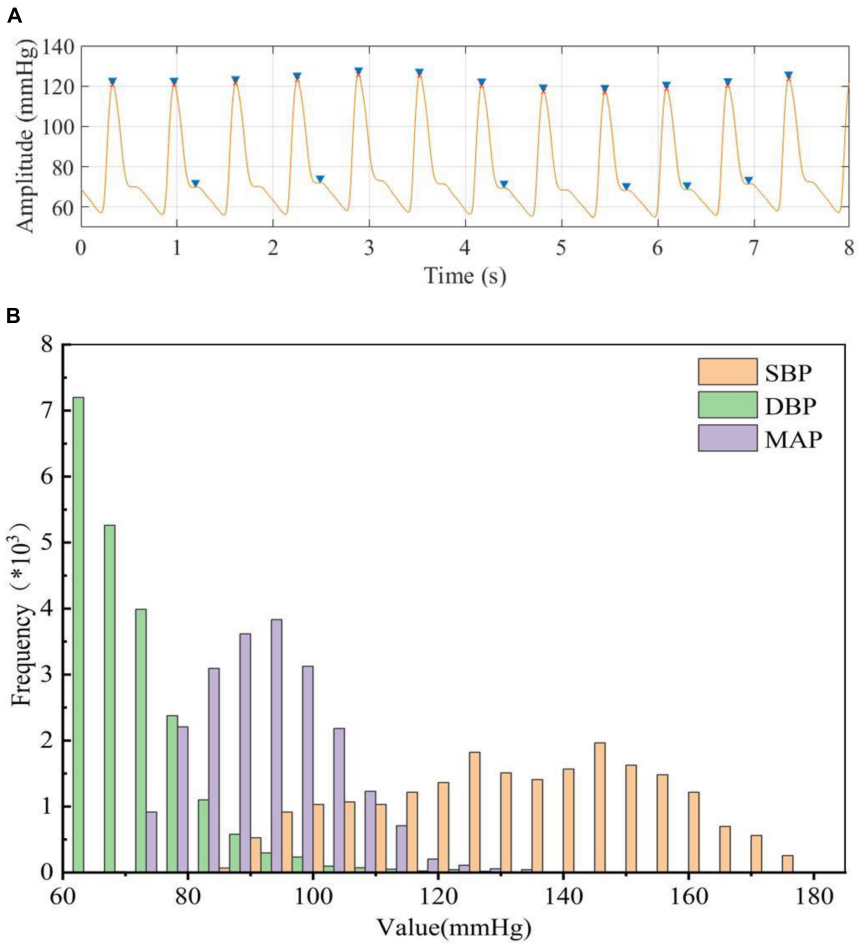


FIGURE 3 | (A) blue "▶" represents peak detection before processing, red "x" represents peak detection after processing, **(B)** Statistical histogram of BP data extracted from ABP.

which could provide the reference BP values for our designed neural network to estimate BP values.

Estimation Performance

To judge whether there is overfitting in our model, we used training loss and validation loss during the training process, and the training loss and validation loss were in a stable state after 100 epochs in the model training process, that is, the model had converged.

In order to investigate the performance of the designed neural network, the BP estimation accuracy of the network was evaluated according to *r*, ME, MAE, and SD. As result, the r of SBP, DBP, and MAP estimated by the MST-net model were 0.96, 0.92, and 0.94, respectively, and ME ± SD were 0.01 ± 5.81, 0.02 ± 3.55, and 0.01 ± 3.58 mmHg, respectively, (**Figure 4**). It can be observed that all reference values have a strong linear relationship with BP estimates (SBP, DBP, and MAP). The *p*-value for SBP, DBP, and MAP were 0.972, 0.796, and 0.969, respectively. It implied that the population mean of the samples was equal. These data points fell on both sides of the regression line and were close to the regression line (**Figures 4A–C**), indicating estimated BP data with high accuracy. The average values and difference values of reference and estimated BP values were the horizontal axis and vertical axis of the Bland-Altman plot (**Figures 4D–F**), respectively, these data points fall within the 95% confidence interval [-1.96 × SD, 1.96 × SD], indicating a good level of consistency of the reference and estimated BP data. Also, the average error between the reference and the estimate (red line) is very close to zero mmHg, indicating a high degree of consistency between the reference and estimated BP data. In addition, we provided the histogram of the error distribution between the estimate and the reference value, and we can observe that most of the errors are concentrated around 0 (**Figures 4G–I**). On the other hand, our network had fewer parameters than previous network models (Biswas et al., 2019; Panwar et al., 2020). That is to say, our network with an optimized algorithm was of lower complexity which can contribute to avoiding the constraints of computing power and memory for platform deployments (e.g., mobile devices, wearable devices). From the above analysis, it can be noted that the method of this study can achieve a precision estimation of SBP, DBP, and MAP. It is noted that the estimated SBP values through our model were not limited to less than

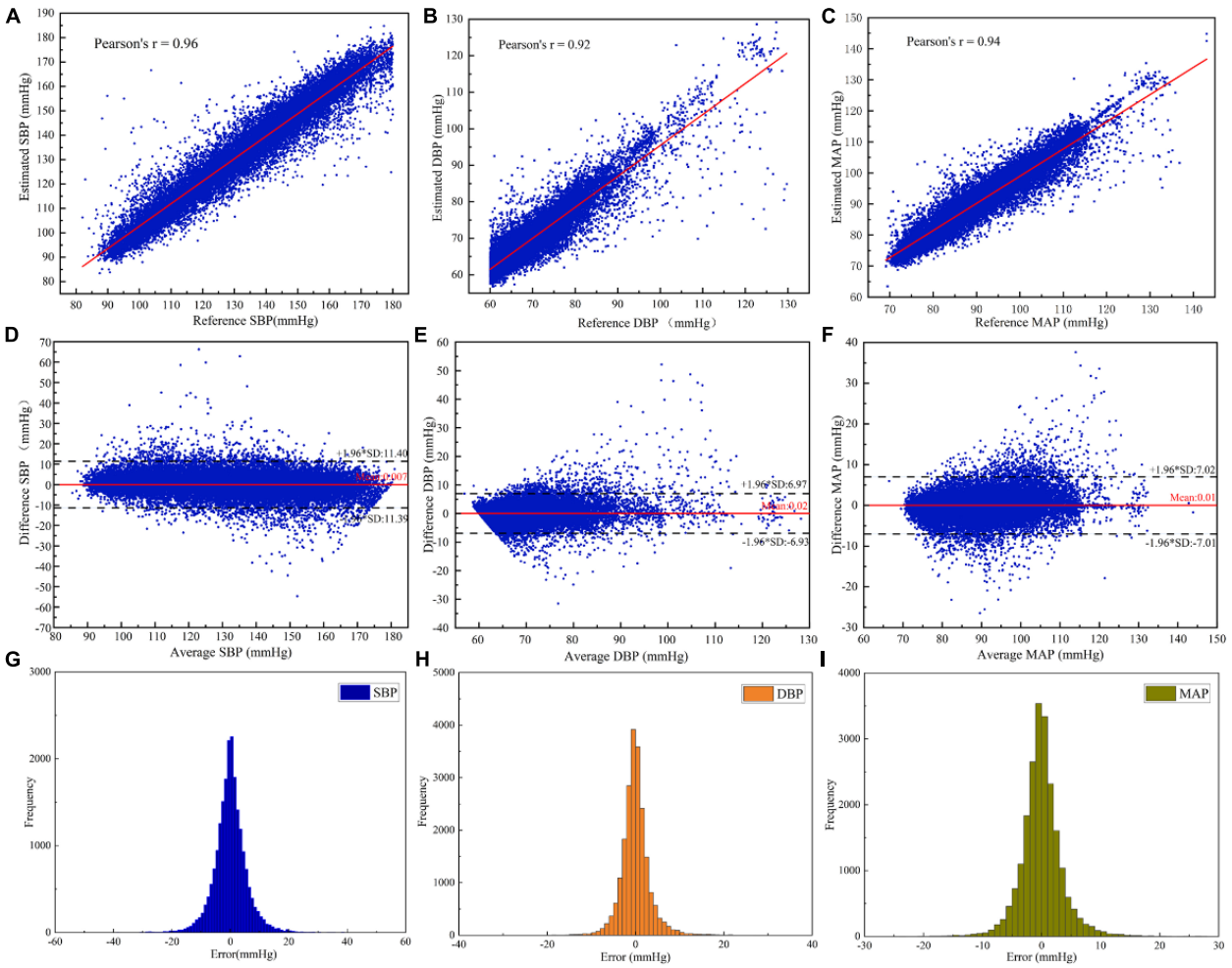


FIGURE 4 | Evaluation of the estimated BP performance of the MST-net model: **(A)** SBP correlation coefficient plot; **(B)** DBP correlation coefficient plot; **(C)** MAP correlation coefficient plot; **(D)** Bland-Altman plot of SBP; **(E)** Bland-Altman plot of DBP; **(F)** Bland-Altman plot of MAP; **(G)** Error histogram for SBP; **(H)** Error histogram for DBP; and **(I)** Error histogram for MAP.

180 mmHg, and the SBP values beyond 180 mmHg can be also predicted through our model. However, we just predicted SBP values within 180 mmHg same as many references (Kachuee et al., 2017; Baek et al., 2019; Thambiraj et al., 2020, etc.). The reasons were as follows: First, there were a few cases of SBP reaching 180 mmHg in the database; Secondly, when analyzing the reference signal and input signal values, the SBP values greater than 180 mmHg were calculated from the reference signal, which was generally caused by irregular noise signals.

In order to evaluate the accuracy of BP estimation based on this study, the BP estimation results of this study were compared with international AAMI and BHS standards. According to the

TABLE 2 | Comparison of estimated BP values between our work and AAMI standard.

		ME (mmHg)	SD (mmHg)	Subjects	Assessment result
Our results	SBP	0.007	5.81		
	DBP	0.022	3.55	514	Satisfied
	MAP	0.009	3.58		
AAMI	SBP	≤5	≤8	≥85	
(AAMI, 2002)	DBP				
	MAP				

TABLE 3 | Comparison of estimated BP values between our work and BHS standard.

		Cumulative error percentage			
		C. P. 5	C. P. 10	C. P. 15	Assessment result
Our result	SBP	71.56%	92.28%	97.66%	A
	DBP	89.88%	98.25%	99.40%	A
	MAP	87.89%	98.05%	99.52%	A
BHS (O'Brien et al., 1990)	Grade A	60%	85%	95%	
	Grade B	50%	75%	90%	
	Grade C	40%	65%	80%	

TABLE 4 | Comparison with other experimental performance.

Work	Dataset	Method	SBP		DBP		MAP	
			MAE	SD	MAE	SD	MAE	SD
Miao et al., 2020a	Own dataset	Feature extraction	6.13	7.76	4.54	5.52	4.81	6.03
Ding et al., 2016			4.09	5.21	3.18	4.13	3.18	4.06
Sharifi et al., 2019	MIMIC II	Deep learning algorithm	7.83	9.10	4.86	5.21	3.63	4.60
Baker et al., 2021	MIMIC III		4.41	6.11	2.91	4.23	2.77	3.88
This work	MIMIC II	MST deep learning algorithm	4.04	5.81	2.29	3.55	2.46	3.39

AAMI standard (**Table 2**), the target subjects of this study were 514, far more than the 85 required by the AAMI standard. the ME of SBP, DBP, and MAP were 0.007, 0.022, and 0.009 mmHg, respectively, which were far lower than the ME $\leq$ 5 mmHg required by the AAMI standard. SD of SBP, DBP, and MAP estimated were 5.81, 3.55, and 3.58 mmHg, respectively, which were far lower than SD $\leq$ 8 mmHg required by the AAMI standard. It showed that the estimated BP values by our customized model met the AAMI standard. The cumulative error percentage estimated by our model of the SBP reached 71.56, 92.28, and 97.66%, respectively; the DBP reached 89.88, 98.25, and 99.40%, respectively, and the MAP reached 87.89, 98.03, and 99.52%, respectively, (**Table 3**), which all showed much higher than the A grade standard of BHS (60, 85, and 95%). To sum up, BP values (SBP, DBP, and MAP) estimated by our customized model reached a small error and achieved good results.

In order to verify the effectiveness of the designed network, the proposed BP estimation method was compared with previous work. In general, it is difficult to make a fair comparison of BP estimation work due to different evaluation metrics and inadequately specified datasets. For example, for some BP estimation work based on ECG and PPG, Miao et al. (2020a) and Ding et al. (2016) used their own data sets, Baker et al. (2021) used the MIMIC III database, Sharifi et al. (2019) used the same database like ours. From the results, our proposed method performed better than these studies (**Table 4**). For a fair comparison, we selected two popular machine learning methods, VGG network (Simonyan and Zisserman, 2015) and Resnet (He et al., 2016), to compare BP estimation results using the same dataset. We can notice that the MAE $\pm$ SD of SBP, DBP, and MAP in our study (SBP: 4.04 $\pm$ 5.81, DBP: 2.29 $\pm$ 3.55, and MAP: 2.46 $\pm$ 3.39) all were smaller than the popular machine learning methods results (VGG, SBP: 8.47 $\pm$ 11.45, DBP: 4.70 $\pm$ 6.70, MAP: 5.09 $\pm$ 6.94; Resnet, SBP: 4.12 $\pm$ 5.97, DBP: 2.31 $\pm$ 3.60, and MAP: 2.50 $\pm$ 3.67), indicating that the network model we designed performed better at a BP estimated work (**Table 5**). It should be noted that the setting of hyperparameters in the two popular machine learning methods is the same as our proposed method (except that the LR of the VGG network is set to 0.005). In addition, we noticed that compared with previous studies that used a separate training method to estimate a BP value, our model can simultaneously estimate multiple BP values by using only one model, which not only reduced the complexity but also improved the work efficiency (**Table 6**).

Overall, the MST-net model proposed in this study had a relatively simple structure and achieved good accuracy in the field of continuous BP estimation, which was a very competitive method and made contributions to the improvement of BP estimation accuracy.

TABLE 5 | Comparison of predicted BP values between our and previous work based on our dataset from MIMIC-II.

Work	MAE ± SD (mmHg)		
	SBP	DBP	MAP
Resnet (He et al., 2016)	4.12 ± 5.97	2.31 ± 3.60	2.50 ± 3.67
VGG (Simonyan and Zisserman, 2015)	8.47 ± 11.45	4.70 ± 6.70	5.09 ± 6.94
This work	4.04 ± 5.81	2.29 ± 3.55	2.46 ± 3.39

TABLE 6 | Comparison of No. model for BP evaluation between our and previous work.

Work	Subjects	Model	SBP (mmHg)	DBP (mmHg)
Rong and Li, 2021	11,546 samples	2	5.59 ± 7.25	3.36 ± 4.48
Kachuee et al., 2017	942 subjects	2	11.17 ± 10.09	5.35 ± 6.14
Gaurav et al., 2016	3,000* subjects	2	4.47 ± 6.85	3.21 ± 4.72
This work	21,334 samples	1	4.04 ± 5.81	2.29 ± 3.55

**Number of subjects before signal processing.*

Impacts of Model Architecture

In order to determine the effects of network structure on the performance of BP estimation, the effects of the number of network channels and the size of the convolution kernel of network channels on BP estimation performance were investigated (**Table 7**). Compared to the BP estimation results of the single-channel model with convolution kernels of 7, the

TABLE 7 | Impacts of that number of network channels and the size of channel convolution kernel on the performance of BP estimation.

Kernel size	SBP (mmHg)	DBP (mmHg)	MAP (mmHg)
MST-net (3)	4.40	2.50	2.68
MST-net (5)	4.20	2.41	2.58
MST-net (7)	4.13	2.33	2.52
MST-net (9)	4.07	2.31	2.48
MST-net (3, 5, 7)	4.10	2.28	2.47
MST-net (5, 7, 9)	4.04	2.29	2.46

estimation errors of SBP, DBP, and MAP *via* the three-channel model with convolution kernels of (3, 5, 7) were reduced by 0.03, 0.05, and 0.05 mmHg, respectively. The reason was that multi-channel can extract more abundant features than single-channel. In the single-channel model, when the size of the convolution kernel increased from 3 to 9, the errors of SBP, DBP, and MAP decreased significantly, indicating that the increase of the convolution kernel in the model could improve BP estimation performance. The reason was that the larger convolution kernel has a larger receptive field which contributes to extracting the features related to BP from time-series signals with periodic patterns. These results also showed that the features related to BP have a larger span of time. At the same time, with the increase in the size of the convolution kernel, the increase of BP estimation performance gradually decreased, indicating that the size of the convolution kernel used to extract features was limited, and cannot be infinite. When the size of the multi-channel convolution kernel was increased from (3, 5, 7) to (5, 7, 9), the estimated errors of SBP, DBP, and MAP were improved by 0.06, 0.01, and 0.01 mmHg, respectively. This also showed that the larger convolution kernel could improve the performance of the BP estimation model.

CONCLUSION

In this article, a novel continuous BP estimation based on multi-scale feature extraction by the neural network with multi-task learning was proposed to estimate BP continuously and accurately without calibration using PPG and ECG signals. This research was a step toward developing an efficient and lightweight solution. We adopted segmenting, denoising, and normalizing to preprocessed target and label signals and then extracted the reference BP value from the preprocessed label signals, especially using peak-to-peak timing limits of signals to eliminate the interference of the false peak of the wave; we designed a neural network with multi-task learning to extract multi-scale features related to BP from preprocessed target signals, fully excavated and learned the relationship between the multi-scale features and BP, and then estimated three BP values simultaneously through multi-task learning. The results showed that the errors of MAE $\pm$ SD for SBP, DBP, and MAP were 4.04 $\pm$ 5.81, 2.29 $\pm$ 3.55, and 2.46 $\pm$ 3.58 mmHg, respectively, and the correlation coefficients were 0.96, 0.92, and 0.94, respectively. These results met the AAMI standard and reached A level of the BHS standard, and showed better BP continuous monitoring results than other previous works, and without calibration.

In conclusion, our study provided convincing evidence that our method can achieve high precision continuous BP estimation and had a relatively simple structure, which can be further applied to wearable devices.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material; further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

HJ, LZ, and QF: conceptualization. HJ: methodology, software, and data curation. HJ and LZ: formal analysis and original draft preparation. LZ, QF, and DH: review and editing, supervision, and funding acquisition. All authors have read and agreed to the published version of the manuscript.

FUNDING

This work was partially supported by Program of the Guangdong Academy of Sciences (Grant Nos. 2019GDASYL-0103016, 2021GDASYL-20210302002, 2018GDASCX-0103, and 2021GDASYL-20210103029) and the National Natural Science Foundation of China (Grant Nos. 81974275 and 12126603).

REFERENCES

Association for the Advancement of Medical Instrumentation (2002). *American National Standard for Electronic or Automated Sphygmomanometers: ANSI/AAMI SP 10: 2002*. Arlington, VA: *AAMI*.

Argha, A., Celler, B. G., and Lovell, N. H. (2022). Artificial Intelligence Based Blood Pressure Estimation From Auscultatory and Oscillometric Waveforms: a Methodological Review. *IEEE Rev. Biomed. Eng.* 15, 152–168. doi: 10.1109/RBME.2020.3040715

Baker, S., Xiang, W., and Atkinson, I. (2021). A hybrid neural network for continuous and non-invasive estimation of blood pressure from raw electrocardiogram and photoplethysmogram waveforms. *Comput. Meth. Prog. Bio.* 207:106191. doi: 10.1016/j.cmpb.2021.106191

Baek, S., Jang, J., and Yoon, S. (2019). End-to-End Blood Pressure Prediction via Fully Convolutional Networks. *IEEE Access* 7, 185458–185468. doi: 10.1109/access.2019.2960844

Biswas, D., Everson, L., Liu, M., Panwar, M., Verhoef, B. E., Patki, S., et al. (2019). CorNET: Deep Learning Framework for PPG-Based Heart Rate Estimation and Biometric Identification in Ambulant Environment. *IEEE Trans. Biomed. Circuits Syst.* 13, 282–291. doi: 10.1109/TBCAS.2019.2892297

Chung, J., Gulcehre, C., Cho, K., and Bengio, Y. (2015). Gated Feedback Recurrent Neural Networks. *arXiv* 1502:02367. doi: 10.48550/arXiv.1502.02367

Chung, E., Chen, G., Alexander, B., and Cannesson, M. (2013). Non-invasive continuous blood pressure monitoring: a review of current applications. *Front. Med.* 7, 91–101. doi: 10.1007/s11684-013-0239-5

Ding, X. R., Zhang, Y. T., Liu, J., Dai, W. X., and Tsang, H. K. (2016). Continuous Cuffless Blood Pressure Estimation Using Pulse Transit Time and Photoplethysmogram Intensity Ratio. *IEEE Trans. Biomed. Eng.* 63, 964–972. doi: 10.1109/TBME.2015.2480679

El-Hajj, C., and Kyriacou, P. A. (2020). A review of machine learning techniques in photoplethysmography for the non-invasive cuff-less measurement of blood pressure. *Biomed. Signal Process. Control.* 58:101870. doi: 10.1016/j.bspc.2020.101870

Eom, H., Lee, D., Han, S., Hariyani, Y. S., Lim, Y., Sohn, I., et al. (2020). End-to-End Deep Learning Architecture for Continuous Blood Pressure Estimation Using Attention Mechanism. *Sensors* 20:2338. doi: 10.3390/s20082338

Gaurav, A., Maheedhar, M., Tiwari, V. N., and Narayanan, R. (2016). Cuff-less PPG based continuous blood pressure monitoring: a smartphone based approach. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* 2016, 607–610. doi: 10.1109/EMBC.2016.7590775

Goldberger, A. L., Amaral, L. A., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., et al. (2000). PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation* 101, e215–e220. doi: 10.1161/01.cir.101.23.e215

Han, J., and Moraga, C. (1995). "The Influence of the Sigmoid Function Parameters on the Speed of Backpropagation Learning," in *From Natural to Artificial Neural Computation. IWANN 1995. Lecture Notes in Computer Science*, Vol. 930, eds J. Mira and F. Sandoval (Berlin: Springer), doi: 10.1007/3-540-59497-3_175

He, K., Zhang, X., Ren, S., and Sun, J. (2016). "Deep Residual Learning for Image Recognition," in *2016 IEEE Conf. Comput. Vision Pattern Recognit*, (*CVPR*), 770–778. doi: 10.1109/cvpr.2016.90

Huynh, T. H., Jafari, R., and Chung, W. Y. (2018). Noninvasive Cuffless Blood Pressure Estimation Using Pulse Transit Time and Impedance Plethysmography. *IEEE Trans. Biomed. Eng.* 66, 967–976. doi: 10.1109/TBME.2018.2865751

Kachuee, M., Kiani, M. M., Mohammadzade, H., and Shabany, M. (2015). Cuff-Less High-Accuracy Calibration-Free Blood Pressure Estimation Using Pulse Transit Time. *IEEE Int. Symp. Circuits Syst.* 2015, 1006–1009. doi: 10.1109/ISCAS.2015.7168806

Kachuee, M., Kiani, M. M., Mohammadzade, H., and Shabany, M. (2017). Cuff-Less Blood Pressure Estimation Algorithms for Continuous Health-Care Monitoring. *IEEE Trans. Biomed. Eng.* 64, 859–869. doi: 10.1109/TBME.2016.2580904

Simonyan, K., and Zisserman, A. (2015). "Very deep convolutional networks for large-scale image recognition," in *The 3rd International Conference on Learning Representations*, (*ICLR2015*). doi: 10.48550/arXiv.1409.1556

Kingma, D. P., and Ba, J. (2014). Adam: A Method for Stochastic Optimization. *arXiv* 1412:6980. doi: 10.48550/arXiv.1412.6980

Lázaro, J., Gil, E., Orini, M., Laguna, P., and Bailón, R. (2019). Baroreflex Sensitivity Measured by Pulse Photoplethysmography. *Front. Neurosci.* 13:339. doi: 10.3389/fnins.2019.00339

Li, Y. H., Harfiya, L. N., Purwandari, K., and Lin, Y. D. (2020). Real-Time Cuffless Continuous Blood Pressure Estimation Using Deep Learning Model. *Sensors* 20:5606. doi: 10.3390/s20195606

Maqsood, S., Xu, S., Springer, M., and Mohawesh, R. (2021). A Benchmark Study of Machine Learning for Analysis of Signal Feature Extraction Techniques for Blood Pressure Estimation Using Photoplethysmography (PPG). *IEEE Access* 9, 138817–138833. doi: 10.1109/access.2021.3117969

Miao, F., Liu, Z. D., Liu, J. K., Wen, B., He, Q. Y., and Li, Y. (2020a). Multi-Sensor Fusion Approach for Cuff-Less Blood Pressure Measurement. *IEEE J. Biomed. Health Inform.* 24, 79–91. doi: 10.1109/JBHI.2019.2901724

Miao, F., Wen, B., Hu, Z., Fortino, G., Wang, X. P., Liu, Z. D., et al. (2020b). Continuous blood pressure measurement from one-channel electrocardiogram signal using deep-learning techniques. *Artif. Intell. Medicine.* 108:101919. doi: 10.1016/j.artmed.2020.101919

Miotto, R., Wang, F., Wang, S., Jiang, X., and Dudley, J. T. (2018). Deep learning for healthcare: review, opportunities and challenges. *Briefings Bioinf.* 19, 1236–1246. doi: 10.1093/bib/bbx044

Mukkamala, R., Hahn, J. O., Inan, O. T., Mestha, L. K., Kim, C. S., Töreyin, H., et al. (2015). Toward Ubiquitous Blood Pressure Monitoring via Pulse Transit Time: theory and Practice. *IEEE Trans. Biomed. Eng.* 62, 1879–1901. doi: 10.1109/TBME.2015.2441951

Nair, V., and Hinton, G. E. (2010). "Rectified Linear Units Improve Restricted Boltzmann Machines," in *Proc. 27th Int. Conf. Machine Learn. (ICML-10), June 21-24, 2010* (Haifa), 807–814.

O'brien, E., Asmar, R., Beilin, L., Imai, Y., Mallion, J. M., Mancia, G., et al. (2003). European Society of Hypertension recommendations for conventional, ambulatory and home blood pressure measurement. *J. Hyperten.* 21, 821–848. doi: 10.1097/01.hjh.0000059016.82022.ca

O'Brien, E., Petrie, J., Littler, W., de Swiet, M., Padfield, P. L., O'Malley, K., et al. (1990). The British Hypertension Society protocol for the evaluation of automated and semi-automated blood pressure measuring devices with special reference to ambulatory systems. *J. Hyperten.* 8, 607–619. doi: 10.1097/00004872-199007000-00004

Panwar, M., Gautam, A., Biswas, D., and Acharyya, A. (2020). PP-Net: A Deep Learning Framework for PPG-Based Blood Pressure and Heart Rate Estimation. *IEEE Sens. J.* 20, 10000–10011. doi: 10.1109/JSEN.2020.2990864

Paviglianiti, A., Randazzo, V., Villata, S., Cirrincione, G., and Pasero, E. (2021). A Comparison of Deep Learning Techniques for Arterial Blood Pressure Prediction. *Cognit. Comput.* 27, 1–22. doi: 10.1007/s12559-021-09910-0

Penaz, J. (1973). "Photoelectric measurement of blood pressure, volume and flow in the finger," in *Dig. 10th Int. Conf. Med. Biol. Eng* (Dresden).

Pressman, G., and Newgard, P. (1963). A Transducer for the Continuous External Measurement of Arterial Blood Pressure. *IEEE Trans. Biomed. Electronics* 10, 73–81. doi: 10.1109/TBMEL.1963.4322794

Radha, M., de Groot, K., Rajani, N., Wong, C. C. P., Kobold, N., Vos, V., et al. (2019). Estimating blood pressure trends and the nocturnal dip from photoplethysmography. *Physiol. Meas.* 40:025006. doi: 10.1088/1361-6579/ab030e

Ravì, D., Wong, C., Deligianni, F., Berthelot, M., Andreu-Perez, J., Lo, B., et al. (2017). Deep Learning for Health Informatics. *IEEE J. Biomed. Health Inform.* 21, 4–21. doi: 10.1109/JBHI.2016.2636665

Rong, M., and Li, K. (2021). A multi-type features fusion neural network for blood pressure prediction based on photoplethysmography. *Biomed. Signal Process Control.* 68:102772. doi: 10.1016/j.bspc.2021.102772

Ruder, S. (2017). An Overview of Multi-Task Learning in Deep Neural Networks. *arXiv* 1706:05098. doi: 10.48550/arXiv.1706.05098

Sharifi, I., Goudarzi, S., and Khodabakhshi, M. B. (2019). A novel dynamical approach in continuous cuffless blood pressure estimation based on ECG and PPG signals. *Artif. Intell. Med.* 97, 143–151. doi: 10.1016/j.artmed.2018.12.005

Singh, B. N., and Tiwari, A. K. (2006). Optimal selection of wavelet basis function applied to ECG signal denoising *Dig. Signal Process.* 16, 275–287. doi: 10.1016/j.dsp.2005.12.003

Song, K., Chung, K. Y., and Chang, J. H. (2020). Cuffless Deep Learning-Based Blood Pressure Estimation for Smart Wristwatches. *IEEE Trans. Instrum. Meas.* 69, 4292–4302. doi: 10.1109/TIM.2019.2947103

Thambiraj, G., Gandhi, U., Mangalanathan, U., Jose, V. J. M., and Anand, M. (2020). Investigation on the effect of Womersley number, ECG and PPG features for cuff less blood pressure estimation using machine learning. *Biomed. Signal Process. Control.* 60:101942. doi: 10.1016/j.bspc.2020.101942

Unser, M. (2000). Sampling—50 Years After Shannon. *Proc. IEEE* 88, 569–587. doi: 10.1109/5.843002

Xu, Z., Liu, J., Chen, X., Wang, Y., and Zhao, Z. (2017). Continuous blood pressure estimation based on multiple parameters from eletrocardiogram and photoplethysmogram by Back-propagation neural network. *Comput. Indust.* 89, 50–59. doi: 10.1016/j.compind.2017.04.003

World Health Organization (2020). *World health statistics 2020:monitoring health for the SDGs,sustainable development goals*. Geneva: World Health Organization.

Yan, W. R., Peng, R. C., Zhang, Y. T., and Ho, D. (2019). Cuffless Continuous Blood Pressure Estimation From Pulse Morphology of Photoplethysmograms. *IEEE Access.* 7, 141970–141977. doi: 10.1109/ACCESS.2019.2942936

Yang, S., Zhang, Y., Cho, S. Y., Correia, R., and Morgan, S. P. (2021). Non-invasive cuff-less blood pressure estimation using a hybrid deep learning model. *Opt. Quant. Electron.* 53, 93. doi: 10.1007/s11082-020-02667-0

Yen, C. T., and Liao, C. H. (2022). Blood Pressure and Heart Rate Measurements Using Photoplethysmography with Modified LRCN Comput. *Mater. Continua.* 71, 1973–1986. doi: 10.32604/cmc.2022.022679

Yoon, Y. Z., Kang, J. M., Kwon, Y., Park, S., Noh, S., Kim, Y., et al. (2018). Cuff-Less Blood Pressure Estimation Using Pulse Waveform Analysis and Pulse Arrival Time. *IEEE J. Biomed. Health inform.* 22, 1068–1074. doi: 10.1109/JBHI.2017.2714674

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Neuroscience

ORIGINAL RESEARCH
published: 25 May 2022
doi: 10.3389/fnins.2022.904623

Real-World Visual Experience Alters Baseline Brain Activity in the Resting State: A Longitudinal Study Using Expertise Model of Radiologists

Jiaxi Su[1,2†], Xiaoyan Zhang[1,2†], Ziyuan Zhang[1,2], Hongmei Wang[3], Jia Wu[4], Guangming Shi[5], Chenwang Jin[3] and Minghao Dong[1,2,5]**

[1] Engineering Research Center of Molecular and Neuro Imaging of Ministry of Education, School of Life Sciences and Technology, Xidian University, Xi'an, China, [2] Xi'an Key Laboratory of Intelligent Sensing and Regulation of Trans-Scale Life Information, School of Life Sciences and Technology, Xidian University, Xi'an, China, [3] Department of Medical Imaging, First Affiliated Hospital of Medical College, Xi'an Jiaotong University, Xi'an, China, [4] School of Foreign Languages, Northwestern Polytechnical University, Xi'an, China, [5] Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education, School of Artificial Intelligence, Xidian University, Xi'an, China

OPEN ACCESS

Edited by:
Yao Wu,
Children's National Hospital,
United States

Reviewed by:
Yuanqiang Zhu,
Fourth Military Medical University,
China
Bingsheng Huang,
Shenzhen University, China
Bihong T. Chen,
City of Hope National Medical Center,
United States

***Correspondence:**
Chenwang Jin
jin1115@xjtu.edu.cn
Minghao Dong
dminghao@xidian.edu.cn

[†] These authors have contributed equally to this work

Specialty section:
This article was submitted to Brain Imaging Methods, a section of the journal Frontiers in Neuroscience

Received: 25 March 2022
Accepted: 25 April 2022
Published: 25 May 2022

Citation:
Su J, Zhang X, Zhang Z, Wang H, Wu J, Shi G, Jin C and Dong M (2022) Real-World Visual Experience Alters Baseline Brain Activity in the Resting State: A Longitudinal Study Using Expertise Model of Radiologists. Front. Neurosci. 16:904623. doi: 10.3389/fnins.2022.904623

Visual experience modulates the intensity of evoked brain activity in response to training-related stimuli. Spontaneous fluctuations in the restful brain actively encode previous learning experience. However, few studies have considered how real-world visual experience alters the level of baseline brain activity in the resting state. This study aimed to investigate how short-term real-world visual experience modulates baseline neuronal activity in the resting state using the amplitude of low-frequency (<0.08 Hz) fluctuation (ALFF) and a visual expertise model of radiologists, who possess fine-level visual discrimination skill of homogeneous stimuli. In detail, a group of intern radiologists (n = 32) were recruited. The resting-state fMRI data and the behavioral data regarding their level of visual expertise in radiology and face recognition were collected before and after 1 month of training in the X-ray department in a local hospital. A machine learning analytical method, i.e., support vector machine, was used to identify subtle changes in the level of baseline brain activity. Our method led to a superb classification accuracy of 86.7% between conditions. The brain regions with highest discriminative power were the bilateral cingulate gyrus, the left superior frontal gyrus, the bilateral precentral gyrus, the bilateral superior parietal lobule, and the bilateral precuneus. To the best of our knowledge, this study is the first to investigate baseline neurodynamic alterations in response to real-world visual experience using longitudinal experimental design. These results suggest that real-world visual experience alters the resting-state brain representation in multidimensional neurobehavioral components, which are closely interrelated with high-order cognitive and low-order visual factors, i.e., attention control, working memory, memory, and visual processing. We propose that our findings are likely to help foster new insights into the neural mechanisms of visual expertise.

Keywords: SVM, recursive feature elimination, visual expertise, resting state fMRI, ALFF

INTRODUCTION

Visual experts refer to individuals with superior behavioral performance in visual recognition in specific domains (Curby and Gauthier, 2010). To become a visual expert, it requires visual learning with at least hundreds of cases of samples (Clark et al., 2012). A few real-world visual expertise models have been used to study the neural substrate underlying this behavioral expertise (Rossignoli-Palomeque et al., 2018). Increased level of activation was found in the left superior frontal gyrus and left cingulate cortex in radiologists, which is responsible for better working memory capability (Haller and Radue, 2005). Harel et al. (2010) demonstrated enhanced neuronal activity in the right precuneus of a group of car experts related to better memory retrieval strategies. Song et al. (2021) reported increased activation in the right anterior cingulate gyrus, but decreased activation in the superior parietal lobule in chess players (Song et al., 2020), which is closely related to improved visual processing and better attention control. These results derived from cross-sectional studies demonstrated that real-world visual experience alters the strength of evoked brain activity across widely distributed regions, which are supportive of high-order cognitive, such as attention control, working memory, and memory, and low-order visual components, such as visual processing (Khader et al., 2005; Cavanna and Trimble, 2006; Schipul and Just, 2016).

Low-frequency spontaneous fluctuations (0.01–0.1 Hz) of the restful brain play an important role in maintaining the ongoing internal brain representations (Oldfield, 1971; Tang et al., 2010; Evans et al., 2011), which are involved in the coding of previous experience and support learned skills (Dong et al., 2014). Particularly, patterns of spontaneous activity within the resting brain are shaped by experience-dependent changes in neural plasticity (Chakraborty, 2006). However, less attention has been paid to analyze how real-world visual experience alters the patterns of resting-state brain activity using longitudinal experimental design. In this regard, the baseline spontaneous neuronal activity reflects cortical excitability (Logothetis et al., 2001; Boly et al., 2007), and the level of cortical excitability may smear the spatial patterns of evoked brain activity (Di et al., 2013; Dong et al., 2015). We propose that the level of baseline brain activity is fundamentally important; therefore, this study aimed at investigating how short-term real-world visual experience modulates baseline neuronal activity in the resting state.

The amplitude of low-frequency fluctuations (ALFF) serves as an indicator of cortical excitability (Duff et al., 2008). Previous studies have used ALFF to measure the level of baseline brain activity in healthy subjects (Yang et al., 2007; Dong et al., 2015). In our study, a group of 32 radiology interns were recruited from our collaborative hospital. The resting-state MRI data were collected before and after 1 month of training in the X-ray department, and ALFF was calculated to quantify the level of baseline brain activity. To better capture the subtle changes in the strength of neuronal activity, a novel but sensitive machine learning analytical framework, support vector machine (SVM), was employed (Xu et al., 2019). We expected to see an altered level of activity in brain regions related to the multidimensional neurobehavioral component that supports visual recognition, such as attention control, working memory, memory extraction, and visual processing (Humphreys et al., 1999). To the best of our knowledge, this study is the first to investigate baseline neurodynamic alterations in response to short-term real-world visual experience using longitudinal experimental design.

MATERIALS AND METHODS

Subjects

The subjects in this study consisted of a cohort group of radiology interns, who were medical students in the undergraduate program in national medical schools. They were recruited from the First Affiliated Hospital of Medical College, Xi'an Jiaotong University. Thirty-two radiology interns [15 male participants, age: 22.47 $\pm$ 1.02 years old, mean $\pm$ standard deviation (SD)]. The participants were aware of the purpose of the study and the reason why they were recruited. All the subjects are medical students in the undergraduate program in national medical schools, and they underwent a 4-week rotation in the First Affiliated Hospital of Medical College, Xi'an Jiaotong University. The subjects worked in the X-ray department during the rotation and reviewed approximately 35 cases per day, 6 days a week. Their daily training included interpreting X-ray images and drafting reports for each case. Senior radiologists were assigned to these interns as mentors and provided response to their clinical reports each day. The intern radiologists reviewed a minimum of 831 cases during the rotation period, as recorded in the Picture Archiving and Communication System (PACS). Moreover, all the subjects had no neurological or psychiatric brain diseases, had no history of head trauma, and had not taken recreational drugs or drugs that influence brain function during the time window of this study (Oldfield, 1971).

Behavioral Tasks

This study employed a longitudinal experimental design, which is rare in visual expertise studies. Basically, the subjects underwent the behavioral assessment (including prescreening tasks and behavioral tasks) and MRI scanning before training, and behavioral assessment (only behavioral tasks) and MRI scanning after a 4-week visual training. Note that the purpose of prescreening and the behavior tasks were different. The prescreening procedures were conducted before MRI scanning, aiming to exclude confounding factors such as visual expertise from other known domains (e.g., cars, chess, birds, and mushrooms). Specifically, we used questionnaires to ensure that the subjects had no visual expertise of other known domains, such as aircrafts, animals, and plants. The behavioral tasks were conducted after MRI scanning, aiming to quantify the level of face expertise and radiological expertise, using the Cambridge face memory test (CFMT) (Duchaine and Nakayama, 2006) and radiological expertise task (Evans et al., 2011) respectively, as introduced in our previous studies (Wang et al., 2021; Zhang et al., 2022).

A standard behavioral task, i.e., radiological expertise task (Evans et al., 2011), was used to quantify the radiological expertise of the subjects before and after radiological training. The images

selected for RET were identical for both tests. Basically, the subjects were shown 100 standard lung X-ray images and were asked to render a diagnostic decision (e.g., tumor present or absent) and prognosis (e.g., malignant vs. benign) for each image in RET. The 100 standard lung X-ray images were carefully chosen from the X-ray image library in the Medical Imaging Department of the First Affiliated Hospital of the College of Medicine under the guidance of three independent senior radiologists with more than 15 years of diagnostic radiology experience. These 100 images used for RET consisted of three ascending levels of difficulty with a portion of 50, 30, and 20%, respectively. Each lung X-ray image contained 0∼N nodules, and there was no mention of a diagnosis unrelated to pulmonary nodules. Sixty-five X-ray images containing only 1∼3 nodules were selected as positive cases, and 35 X-ray images without tumors were selected as negative cases. The pathologies in the images were carefully examined and reconfirmed by these three experts. The detailed procedure of CFMT and RET was introduced in our previous study (Zhang et al., 2022).

MRI Data Acquisition

Before MRI scanning, all subjects underwent complete physical and neurological examination. Note that, to eliminate the potential influence of behavioral tasks on central representation, the behavioral task took place after MRI data acquisition. The MRI scanning was performed on the 3 Telsa MRI system (EXCITE, General Electric, Milwaukee, Wisc.) at the First Affiliated Hospital of Medical College, Xi'an Jiaotong University in Xi'an, China. To eliminate the time-of-day effect, the scanning was performed from 8:30 to 12:30 a.m. (Hasler et al., 2014). A resting scan and a structural scan were conducted. A standard birdcage head coil was used, along with restraining foam pads to minimize head motion and to diminish scanner noise. Prior to the scan, subjects were instructed to close their eyes, keep their heads still, and stay awake during the scanning process. After scanning, the subjects would be asked whether they had fallen asleep during the process.

For the resting MRI scanning, the following parameters were used. Each volume contains 35 axial slices, scan duration = 370 s, repetition time (TR)/echo time (TE) = 2,000 ms/30 ms, field of view = 240 mm, total brain volume collection = 185, matrix = 64 × 64, flip angle = 90°, voxel size = 3.8 × 3.8 × 5.0 mm^3, gap = 0 mm, thickness = 4 mm, layer spacing = 0 mm. The resting-state fMRI scans lasted for 8 min and 20 s. High-resolution T1-weighted structural imaging data used 3D magnetization preparation to quickly acquire gradient echo sequence for acquisition.

MRI Data Preprocessing

Statistical Parametric Mapping (SPM12)[1] and the Data Processing Assistant for Resting-State fMRI (DPARSF 4.5)[2] were used for MRI data preprocessing. The first 10 images were deleted to eliminate non-equilibrium effects of magnetization and allow the participants to adapt to the experimental environment. The images were corrected for the acquisition delay between slices, motion corrected and co-registered to the subject's anatomical images in native space. Two subjects had head motion exceeding the threshold of 0.2 mm (frame-wise displacement, i.e., Power FD). For the remaining 30 subjects, a two-sample *t*-test was used to verify that there was no significant difference in head movement between the two groups for the remaining subjects. Next, all the functional images were normalized to the MRI space using the deformation field maps obtained from structural image segmentation, following the segmentation routine in SPM12. The normalized images were resampled to 3 mm isotropic voxels, which were then spatially smoothed with a 6-mm full width-at-half-maximum Gaussian kernel. Finally, the linear trend was removed (Dale et al., 2000), and temporal filtering (0.01–0.08 Hz) was performed on the time series of each voxel to reduce the effect of low-frequency drifts and high-frequency noise (Zou et al., 2008).

Feature Extraction

Generation of Voxel-Wise Amplitude of Low-Frequency Fluctuations Map

Resting-State fMRI Data Analysis Toolkit (REST)[3] was used to compute ALFF (Song et al., 2011). ALFF measures the level of intrinsic or spontaneous neuronal activity in a given voxel (Jiang et al., 2004). The ALFF serves as an indicator of cortical excitability (Duff et al., 2008), and the volume of regional cerebral blood flow is correlated with ALFF in the brain region from the resting-state data (Li et al., 2012); therefore, it is taken as the index for the level of baseline brain activity. To calculate ALFF, after preprocessing, a fast Fourier transform (FFT) was used to transform from time domain to frequency domain for a given voxel, and the specific parameters are as follows: the taper percentage was zero, and the FFT length was set to short. Then, the square root of the power spectrum at each frequency was calculated, and the average value was taken in the range of 0.01–0.08 Hz. The average square root of a given voxel was taken as ALFF (Jia et al., 2020). To minimize the impact of variability among participants and reduce noise interference, we divided the ALFF of a given voxel by the average ALFF value of whole brain voxels to obtain the standardized value.

Generation of Region-Wise Amplitude of Low-Frequency Fluctuations Map

The voxel-wise ALFF map was averaged into a region-wise ALFF map. The Brainnetome atlas was used to divide the ALFF map into 246 regions of interest (ROIs) (Fan et al., 2016), and the average ALFF value of each region was obtained by averaging the ALFF value in this region (Li et al., 2012). Mean ALFF values from the 246 ROIs then served as the input vector to the classification procedure.

Feature Selection

Feature selection is necessary in MRI data analysis to avoid dimension disaster (Mladenić, 2006), reduce training time, and increase classification performance (Jiang et al., 2004;

[1] http://www.fil.ion.ucl.ac.uk/spm

[2] http://rfmri.org/DPARSF

[3] http://rest.restfmri.net

Dosenbach et al., 2010). A two-stage feature selection procedure was conducted, identifying features with the highest discriminative power. For the first-level selection, the paired sample *t*-test was performed between the region-wise post- and pre-training ALFF maps in a leave-one-out fashion. The combined region-wise features that survived the statistical threshold ($p < 0.05$) from each iteration were used as the input for a second-level feature elimination. Note that the remaining ALFF was regressed against the outcome of CFMT individually to remove the confounding effect from other domains of visual expertise, i.e., face in this study. Second, a recursive feature elimination-support vector machine (RFE-SVM) approach was used. This process recursively eliminates the least useful features until further elimination reduces the accuracy (Ding et al., 2015). The specific steps were as follows:

1. The training set was regressed against the outcome of CFMT.
2. The resulting beta-maps were normalized across all brain feature data between 0 and 1 through normalization of mean variance.
3. RFE reduced the dimension of features again and used the classifier itself to discard irrelevant features (**Figure 1**). Our implementation of RFE is described by the following pseudo-code:
    a. Input all training samples and class labels, train SVM classifier, calculate the classification accuracy of the model $accuracy_0$;
    b. Sequentially subtract one feature, inputting the other into LOOCV-SVM, calculating the classification $accuracy_i$ of the model, finding all $accuracy_i$ greater than or equal to $accuracy_0$, and determining the corresponding removed feature $feature_i$;
    c. Delete these features and update the feature set; and
    d. Repeat the above steps until further elimination reduces the accuracy.

As a result, we were able to identify a set of brain regions of the highest discriminative power.

Support Vector Machine

Basically, SVM is a binary classification model (Cortes and Vapnik, 1995). The basic idea is to find the separation hyperplane with the largest interval in the feature space to make the data binary classification efficiently (Li et al., 2007). Linear SVM is often used in neuroimaging data in that it produces interpretable results (Rasmussen et al., 2011). Therefore, this study adopted the linear SVM classifier model of soft interval separation and hinge loss function. LIBSVM toolbox[4] was used in this study (Chang and Lin, 2011).

The leave-one-out cross-validation (LOOCV) was used to assess the performance of the classifier (Dai et al., 2012). In LOOCV, each sample was designated as the test sample, while the remaining samples were used to train the multi-classifier. To quantify the performance of the classifier, according to the prediction results of LOOCV, the accuracy, sensitivity, and specificity were defined as follows:

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP} \quad (1)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (2)$$

$$Specificity = \frac{TN}{TN + FP} \quad (3)$$

where TP, FN, TN, and FP denoted the number of samples correctly predicted, the number of trained subjects classified as untrained ones, the number of untrained subjects correctly predicted, and the number of untrained subjects classified as trained ones, respectively. In this study, the area under the curve (AUC) was also used to represent the classification ability of SVM. A greater AUC value also represented a better classification ability.

Statistical Analysis

The non-parametric permutation test (Filgueiras et al., 2014) was used to evaluate the statistical significance of the classification results. The features with the highest discriminative power were used in this step, i.e., the 10 features after feature selection. Each subject was treated as an independent sample. For a given sample, the label was randomly set to 1/-1 (1: post-training data, -1: pre-training data), while the label of the testing sample remained unchanged to determine the outcome of SVM. The procedure was repeated 1,000 times. Accordingly, the statistical significance of the original accuracy was calculated as the probability that the SVM classification result was greater than or equal to the original accuracy in the 1,000 replacement. The average accuracy was obtained in all permutations, and the *p*-value was calculated as a proportion larger than the average accuracy obtained by our method. The threshold of $p < 0.05$ was used to determine the significance.

Regression Analysis

To assess the relationship between behavioral measurement and brain activity, Pearson's correlation analysis was conducted between alterations in outcomes of CFMT and RET and alterations in region-wise ALFF. The significance level was set at $p < 0.05$ after multiple comparison correction (false discovery rate, FDR).

RESULTS

Results of Behavioral Tasks

During 1 month of training, the subjects reviewed at least 831 cases (926 ± 73, mean ± *SD*). As shown in **Table 1**, after 1 month of training, the performance of the radiologist interns significantly improved as revealed by higher scores in RET ($p < 0.001$, Mann-Whitney *U*-test) and shorter response time

[4]https://ww2.mathworks.cn/matlabcentral/fileexchange/75567-svm-boundary-libsvm?s_tid=srchtitle_libsvm_4

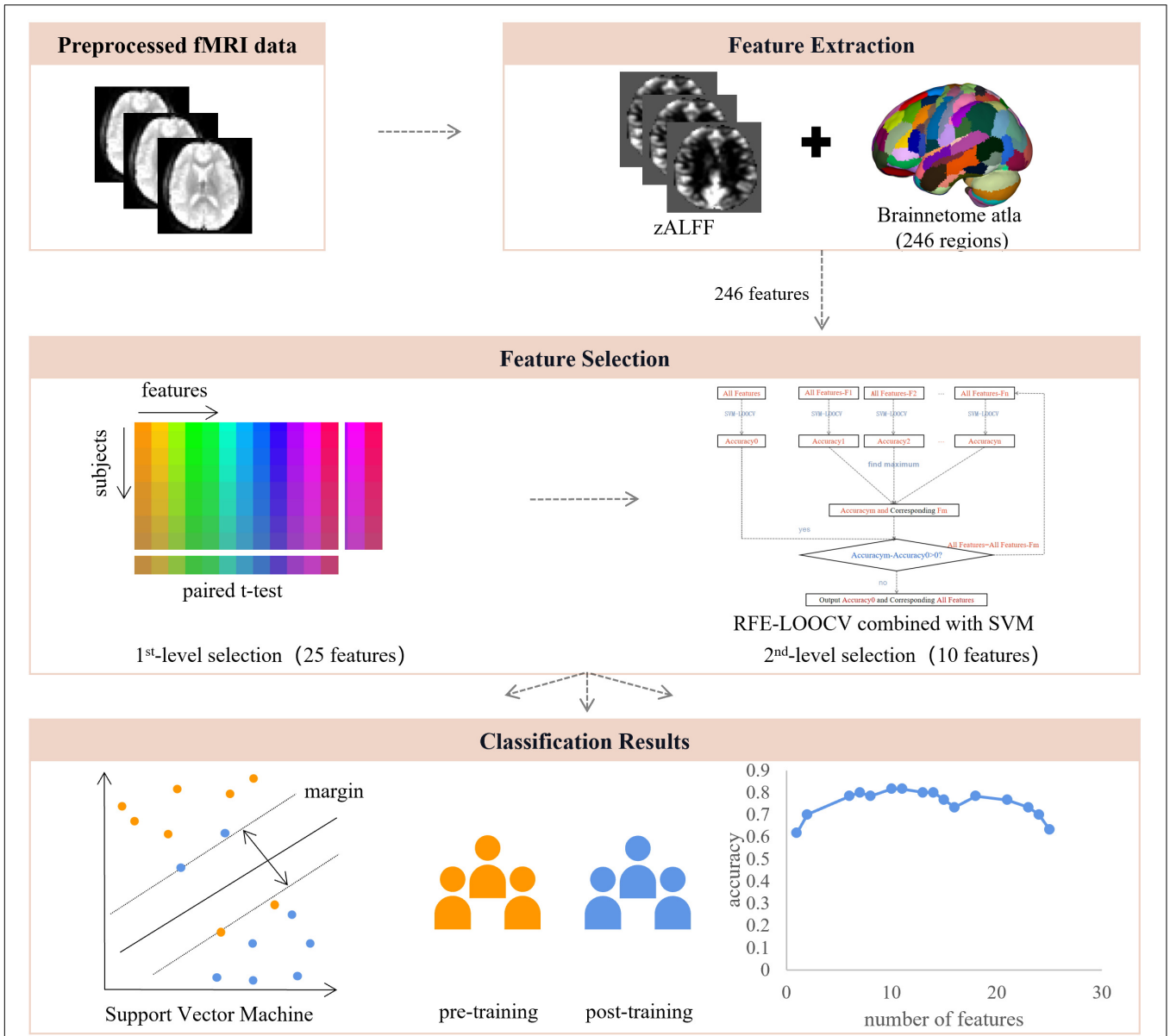


FIGURE 1 | The pipeline of data analysis. After the resting-state fMRI data were preprocessed, voxel-wise and region-wise amplitudes of low-frequency fluctuations were extracted and used for feature selection, which consisted of two steps, including region-wise paired *t*-test and recursive feature elimination embedded in a leave-one-out cross-validation framework, resulting in 10 features of highest discriminative power. These features were used for SVM modeling with LOOCV. ALFF, amplitude of low-frequency fluctuations; RFE, recursive feature elimination; LOOCV, leave-one-out cross-validation.

in RET ($p < 0.001$, Mann-Whitney *U*-test). Whereas the level of face expertise remained the same after 1 month of training in the domain of radiology ($p = 0.19$, Mann-Whitney *U*-test) (**Figure 2**).

Performance of Support Vector Machine

After feature selection, 10 features remained corresponding to the highest accuracy (**Figure 1**). The classification accuracy of SVM after LOOCV reached 86.7% (**Figure 3A**), and the AUC was 0.8244 (**Figure 3B**). The specificity and sensitivity of SVM after LOOCV were 80.00 and 83.33%, respectively. The classification results were tested 1,000 times, and no repetition reached the classification accuracy of 86.7%. Thus, the statistical significance was $p < 0.001$, indicating that the results of our study were significantly higher than the chance value.

As for the brain regions, 10 regions were identified with the highest discriminative power, including the left cingulate cortex (CG_L_7_4), the right cingulate cortex (CG_R_7_2), the left superior frontal gyrus (SFG_L_7_2), the right precentral gyrus (PrG_R_6_4), the left precentral gyrus (PrG_L_6_4), the right superior parietal lobule (SPL_R_5_4), the right superior parietal lobule (SPL_R_5_1), the left superior parietal lobule

TABLE 1 | The results of behavioral tasks within the subjects pre- and post-training.

	Pre-training (n = 30)		Post-training (n = 30)		p-values
	Mean	***SD***	**Mean**	***SD***	
Cases reviewed	N/A	N/A	926	73	–
RET	0.61	0.05	0.84	0.04	<0.001
*RT of RET(s)	3.08	0.30	2.53	0.34	<0.001
CFMT	56.90	4.29	57.30	4.67	0.19

*Note that the Mann-Whitney U-test was used to investigate group difference between groups. *Denotes the items showing significant difference between groups after Mann-Whitney U-test ($p < 0.001$).*
SD, standard deviation; s, seconds; RET, radiological expertise task; RT, response time; CFMT, Cambridge face memory test.

(SPL_L_5_4), the right precuneus (PCun_R_4_4), and the left precuneus (PCun_L_4_3) (**Figure 4** and **Table 2**).

Results of Regression Analysis

No significant correlations were found between alterations in the outcomes of CFMT and RET and alterations in region-wise ALFF after multiple comparisons.

DISCUSSION

The acquisition of visual expertise requires at least hundreds of cases of training within a specific domain (Annis and Palmeri, 2018). In real-world visual learning, several behavioral components, including high-order cognitive, such as memory (Viggiano et al., 2006), attention (Rose et al., 2004), and working memory (Ennaceur, 2010), and low-order visual factors, such as visual processing (Binder and Desai, 2011), are required. Existing neuroimaging studies demonstrated differentiated patterns of brain response in visual experts under tasks, which are modulated by their accumulated experience in a given domain. Resting-state spontaneous brain fluctuations actively encode previous learning experience. However, few studies have considered how real-world visual experience alters the level of baseline brain activity in the resting state. This study aimed to investigate how short-term real-world visual experience modulates baseline neuronal brain activity in the resting state using the amplitude of low frequency (<0.08 Hz) and a group of intern radiologists (n = 32). The resting-state fMRI data and the behavioral data regarding their level of visual expertise in radiology and face recognition were collected before and after 1 month of training in the X-ray department. A novel machine learning analytical method, i.e., recursive feature elimination SVM embedded in LOOCV, was used to identify subtle changes in the level of baseline brain activity (**Figure 1**). With a superb classification accuracy of 86.7% (**Figure 3A**), the results demonstrated that the left posterior cingulate cortex (CG_L_7_4), the right anterior cingulate cortex (CG_R_7_2), the left superior frontal gyrus (SFG_L_7_2), the bilateral precentral gyrus (PrG_L_6_4 and PrG_R_6_4), the bilateral superior parietal lobule (SPL_R_5_4, SPL_R_5_1, and SPL_L_5_4), the bilateral precuneus (PCun_R_4_4 and PCun_L_4_3) showed highest discriminative power after short-term visual learning (**Figure 4** and **Table 2**). To the best of our knowledge, this study is the first to investigate the baseline neurodynamic alterations in response to real-world visual experience using longitudinal experimental design. Our findings may help develop new insights into the neural mechanism of visual experts and provide new ideas for the cultivation of visual experts.

Increased Level of Activity in Brain Regions Supporting Working Memory

Working memory (WM) supports the online maintenance and manipulation of information without external stimulation (Baddeley, 1987). The capacity of WM serves as a reliable predictor for the performance of visual experts (Sohn and Doane, 2004). In this study, after training, the radiology interns had increased ALFF in the anterior cingulate gyrus, the posterior cingulate gyrus, and the superior frontal gyrus (**Figure 4** and **Table 2**). Jonides (2004) reported deactivation in the anterior cingulate gyrus, which supported increased WM load under task condition. Duan et al. (2012) found that the activation of posterior cingulate gyrus was enhanced in professional chess players in the game, which was related to enhanced requirement in the WM. Teresa et al. (2018) found increased activation in the superior frontal gyrus under the visual tasks, which required online monitoring and manipulation of task-related information. In sum, all these regions, i.e., the anterior cingulate gyrus, the posterior cingulate gyrus, and the superior frontal gyrus, are closely related to the WM process. The increased level of baseline brain activity in these regions might reflect tuning with training, which in turn decreases the need for executive control in the maintenance of task-relevant information. We propose that these alterations during expertise acquisition might support more automated encoding and maintenance of objects in their expert domain, indicating a more efficient mechanism subserving visual expertise.

Decreased Level of Activity in Brain Regions Underlying Memory Extraction

In our study, compared with the pre-training condition, the radiology interns had decreased level of ALFF in the bilateral precuneus (**Figure 4** and **Table 2**). Visual recognition intensively depends on the retrieval of conceptual knowledge (Binder and Desai, 2011). The difference in memory extraction predicts the performance difference between visual experts and novices (Binder and Desai, 2011). Assaf et al. (2013) reported the involvement of the right precuneus in memory extraction using the visual expertise model of car experts. While in the resting-state study, Duan et al. reported the reduction of default mode network activity, including left precuneus in the professional chess players, which is closely related to episodic memory extraction. In this study, the bilateral precuneus explicated decreased level of activity after short-term visual training. Given the fact that the resting-state brain activity is involved in the coding of expected sensory stimuli (Jin et al., 2017), we propose that the tuning in these regions is likely to reflect the optimal

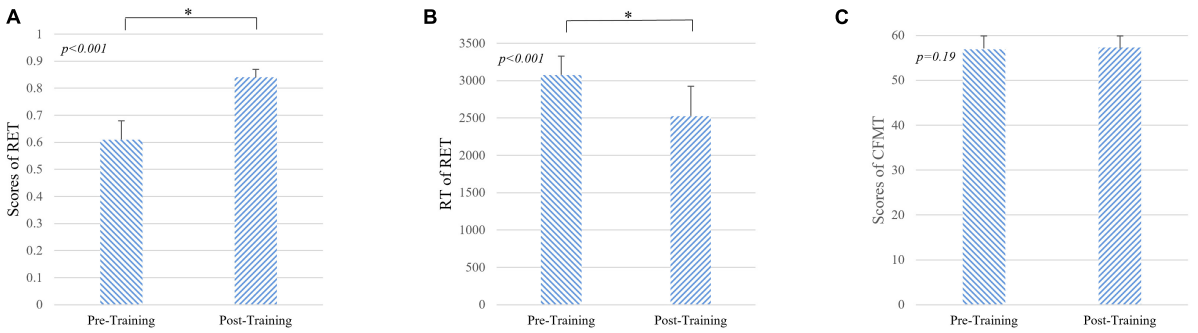


FIGURE 2 | Results of behavioral tasks pre- and post-training. **(A)** The level of radiological expertise assessed by the radiological expertise task. The radiology interns had a significantly greater scores after training compared with scores before training ($p < 0.001$, Mann-Whitney *U*-test), indicating improved performance in visual recognition of radiological images. **(B)** Response time of radiological expertise task pre- and post-training. The radiology interns had a significantly faster in behavioral response after training compared with that before training ($p < 0.001$, Mann-Whitney *U*-test). **(C)** The level of face expertise measured by the Cambridge face memory test. No significant differences were found ($p = 0.19$, Mann-Whitney *U*-test). RET, radiological expertise task; RT, response time; CMFT, Cambridge face memory test. Error bars indicate the standard deviation. * indicats the significant differences between groups ($p < 0.001$).

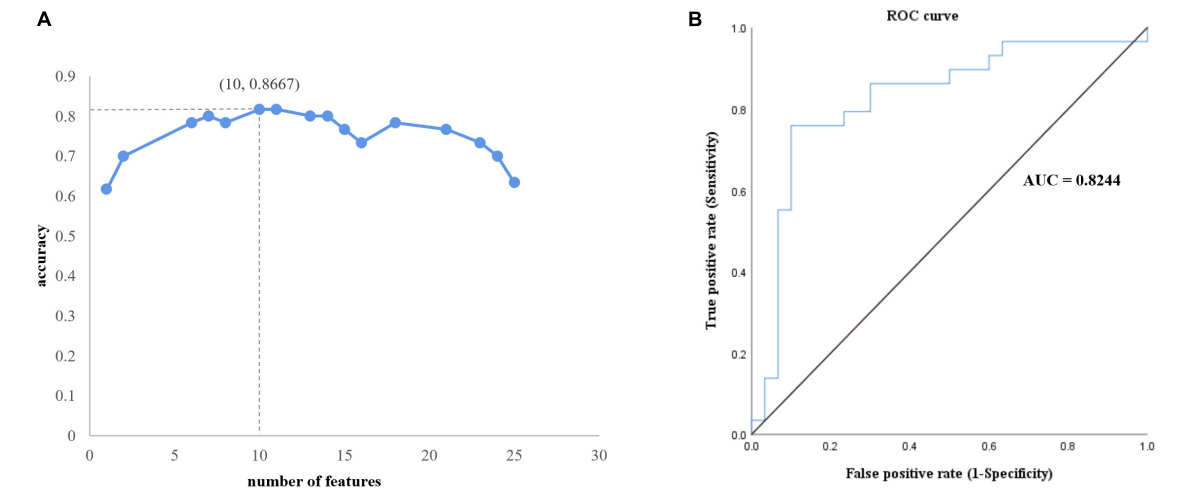


FIGURE 3 | Performance of the proposed analytical framework. **(A)** Ten features corresponding to best classification accuracy. **(B)** The receiver operating characteristic curve. The area under the curve is 0.8244, which indicates outstanding performance.

internal coping mechanism that supports the redistribution of cognitive resources into more demanding brain process (Fox et al., 2005).

Decrement in the Level of Activity in Brain Regions Underlying Attention Control

Visual attention is a critical component in visual recognition, which facilitates subjects to focus on the target objects in a more efficient way when dealing with complex visual scenes and gives priority to the target visual objects to ensure task completion (Cohen and Lefebvre, 2005). Therefore, the difference in the brain representation underlying attention control may serve to distinguish the brain states of experts and novices (Memmert et al., 2009). In this study, the radiology interns had decreased ALFF in the superior parietal lobule after training (**Figure 4** and **Table 2**). Reilhac et al. (2013) reported deactivation in the right superior parietal lobules, which was closely related to visual attention in radiologists. Ouellette et al. (2020) also found deactivation in the left SPL in radiologists, which was attributed to more efficient control of visual attention supported by accelerated eye-tracking data. We propose that decreased ALFF in SPL also reflects a similar trend. After visual training, the attention control is more efficient, which gives the subjects more flexibility in manipulating attentional resources, so that the resource allocated to attention before training might be allocated later to other brain regions supporting more demanding tasks.

Enhanced Level of Activity in Brain Regions Supporting Visual Recognition

In our study, the bilateral precentral (the PrG_L_6_4 and the PrG_R_6_4) showed enhanced ALFF after short-term

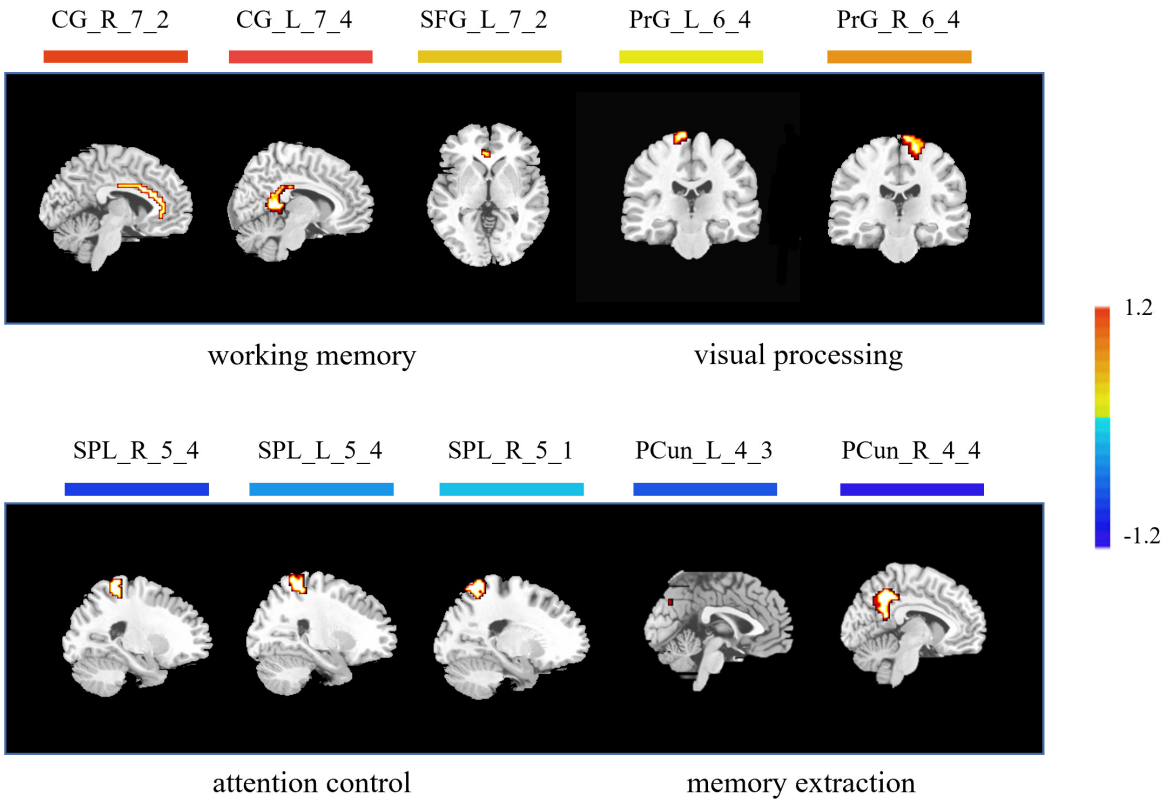


FIGURE 4 | Brain regions with highest discriminative power pre- and post-training. The color bar indicates the weight of the feature. Note that positive weights refer to higher level of ALFF after training, and negative weights refer to lower level of ALFF after training. CG, cingulate cortex; SFG, superior frontal gyrus; PrG, precentral gyrus; SPL, superior parietal lobule; PCun, precuneus; L, left; R, right.

TABLE 2 | Brain regions that show highest discriminative power pre- and post-training.

Cognitive component	Labels	Brain region	Brodmann area	Hemisphere	Weight
Working memory	CG_L_7_4	Posterior cingulate cortex	BA23	L	0.59
	CG_R_7_2	Anterior cingulate cortex	BA24	R	0.99
	SFG_L_7_2	Superior frontal gyrus	BA8	L	0.26
Memory	PCun_L_4_3	Precuneus	-	L	-0.85
	PCun_R_4_4	Precuneus	BA31	R	-1.02
Attention control	SPL_L_5_4	Postcentral area	BA7	L	-0.55
	SPL_R_5_4	Superior parietal lobule	BA7	R	-0.92
	SPL_R_5_1	Postcentral area	BA7	R	-0.54
Visual processing	PrG_L_6_4	Precentral gyrus	BA4	L	0.12
	PrG_R_6_4	Precentral gyrus	BA4	R	0.44

Note that positive weights refer to higher level of ALFF after training, and negative weights refer to lower level of ALFF after training.
L, left; R, right.

visual training (**Figure 4** and **Table 2**). Activations were found in the bilateral anterior central gyrus when visual stimuli were shown to subjects (Marks et al., 2019) and the level of brain activity increased with the number of stimuli (Mechelli et al., 2014). Studies using car experts reported an increase in gray matter volumes in this region (Gilaie-Dotan et al., 2012) and an increased level of evoked brain response to expertise-related visual stimuli in this region (Bentin, 2010). We suggest that our finding also reflects similar changes, but the exact nature of the alteration remained to be elucidated.

LIMITATIONS

Several issues should be mentioned when the findings from this study are considered. First, the sample size is not optimal. Given the longitudinal design and the COVID pandemic, the

current size is the best that can be achieved. Second, given the ratio between the number of discriminative features and the number of samples, this study faced an overfitting issue, which is quite common in MRI studies using a machine learning analytical framework. But it should be noted that three steps were taken to minimize the possibility of overfitting in our study. Particularly, a region-wise feature extraction strategy was used, which decreased the number of features from tens of thousands to 246. Then, a two-step feature selection was conducted, which decrease the number of features from 246 to 25. At last, an RFE-SVM analytical framework was employed to cut off the number of features to an optimal level, resulting in 10 features, i.e., 10 brain regions. Taken together, we do recommend further studies to repeat the current findings using larger samples. Third, for the behavioral tasks, only visual tasks were used. Tasks for WM, visual attention, and memory should be taken into consideration in future studies.

CONCLUSION

Our results suggest that real-world visual experience alters the resting-state brain representation in multidimensional neurobehavioral components, which are closely interrelated with high-order cognitive and low-order visual factors, i.e., attention control, WM, memory, and visual processing. We propose that our findings are likely to help foster new insights into the neural mechanisms of visual expertise.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of the First Affiliated Hospital of Xi'an Jiaotong University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

GS, CJ, and MD: conception and study design. HW, GS, and MD: data collection and acquisition. XZ and ZZ: statistical analysis. JS and GS: interpretation of results. JS, JW, and MD: drafting the manuscript work and revising it critically for important intellectual content. All authors approval of final version to be published and agreement to be accountable for the integrity and accuracy of all aspects of the work.

FUNDING

This manuscript was supported by the National Natural Science Foundation of China under Grant no. U19B2030, Science and Technology Projects of Xi'an, China (201809170CX11JC12), and Fundamental Research Funds for the Central Universities (No. JB191206).

ACKNOWLEDGMENTS

We would like to thank Dr. Karen M. von Deneen for her professional assistance in language editing. Also, professional and helpful advice from dedicated reviewers is truly appreciated.

REFERENCES

Annis, J., and Palmeri, T. J. (2018). Modeling Memory Dynamics in Visual Expertise. *J. Exp. Psychol. Learn Mem. Cogn.* 45, 1599–1618. doi: 10.1037/xlm0000664

Assaf, H., Dwight, K., and Baker, C. I. (2013). Beyond perceptual expertise: revisiting the neural substrates of expert object recognition. *Front. Hum. Neurosci.* 7:885. doi: 10.3389/fnhum.2013.00885

Baddeley, A. D. (1987). *Working Memory*. Oxford: Oxford University Press.

Bentin, S. (2010). Top-Down Engagement Modulates the Neural Expressions of Visual Expertise. *J. Vis.* 10, 945–945. doi: 10.1093/cercor/bhp316

Binder, J. R., and Desai, R. H. (2011). The neurobiology of semantic memory. *Trends Cogn. Sci.* 15, 527–536. doi: 10.1016/j.tics.2011.10.001

Boly, M., Balteau, E., Schnakers, C., Degueldre, C., Moonen, G., Luxen, A., et al. (2007). Baseline brain activity fluctuations predict somatosensory perception in humans. *Proc. Natl. Acad. Sci. U. S. A.* 104, 12187–12192.

Cavanna, A. E., and Trimble, M. R. (2006). The precuneus: a review of its functional anatomy and behavioural correlates. *Brain* 129, 564–583. doi: 10.1093/brain/awl004

Chakraborty, D. P. (2006). Analysis of location specific observer performance data: validated extensions of the jackknife free-response (JAFROC) method. *Acad. Radiol.* 13, 1187–1193. doi: 10.1016/j.acra.2006.06.016

Chang, C.-C., and Lin, C.-J. (2011). LIBSVM: a library for support vector machines. *ACM Trans. Intell. Syst. Technol.* 2, 1–27. doi: 10.1145/1961189.1961199

Clark, K., Cain, M. S., Adamo, S. H., and Mitroff, S. R. (2012). Overcoming hurdles in translating visual search research between the lab and the field. *Nebr. Symp. Motiv.* 59, 147–181. doi: 10.1007/978-1-4614-4794-8_7

Cohen, H., and Lefebvre, C. (2005). *Handbook of Categorization in Cognitive Science*. Netherland: Elsevier Science.

Cortes, C., and Vapnik, V. (1995). Support-vector networks. *Mach. Learn.* 20, 273–297.

Curby, K. M., and Gauthier, I. (2010). To the trained eye: perceptual expertise alters visual processing. *Top. Cogn. Sci.* 2, 189–201. doi: 10.1111/j.1756-8765.2009.01058.x

Dai, Z., Yan, C., Wang, Z., Wang, J., Xia, M., Li, K., et al. (2012). Discriminative analysis of early Alzheimer's disease using multi-modal imaging and multi-level characterization with multi-classifier (M3). *Neuroimage* 59, 2187–2195. doi: 10.1016/j.neuroimage.2011.10.003

Dale, A. M., Liu, A. K., Fischl, B. R., Buckner, R. L., Belliveau, J. W., Lewine, J. D., et al. (2000). Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity. *Neuron* 26, 55–67. doi: 10.1016/s0896-6273(00)81138-1

Di, X., Kannurpatti, S. S., Rypma, B., and Biswal, B. B. (2013). Calibrating BOLD fMRI activations with neurovascular and anatomical constraints. *Cereb. Cortex* 23, 255–263. doi: 10.1093/cercor/bhs001

Ding, X., Yang, Y., Stein, E. A., and Ross, T. J. (2015). Multivariate classification of smokers and nonsmokers using SVM-RFE on structural MRI images. *Hum. Brain Mapp.* 36, 4869–4879. doi: 10.1002/hbm.22956

Dong, M., Li, J., Shi, X., Gao, S., Fu, S., Liu, Z., et al. (2015). Altered baseline brain activity in experts measured by amplitude of low frequency fluctuations (ALFF): a resting state fMRI study using expertise model of acupuncturists. *Front. Hum. Neurosci.* 9:99. doi: 10.3389/fnhum.2015.00099

Dong, M., Qin, W., Zhao, L., Yang, X., Yuan, K., Zeng, F., et al. (2014). Expertise modulates local regional homogeneity of spontaneous brain activity in the resting brain: an fMRI study using the model of skilled acupuncturists. *Hum. Brain Mapp.* 35, 1074–1084. doi: 10.1002/hbm.22235

Dosenbach, N. U., Nardos, B., Cohen, A. L., Fair, D. A., Power, J. D., Church, J. A., et al. (2010). Prediction of individual brain maturity using fMRI. *Science* 329, 1358–1361. doi: 10.1126/science.1194144

Duan, X., He, S., Liao, W., Liang, D., Qiu, L., Wei, L., et al. (2012). Reduced caudate volume and enhanced striatal-DMN integration in chess experts. *NeuroImage* 60, 1280–1286. doi: 10.1016/j.neuroimage.2012.01.047

Duchaine, B., and Nakayama, K. (2006). The Cambridge Face Memory Test: results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. *Neuropsychologia* 44, 576–585. doi: 10.1016/j.neuropsychologia.2005.07.001

Duff, E. P., Johnston, L. A., Xiong, J., Fox, P. T., Mareels, I., and Egan, G. F. (2008). The power of spectral density analysis for mapping endogenous BOLD signal fluctuations. *Hum. Brain Mapp.* 29, 778–790. doi: 10.1002/hbm.20601

Ennaceur, A. (2010). One-trial object recognition in rats and mice: methodological and theoretical issues ScienceDirect. *Behav. Brain Res.* 215, 244–254. doi: 10.1016/j.bbr.2009.12.036

Evans, K. K., Cohen, M. A., Tambouret, R., Horowitz, T., Kreindel, E., and Wolfe, J. M. (2011). Does visual expertise improve visual recognition memory? *Atten. Percept. Psychophys.* 73, 30–35. doi: 10.3758/s13414-010-0022-5

Fan, L., Li, H., Zhuo, J., Zhang, Y., Wang, J., Chen, L., et al. (2016). The human brainnetome atlas: a new brain atlas based on connectional architecture. *Cereb. Cortex* 26, 3508–3526. doi: 10.1093/cercor/bhw157

Filgueiras, P. R., Alves, J., Sad, C., Castro, E., Dias, J., and Poppi, R. J. (2014). Evaluation of trends in residuals of multivariate calibration models by permutation test. *Chemom. Intell. Lab. Syst.* 133, 33–41. doi: 10.1016/j.chemolab.2014.02.002

Fox, M. D., Snyder, A. Z., and Vincent, J. L. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional network. *Proc. Natl. Acad. Sci. U. S. A.* 102, 9673–9678. doi: 10.1073/pnas.0504136102

Gilaie-Dotan, S., Harel, A., Bentin, S., Kanai, R., and Rees, G. (2012). Neuroanatomical correlates of visual car expertise. *NeuroImage* 62, 147–153. doi: 10.1016/j.neuroimage.2012.05.017

Haller, S., and Radue, E. W. (2005). What is different about a radiologist's brain? *Radiology* 236, 983–989. doi: 10.1148/radiol.2363041370

Harel, A., Gilaie-Dotan, S., Malach, R., and Bentin, S. (2010). Top-down engagement modulates the neural expressions of visual expertise. *Cereb. Cortex* 20, 2304–2318.

Hasler, B. P., Forbes, E. E., and Franzen, P. L. (2014). Time-of-day differences and short-term stability of the neural response to monetary reward: a pilot study. Psychiatry Research. *Neuroimaging* 224, 22–27. doi: 10.1016/j.pscychresns.2014.07.005

Humphreys, G. W., Price, C. J., and Riddoch, M. J. (1999). From objects to names: a cognitive neuroscience approach. *Psychol. Res.* 62, 118–130. doi: 10.1007/s004260050046

Jia, X.-Z., Sun, J.-W., Ji, G.-J., Liao, W., Lv, Y.-T., Wang, J., et al. (2020). Percent amplitude of fluctuation: a simple measure for resting-state fMRI signal at single voxel level. *PLoS One* 15:e0227021. doi: 10.1371/journal.pone.0227021

Jiang, T., He, Y., Zang, Y., and Weng, X. (2004). Modulation of functional connectivity during the resting state and the motor task. *Hum. Brain Mapp.* 22, 63–71. doi: 10.1002/hbm.20012

Jin, C., Guan, M., Dong, M., Wu, J., He, Z., Chen, X., et al. (2017). Aberrant baseline brain activity in psychogenic erectile dysfunction patients: a resting state fMRI study. *Brain. Imag. Behav.* 12, 1393–1404. doi: 10.1007/s11682-017-9805-9

Jonides, J. (2004). How does practice make perfect? *Nat. Neurosci.* 7, 10–11.

Khader, P., Burke, M., Bien, S., Ranganath, C., and Rösler, F. (2005). Content-specific activation during associative long-term memory retrieval. *Neuroimage* 27, 805–816. doi: 10.1016/j.neuroimage.2005.05.006

Li, H., Gu, H., Han, Y., and Yang, J. (2007). "Object-oriented classification of very high-resolution remote sensing imagery based on improved CSC and SVM," in *Geoinformatics 2007*, (Nanjing: Remotely Sensed Data and Information), 67523I.

Li, Z., Zhu, Y., Childress, A. R., Detre, J. A., and Wang, Z. (2012). Relations between BOLD fMRI-derived resting brain activity and cerebral blood flow. *PLoS One* 7:e44556. doi: 10.1371/journal.pone.0044556

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., and Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature* 412, 150–157. doi: 10.1038/35084005

Marks, R. A., Kovelman, I., Kepinska, O., Oliver, M., Xia, Z., Haft, S. L., et al. (2019). Spoken language proficiency predicts print-speech convergence in beginning readers. *Neuroimage* 201:116021. doi: 10.1016/j.neuroimage.2019.116021

Mechelli, A., Friston, K. J., and Price, C. J. (2014). The effects of presentation rate during word and pseudoword reading: a comparison of PET and fMRI.[J]. *J. Cogn. Neurosci.* 12, 145–156. doi: 10.1162/089892900564000

Memmert, D., Simons, D. J., and Grimme, T. (2009). The relationship between visual attention and expertise in sports[J]. *Psychol. Sport Exerc.* 10, 146–151. doi: 10.1016/j.psychsport.2008.06.002

Mladenić, D. (2006). "Feature selection for dimensionality reduction," in *Subspace, Latent Structure and Feature Selection. SLSFS 2005. Lecture Notes in Computer Science*, Vol. 3940, eds C. Saunders, M. Grobelnik, S. Gunn, and J. Shawe-Taylor (Berlin: Springer). doi: 10.1007/11752790_5

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4

Ouellette, D. J., Staalduinen, E. V., Hussaini, S. H., Govindarajan, S. T., Stefancin, P., Hsu, D. L., et al. (2020). Functional, anatomical and diffusion tensor MRI study of radiology expertise[J]. *PLoS One* 15:e0231900. doi: 10.1371/journal.pone.0231900

Rasmussen, P. M., Madsen, K. H., Lund, T. E., and Hansen, L. K. (2011). Visualization of nonlinear kernel models in neuroimaging by sensitivity maps. *NeuroImage* 55, 1120–1131. doi: 10.1016/j.neuroimage.2010.12.035

Reilhac, C., Peyrin, C., and Démonet, J. F. (2013). Role of the superior parietal lobules in letter-identity processing within strings: FMRI evidence from skilled and dyslexic readers.[J]. *Neuropsychologia* 51, 601–612. doi: 10.1016/j.neuropsychologia.2012.12.010

Rose, S. A., Feldman, J. F., and Jankowski, J. J. (2004). Infant visual recognition memory. *Dev. Rev.* 24, 74–100. doi: 10.1016/j.dr.2003.09.004

Rossignoli-Palomeque, T., Perez-Hernandez, E., and Gonzalez-Marques, J. (2018). Brain training in children and adolescents: is it scientifically valid? *Front. Psychol.* 9:565. doi: 10.3389/fpsyg.2018.00565

Schipul, S. E., and Just, M. A. (2016). Diminished neural adaptation during implicit learning in autism. *NeuroImage* 125, 332–341. doi: 10.1016/j.neuroimage.2015.10.039

Sohn, Y. W., and Doane, S. M. (2004). Memory processes of flight situation awareness: interactive roles of working memory capacity, long-term working memory, and expertise. *Hum. Factors* 46:461. doi: 10.1518/hfes.46.3.461.50392

Song, L., Peng, Q., Liu, S., and Wang, J. (2020). Changed hub and functional connectivity patterns of the posterior fusiform gyrus in chess experts. *Brain Imaging Behav.* 14, 797–805. doi: 10.1007/s11682-018-0020-0

Song, L., Yang, H., Yang, M., Liu, D., Ge, Y., Long, J., et al. (2021). Professional chess expertise modulates whole brain functional connectivity pattern homogeneity and couplings. *Brain Imaging Behav.* 16, 587–595. doi: 10.1007/s11682-021-00537-1

Song, X.-W., Dong, Z.-Y., Long, X.-Y., Li, S.-F., Zuo, X.-N., Zhu, C.-Z., et al. (2011). REST: a toolkit for resting-state functional magnetic resonance imaging data processing. *PLoS One* 6:e25031. doi: 10.1371/journal.pone.0025031

Tang, Z., Varghese, R. S., Bekesova, S., Loffredo, C. A., Hamid, M. A., Kyselova, Z., et al. (2010). Identification of N-glycan serum markers associated with hepatocellular carcinoma from mass spectrometry data. *J. Proteome Res.* 9, 104–112. doi: 10.1021/pr900397n

Teresa, R. P., Elena, P. H., and Javier, G.-M. (2018). Brain Training in Children and Adolescents: Is It Scientifically Valid? *Front. Psychol.* 9:565.

Viggiano, M. P., Righi, S., and Galli, G. (2006). Category-specific visual recognition as affected by aging and expertise. *Arch. Gerontol. Geriatr.* 42, 329–338. doi: 10.1016/j.archger.2005.08.003

Wang, Y., Jin, C., Yin, Z., Wang, H., Ji, M., Dong, M., et al. (2021). Visual experience modulates whole-brain connectivity dynamics: a resting-state fMRI study using

the model of radiologists. *Hum. Brain Mapp.* 42, 4538–4554. doi: 10.1002/hbm.25563

Xu, X., Zhang, X., Tian, Q., Wang, H., Cui, L. B., Li, S., et al. (2019). Quantitative identification of nonmuscle-invasive and muscle-invasive bladder carcinomas: a multiparametric MRI radiomics analysis. *J. Magn. Reson. Imaging* 49, 1489–1498. doi: 10.1002/jmri.26327

Yang, H., Long, X.-Y., Yang, Y., Yan, H., Zhu, C.-Z., Zhou, X.-P., et al. (2007). Amplitude of low frequency fluctuation within visual areas revealed by resting-state functional MRI. *Neuroimage* 36, 144–152. doi: 10.1016/j.neuroimage.2007.01.054

Zhang, T., Dong, M., Wang, H., Jia, R., Li, F., Ni, X., et al. (2022). Visual expertise modulates baseline brain activity: a preliminary resting-state fMRI study using expertise model of radiologists. *BMC Neurosci.* 23:24. doi: 10.1186/s12868-022-00707-x

Zou, Q.-H., Zhu, C.-Z., Yang, Y., Zuo, X.-N., Long, X.-Y., Cao, Q.-J., et al. (2008). An improved approach to detection of amplitude of low-frequency fluctuation (ALFF) for resting-state fMRI: fractional ALFF. *J. Neurosci. Methods* 172, 137–141. doi: 10.1016/j.jneumeth.2008.04.012

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Neuroscience

ORIGINAL RESEARCH
published: 25 May 2022
doi: 10.3389/fnins.2022.832276

Decomposition-Based Correlation Learning for Multi-Modal MRI-Based Classification of Neuropsychiatric Disorders

*Liangliang Liu[1], Jing Chang[1], Ying Wang[1], Gongbo Liang[2], Yu-Ping Wang[3] and Hui Zhang[1]**

[1] College of Information and Management Science, Henan Agricultural University, Zhengzhou, China, [2] Department of Computer Science, Eastern Kentucky University, Richmond, KY, United States, [3] Biomedical Engineering Department, Tulane University, New Orleans, LA, United States

Multi-modal magnetic resonance imaging (MRI) is widely used for diagnosing brain disease in clinical practice. However, the high-dimensionality of MRI images is challenging when training a convolution neural network. In addition, utilizing multiple MRI modalities jointly is even more challenging. We developed a method using decomposition-based correlation learning (DCL). To overcome the above challenges, we used a strategy to capture the complex relationship between structural MRI and functional MRI data. Under the guidance of matrix decomposition, DCL takes into account the spike magnitude of leading eigenvalues, the number of samples, and the dimensionality of the matrix. A canonical correlation analysis (CCA) was used to analyze the correlation and construct matrices. We evaluated DCL in the classification of multiple neuropsychiatric disorders listed in the Consortium for Neuropsychiatric Phenomics (CNP) dataset. In experiments, our method had a higher accuracy than several existing methods. Moreover, we found interesting feature connections from brain matrices based on DCL that can differentiate disease and normal cases and different subtypes of the disease. Furthermore, we extended experiments on a large sample size dataset and a small sample size dataset, compared with several other well-established methods that were designed for the multi neuropsychiatric disorder classification; our proposed method achieved state-of-the-art performance on all three datasets.

Keywords: multi-modal, decomposition-based, matrix decomposition, canonical correlation analysis, neuropsychiatric disorders

OPEN ACCESS

Edited by:
Feng Liu,
Tianjin Medical University General Hospital, China

Reviewed by:
Tuo Zhang,
Northwestern Polytechnical University, China
Junhai Xu,
Tianjin University, China
Xiaoke Hao,
Hebei University of Technology, China

***Correspondence:**
Hui Zhang
zhhnau@163.com

Specialty section:
This article was submitted to Brain Imaging Methods, a section of the journal Frontiers in Neuroscience

Received: *09 December 2021*
Accepted: *21 April 2022*
Published: *25 May 2022*

Citation:
Liu L, Chang J, Wang Y, Liang G, Wang Y-P and Zhang H (2022) Decomposition-Based Correlation Learning for Multi-Modal MRI-Based Classification of Neuropsychiatric Disorders. Front. Neurosci. 16:832276. doi: 10.3389/fnins.2022.832276

1. INTRODUCTION

Many neuropsychiatric disorders (NDs) not only result in a huge socioeconomic burden but are also accompanied by several comorbidities (Kessler et al., 2012). Although NDs arise from physical defects or injuries, they are usually considered a chronic course of mental disease, resulting in the collapse of an understanding of the real world, cognitive problems, and persistent damage (Heinrichs and Zakzanis, 1998). Diagnosis of NDs is important for tracking the development of the disease and for choosing and evaluating the effects of an intervention such as drug treatment. Furthermore, subtyping an ND can help in personalizing treatment. As a result, increasing

attention has been paid to the identification of the subtypes of the ND, such as schizophrenia (SZ), bipolar disorder (BD), and attention deficit hyperactivity disorder (ADHD). However, it is difficult to distinguish these subtypes due to a lack of standard clinical criteria (McIntosh et al., 2005; Strasser et al., 2005; Finn et al., 2015; Liu Z. et al., 2018; Hu et al., 2019; Lake et al., 2019; Jiang et al., 2020).

Multi-modal magnetic resonance imaging (MRI) is a useful tool for clinical diagnosis of ND. It can provide information on different aspects of the brain. Functional MRI (fMRI) can be used to analyze the functional connections (FCs) between different brain regions. These FCs reveal individual differences in neural activity patterns, which can predict continuous phenotypic measurements (Dubois and Adolphs, 2016; Rosenberg et al., 2018; Hu et al., 2021). On the other hand, structural MRI (sMRI) reflects the location, volume, and lesions of brain tissue (McIntosh et al., 2005; Liu et al., 2019), in addition to providing information about structural connections among brain regions (Wang et al., 2009). A number of MRI studies have been conducted on ND classification, including Alzheimer's disease (Fan et al., 2020), ADHD (Connaughton et al., 2022), SZ (de Filippis et al., 2019), BD (Madeira et al., 2020), depression (Han et al., 2019), and autism (Rakić et al., 2020). However, most of these studies focus only on one type of MRI image or one type of ND. They overlook complementary information, resulting in lower classification accuracy.

Compared to natural image studies, the limited number of medical MRI samples is a challenge for the state-of-the-art convolutional neural networks and graph convolutional networks (Yu et al., 2019; Willemink et al., 2020). In particular, the high-dimensionality of MRI and nonlinear relations between the matrices of MRIs pose challenges for these machine learning methods. In addition, the imaging principles of sMRI and fMRI are different, and there is no direct correlation between them. Exploring the relationship between them is itself challenging.

Previous multi-modal MRI studies have demonstrated the potential of a multi-modal fusion approach in studying the relationship between fMRI and sMRI images (Qiao et al., 2019; Gao et al., 2020; Jiang et al., 2021; Mill et al., 2021). For example, Qiao et al. (2019) proposed a hybrid feature selection method based on statistical approaches and machine learning. This method explored the brain abnormalities in SZ using both fMRI and sMRI images. A multi-kernel support vector machine (SVM) was used for SZ classification, which was based on the similarity of the decomposed components from multi-modal MRI (Gao et al., 2020). Jiang et al. (2021) combined the multi-dimensional features of sMRI and fMRI to predict the state of SZ and guide medication. Different modalities contain complementary information, which can improve the performance of the model (Jiang et al., 2021; Mill et al., 2021). However, the poor interpretability of some models has become an issue when identifying significant biomarkers (Olesen et al., 2003; Seghier et al., 2004). Various strategies are widely used in multi-modal data analysis, including multi-modal canonical correlation analysis (CCA) (Correa et al., 2010), deep collaborative learning (Hu et al., 2019), parallel independent component analysis (Liu et al., 2008), and methods similar to independent component analysis (Sui et al., 2009; Calhoun et al., 2010; Groves et al., 2011).

Some previous studies have identified a correlation between fMRI and sMRI images in ND groups (Sui et al., 2011; Qiao et al., 2019; Su et al., 2020). Therefore, we propose a prediction method, called decomposition-based correlation learning (DCL), for the multi-modal MRI-based classification of NDs. We first used the shrinkage principal orthogonal complement thresholding method (S-POET) (Fan and Wang, 2015) to estimate spiked fMRI and sMRI matrices. Subsequently, in the DCL method, we use decomposition-based CCA to decompose each pair of matrices into two common matrices and two orthogonal distinctive matrices. Finally, we computed the correlation between the common matrices and the distinctive matrices. We validated the DCL method on the Consortium for Neuropsychiatric Phenomics (CNP) dataset. Our results demonstrate that the proposed DCL model outperforms several other methods. We also discovered interesting feature connections when identifying significant features in fMRI data.

The rest of this paper is organized as follows. Section 2 describes the DCL pipeline and provides a quantitative evaluation of our method. The dataset and experiments in applying DCL to NDs are presented in Sections 3, 4. A discussion and analysis of the results are in Section 5. Section 6 concludes this paper.

2. METHODOLOGY

The DCL pipeline is shown in **Figure 1**. DCL has three steps: data processing (feature extraction), S-POET (spiked covariance matrix estimation), and CCA (canonical correlation and matrix construction).

2.1. Overview of Principal Component Analysis (PCA)

Principal component analysis is a powerful tool for feature extraction and data visualization. PCA can extract principal components from multivariate data by maximizing the variance of the features while minimizing the reconstruction error.

Let $X \in \mathbb{R}^{m \times n}$ be a matrix, where m and n are the size of the matrix. Hence,

$$X = [x_1, x_2, x_3, \ldots, x_m]. \tag{1}$$

Let $\widehat{X}$ be the average signal, which is defined as follows:

$$\widehat{X} = \frac{1}{m} \sum_{n=1}^{m} x_n. \tag{2}$$

The normalized vectors are computed by subtracting the average signal from each training vector. They are defined as follows:

$$\phi_i = x_i - \widehat{X}. \tag{3}$$

These vectors go through PCA. Let C be a covariance matrix:

$$C = \frac{1}{m} \sum_{n=1}^{m} \phi_n \phi_i^{\top}. \tag{4}$$

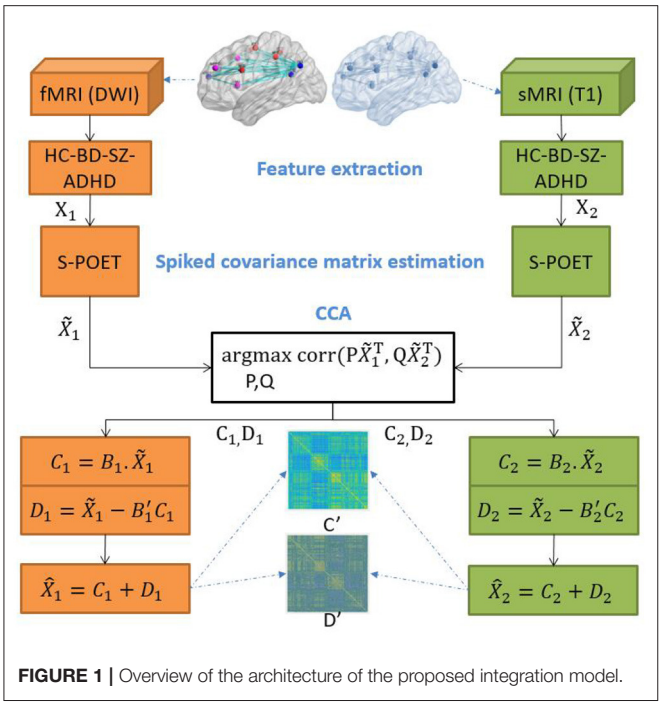


FIGURE 1 | Overview of the architecture of the proposed integration model.

2.2. Overview of S-POET

The shrinkage principal orthogonal complement thresholding method (Fan and Wang, 2015) is a covariance estimator with an approximate factor model. It is based on sparse PCA. Feature matrices from fMRI and sMRI data are input into S-POET, which calculates an asymptotic first-order distribution for the eigenvalues and eigenvectors of the sample correlation matrices.

Specifically, let k be the number of datasets and n be the number of samples in the k-th dataset. A high-dimensional dataset can be written as matrix $\widetilde{X} \in \mathbb{R}^{p_k \times n}$. In our experiment, we have two matrices, one from fMRI and one from sMRI, so we set $k = 2$. p_k is a row, which corresponds to a mean-zero variable. S-POET constructs $\widetilde{X}_k$, which is the estimate of matrix X_k. Before defining $\widetilde{X}_k$, we let the full singular value decomposition of Y_k be as follows:

$$Y_k = V_{k1}\lambda_{yk}V_{k2}^{\top}, \tag{5}$$

where V_{k1} and V_{k2} are two orthogonal matrices. λ_{yk} is a rectangular diagonal matrix whose singular values on the main diagonal are arranged in descending order. $\widetilde{X}_k$ is a matrix:

$$\widetilde{X}_k = V_{k1}^{[:,1:r_k]}\,\mathrm{diag}(\widehat{\sigma}_1^S(Y_k),\ldots,\widehat{\sigma}_{r_k}^S(Y_k))(V_{k2}^{[:,1:r_k]})^{\top}, \tag{6}$$

$$\widehat{\sigma}_l^S(Y_k) = \sqrt{\max\{\sigma_l^2(Y_k) - \tau_k p_k, 0\}}, \tag{7}$$

$$\tau_k = \sum_{l=r_k+1}^{p_k} \sigma_l^2(Y_k)/(np_k - nr_k - p_k r_k), \tag{8}$$

where $\widetilde{r}_k = \mathrm{rank}(\widetilde{X}_k)$ and $\widetilde{r}_k = r_k$.

We summarize the S-POET method in **Algorithm 1**.

Algorithm 1 | S-POET

Input: $\boldsymbol{X} \in \mathbb{R}^{p_k \times n}$
Output: $\widetilde{\boldsymbol{X}}_k$

$K \leftarrow \mathrm{rank\,cov}(\boldsymbol{X})$ *//Covariance estimator*
$p, n \leftarrow \mathrm{shape}(\boldsymbol{X})$
$V, S, U_t \leftarrow \mathrm{SVD}(\boldsymbol{X}, full_{matrices} = False)$
$S \leftarrow \mathrm{diag}(S)$
$Lambda \leftarrow S * *2/n$ *//lambda expression*
$\widetilde{c} \leftarrow Sum(Lambda.diagonal()[K:])/(p - K - p * K/n)$
$Lambda_s \leftarrow Maximum(Lambda[:K, :K] - \widetilde{c} * p/n, 0)$
$\widetilde{\boldsymbol{X}}_k \leftarrow V[:, :K]@\,\mathrm{Sqrt}(Lambda_s * n)@U_t[:K, :]$
return $\widetilde{\boldsymbol{X}}_k$, $Lambda_s$, $V[:, :K]$, K

2.3. Overview of CCA

Canonical correlation analysis is a multivariate statistical analysis method. It determines the overall correlation between two groups of indicators. We use CCA to examine the cross-covariances of multi-modal MRI data.

Let $\widetilde{X}_1 \in \mathbb{R}^{n \times r}$ and $\widetilde{X}_2 \in \mathbb{R}^{n \times s}$ be two matrices, where n is the number of samples, and r and s are the feature sizes of the two matrices, respectively. CCA is used to find two coefficient vectors $v_1 \in \mathbb{R}^{r \times 1}$ and $v_2 \in \mathbb{R}^{s \times 1}$ by optimizing the Pearson correlation between $\widetilde{X}_1 v_1$ and $\widetilde{X}_2 v_2$, which is defined as follows:

$$(v_1^*, v_2^*) = \underset{v_1, v_2}{\arg\max}\, v_1' \Phi_{12} v_2, \tag{9}$$

where $v_1' \Phi_{11} v_1 = 1$, $v_2' \Phi_{22} v_2 = 1$, $v_1 \in \mathbb{R}^{r \times 1}$, $v_2 \in \mathbb{R}^{s \times 1}$, and $\Phi_{ij} = \widetilde{X}_i' \widetilde{X}_j$. $\widetilde{X}_1 v_1$ and $\widetilde{X}_2 v_2$ are two identified canonical vectors, both of which are linear combinations of raw features in the original data, $\widetilde{X}_1$ and $\widetilde{X}_2$, respectively. $\widetilde{X}_1 v_1$ and $\widetilde{X}_2 v_2$ facilitate the interpretation of multi-omics associations by reducing the dimensionality ($\widetilde{X}_1 v_1, \widetilde{X}_2 v_2 \in \mathbb{R}^{n \times 1}$). We use Equation (9) as a constraint, and $v_1' \Phi_{12} v_2$ can be used as the cross-data correlation, i.e.,

$$v_1' \Phi_{12} v_2 = \frac{v_1' \Phi_{12} v_2}{\sqrt{v_1' \Phi_{11} v_1 v_2' \Phi_{22} v_2}}.$$

Canonical correlation analysis is used to guarantee the highest total correlation of the pair-wise independent canonical vectors, which is defined as follows:

$$(V_1^*, V_2^*) = \underset{V_1, V_2}{\arg\max}\, \mathrm{trace}(V_1' \Phi_{12} V_2), \tag{10}$$

where $V_1' \Phi_{11} V_1 = V_2' \Phi_{22} V_2 = I_n$, $V_1 \in \mathbb{R}^{r \times k}$, $V_2 \in \mathbb{R}^{s \times k}$, and $k = \min[\mathrm{rank}(\widetilde{X}_1), \mathrm{rank}(\widetilde{X}_2)]$. Since Φ_{11} and Φ_{22} may be singular when calculating the loading vectors, matrix regularization is usually enforced on them to ensure that they are positive definite:

$$\begin{aligned} \widehat{\Phi_{11}} &= \Phi_{11} + r_1 I_r, \\ \widehat{\Phi_{22}} &= \Phi_{22} + r_2 I_r. \end{aligned} \tag{11}$$

2.4. Decomposition-Based Correlation Learning

Let X_1 and X_2 be paired matrices of fMRI and sMRI, which are the input of S-POET methods. We use the DCL method to decompose this pair of matrices into two common matrices and two orthogonal distinctive matrices. Then, we collect these two types of matrices into a common matrix (C_k) and a distinctive matrix (D_k), respectively. Based on the output ($\widetilde{X}_k$) of S-POET, we use $\widetilde{X}_k$ to develop two estimators for C_k and D_k. First, we define the common variable c_{base} as follows:

$$c_{base} \propto n^{-1} \underset{w\in(l_0^2,cov)}{\arg\max}\{\mathrm{corr}^2(X_1, w) + \mathrm{corr}^2(X_2, w)\}, \tag{12}$$

where the constraints $X_1 = C_1 + D_1$, $X_2 = C_2 + D_2$, $\mathrm{corr}(D_1, D_2) = 0$, and $c_{base} \in [0, 1]$.

Then, the estimator of C_k can be defined as follows:

$$\widehat{C}_k = n^{-1}\widetilde{X}_k(\widehat{V}_k^{[1:r_{12},:]})^\top \widehat{A}_C^{(r_{12})} \sum_{j=1}^{2} \widehat{V}_k^{[1:r_{12},:]} c_{base}, \tag{13}$$

where $\widehat{A}_C^{(r)} = \mathrm{diag}(\widehat{a}_1, \ldots \widehat{a}_r)$, C_1 and C_2 have the maximum correlation between each other, while the vectors within each are uncorrelated and whitened. Their correlation vectors $\widehat{a}_1$, $\widehat{a}_2$,..., $\widehat{a}_r$ are called the canonical correlation coefficients.

The estimator of D_k is defined as follows:

$$\widehat{D}_k = \widetilde{X}_k - n^{-1}\widetilde{X}_k(\widehat{V}_k^{[1:\widetilde{r}_{12},:]})^\top \widehat{A}_C^{(\widetilde{r}_{12})} \sum_{j=1}^{2} \widehat{V}_k^{[1:\widetilde{r}_{12},:]} c_{base}. \tag{14}$$

In our experiment, we use the relationship between $\widehat{D}_1$ and $\widehat{D}_2$ to represent the orthogonal relationship between two distinctive matrices, and $\widehat{D}_1\widehat{D}_2 = 0_{p1\times p2}$. Finally, $\widehat{X}_k$, the estimator of X_k, is defined as follows:

$$\widehat{X}_k = \widehat{C}_k + \widehat{D}_k. \tag{15}$$

We summarize DCL in **Algorithm 2**.

Algorithm 2 | DCL

Input: $\boldsymbol{X}_1 \in \mathbb{R}^{p\times n}$, $\boldsymbol{X}_2 \in \mathbb{R}^{s\times n}$ //Input of sMRI and fMRI, respectively.

Output: $\hat{\boldsymbol{X}}_1$,$\hat{\boldsymbol{X}}_2$

$\widetilde{\boldsymbol{X}}_1, Lambda_1, U_1 \leftarrow$ S-POET($\boldsymbol{X}_1$) //processed by S-POET method

$\widetilde{\boldsymbol{X}}_2, Lambda_2, U_2 \leftarrow$ S-POET($\boldsymbol{X}_2$) //processed by S-POET method

$Lambda_{11} \leftarrow$ Construct diag($Lambda_1$)

$Lambda_{22} \leftarrow$ Construct diag($Lambda_2$)

$Theta \leftarrow (Lambda_{11}@U_1.T@\widetilde{\boldsymbol{X}}_1)@(\widetilde{\boldsymbol{X}}_2.T@U_2@Lambda_{22})/n$

$V_{theta}, D_{theta} \leftarrow$ SVD($Theta, full_{matrices} = True$) //Singular Value Decomposition

$Gamma_1 \leftarrow U_1@Lambda_{11}@V_{theta}$

$Gamma_2 \leftarrow U_2@Lambda_{22}@V_{theta}$

$\boldsymbol{A}_{mat} \leftarrow \mathrm{diag}(D_{theta})$ //Diagonal matrix

$\boldsymbol{C}_{base} \leftarrow$ Common variables corr($\widetilde{\boldsymbol{X}}_1, \widetilde{\boldsymbol{X}}_2$)

$\widetilde{\boldsymbol{C}}_1 \leftarrow$ Common matrix ($\widetilde{\boldsymbol{X}}_1, C_{base}, A_{mat}$)

$\widetilde{\boldsymbol{C}}_2 \leftarrow$ Common matrix ($\widetilde{\boldsymbol{X}}_2, C_{base}, A_{mat}$)

$\widetilde{\boldsymbol{D}}_1 \leftarrow$ Distinctive matrix ($\widetilde{\boldsymbol{X}}_1, \widetilde{C}_1$)

$\widetilde{\boldsymbol{D}}_2 \leftarrow$ Distinctive matrix ($\widetilde{\boldsymbol{X}}_2, \widetilde{C}_2$)

$\widetilde{\boldsymbol{X}}_1 \leftarrow$ Combination of common and distinctive matrices

$\widetilde{\boldsymbol{X}}_2 \leftarrow$ Combination of common and distinctive matrices

return $\hat{\boldsymbol{X}}_1, \hat{\boldsymbol{X}}_2$

3. METHODS

3.1. CNP Dataset

We evaluated the proposed DCL method in classifying NDs in the CNP dataset (Poldrack et al., 2016). The CNP dataset was collected by a consortium at the University of California, Los Angeles (UCLA), with financial support provided by the National Institutes of Health. This dataset has been used to elucidate the association between the human genome and complex psychological syndromes and promote the development of new therapies for NDs. All of this research was based on image phenotypic features in the mental disease.

The consortium for neuropsychiatric dataset was obtained from the OpenfMRI project (Gorgolewski et al., 2016). It includes sMRI data, task-based fMRI data, and resting-state fMRI data. These MRI images were acquired on one of two 3T Siemens Trio scanners at UCLA. The database contains extensive details of neuropsychologic assessments, neurocognitive tasks, and demographic information (including biological sex, age, and education). In addition, there are also details of the medication taken by those in ND groups.

The present study includes 272 images of subjects in one of four categories: 130 healthy controls (HCs), 50 SZ subjects, 49 BD subjects, and 43 ADHD subjects. These 272 images were from people in the Los Angeles area aged between 21 and 50 years old who were recruited through community advertisements. The details of the CNP dataset are listed in **Table 1**.

3.2. Brain Connectivity Data

Brain connectivity information may be reflected in fMRI images. In the CNP dataset, each sample has seven fMRI modalities, which were collected during different task states: BOLD contrast, resting state (with physiological monitoring), breath-holding tasks (with physiological monitoring), balloon analog risk tasks, stop-signal tasks, task switching, and spatial working memory capacity tasks. In this study, we attempted to classify NDs using resting-state fMRI images.

Resting-state fMRI is an imaging technique that obtains a brain activity function map when the subject is in a resting state undisturbed by other activities, which is better for distinguishing ND groups. The CNP dataset has resting-state fMRI images with scans lasting 304 s. The participants were relaxed with their eyes open. They were not stimulated or asked to respond during scanning (Poldrack et al., 2016). The fMRI data were collected under the following parameters: the slice thickness was 4 mm, 34 slices were taken, TR was 2 s, TE was 30 ms, the flip angle was 90°, the matrix size was 64 × 64, the field of view was 192 mm, and

TABLE 1 | Details of the Consortium for Neuropsychiatric Phenomics (CNP) dataset.

ID	Subtype	Number	Details
0	Healthy controls (HC)	130	–
1	Schizophrenia (SZ)	50	Disorganized, paranoid, or residual types
2	Bipolar disorder (BD)	49	Most recent hypomanic or manic episode, mild or moderate
3	Attention deficit hyperactivity disorder (ADHD)	43	Predominantly inattentive, combined, or predominantly hyperactive-impulsive types

the orientation was an oblique slice. In addition, high-resolution anatomical MP-RAGE data were collected under the following parameters: TR was 1.9 s, TE was 2.26 ms, the field of view was 250 mm, the matrix size was 256 × 256, the slices were in the sagittal plane, the slice thickness was 1 mm, and 176 slices were taken. We excluded 24 samples for which the whole-brain image volumes were unavailable or the head had moved excessively. Finally, we had 248 samples.

Before subsequent experiments, we preprocessed the fMRI data according to Gorgolewski et al. (2017), including slice timing, head motion corrections, spatial smoothing, band-pass filtering (0.01–0.1 Hz), nuisance signal regression, and Montreal Neurological Institute (MNI) space normalization and so on. Then, we used FSL to skull stripped and co-registered fMRI to the corresponding T1 weighted volume using boundary based registration with 9 degrees of freedom implemented in FreeSurfer. Finally, we obtained the functional connectivity matrix of the brain through the following steps: first, we used the BioImage Suite (Joshi et al., 2011) to calculate connectivity matrices for the fMRI images. We then used the Anatomical Automatic Labeling 90 (AAL90) brain atlas, which divided the brain images into 90 regions. The Pearson correlation coefficient was used to calculate the node values. The Fisher transformation was used to normalize the z scores. Finally, we obtained a 90×90 symmetric connectivity matrix for each sample. These connectivity matrices were not thresholded or binarized.

3.3. Brain Structure Data

Structural MRI are also used as inputs to the DCL method. It was obtained with the same parameter values used for the fMRI images. We used the open-source software FreeSurfer to process and analyze these sMRI images. FreeSurfer is used to analyze and visualize cross-sectional structural images. It can be used for stripping the skull, correcting the B1 bias field, registering an image, reconstructing the cortical surface, and estimating the cortical thickness.

We used FreeSurfer to generate high-precision gray and white matter segmentation surfaces and gray matter and cerebrospinal fluid segmentation surfaces. From these two surfaces, we calculated the cortical thickness and other surface features, such as the cortical surface area, curvature, and gray matter volume. Overall, there were 248 subjects, we obtained 2,196 features from the sMRI image of a subject. Finally, we constructed a $248 \times 2,196$ matrix from the sMRI image of 248 subjects.

4. EXPERIMENTS AND RESULTS

4.1. Experimental Design and Metrics

In our experiments, we focused on two aspects of brain connectivity: (1) classifying NDs into different subtypes using fMRI and sMRI data and (2) extracting important features from the fMRI and sMRI images. The classification task was to validate the performance of the DCL method for the different ND groups, whereas the feature extraction task was used to assess the capability of the method in detecting correlated features.

We obtained the correlation matrices by inputting the 248 fMRI (90×90) and sMRI (248×2196)matrices into S-POET. Then, we decomposed each pair of canonical matrices and computed their correlations. Finally, we used the leave-one-out (LOO) method to select the important features in the test sample matrix. For a dataset with n samples, verification based on LOO is carried out over n iterations. In each iteration, the classifier uses $n-1$ samples as training samples and uses the remaining sample as testing samples.

In our experiments, accuracy (ACC), precision (PRE), recall (REC), and F-score (F_1) are used to measure the classification performance. They are defined as follows:

$$\begin{aligned} \mathrm{ACC} &= \frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{FN}}, \\ \mathrm{PRE} &= \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}}, \\ \mathrm{REC} &= \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}, \\ F_1 &= 2\frac{\mathrm{PRE}\times\mathrm{REC}}{\mathrm{PRE}+\mathrm{REC}}, \end{aligned} \tag{16}$$

where TP is the number of true positives, TN is the number of true negatives, FP is the number of false positives, and FN is the number of false negatives. The values of these metrics were obtained from a LOO-based cross-validation.

Our experiments were implemented in Python on an NVIDIA Titan X Pascal CUDA GPU processor.

4.2. LOO Classification Method

We compared the performance of the DCL method with other methods: SVM, random forest (RF), XGBoost, PCA+SVM, PCA+RF, PCA+XGBoost, CCA+SVM, CCA+RF, and CCA+XGBoost. The linear kernel in the SVM classifier was used, as it provides better experimental performance than other kernels. As a trade-off between performance and computational cost, we set the number of trees in RF to 100. To prevent overfitting by XGBoost, we set the maximum tree depth for base

TABLE 2 | Mean values in the evaluation of the classification performance on the CNP dataset.

Classifier	ACC (%)	PRE (%)	REC (%)	F_1 (%)
SVM	38.00 (4.00)	40.00 (10.00)	39.00 (5.00)	37.00 (6.00)
RF	41.00 (10.00)	32.00 (11.00)	42.00 (7.00)	35.00 (9.00)
XGBoost	45.00 (6.00)	32.00 (9.00)	46.00 (4.00)	36.00 (5.00)
PCA+SVM	46.00 (2.00)	43.00 (7.00)	50.00 (7.00)	40.00 (3.00)
PCA+RF	47.00 (9.00)	49.00 (7.00)	46.00 (6.00)	44.00 (3.00)
PCA+XGBoost	49.00 (11.00)	45.00 (8.00)	49.00 (8.00)	45.00 (7.00)
CCA+SVM	45.00 (9.00)	42.00 (18.00)	49.00 (15.00)	38.00 (12.00)
CCA+RF	47.00 (13.00)	48.00 (11.00)	48.00 (10.00)	43.00 (10.00)
CCA+XGBoost	49.00 (8.00)	46.00 (14.00)	49.00 (12.00)	44.00 (14.00)
DCL+SVM	64.00 (9.00)	69.00 (7.00)	66.00 (6.00)	65.00 (8.00)
DCL+RF	68.00 (10.00)	73.00 (3.00)	72.00 (4.00)	72.00 (4.00)
DCL+XGBoost	72.00 (8.00)	81.00 (2.00)	70.00 (3.00)	75.00 (3.00)

learners and the turning parameter for the L_2 regularization term to 10 and 5, respectively. In the experiments, SVM, RF, and XGBoost use concatenated fMRI and sMRI matrices as their input, while the fMRI and sMRI matrices input to the other methods were first processed by the PCA, CCA, or DCL methods.

The classification results for the DCL method and the other classifiers are shown in **Table 2**. Each experiment was verified with 10-fold cross-validation. The conventional machine learning classifiers (SVM, RF, and XGBoost) had the lowest accuracy. These classifiers cannot capture distinguishable information from the union matrix. Compared with SVM, RF, and XGBoost, the PCA and CCA classifiers achieved better classification results. The best accuracy for both was 49.00%, which demonstrates that correlation information can be incorporated to improve the classification. The classifiers based on DCL had much better performance than those based on PCA or CCA. The best accuracy was 72.00%. Our proposed DCL method is a natural extension of the traditional CCA method. Based on the CCA decomposition, DCL determines the common and discernibility matrices and establishes an orthogonal relationship between the two discernibility matrices.

In addition, our comparative experiment was based on a sample size of 248. As shown in **Table 2**, we used three typical machine learning methods (SVM, RF, and XGBoost) as the baseline. The performance of these three machine learning methods was very different from that based on the PCA, CCA, or DCL methods. There are two reasons:

1. Machine learning methods can be effective for classifying simple images, but because medical images are very complex, these three machine learning methods were overwhelmed.
2. The limited sample size does not meet the training requirements of the three machine learning methods. The multi-class classification task increased the imbalance for the samples, making it difficult for these methods to obtain key feature information from the high latitude and limited samples.

Therefore, unlike the other methods, the DCL method first preprocesses the complex relationship between the sMRI and fMRI data, which reduces the complexity of the input data. **Table 2** shows that, despite the limited sample size, DCL can better deal with the relations in high latitude data and improve the performance of machine learning.

Of the DCL-based classifiers, XGBoost had the best results in the multi-class classification task. The best accuracy was 72.00%. The receiver operating characteristic (ROC) curves for XGBoost in multi-class classification is plotted in **Figure 2**. The areas under the micro-averaged and macro-averaged ROC curves in **Figures 2B,C** are much larger than those in **Figure 2A**. Moreover, the areas under the curves for the four subtypes in **Figures 2B,C** are much larger than those in **Figure 2A**. These results indicate that the correlation information obtained by PCA or CCA can improve the performance of a classifier. The classification results for DCL are much better than those for PCA or CCA. The areas under all the ROC curves in **Figure 2D** are larger than those in **Figures 2B,C**. This indicates that our DCL method can better describe brain connection networks and thus improve the performance of the classifiers.

4.3. Feature Selection Based on the LOO Method

Besides assessing the performance of the DCL method, we also identified the important features with the DCL+XGBoost method. The aim was to find which edges contribute to brain connectivity. The extracted features are mapped back into the brain space, which facilitates the interpretation of the known relationship between brain structure and function. However, due to the dimensionality of the connectivity network, the visualization is challenging. In the LOO method, we used a weight-based method to evaluate the importance of features in the test sample matrix. The weight in XGBoost is used to calculate the number of times a feature is used as a split point across all trees. Finally, we counted the number of samples whose feature weights were >0. We visualized the representations of all important features for both the sMRI and fMRI data.

4.4. Visualization of FCs

It is interesting to investigate how different brain networks cooperate and connect with each other. We found that there were significant differences between the FCs of each group, which indicates that these FCs not only reflect the information common to the different groups but also the differences among them. We used the BrainnetViewer software (https://www.nitrc.org/projects/bnv/) to visualize which FCs have the strongest relationships in the brain network.

The first row in **Figure 3** is for the HC group, whereas the second row is for the ND group. **Figure 3A** shows 3D plots of the brain network to visualize the selected edges. A sphere denotes the center of a node. Different colors denote different brain regions. If two brain regions are functionally related, they are connected by a colored line. The colors of the lines indicate the edge strength and whether there is a positive correlation between the behaviors and the FCs. The

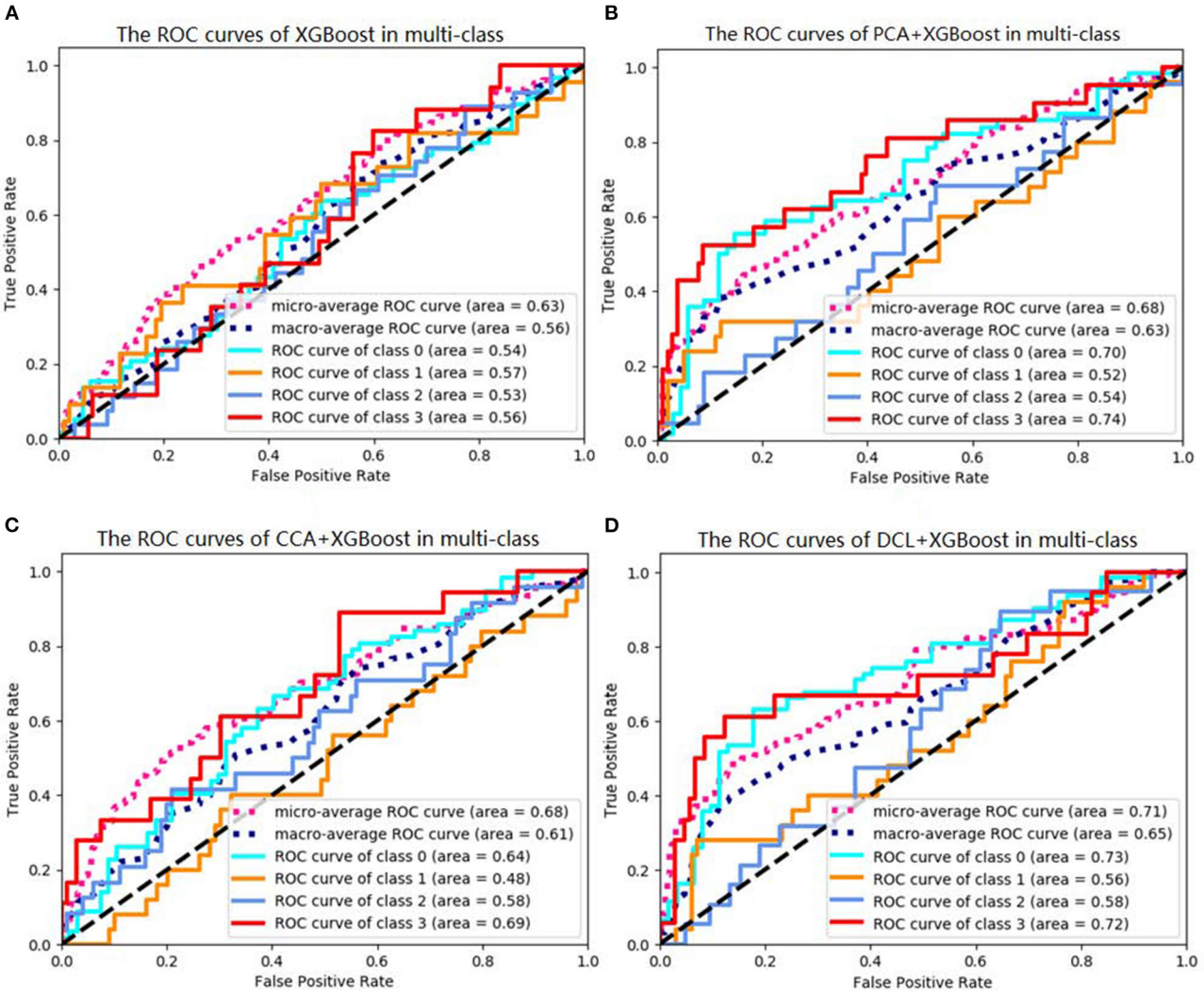


FIGURE 2 | Receiver operating characteristics (ROC) curves of XGBoosts with different pretreatment methods. **(A)** XGBoost method is used in classification task. **(B)** The PCA-based XGBoost is used in classification task. **(C,D)** CCA and DCL-based XGBoosts are used in classification task.

brain network visualization has a small number of edges, which demonstrate the degree of the distribution across the whole brain network.

The 2D circle plots in **Figure 3B** are also used to visualize relationships between pairs of brain regions. The wider the edge between two regions, the closer their relationship is. These circle plots indicate how many FCs a region has with other brain regions.

Figure 3C has mappings of the 90 × 90 connectivity matrices, which are used to visualize aggregate statistics within and between predefined regions or networks. In a connectivity matrix, nodes represent brain regions and links measure conditional dependence between the brain regions. Brain connectivity analysis is equivalently transformed into the estimation of a spatial partial correlation matrix.

4.5. Analysis of HCs and NDs

In both HC group (the first row in **Figure 3**) and ND group (the second row in **Figure 3**), most of the FCs are common to both groups. These overlapping FCs are mainly within or across the temporal lobes or across the frontal, occipital, and parietal lobes, which confirm the results of previous studies. For instance, Haier et al. (2005) and Rubia et al. (2007) showed that temporal lobe dysfunction is strongly correlated with ADHD. Several brain regions in the frontal, parietal, temporal, and occipital lobes have been identified as significant predictors of ND (Gaudio et al., 2019; Zhang et al., 2020).

Furthermore, **Figures 3A,B** show that there are significant differences between the FCs of the two groups. Compared with the HCs, the ND group has abnormal brain regions, mainly in the supramarginal gyrus, cingulate gyrus, middle frontal gyrus, etc. Other studies have also found that there are fewer FCs in

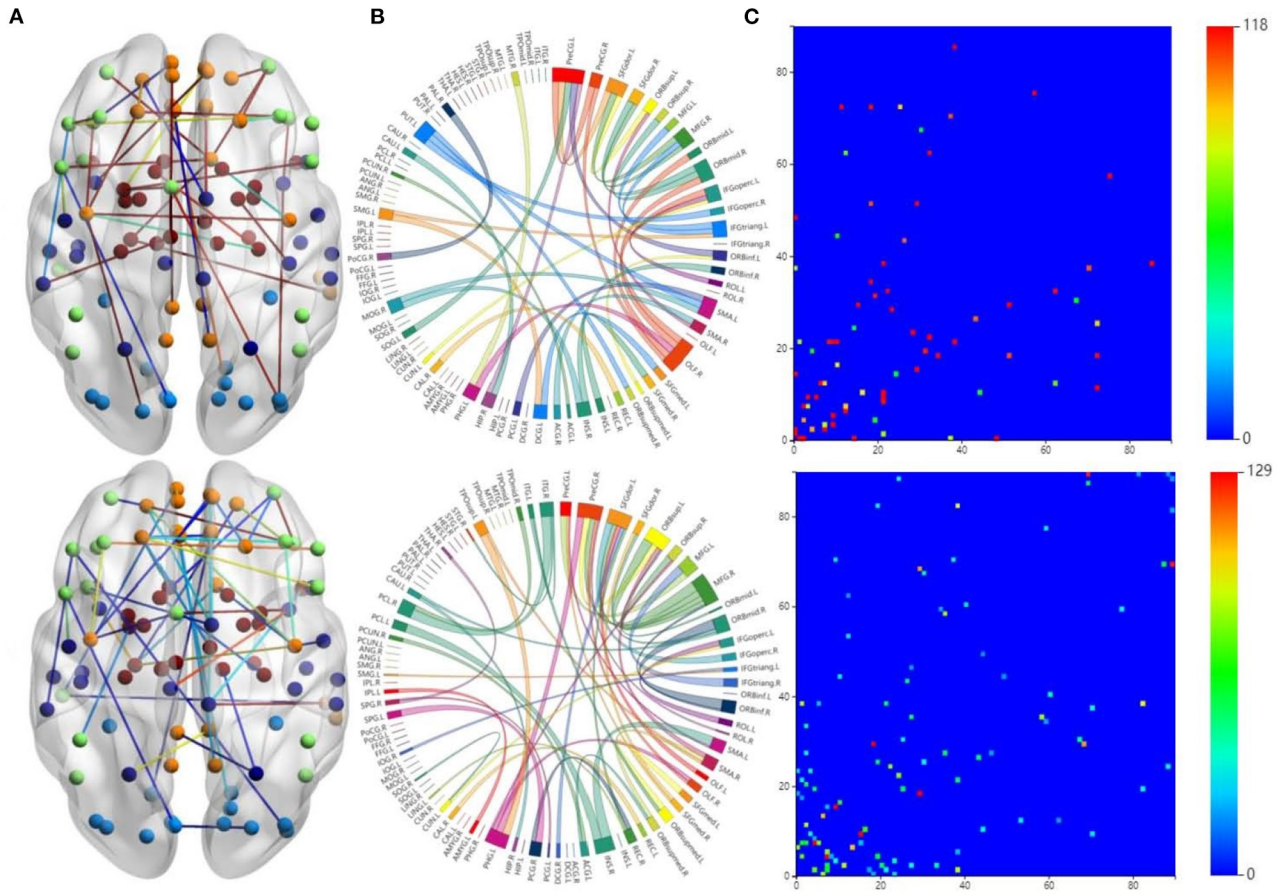


FIGURE 3 | Visualizations of the connectivity of HC and neuropsychiatric disorders (NDs) in different manners on the CNP dataset. The first row and second row show the HC group and ND group, respectively. **(A)** Shows the connectivity in glass brain plots. **(B)** Shows the connectivity in circle plots. **(C)** Shows the connectivity in symmetric matrices.

the middle frontal gyrus and anterior cingulate regions in SZ brains compared to HCs (Camchong et al., 2011; Liu et al., 2011). However, the FCs in the ND group are more complicated than those in the HC group, which may be due to their mental illness. These differences may affect the behaviors and mental states of the ND group. There are many highlighted cells in the HC matrix in **Figure 3C**, whereas the highlighted cells in the ND matrix are more dispersed. This also indicates that NDs may affect the FCs between brain regions.

4.6. Analysis of Different NDs

To study the specificity of subtypes in NDs, we visualized the FCs of the three ND subtypes in **Figure 4**. **Figure 4A** is for all the ND subtypes. **Figure 4B** is for the SZ subtype. **Figures 4C,D** are for the BD and ADHD subtypes, respectively.

The brain networks clearly suggest that the FCs of these diseases are very similar, but their differences are also very obvious. In particular, the FCs in the ADHD plots are obviously different from those in the SZ and BD plots. This is why classifying ADHD is usually a separate task in most approaches to classifying NDs. Moreover, the connections between brain regions shown in the circle plots in the second column are obviously different for the three diseases.

4.7. Features Distribution of PCA and DCL

Figure 5 compares the principal components found by the PCA method with those found by the proposed DCL method. **Figures 5A,B** visualize the fMRI and sMRI feature matrices found by PCA. **Figure 5C** is the visualization of the combined feature matrix for the fMRI and sMRI images for PCA. **Figure 5D** is the feature matrix produced by DCL.

As shown in **Figure 5**, the figure shows that the three distributions of features produced by PCA are disordered (**Figures 5A–C**). Although the distributions of the PCA-processed fMRI and sMRI matrices (**Figure 5C**) are relatively concentrated, the four icons of subtypes are still indistinguishable. It would be difficult for classifiers to distinguish the features of the four subtypes. In contrast, the distribution of fMRI and sMRI matrices after DCL processing shows the effect of aggregation, which is shown in **Figure 5D**. The features of

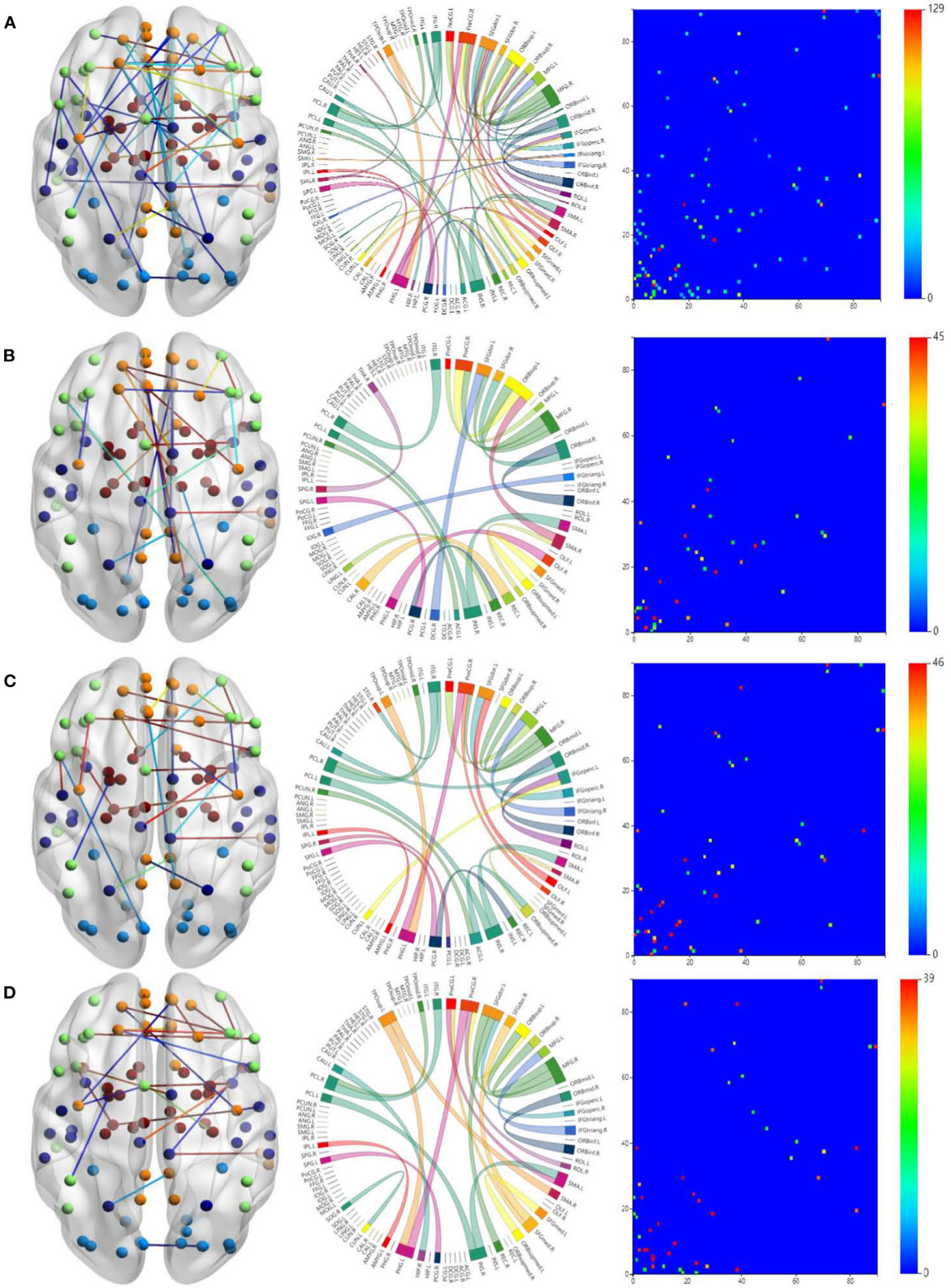


FIGURE 4 | Visualizations of the connectivity of three ND subtypes in glass brain plot graph, circle plot graph, and symmetric matrix graph on CNP dataset. **(A)** Shows all the ND subtypes. **(B)** Shows the SZ subtype. **(C,D)** Show the BD and ADHD subtypes, respectively.

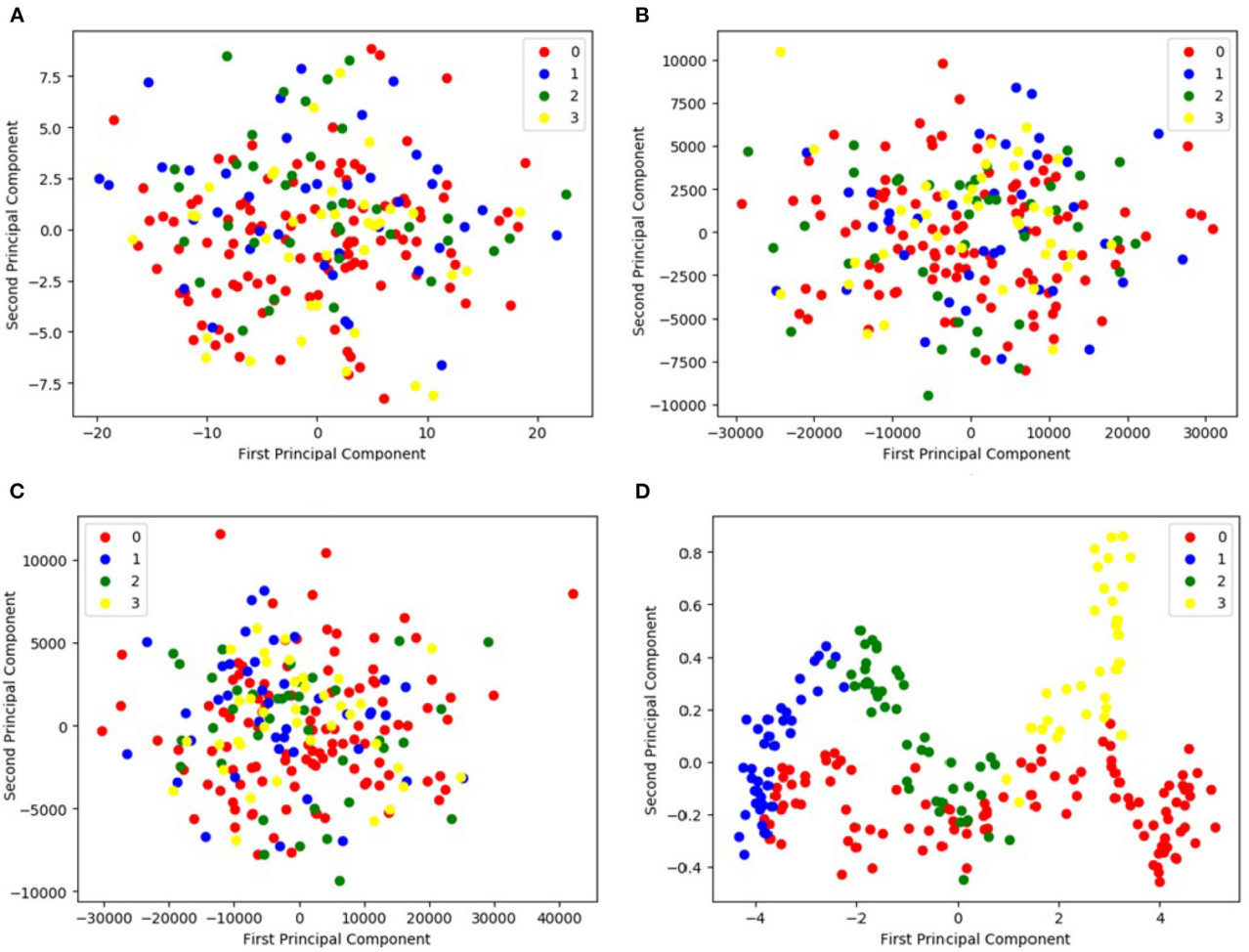


FIGURE 5 | Representation of feature distribution on CNP dataset. **(A,B)** Visualize the fMRI and sMRI feature matrices processed by PCA, respectively. **(C)** Visualizes the combined feature matrix for the fMRI and sMRI images processed by PCA. **(D)** Visualizes the feature matrix produced by DCL. In the legend, 0 represents HC, 1 represents SZ, 2 represents BD, and 3 represents ADHD.

the four subtypes can be clearly distinguished. Therefore, the performance of a classifier would be greatly improved by using a feature matrix produced by the DCL method. At the same time, in order to eliminate the difference in the distribution of subtypes, we normalized the matrices in the DCL method, so that the subtypes are distributed in a smaller range.

5. ABLATION EXPERIMENTS AND DISCUSSION

We proposed the DCL framework to classify psychiatric disorders using fMRI and sMRI. In this section, we discussed several factors that influence the experimental results. To validate the performance of DCL on different size of datasets, we extended experiments on a larger sample size dataset (a subset of ADNI) and a small sample size dataset (a subset of OpenfMRI), respectively.

5.1. Influence of S-POET

The shrinkage principal orthogonal complement thresholding method is a covariance estimator with the approximate factor model, which is based on sparse PCA. In our method, we used the S-POET method to obtain asymptotic first-order distribution for the eigenvalues and eigenvectors of the fMRI and sMRI correlation matrices, respectively. To verify the effect of the S-POET method in our proposed DCL method, we extended two different DCL methods on XGBoost: one is based on PCA[DCL(PCA)] and another is based on S-POET[DCL(S-POET)].

As shown in **Table 3**, we extended the experiments on CNP dataset. For both datasets, compared with the DCL(PCA)-based XGBoost, the DCL(S-POET)-based XGBoost obtained the super performance. The accuracy was almost improved by 13% on CNP. Although S-POET is obtained by sparse PCA extension, S-POET is more suitable for sparse high-latitude data. PCA has widely been proved that it is a powerful tool for

dimensionality reduction and data visualization. Its theoretical properties such as the consistency and asymptotic distributions of empirical eigenvalues and eigenvectors are challenging especially in the high dimensional regime. While, in the method S-POET, the spike magnitude of leading eigenvalues, sample size, and dimensionality of the leading eigenvalues are considered. In addition, a new covariance estimator is introduced in S-POET to correct the bias of PCA estimation of leading eigenvalues and eigenvectors. Therefore, S-POET is more advantageous in the process of fMRI and sMRI matrices analysis with high dimensionality and sparse features (Fan and Wang, 2015). Therefore, in the end, we build the DCL method with S-POET.

5.2. Effectiveness of Different Inputs on XGBoosts

To verify the influence of different MRI modalities on model, we separately used fMRI, sMRI, and fMRI+sMRI matrices as inputs to three types of XGBoosts, namely PCAXGBoost, CCA-XGBoost, and DCL-XGBoost.

The results are shown in **Table 4**. The classification results of three XGBoost-based methods, using a single fMRI or sMRI matrix as input, are similar. However, the results of using PCA, CCA, and DCL processed fMRI and sMRI matrices as input to the XGBoost classifier have greatly improved. Especially for the DCL-XGBoost method, the accuracy is improved by almost 14% on the CNP dataset. As the two modalities complement each other, their combination results in higher classification accuracy. Furthermore, the performance of PCA and CCA-processed matrices is not as good as when using DCL-processed matrices as the XGBoost input.

TABLE 3 | Influence of shrinkage principal orthogonal complement thresholding method (S-POET) on XGBoost with CNP dataset.

Method	ACC (%)	PRE (%)	REC (%)	F_1 (%)
DCL(PCA)	59.00 (5.00)	60.00 (9.00)	61.00 (11.00)	59.00 (7.00)
DCL(S-POET)	72.00 (8.00)	81.00 (2.00)	70.00 (3.00)	75.00 (3.00)

5.3. Influence of Medication Taken

Some patients in the ND group had taken medication for their mental illness. To analyze the impact of these medications on the patients, we visualized the selected FCs for a group who had taken medication and for a group who had not. There are significant differences between these two groups, as shown in **Figure 6**. **Figure 6A** shows NDs without medication. **Figure 6B** shows NDs with medication. The representations of the FCs over the whole brain are similar, but for the group who had not used medication, there are more edges over the boundary of the brain. This may be due to the fact that some FCs are interrupted by the patient taking certain medication, resulting in remission or deepening of mental illness.

5.4. Extend Experiments

To verify the performance of DCL on different datasets, we extended experiments on a larger sample size dataset (a subset of ADNI) and a small sample size dataset (a subset of OpenfMRI), respectively. The Alzheimer's Disease Neuroimaging Initiative (ADNI) (Carrillo et al., 2012) is a large dataset including Alzheimer's disease (AD) and mild cognitive impairment (MCI). We selected a subset of the ADNI dataset to evaluate our

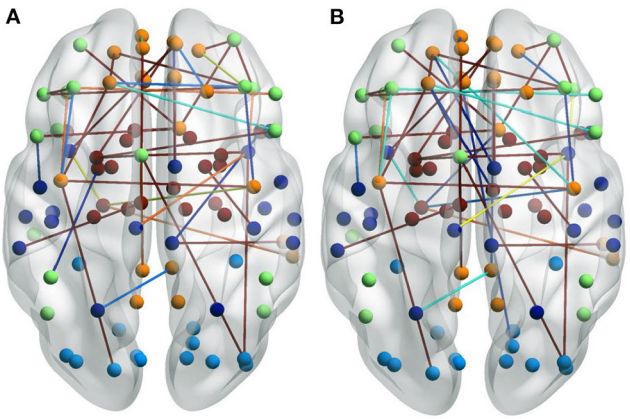


FIGURE 6 | Visualizations of the connectivity of NDs who took medicine or not in the glass brain plot graph on the CNP dataset. **(A)** Shows NDs without medication. **(B)** Shows NDs with medication.

TABLE 4 | Evaluation of different inputs to the different combinations of XGBoost on the CNP dataset.

Method	Input	ACC (%)	PRE (%)	REC (%)	F_1 (%)
PCA+XGBoost	fMRI	38.00 (7.00)	35.00 (4.00)	36.16 (3.00)	33.00 (2.00)
	sMRI	37.00 (8.00)	35.00 (3.00)	36.00 (10.00)	32.00 (9.00)
	fMRI+sMRI	49.00 (11.00)	45.00 (8.00)	49.00 (8.00)	45.00 (7.00)
CCA+XGBoost	fMRI	36.00 (10.00)	34.00 (4.00)	36.00 (7.00)	35.00 (8.00)
	sMRI	38.20 (1.00)	37.06 (7.00)	35.00 (9.00)	36.00 (4.00)
	fMRI+sMRI	49.00 (8.00)	46.00 (14.00)	49.00 (12.00)	44.00 (14.00)
DCL-XGBoost	fMRI	56.00 (2.00)	58.00 (8.00)	60.00 (3.00)	53.00 (9.00)
	sMRI	58.00 (6.00)	62.00 (8.00)	52.00 (11.00)	55.00 (6.00)
	fMRI+sMRI	72.00 (8.00)	81.00 (2.00)	70.00 (3.00)	75.00 (3.00)

proposed DCL method. This subset includes 420 samples with sMRI (T1w MRI) and fMRI (rs-fMRI). It consists of 105 subjects with AD, 105 late mild cognitive impairment (LMCI) subjects, 105 early mild cognitive impairment (EMCI) subjects, and 105 HC subjects. The OpenfMRI Poldrack et al. (2013) was designed to serve as a repository for the open sharing and dissemination of task-based fMRI data. As it has grown, it has broadened to encompass other data types as well, including EEG, MEG, rs-fMRI (fMRI), and diffusion MRI (sMRI), which were acquired on both healthy and clinical populations. We selected a small subset of OpenfMRI dataset with the resting state. This subset includes 93 samples with sMRI and fMRI. It consists of 20 HC subjects, 16 BD subjects, 28 SC subjects, and 29 ADHD subjects.

In our study, the subsets of ADNI and OpenfMRI are used as the external datasets to evaluate the performance of DCL. The data processing steps followed the manner in Section 3. The experimental design and metrics follow the design in Section 4. The classification results of this subset are show in **Table 5**. We also used three typical machine learning methods (SVM, RF, and XGBoost) as the baseline. As shown in **Tables 5**, **6**, the accuracy trend of the experimental results is similar to that in **Table 2**. The DCL-based classifiers achieve much better classification results, which further proves that the DCL method can reduce the complexity of the data by preprocessing the two types of MRI, thereby improving the classification performance of the classifiers. By comparing **Tables 2**, **5**, **6**, it can be found that the classification results of the three classifiers on the subset of ADNI achieve the best performance and that on the subset of OpenfMRI achieve the worst performance. In addition to the reasons for the samples themselves, in these three datasets, the subset of ADNI has the largest sample size, which can lead to better training and prediction of the machine learning methods. While the subset of OpenfMRI has the smallest sample size, which limits the training and prediction of the machine learning methods. Furthermore, in the case of a limited sample size on the subset of OpenfMRI, the performance of DCL-based methods got obvious advantages compared to other methods.

We compared DCL+XGBoost with several other well-established methods that were designed for the multi neuropsychiatric disorders classification: mMLDA (Janousova et al., 2015), MFMK-SVM (Liu J. et al., 2018), KFCM (Baskar et al., 2019), MK-SVM (Zhuang et al., 2019), and mRMR-SVM (Zhang et al., 2021). These methods used one or both types of MRI data as input of the model for multi neuropsychiatric disorder classification. These methods were trained using different datasets and utilize very different predictive architectures. We either re-implemented them exactly as described by the authors or used the code released by the author. To ensure that the comparative evaluation is fair, we used the same training data and test data for all considered methods on tree datasets. The results are shown in **Table 7**, it can be found that our proposed method achieves state-of-the-art performance on all three datasets. These methods needed much more feature selection work and parameter settings, for example, mRMR-SVM needs mutual selected information as a measure to solve the trade-off between feature redundancy and relevance (Morgado et al., 2015). It increases the difficulty of model optimization. In addition, the performance of these methods improved as the sample size increased. This means that sample size and model performance are positively correlated.

TABLE 5 | Mean values in the evaluation of the classification performance on the subset of Alzheimer's Disease Neuroimaging Initiative (ADNI).

Classifier	ACC (%)	PRE (%)	REC (%)	F_1 (%)
SVM	54.00 (8.00)	53.00 (5.00)	59.00 (4.00)	57.00 (7.00)
RF	54.00 (4.00)	52.00 (10.00)	58.00 (4.00)	55.00 (8.00)
XGBoost	55.00 (10.00)	52.00 (15.00)	58.00 (7.00)	56.00 (9.00)
PCA+SVM	60.00 (10.00)	62.00 (4.00)	61.00 (7.00)	60.00 (6.00)
PCA+RF	65.00 (7.00)	61.00 (8.00)	63.00 (10.00)	63.00 (3.00)
PCA+XGBoost	72.00 (3.00)	68.00 (10.00)	73.00 (13.00)	72.00 (9.00)
CCA+SVM	62.00 (10.00)	63.00 (11.00)	65.00 (8.00)	63.00 (7.00)
CCA+RF	62.00 (4.00)	64.00 (3.00)	66.00 (6.00)	62.00 (6.00)
CCA+XGBoost	75.00 (4.00)	73.00 (6.00)	76.00 (7.00)	75.00 (9.00)
DCL+SVM	77.00 (12.00)	78.00 (3.00)	77.00 (9.00)	79.00 (10.00)
DCL+RF	78.00 (6.00)	79.00 (4.00)	78.00 (10.00)	80.00 (13.00)
DCL+XGBoost	80.00 (9.00)	79.00 (9.00)	80.00 (5.00)	82.00 (7.00)

TABLE 6 | Mean values in the evaluation of the classification performance on the subset of OpenfMRI.

Classifier	ACC (%)	PRE (%)	REC (%)	F_1 (%)
SVM	33.00 (7.00)	33.00 (3.00)	35.00 (9.00)	34.00 (8.00)
RF	34.00 (6.00)	34.00 (11.00)	33.00 (2.00)	35.00 (7.00)
XGBoost	35.00 (6.00)	35.00 (7.00)	34.00 (9.00)	36.00 (2.00)
PCA+SVM	39.00 (7.00)	38.00 (10.00)	39.00 (10.00)	40.00 (4.00)
PCA+RF	41.00 (2.00)	40.00 (10.00)	41.00 (6.00)	41.00 (13.00)
PCA+XGBoost	43.00 (9.00)	44.00 (8.00)	45.00 (8.00)	42.00 (7.00)
CCA+SVM	43.00 (2.00)	44.00 (5.00)	46.00 (9.00)	45.00 (2.00)
CCA+RF	45.00 (7.00)	46.00 (7.00)	47.00 (3.00)	46.00 (10.00)
CCA+XGBoost	51.00 (14.00)	53.00 (7.00)	56.00 (7.00)	50.00 (5.00)
DCL+SVM	55.00 (6.00)	57.00 (8.00)	56.00 (8.00)	57.00 (6.00)
DCL+RF	62.00 (10.00)	64.00 (3.00)	64.00 (7.00)	63.00 (9.00)
DCL+XGBoost	67.00 (8.00)	69.00 (10.00)	68.00 (9.00)	68.00 (10.00)

5.5. Limitations

There are several limitations to this study. (1) We used only MRI data as the input. However, the classification of complex disorders could be made more accurate by including phenotypic information. (2) The amount and uneven quality of the MRI data have a significant influence on the performance of a model and reduce the accuracy of classification.

6. CONCLUSION

This work demonstrated that the DCL method can effectively combine different information from fMRI and sMRI images. DCL identifies both the common and distinct information

TABLE 7 | Comparison results with other methods on tree datasets.

Dataset	Classifier	MRIs	ACC (%)	PRE (%)	REC (%)	F_1 (%)
CNP	mMLDA (Janousova et al., 2015)	sMRI	65.00 (7.00)	65.00 (6.00)	67.00 (9.00)	64.00 (6.00)
	MFMK-SVM (Liu J. et al., 2018)	sMRI, DTI	67.00 (9.00)	64.00 (12.00)	65.00 (7.00)	68.00 (9.00)
	KFCM (Baskar et al., 2019)	sMRI	70.00 (7.00)	71.00 (7.00)	70.00 (6.00)	69.00 (10.00)
	MK-SVM (Zhuang et al., 2019)	sMRI, fMRI	70.00 (11.00)	75.00 (4.00)	72.00 (4.00)	74.00 (7.00)
	mRMR-SVM (Zhang et al., 2021)	sMRI, fMRI	71.00 (9.00)	78.00 (7.00)	71.00 (6.00)	72.00 (10.00)
	DCL+XGBoost	sMRI, fMRI	72.00 (8.00)	81.00 (2.00)	70.00 (3.00)	75.00 (3.00)
ADNI	mMLDA (Janousova et al., 2015)	sMRI	70.00 (8.00)	72.00 (8.00)	70.00 (10.00)	69.00 (9.00)
	MFMK-SVM (Liu J. et al., 2018)	sMRI, DTI	73.00 (9.00)	72.00 (10.00)	74.00 (6.00)	75.00 (7.00)
	KFCM (Baskar et al., 2019)	sMRI	75.00 (9.00)	74.00 (1.00)	76.00 (4.00)	74.00 (8.00)
	MK-SVM (Zhuang et al., 2019)	sMRI, fMRI	75.00 (11.00)	74.00 (9.00)	75.00 (8.00)	75.00 (2.00)
	mRMR-SVM (Zhang et al., 2021)	sMRI, fMRI	79.00 (12.00)	82.00 (10.00)	79.00 (6.00)	81.00 (7.00)
	DCL+XGBoost	sMRI, fMRI	80.00 (9.00)	79.00 (9.00)	80.00 (5.00)	82.00 (7.00)
OpenfMRI	mMLDA (Janousova et al., 2015)	sMRI	54.00 (7.00)	53.00 (10.00)	55.00 (7.00)	53.00 (9.00)
	MFMK-SVM (Liu J. et al., 2018)	sMRI, DTI	57.00 (5.00)	58.00 (7.00)	56.00 (10.00)	57.00 (6.00)
	KFCM (Baskar et al., 2019)	sMRI	63.00 (11.00)	64.00 (8.00)	64.00 (4.00)	64.00 (9.00)
	MK-SVM (Zhuang et al., 2019)	sMRI, fMRI	66.00 (8.00)	65.00 (12.00)	67.00 (8.00)	64.00 (10.00)
	mRMR-SVM (Zhang et al., 2021)	sMRI, fMRI	67.00 (5.00)	70.00 (7.00)	71.00 (6.00)	72.00 (10.00)
	DCL+XGBoost	sMRI, fMRI	67.00 (8.00)	69.00 (10.00)	68.00 (9.00)	68.00 (10.00)

between the two input MRI matrices. The decomposition-based CCA is used to analyze the correlation and construct the required matrices. Thus, DCL has better performance in both classification and identifying FCs. The DCL method can be used to detect complex and nonlinear relationships between the two types of MRI images. Our experiments showed that the DCL method can improve classification performance so that it is a suitable method for classifying mental illnesses.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

LL, HZ, and GL contributed to the conception and design of the study. JC organized the database. YW performed the statistical analysis. LL wrote the first draft of the manuscript. Y-PW modified the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This work was funded partially by the National Natural Science Foundation of China under Grant Nos. 62172444, 61877059, and 62102454, the 111 Project (No. B18059), and the Henan Provincial Key Research and Promotion Projects (No. 222102310085).

REFERENCES

Baskar, D., Jayanthi, V., and Jayanthi, A. (2019). An efficient classification approach for detection of Alzheimer's disease from biomedical imaging modalities. *Multim. Tools Appl.* 78, 12883–12915. doi: 10.1007/s11042-018-6287-8

Calhoun, V. D., Adali, T., Kiehl, K. A., Astur, R., and Pearlson, G. D. (2010). A method for multitask fMRI data fusion applied to schizophrenia. *Hum. Brain Mapp.* 27, 598–610. doi: 10.1002/hbm.20204

Camchong, J., MacDonald III, A. W., Bell, C., Mueller, B. A., and Lim, K. O. (2011). Altered functional and anatomical connectivity in schizophrenia. *Schizophr. Bull.* 37, 640–650. doi: 10.1093/schbul/sbp131

Carrillo, M. C., Bain, L. J., Frisoni, G. B., and Weiner, M. W. (2012). Worldwide Alzheimer's disease neuroimaging initiative. *Alzheimer's Dement.* 8, 337–342. doi: 10.1016/j.jalz.2012.04.007

Connaughton, M., Whelan, R., O'Hanlon, E., and McGrath, J. (2022). White matter microstructure in children and adolescents with ADHD. *NeuroImage* 2022, 102957. doi: 10.1016/j.nicl.2022.102957

Correa, N. M., Eichele, T., Adali, T., Li, Y. O., and Calhoun, V. D. (2010). Multi-set canonical correlation analysis for the fusion of concurrent single trial ERP and functional MRI. *Neuroimage* 50, 1438–1445. doi: 10.1016/j.neuroimage.2010.01.062

de Filippis, R., Carbone, E. A., Gaetano, R., Bruni, A., Pugliese, V., Segura-Garcia, C., et al. (2019). Machine learning techniques in a structural and functional MRI diagnostic approach in schizophrenia: a systematic review. *Neuropsychiatr. Dis. Treat.* 15, 1605. doi: 10.2147/NDT.S202418

Dubois, J., and Adolphs, R. (2016). Building a science of individual differences from fMRI. *Trends Cogn. Sci.* 20, 425–443. doi: 10.1016/j.tics.2016.03.014

Fan, J., and Wang, W. (2015). Asymptotics of empirical Eigen-structure for ultra-high dimensional spiked covariance model. *arXiv[Preprint].arXiv:1502.04733.*

Fan, Z., Xu, F., Qi, X., Li, C., and Yao, L. (2020). Classification of Alzheimer disease based on brain MRI and machine learning. *Neural Comput. Appl.* 32, 1927–1936. doi: 10.1007/s00521-019-04495-0

Finn, E. S., Shen, X., Scheinost, D., Rosenberg, M. D., Huang, J., Chun, M. M., et al. (2015). Functional connectome fingerprinting: identifying

individuals using patterns of brain connectivity. *Nat. Neurosci*. 18, 1664–1671. doi: 10.1038/nn.4135

Gao, S., Calhoun, V. D., and Sui, J. (2020). "Multi-modal component subspace-similarity-based multi-kernel SVM for schizophrenia classification," in *Medical Imaging 2020: Computer-Aided Diagnosis: International Society for Optics and Photonics 113143X* (Houston, TX: SPIE Medical Imaging). doi: 10.1117/12.2550339

Gaudio, S., Carducci, F., Piervincenzi, C., Olivo, G., and Schioth, H. B. (2019). Altered thalamo-cortical and occipital-parietal-temporal-frontal white matter connections in patients with anorexia and bulimia nervosa: a systematic review of diffusion tensor imaging studies. *J. Psychiatry Neurosci*. 44, 324–339. doi: 10.1503/jpn.180121

Gorgolewski, K. J., Auer, T., Calhoun, V. D., Craddock, R. C., Das, S., Duff, E. P., et al. (2016). The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Sci. Data* 3, 160044–160044. doi: 10.1038/sdata.2016.44

Gorgolewski, K. J., Durnez, J., and Poldrack, R. A. (2017). Preprocessed consortium for neuropsychiatric phenomics dataset. *F1000Research* 6, 1262. doi: 10.12688/f1000research.11964.1

Groves, A. R., Beckmann, C. F., Smith, S. M., and Woolrich, M. W. (2011). Linked independent component analysis for multimodal data fusion. *Neuroimage* 54, 2198–2217. doi: 10.1016/j.neuroimage.2010.09.073

Haier, R. J., Jung, R. E., Yeo, R. A., Head, K., and Alkire, M. T. (2005). The neuroanatomy of general intelligence: sex matters. *NeuroImage* 25, 320–327. doi: 10.1016/j.neuroimage.2004.11.019

Han, K.-M., De Berardis, D., Fornaro, M., and Kim, Y.-K. (2019). Differentiating between bipolar and unipolar depression in functional and structural MRI studies. *Prog. Neuro Psychopharmacol. Biol. Psychiatry* 91, 20–27. doi: 10.1016/j.pnpbp.2018.03.022

Heinrichs, R. W., and Zakzanis, K. K. (1998). Neurocognitive deficit in schizophrenia: a quantitative review of the evidence. *Neuropsychology* 12, 426. doi: 10.1037/0894-4105.12.3.426

Hu, W., Cai, B., Zhang, A., Calhoun, V. D., and Wang, Y.-P. (2019). Deep collaborative learning with application to the study of multimodal brain development. *IEEE Trans. Biomed. Eng.* 66, 3346–3359. doi: 10.1109/TBME.2019.2904301

Hu, W., Meng, X., Bai, Y., Zhang, A., Qu, G., Cai, B., et al. (2021). Interpretable multimodal fusion networks reveal mechanisms of brain cognition. *IEEE Trans. Med. Imaging* 40, 1474–1483. doi: 10.1109/TMI.2021.3057635

Janousova, E., Schwarz, D., and Kasparek, T. (2015). Combining various types of classifiers and features extracted from magnetic resonance imaging data in schizophrenia recognition. *Psychiatry Res.* 232, 237–249. doi: 10.1016/j.pscychresns.2015.03.004

Jiang, R., Calhoun, V. D., Cui, Y., Qi, S., Zhuo, C., Li, J., et al. (2020). Multimodal data revealed different neurobiological correlates of intelligence between males and females. *Brain Imaging Behav.* 14, 1979–1993. doi: 10.1007/s11682-019-00146-z

Jiang, Y., Duan, M., He, H., Yao, D., and Luo, C. (2021). Structural and functional MRI brain changes in patients with schizophrenia following electroconvulsive therapy: a systematic review. *Curr. Neuropharmacol.* doi: 10.2174/1570159X19666210809101248

Joshi, A., Scheinost, D., Okuda, H., Belhachemi, D., Murphy, I., Staib, L. H., et al. (2011). Unified framework for development, deployment and robust testing of neuroimaging algorithms. *Neuroinformatics* 9, 69–84. doi: 10.1007/s12021-010-9092-8

Kessler, R. C., Petukhova, M., Sampson, N. A., Zaslavsky, A. M., and Wittchen, H. (2012). Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. *Int. J. Methods Psychiatr. Res.* 21, 169–184. doi: 10.1002/mpr.1359

Lake, E. M., Finn, E. S., Noble, S. M., Vanderwal, T., Shen, X., Rosenberg, M. D., et al. (2019). The functional brain organization of an individual allows prediction of measures of social abilities trans-diagnostically in autism and attention/deficit and hyperactivity disorder. *Biol. Psychiatry* 86, 315–326. doi: 10.1016/j.biopsych.2019.02.019

Liu, H., Fan, G., Xu, K., and Wang, F. (2011). Changes in cerebellar functional connectivity and anatomical connectivity in schizophrenia: a combined resting-state functional MRI and diffusion tensor imaging study. *J. Magn. Reson. Imaging* 34, 1430–1438. doi: 10.1002/jmri.22784

Liu, J., Demirci, O., and Calhoun, V. D. (2008). A parallel independent component analysis approach to investigate genomic influence on brain function. *IEEE Signal Process.Lett.* 15, 413–416. doi: 10.1109/LSP.2008.922513

Liu, J., Wang, X., Zhang, X., Pan, Y., Wang, X., and Wang, J. (2018). MMM: classification of schizophrenia using multi-modality multi-atlas feature representation and multi-kernel learning. *Multim. Tools Appl.* 77, 29651–29667. doi: 10.1007/s11042-017-5470-7

Liu, L., Chen, S., Zhu, X., Zhao, X.-M., and Wang, J. (2019). Deep convolutional neural network for accurate segmentation and quantification of white matter hyperintensities. *Neurocomputing* 384, 231–242. doi: 10.1016/j.neucom.2019.12.050

Liu, Z., Zhang, J., Xie, X., Rolls, E. T., Sun, J., Zhang, K., et al. (2018). Neural and genetic determinants of creativity. *Neuroimage* 174, 164–167. doi: 10.1016/j.neuroimage.2018.02.067

Madeira, N., Duarte, J. V., Martins, R., Costa, G. N., Macedo, A., and Castelo-Branco, M. (2020). Morphometry and gyrification in bipolar disorder and schizophrenia: a comparative MRI study. *NeuroImage* 26, 102220. doi: 10.1016/j.nicl.2020.102220

McIntosh, A. M., Job, D. E., Moorhead, T. W. J., Harrison, L. K., Lawrie, S. M., and Johnstone, E. C. (2005). White matter density in patients with schizophrenia, bipolar disorder and their unaffected relatives. *Biol. Psychiatry* 58, 254–257. doi: 10.1016/j.biopsych.2005.03.044

Mill, R. D., Winfield, E. C., Cole, M. W., and Ray, S. (2021). Structural MRI and functional connectivity features predict current clinical status and persistence behavior in prescription opioid users. *NeuroImage* 30, 102663. doi: 10.1016/j.nicl.2021.102663

Morgado, P. M., Silveira, M., for the Alzheimer's Disease Neuroimaging Initiative (2015). Minimal neighborhood redundancy maximal relevance: application to the diagnosis of Alzheimer' s disease. *Neurocomputing* 155, 295–308. doi: 10.1016/j.neucom.2014.12.070

Olesen, P. J., Nagy, Z., Westerberg, H., and Klingberg, T. (2003). Combined analysis of DTI and fMRI data reveals a joint maturation of white and grey matter in a fronto-parietal network. *Cogn. Brain Res.* 18, 48–57. doi: 10.1016/j.cogbrainres.2003.09.003

Poldrack, R. A., Barch, D. M., Mitchell, J., Wager, T., Wagner, A. D., Devlin, J. T., et al. (2013). Toward open sharing of task-based fMRI data: the openfMRI project. *Front. Neuroinformatics* 7, 12. doi: 10.3389/fninf.2013.00012

Poldrack, R. A., Congdon, E., Triplett, W., Gorgolewski, K. J., Karlsgodt, K. H., Mumford, J. A., et al. (2016). A phenome-wide examination of neural and cognitive function. *Sci. Data* 3, 160110. doi: 10.1038/sdata.2016.110

Qiao, C., Lu, L., Yang, L., and Kennedy, P. J. (2019). Identifying brain abnormalities with schizophrenia based on a hybrid feature selection technology. *Appl. Sci.* 9, 2148. doi: 10.3390/app9102148

Rakić, M., Cabezas, M., Kushibar, K., Oliver, A., and Llado, X. (2020). Improving the detection of autism spectrum disorder by combining structural and functional MRI information. *NeuroImage* 25, 102181. doi: 10.1016/j.nicl.2020.102181

Rosenberg, M. D., Casey, B. J., and Holmes, A. J. (2018). Prediction complements explanation in understanding the developing brain. *Nat. Commun.* 9, 589. doi: 10.1038/s41467-018-02887-9

Rubia, K., Smith, A. B., Brammer, M. J., and Taylor, E. (2007). Temporal lobe dysfunction in medication-nave boys with attention-deficit/hyperactivity disorder during attention allocation and its relation to response variability. *Biol. Psychiatry* 62, 999–1006. doi: 10.1016/j.biopsych.2007.02.024

Seghier, M. L., Lazeyras, F., Zimine, S., Maier, S. E., Hanquinet, S., Delavelle, J., et al. (2004). Combination of event-related fMRI and diffusion tensor imaging in an infant with perinatal stroke. *Neuroimage* 21, 463–472. doi: 10.1016/j.neuroimage.2003.09.015

Strasser, H. C., Lilyestrom, J., Ashby, E. R., Honeycutt, N. A., Schretlen, D. J., Pulver, A. E., et al. (2005). Hippocampal and ventricular volumes in psychotic and nonpsychotic bipolar patients compared with schizophrenia patients and community control subjects: a pilot study. *Biol. Psychiatry* 57, 633–639. doi: 10.1016/j.biopsych.2004.12.009

Su, C., Xu, Z., Pathak, J., and Wang, F. (2020). Deep learning in mental health outcome research: a scoping review. *Transl. Psychiatry* 10, 1–26. doi: 10.1038/s41398-020-0780-3

Sui, J., Adali, T., Pearlson, G. D., and Calhoun, V. D. (2009). An ICA-based method for the identification of optimal fMRI features and components using combined group-discriminative techniques. *Neuroimage* 46, 73–86. doi: 10.1016/j.neuroimage.2009.01.026

Sui, J., Pearlson, G., Caprihan, A., Adali, T., Kiehl, K. A., Liu, J., et al. (2011). Discriminating schizophrenia and bipolar disorder by fusing fMRI and DTI in a multimodal CCA+ joint ICA model. *Neuroimage* 57, 839–855. doi: 10.1016/j.neuroimage.2011.05.055

Wang, F., Kalmar, J. H., He, Y., Jackowski, M., Chepenik, L. G., Edmiston, E. E., et al. (2009). Functional and structural connectivity between the perigenual anterior cingulate and amygdala in bipolar disorder. *Biol. Psychiatry* 66, 516–521. doi: 10.1016/j.biopsych.2009.03.023

Willemink, M. J., Koszek, W. A., Hardell, C., Wu, J., Fleischmann, D., Harvey, H., et al. (2020). Preparing medical imaging data for machine learning. *Radiology* 295, 4–15. doi: 10.1148/radiol.2020192224

Yu, Y., Li, M., Liu, L., Li, Y., and Wang, J. (2019). Clinical big data and deep learning: applications, challenges and future outlooks. *Big Data Mining Analyt*. 2, 288–305. doi: 10.26599/BDMA.2019.9020007

Zhang, L., Ai, H., Opmeer, E. M., Marsman, J. C., Der Meer, L. V., Ruhe, H. G., et al. (2020). Distinct temporal brain dynamics in bipolar disorder and schizophrenia during emotion regulation. *Psychol. Med*. 50, 413–421. doi: 10.1017/S0033291719000217

Zhang, T., Liao, Q., Zhang, D., Zhang, C., Yan, J., Ngetich, R., et al. (2021). Predicting MCI to ad conversation using integrated sMRI and RS-fMRI: machine learning and graph theory approach. *Front. Aging Neurosci.* 13, 688926. doi: 10.3389/fnagi.2021.688926

Zhuang, H., Liu, R., Wu, C., Meng, Z., Wang, D., Liu, D., et al. (2019). Multimodal classification of drug-naïve first-episode schizophrenia combining anatomical, diffusion and resting state functional resonance imaging. *Neurosci. Lett*. 705, 87–93. doi: 10.1016/j.neulet.2019.04.039

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Neuroscience

ORIGINAL RESEARCH
published: 02 June 2022
doi: 10.3389/fnins.2022.918623

CAU-Net: A Deep Learning Method for Deep Gray Matter Nuclei Segmentation

Chao Chai[1], Mengran Wu[2], Huiying Wang[3], Yue Cheng[1], Shengtong Zhang[3], Kun Zhang[1], Wen Shen[1], Zhiyang Liu[2,4] and Shuang Xia[1]**

[1] Department of Radiology, Tianjin Institute of Imaging Medicine, Tianjin First Central Hospital, School of Medicine, Nankai University, Tianjin, China, [2] College of Electronic Information and Optical Engineering, Nankai University, Tianjin, China, [3] School of Medicine, Nankai University, Tianjin, China, [4] Tianjin Key Laboratory of Optoelectronic Sensor and Sensing Network Technology, Tianjin, China*

The abnormal iron deposition of the deep gray matter nuclei is related to many neurological diseases. With the quantitative susceptibility mapping (QSM) technique, it is possible to quantitatively measure the brain iron content *in vivo*. To assess the magnetic susceptibility of the deep gray matter nuclei in the QSM, it is mandatory to segment the nuclei of interest first, and many automatic methods have been proposed in the literature. This study proposed a contrast attention U-Net for nuclei segmentation and evaluated its performance on two datasets acquired using different sequences with different parameters from different MRI devices. Experimental results revealed that our proposed method was superior on both datasets over other commonly adopted network structures. The impacts of training and inference strategies were also discussed, which showed that adopting test time augmentation during the inference stage can impose an obvious improvement. At the training stage, our results indicated that sufficient data augmentation, deep supervision, and nonuniform patch sampling contributed significantly to improving the segmentation accuracy, which indicated that appropriate choices of training and inference strategies were at least as important as designing more advanced network structures.

Keywords: convolutional neural network (CNN), deep learning, medical image segmentation, gray matter nuclei, quantitative susceptibility mapping, strategically acquired gradient echo (STAGE) imaging

OPEN ACCESS

Edited by:
Yao Wu,
Children's National Hospital, United States

Reviewed by:
Xiaojun Guan,
Zhejiang University, China
Mingyuan Jiu,
Zhengzhou University, China

****Correspondence:***
Zhiyang Liu
liuzhiyang@nankai.edu.cn
Shuang Xia
xiashuang77@163.com

Specialty section:
This article was submitted to Brain Imaging Methods, a section of the journal Frontiers in Neuroscience

Received: *12 April 2022*
Accepted: *03 May 2022*
Published: *02 June 2022*

Citation:
Chai C, Wu M, Wang H, Cheng Y, Zhang S, Zhang K, Shen W, Liu Z and Xia S (2022) CAU-Net: A Deep Learning Method for Deep Gray Matter Nuclei Segmentation. Front. Neurosci. 16:918623. doi: 10.3389/fnins.2022.918623

INTRODUCTION

In the last decade, the advent of the quantitative susceptibility mapping (QSM) technique can achieve the quantitative measurement of brain iron content *in vivo* (Langkammer et al., 2010; Liu et al., 2015, 2017). QSM employed the magnetic susceptibility of tissue as the inherent physical magnetic resonance imaging (MRI) parameter, which indicated how the local magnetic field in tissues changes when an external magnetic field is applied (Li et al., 2019). Magnetic susceptibility of tissue can provide unique information of tissue iron composition (Li et al., 2019). Compared with other iron-sensitive techniques, including the transverse relaxation rates (R_2, R_2*, and R_2'), field-dependent rate increase, phase information from susceptibility-weighted imaging (SWI), and magnetic field correlation imaging, QSM can overcome the limitations of these techniques, such as the relatively low accuracy of R_2* due to other confounding factors (water content and calcium),

geometry- and orientation-dependence of phase images, and low sensitivity to small changes in brain iron (Stankiewicz et al., 2007; Bilgic et al., 2012; Deistung et al., 2013; Chai et al., 2019). QSM was more accurate in measuring the iron content and strongly correlated with the iron concentration of postmortem brain tissues (Langkammer et al., 2010).

The quantitative measurement of brain iron content using QSM has brought into focus the role of iron in the brain development, physical function modulation, and aging (Salami et al., 2018; Peterson et al., 2019), as well as in various neurological diseases, including Alzheimer's disease, Parkinson's disease, multiple sclerosis, metabolic diseases (hepatic encephalopathy and renal encephalopathy), sleep disorders, hematological system diseases, and cerebrovascular diseases (Chai et al., 2015a; Xia et al., 2015; Miao et al., 2018; Chai et al., 2019; Valdés Hernández et al., 2019; Pudlac et al., 2020; Cogswell et al., 2021; Thomas et al., 2021; Zhang et al., 2021; Galea et al., 2022). As iron has been proned to accumulate in the gray matter nuclei in normal people and all these neurological diseases have abnormal iron deposition in the gray matter nuclei, the gray matter nuclei are the critical target structures to explore the abnormal iron deposition. Previous studies have found that routine structural MR images such as T_1-weighted images could hardly show iron-rich gray matter nuclei clearly, such as substantia nigra (SN), red nucleus (RN), and dentate nucleus (DN; Beliveau et al., 2021). Therefore, these nuclei were not found in the most popular brain atlas, including FreeSurfer, FMRIB Software Library (FSL), and Statistical Parametric Mapping (SPM). Most segmentation tools cannot extract these nuclei (Beliveau et al., 2021). However, all the gray matter nuclei, including SN, RN, and DN, showed the obvious contrast (high signal) relative to the surrounding brain tissues in the QSM images because QSM was very sensitive to the iron, even when the amount was small and QSM can also enhance the iron-related contrast (Beliveau et al., 2021). The apparent contrast can help to identify the gray matter nuclei clearly and accurately. The measurement of iron content needs to manually outline the volumes of interest (VOIs) of the gray matter nuclei, which heavily depend on the operator's experience and cause some bias (Chai et al., 2022). The manual drawing of VOIs was also a tedious task and consumed an amount of time, which limited the wide application beyond research interest. To date, one study has used the SWI as the target modality because SWI can provide the enhanced contrast to visualize the gray matter nuclei compared to the other iron-sensitive modalities besides QSM and SWI also has a wide range of clinical applications (Beliveau et al., 2021). However, it was not far from enough to visualize and segment the nuclei using SWI, and the quantitative measurement of iron content was also a very critical step for the clinical evaluation of abnormal iron deposition for the diagnosis of neurological diseases. Therefore, QSM as the target modality can provide the enhanced contrast as good as SWI and directly quantitatively provide the information about iron content (Liu et al., 2015).

Deep learning has recently been successfully applied in biomedical image segmentation tasks (Minaee et al., 2021). It has been shown that, in many medical image segmentation tasks, such as tumor segmentation (Menze et al., 2015; Chang et al., 2018), stroke lesion segmentation (Maier et al., 2017; Liu et al., 2018), and organ segmentation (Gibson et al., 2018), deep learning methods were able to significantly exceed the conventional atlas-based methods. Most deep-learning-based medical image segmentation tasks adopted the U-Net (Ronneberger et al., 2015) or its variants (Cicek et al., 2016; Chang et al., 2018; Liu et al., 2018; Meng et al., 2018; Wang et al., 2020). By introducing dense skip connections between the encoder and decoder layers, U-Net like structures were able to effectively fuse the spatial and semantic information even when the training set was small. To further improve the segmentation accuracy of U-Net, some modifications at the encoder part or at the skip connections were proposed in the literature. The modications at the encoder mainly focused on making the encoders wider (Chen et al., 2019; Wang et al., 2019; Ibtehaz and Rahman, 2020), so as to enrich the feature maps from multiple fields of view. At the skip connections, the modifications were applied by incorporating various attention mechanisms to guide the decoder to utilize the most essential features (Oktay et al., 2018; Guo et al., 2021).

When applied to the brain gray matter nuclei segmentation task, deep learning methods have also been more robust and accurate than the atlas-based methods (Guan et al., 2021; Chai et al., 2022). For instance, Chai et al. (2022) proposed a double-branch U-Net structure for gray matter nuclei segmentation in the QSM images, which incorporated the local feature maps from image patches with the original resolutions and the global feature maps from down-sampled image patches and presented high accuracy in nuclei segmentation with a light-weighted neural network. Guan et al. (2021) also developed a segmentation method known as DeepQSMSeg to segment five pairs of nuclei, including CN, PUT, GP, SN, and RN in the QSM images, which incorporated the spatial-wise and channel-wise attention mechanism into the U-Net architecture.

Most deep-learning methods mainly focused on proposing novel network architectures, and most of them were developed based on U-Net. The training strategies, however, were not emphasized. In this study, we attempted to emphasize not only the network structures, but also the importance in fine tuning the networks with appropriate training and inference strategies. In particular, we adopted a minor modification in the U-Net by introducing contrast attention (CA) modules at the skip connections and attempted to improve the segmentation accuracy without introducing additional network parameters. Experiments were conducted on two different datasets (Datasets I and II) with QSM acquired using different MRI sequences with different imaging parameters from different MRI devices. Dataset I was randomly split as a training set with 42 subjects and a test set with 20 subjects. The network was trained on the training set and evaluated on the test set and Dataset II. Experimental results revealed that on both datasets, the proposed method was able to overperform the other popular U-Net-shaped structures, including 3D U-Net (Cicek et al., 2016), Attention U-Net (Oktay et al., 2018), and DeepQSMSeg (Guan et al., 2021), which highlighted the ability of generalization of our proposed method. The effects of various training strategies were also discussed, which implied that data augmentation, deep supervision, and

nonuniform patch sampling were beneficial for improving the segmentation accuracy.

MATERIALS AND METHODS

Datasets

This prospective study was approved by the Tianjin First Central Hospital Review Board and Ethics Committee. The informed consent of all subjects was obtained before the MRI examination. Our study included two datasets acquired using different MRI sequences from different MRI devices, Dataset I with sixty-two healthy subjects (age range 22–60 years, mean age 37.34 ± 11.32 years; male 24 and female 38) and Dataset II with twenty-six healthy subjects (age range 54–72 years, mean age 62.44 ± 4.35 years; male 18 and female 9). All were enrolled from Tianjin First Central Hospital staff or community members by advertisement. The inclusion criteria were as follows: (1) the age of the subjects was 18 years or older; (2) the subjects had no MRI contraindications, including metal implant, pacemaker, or claustrophobia; (3) the subjects had no history of central nervous system diseases, including the cerebral infarction, cerebral hemorrhage, cerebral tumor, traumatic cerebral injury, or contusion, which might affect the segmentation of the cerebral structures. The exclusion criteria were as follows: (1) the subjects cannot finish the MRI scanning and acquire the available SWI images and 3D T_1-weighted images; (2) the subjects had the congenital abnormalities and above central nervous system diseases, which might affect the segmentation of the cerebral structures; (3) the quality of MRI images was not good for the post process and analysis.

Dataset I was randomly split into the training set and test set, with 42 (age range 22–55 years, mean age 36.6 ± 10.94 years; male 15 and female 27) and 20 subjects (age range 25–60 years, mean age 38.9 ± 12.22 years; male 9 and female 11), respectively. The training set was used to train the neural networks, while the test set was used to evaluate the performance. All subjects in Dataset II were used for evaluation.

MRI data of Dataset I included SW images and 3D T_1W images and were collected using a 3.0 T MRI scanner (Magnetom TIM TRIO scanner, Siemens Healthineers, Erlangen, Germany) equipped with an 8-channel phased-array head coil. The acquisition parameters of Dataset I were listed as follows: (1) the parameters of SWI: TR (time repetition)/TE (time echo) = 27/20 ms, number of slices = 56, FoV = 230 mm × 200 mm, voxel resolution = 0.5 mm × 0.5 mm × 2 mm, corresponding matrix sizes = 336 × 448 × 56, receiver bandwidth = 120 Hz/pixel, flip angle = 15°, and acquisition time = 334 s; (2) the parameters of 3D T_1WI: TR/TE = 1,900/2.52 ms, TI (time inversion) = 900 ms, number of slices = 176, FoV = 250 × 250 mm^2, voxel size = 1.0 mm × 1.0 mm × 1.0 mm, corresponding matrix sizes = 256 × 256 × 176, flip angle = 9°, and acquisition time = 258 s. MRI data of Dataset II were collected using another 3.0T MRI scanner (MAGNETOM Prisma, Siemens Healthcare, Erlangen, Germany) equipped with a 20-channel phased-array head coil.

The subjects of Dataset II had strategically acquired the gradient echo (STAGE)-MR angiography and venography (MRAV) sequence instead of the SWI sequence and also 3D T_1WI. The STAGE-MRAV sequence is a multi-parametric MRI sequence, which can be post-processed to acquire the QSM images directly. The acquisition parameters of Dataset II were listed as follows: (1) the parameters of STAGE sequence: TR/TE = 20/(2.5, 12.5) ms, matrix sizes = 384 × 288, flip angle = 12°, number of slices = 64, slice thickness = 2 mm, in-plane spatial resolution = 0.67 mm × 0.67 mm, FoV = 256 mm × 192 mm, receiver bandwidth/pixel = 240 Hz/pixel, and total acquisition time = 368 s; (2) the parameters of 3D T_1WI: TR/TE = 2,000/2.98 ms, TI = 900 ms, number of slices = 176, FoV = 256 mm × 248 mm, voxel size = 1.0 mm × 1.0 mm × 1.0 mm, corresponding matrix sizes = 256 × 248 × 176, flip angle = 9°, and acquisition time = 269 s.

Considering that the SWI and QSM images of STAGE-MRAV and 3D T_1WI were acquired using different parameters and different FoVs, we first registered the T_1WI images and the SWI images or QSM images of STAGE-MRAV using rigid affine transformation with mutual information as the criterion, and then resampled the T_1WI images using linear interpolation, so that the T_1WI image and its corresponding SWI image or QSM images of STAGE-MRAV were with the same spatial resolutions and matrix sizes.

The QSM images were reconstructed from the phase and magnitude images of SWI by employing the SMART software (Susceptibility Mapping and Phase Artifacts Removal Toolbox, Detroit, MI. The QSM images from the STAGE-MRAV sequence were acquired using the STAGE software (SpinTech Inc., MI, United States). The postprocessing steps of reconstruction of QSM have been reported in several studies (Chai et al., 2015b, 2022; Tang et al., 2020; Zhang et al., 2021). First, the elimination of the skull and other regions with low signals was performed using the Brain Extraction Tool (BET) in the FMRIB Software Library (FSL; Smith, 2002). Second, excluding the phase wraps in the original phase images was performed using a 3D best-path algorithm (Abdul-Rahman et al., 2007). Third, the elimination of the background phase information was performed using a sophisticated harmonic artifact reduction for the phase data (SHARP) algorithm (Schweser et al., 2011). Finally, the reconstruction of QSM images was performed using the truncated k-space division algorithm with a k-space threshold of 0.1 (Haacke et al., 2010).

Manual Annotation

The drawing of gray matter nuclei's volume of interest (VOI) in the QSM images was performed using the SPIN software (Signal Processing in Nuclear Magnetic Resonance, Detroit, MI, United States). The gray matter nuclei in our study included the bilateral caudate nuclei (CN), globus pallidus (GP), putamen (PUT), thalamus (THA), red nuclei (RN), substantia nigra (SN), and dentate nuclei (DN), as shown in **Figure 1**. These nuclei showed a high signal in the QSM images. Considering the personal difference in the shape and size of the nuclei in different people and in order to assure that the susceptibility values were assessed as accurately as possible for each subject, the VOIs were outlined manually on the contiguous slices of gray matter nuclei to include the whole volume of each nucleus by two well-trained

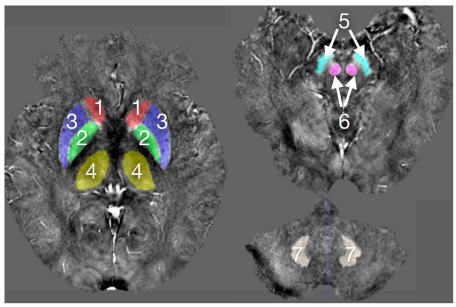


FIGURE 1 | The deep gray matter nuclei of interest outlined in the QSM images.1, CN; 2, GP; 3, PUT, 4, THA; 5, SN; 6, RN; and 7, DN.

neuroradiologists (C.C. and H.Y.W.) with 11 and 6 years of experience in neuroradiology who were blinded to the clinical and epidemiological information. When drawing the VOIs of the nuclei, we also magnified the images to obtain the more precise margin of nuclei. The topmost and lowermost slices of nuclei were excluded to eliminate the influence of edge partial volume effects. The susceptibility values of gray matter nuclei were presented as mean values and standard deviation.

Proposed Method

In this study, we employed both T1WI and QSM for nuclei segmentation, so as to utilize the high structural contrast of T1WI and the enhanced iron-related contrast of QSM. To better segment the nuclei, a contrast-attention U-Net (CAU-Net) was proposed for nuclei segmentation. In the classical U-Net, skip connections were employed to fuse the feature maps hierarchically with the decoder feature maps. In our proposed network, the CA module was added at the skip connections to encourage the network to extract the most prominent features and pass them to the decoder. As shown in **Figure 2**, the proposed CAU-Net employed a U-Net like structure in general, but made several significant modifications. The detailed hyperparameters, such as the numbers of filters and kernel sizes, can be found in **Figures 2**, **3**.

Contrast Attention

U-Net is the most successful network architecture in medical image segmentation, which fuses high-level and low-level features by skipping connections to obtain rich contextual information and precise location information.

Basically, to obtain accurate segmentation results, the network should be able to utilize both semantic and spatial information. In the encoder layers of a U-Net, the semantic information is extracted by many consecutive convolution layers, making it necessary to down-sample the feature maps to enlarge the FoVs of the convolution layers. In the decoder part, to recover the spatial information and generate accurate segmentation, it has to utilize both the semantic information from the deepest layer of the encoder and the spatial information from the shallower layers of the encoder. To generate a fine segmentation map, contour information and local details of images are meaningful for semantic segmentation. For instance, high-pass filters, such as Sobel and Laplacian operators, are widely used to extract the image's contour in image signal processing. Therefore, we assume that it is more important to pass the contour information to the decoder layers, instead of directly passing all output feature maps of the encoder layers to the decoder.

To cope with this problem, we added the CA at the skip connections of the U-Net, which can remove the identical information and extract the local differential information from the feature maps. **Figure 4** shows the structure of the CA module. The CA module does not include any parameter, and it is simply calculated as follows:

$$Y = X - Avg_3(X), \tag{1}$$

where $Avg_3(X)$ denotes the output of the average pooling layer with kernel size 3 and stride 1. It can be easily seen that the CA module works as a high-pass filter, which captures the local differential information and filters out the identical information from each feature map. It can also be interpreted as an implicit edge attention module, making the model better distinguish the edges of different tissues.

Training Strategy

Before training, both the T_1WI and the QSM images were normalized to zero mean and unit variance. The mean and variance values were calculated on all foreground regions of the training set. The T_1WI and corresponding QSM images were then concatenated to a dual-channel 3D image.

Due to limited GPU memory, cutting the whole volume into volumetric patches was necessary and commonly used in training 3D CNN segmentation networks. In our method, the whole volume was split into multiple patches with the size $128 \times 128 \times 32$. The patches were randomly sampled while ensuring that at least 2/3 of the patches were centered at the foreground voxels.

In our study, the training dataset size was significantly limited. Deep supervision was adopted to train the millions of network parameters and force the convolution layers to efficiently extract valuable features. In particular, a convolution layer with softmax activation was used at each stage of the decoder to generate a segmentation map, as shown in **Figure 2**. The deep supervision outputs were then up-sampled to the original size and the losses were computed. All deep supervision losses were summed up with the loss at the final output with equal weights, and the sum loss was used to update the network parameters. We used the same loss function at the deep supervision outputs and the final output, which was the sum of Dice loss and the cross entropy loss given as follows:

$$L\left(y,\hat{y}\right) = -\sum_{k=0}^{7}\sum_{i} y_{i,k}\,\log\hat{y}_{i,k} - \sum_{k=1}^{7}\frac{2\sum_i y_{i,k}\cdot\hat{y}_{i,k}}{\sum_i y_{i,k}+\sum_i \hat{y}_{i,k}}, \tag{2}$$

where $y_{i,k} \in \{0, 1\}$ denotes whether the i-th voxel was classified as the k-th class or not, and $\hat{y}_{i,k}$ denotes the value of the i-th voxel at the k-th channel of the network output. It is noted that we

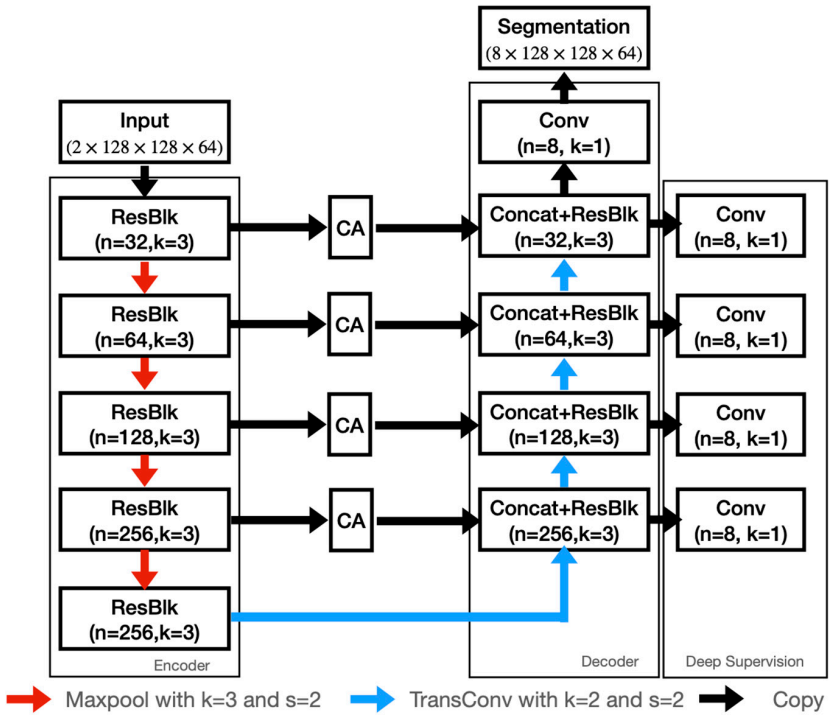


FIGURE 2 | The architecture of our proposed CAU-NET. The red arrows denote maxpooling, and the blue arrows denote the transposed convolution. The black arrows denote the copying of feature maps. "Concat" denotes the channel-wise concatenation. "ResBlk" denotes the residual block, whose structure is shown in **Figure 3**.

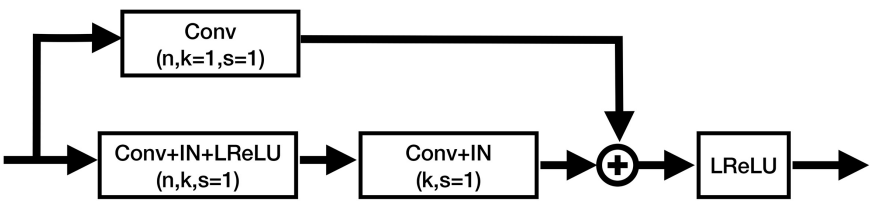


FIGURE 3 | Structure of the residual blocks in CAU-Net. Conv, ConvTrans, IN, and LReLU denote the convolution layer, transposed convolution layer, instance normalization, and Leaky ReLU activation function, respectively. "+" denotes the element-wise addition.

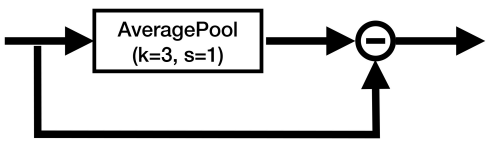


FIGURE 4 | Structure of the CA module. "−" denotes the element-wise subtraction.

TABLE 1 | Data augmentation methods adopted in our proposed method.

Data augmentation method	Probability	Parameter
Random flipping	0.5	Along with X, Y, and Z axes
Random zooming	0.2	Range: 0.7–1.3
Random rotation	0.2	Range: [−30,30]
Gaussian smoothing	0.1	σ=0.125
Random intensity rescale	0.2	Range: [0.9,1.1]

only computed the Dice loss of the foreground voxels because the numbers of the foreground and the background voxels were significantly imbalanced.

Sufficient data augmentation is another appropriate technique in dealing with the small training dataset. In our method, we used random zooming, random rotation within the range [−30, 30], random flipping, and random Gaussian smoothing to improve the data diversity. The detailed parameters of data augmentations are summarized in **Table 1**.

The network parameters were initialized as suggested by He et al. (2015). Stochastic gradient descent (SGD) with Nestrov trick was adopted as the optimizer. The momentum was set to

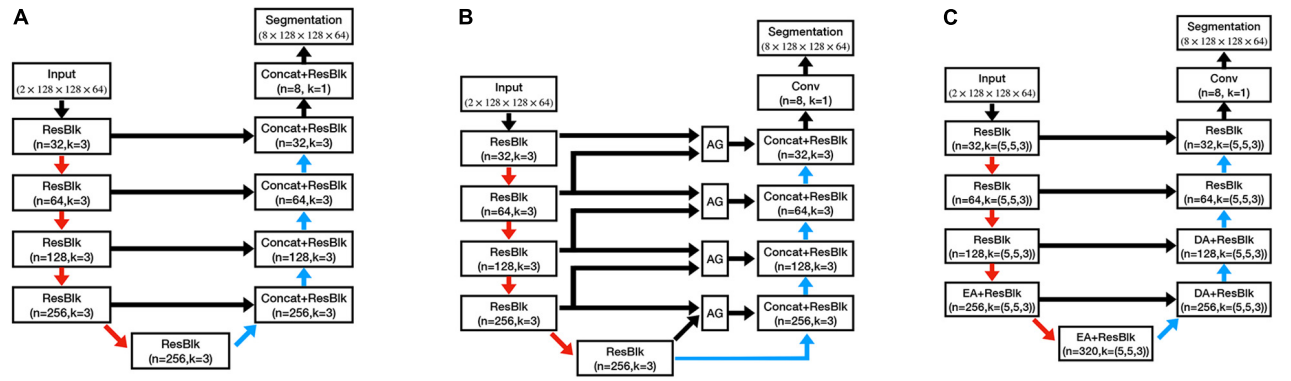


FIGURE 5 | Architectures of the comparative structures. **(A)** U-Net. **(B)** AU-Net. **(C)** DeepQSMSeg. AG denotes the attention gate. EA and DA denote the encoder attention and decoder attention, respectively.

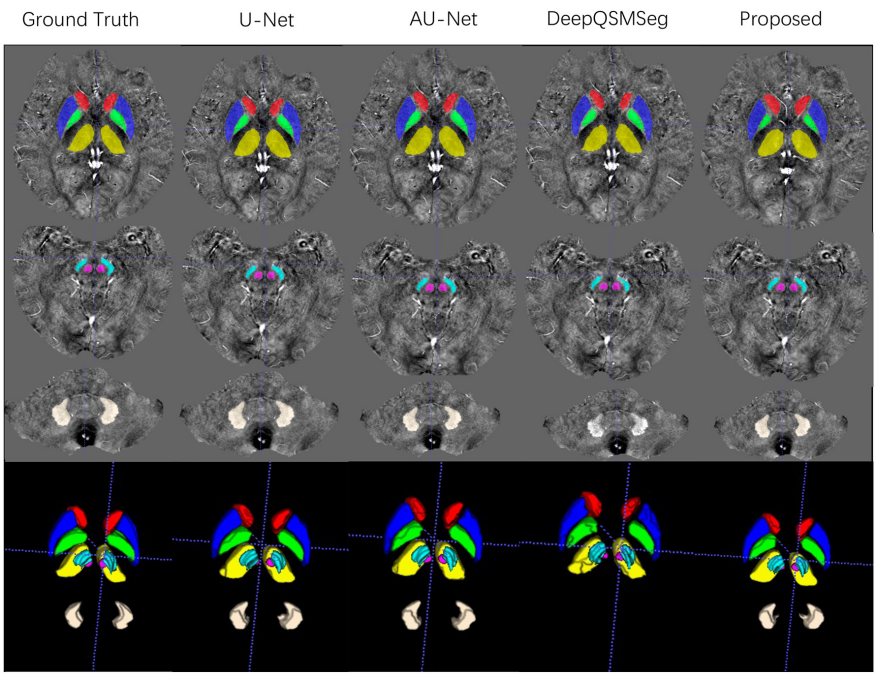


FIGURE 6 | Visualized examples of the segmentations of manual delineation, U-Net, AU-Net, DeepQSMSeg, and our proposed method on the test set of Dataset I. The segmentation results were overlaid on QSM images.

be 0.99, and the initial learning rate was set to be 0.01. To train the network parameters sufficiently, we trained the network for a sufficient number of updates. We defined an epoch as 250 batch iterations and trained the network for 500 epochs. The learning rate was adjusted after each epoch, and reduced in a polynomial way.

Inference Strategy

During inference, the patches of matrix size 128 × 128 × 32 were extracted from the image with the overlapping rate of 0.5. The whole segmentation map was constructed by combining the segmentations of all patches. Test time augmentation (TTA) was also adopted to further improve the segmentation accuracy. The augmentation included four procedures, namely, augmentation, prediction, disaugmentation, and merging. During the inference, to avoid introducing errors on the segmentation maps due to interpolation, we only used the augmentation methods without requiring interpolation. In particular, we adopted mirroring along all 3 axes and rotating ±90° and generated 8 augmented copies of the original image. We predicted on both the original and the augmented images, and then reverted the transformations on the predictions. Finally, we merged the predictions to generate the final prediction. In our study, we used the soft majority voting method to merge the multiple predictions.

Evaluation Metrics

In this study, we adopted both symmetric and surface distance metrics to evaluate the segmentation performance. In particular, the symmetric metric we adopted was Dice Coefficient (DC), which was defined as

$$DC_k = 2 \times \frac{|P_k \bigcap G_k|}{|P_k| + |G_k|},$$

where P_k and G_k denote the regions identified as the k-th class at the prediction and the ground truth, respectively. $|\cdot|$ denotes the area. DC measured how similar the two segmentation maps were.

In addition to DC, we further used surface distance metrics, including surface Dice Coefficient (SDC), Hausdorff distance (HD), and average symmetric surface distance (ASSD) to thoroughly evaluate the segmentation accuracy. These metrics were calculated based on the measurement of the surface distances, i.e., the distances between the surface points of the two segmentation volumes. Similar to the DC, the SDC was also defined as follows:

$$SDC_k = 2 \times \frac{|A \bigcap B|}{|A| + |B|},$$

where A and B are surface point sets of the prediction and the ground truth volumes, respectively, and the intersection between the two sets was measured with a given tolerance. In our study, we set the tolerance as 1 mm.

The HD measured the maximum distance between two volume surface points, which was defined as follows:

$$HD(A, B) = \max\left(\max_{a\in A}\min_{b\in B} d(a, b), \max_{b\in B}\min_{a\in A} d(a, b)\right).$$

To reduce the influence of some rare outliers, we used the 95% HD, denoted as HD95, which was obtained by measuring the 95th percentile value instead of the maximum value. The ASSD denoted the average distance between the volume surface points averaged over both directions, which was given as follows:

$$ASSD(A, B) = \frac{1}{2}\left(\frac{\sum_{a\in A}\min_{b\in B} d(a, b)}{|A|} + \frac{\sum_{b\in B}\min_{a\in A} d(b, a)}{|B|}\right)$$

Both HD95 and ASSD were given in *mm*, and the lower, the better. Unlike DC and SDC, the HD95 and ASSD worked equally well for large and small objects.

RESULTS

Comparative Methods

To evaluate the performance of our proposed method, we further trained three other models on the same training set, which were 3D U-Net (Cicek et al., 2016), 3D Attention U-Net (AU-Net;, and DeepQSMSeg (Guan et al., 2021). The network structures can be found in **Figure 5**, and the detailed structures of the attention modules can be found in their studies.

The U-Net had a similar structure to the CAU-Net, and the only difference was the absence of the CA modules. On the other hand, the AU-Net introduced an additive soft attention mechanism at the attention gates (AGs) at the skip connections of U-Net. The AG fused the feature maps from the current layer and the next lowest layer of the network to generate the attention weights for the most critical positions. DeepQSMSeg was a network structure specifically designed for nuclei segmentation from the QSM. It employed the basic encoder-decoder structure as U-Net, while inserting attention modules between the last two encoder stages and the first two decoder stages to capture

TABLE 2 | Numerical evaluation results on the test set.

Metric		CN	GP	PUT	THA	SN	RN	DN	Average
DC	U-Net	0.8232	0.8620	0.8582	0.8595	0.7142	0.8271	0.8050	0.8213
	AU-Net	0.8128	0.8472	0.8559	0.8563	0.7052	0.8418	**0.8056**	0.8178
	DeepQSMSeg	0.7551	0.8278	0.8137	0.7997	0.6391	0.7966	0	0.6617
	Proposed	**0.8306**	**0.8694**	**0.8629**	**0.8595**	**0.7161**	**0.8465**	0.7950	**0.8257**
SDC	U-Net	0.8169	0.8843	0.8596	0.76	0.8588	0.9354	0.8661	0.8544
	AU-Net	0.8033	0.8677	0.8584	0.7582	0.852	0.9431	**0.8701**	0.8504
	DeepQSMSeg	0.7167	0.8321	0.7968	0.6533	0.7864	0.911	0	0.6709
	Proposed	**0.8240**	**0.8913**	**0.8675**	**0.7648**	**0.8600**	**0.9515**	0.8549	**0.8591**
ASSD (mm)	U-Net	0.5183	0.3272	0.3881	0.5544	0.3794	0.2235	0.3614	0.3932
	AU-Net	0.5548	0.3725	0.3972	0.5620	0.3994	0.1934	**0.3540**	0.4047
	DeepQSMSeg	0.7622	0.4693	0.5448	0.8131	0.5503	0.2705	∞	/
	Proposed	**0.4995**	**0.3055**	**0.3676**	**0.5505**	**0.3763**	**0.1837**	0.3992	**0.3832**
HD95 (mm)	U-Net	2.7923	1.6541	1.9375	2.1780	2.1834	1.2474	**1.9552**	1.9926
	AU-Net	2.9685	1.8428	1.9730	2.1899	2.2985	1.1228	2.0141	2.0585
	DeepQSMSeg	3.5513	2.3229	2.5590	2.9516	2.9403	1.3858	∞	/
	Proposed	**2.6380**	**1.5405**	**1.8432**	**2.1323**	**2.0489**	**1.0143**	2.1697	**1.9124**

The most prominent value for each metric is highlighted in bold font. DeepQSMSeg fails in segmenting DN on all subjects.

the small target structures' semantic features. In each attention module of DeepQSMSeg, the channel-wise attention and spatial-wise attention were consecutively used to exploit both the channel and spatial relationships, and guide the decoder to generate a finer segmentation.

In our study, as the U-Net and AU-Net were not specifically designed for nuclei segmentation, we used the same training strategy as our proposed one. For DeepQSMSeg, we strictly followed the training protocol introduced by Guan et al. (2021).

Implementation Setup

The experiments were performed on a workstation with an Intel Core i7-7700K CPU, 64GB RAM, and Nvidia Geforce GTX 1080Ti GPU with 11GB memory. The workstation operated on Linux (Ubuntu 18.04 LTS) with CUDA 11.1. The networks were implemented on PyTorch (Paszke et al., 2019) v1.9.0 and trained using the framework of monai (MONAI Consortium, 2021) v0.6.0. The MR image files were stored as Neuroimaging Informatics Technology Initiative (NIfTI) format, and processed using a Simple Insight Toolkit (SimpleITK; Lowekamp et al., 2013) v2.1.0. The visualized results were presented using ITK-SNAP (Yushkevich et al., 2006) v3.8.0.

Results on the Test Set

We evaluated the segmentation performance on the test dataset with 20 subjects. **Figure 6** presents the visualized examples on a randomly chosen subject. As we can see, U-Net, AU-Net, and our proposed CAU-Net achieved better performance than DeepQSMSeg. The DeepQSMSeg failed in identifying DNs on all subjects. To further compare the results, we presented the

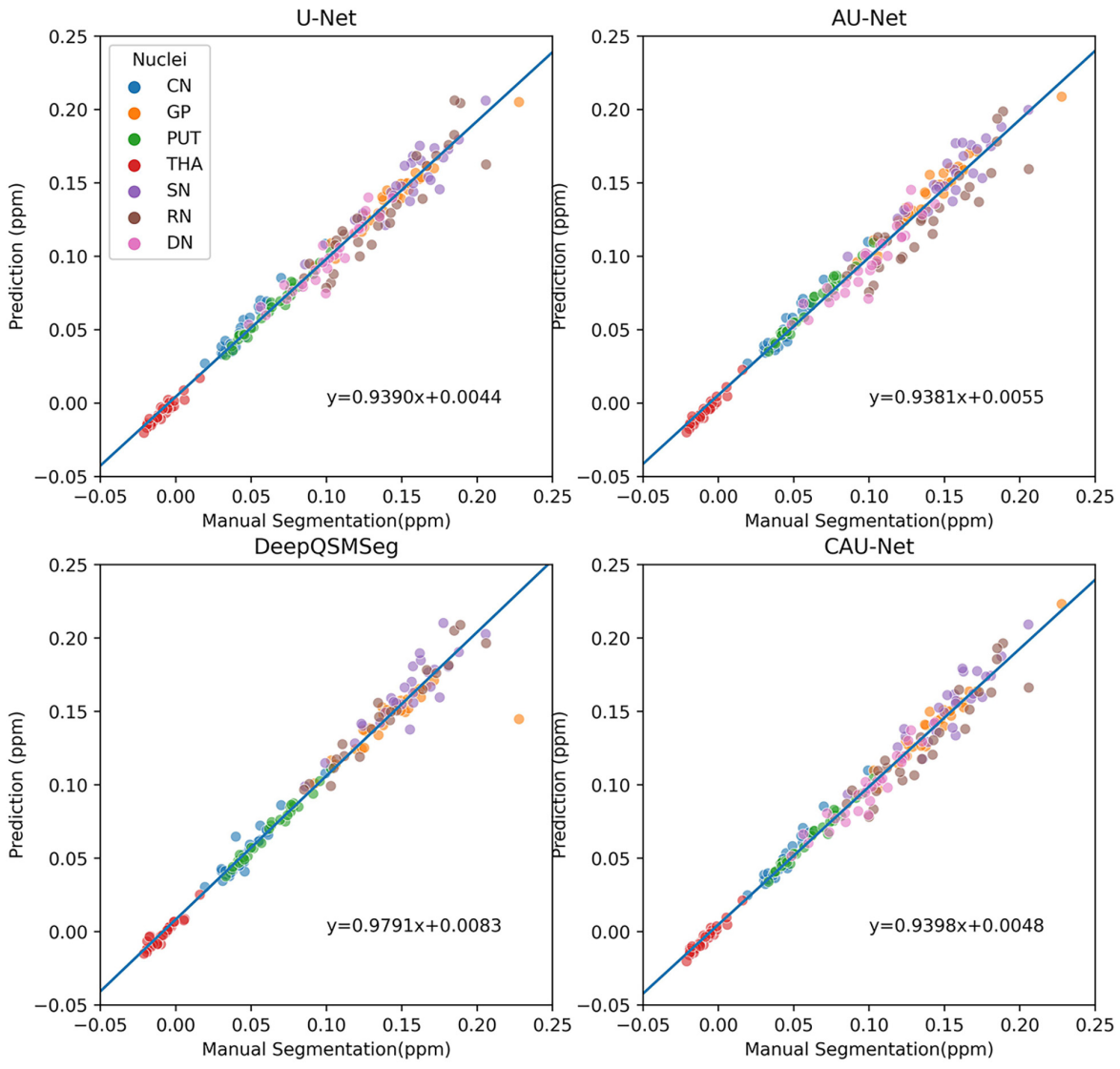


FIGURE 7 | Scatter plot of susceptibility values measured from manual segmentations and automatic segmentations on the subjects of the test set of Dataset I. The correlation lines are also plotted. For DeepQSMSeg, we omitted the results on DN.

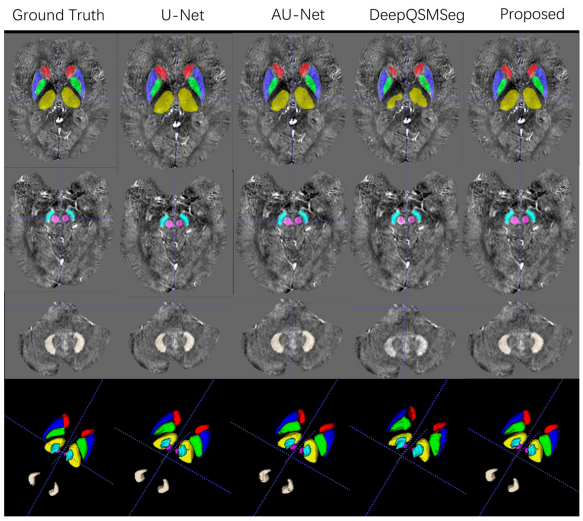


FIGURE 8 | Visualized examples of the segmentations of manual delineation, U-Net, AU-Net, DeepQSMSeg, and our proposed method on Dataset II. The segmentation results were overlaid on QSM images.

DC, SDC, ASSD and HD95 of our interested deep nuclei in **Table 2**. To compare the overall performance, the mean value of the metrics over all 7 nuclei was also presented in the last column.

As we can see from **Table 2**, our proposed CAU-Net achieved the best segmentation accuracy on all nuclei except DN. On DN, our proposed method was slightly worse than AU-Net. Across the nuclei, all methods achieved lower DC and SDC values on SN. The main reason was that the DC was more sensitive to small objects, while ASSD and HD95 were equally sensitive to small and large objects.

To further show the accuracy in segmenting each nucleus of each subject, we plotted scatter maps to show the correlations between the ground truth and the predictions in terms of the measured susceptibility values, as shown in **Figure 7**. Our proposed method presented the highest correlation with the manual delineations, while the DeepQSMSeg presented the lowest. As the DeepQSMSeg was originally developed on a large QSM dataset with 631 subjects, which is much larger than ours, it did not include as many data augmentation approaches as we did, and it may lead to the performance reduction compared with that reported by Guan et al. (2021).

Results on Dataset II

All subjects in Dataset II were used as an additional test set. We adopted the networks trained on the training set of Dataset I to generate the segmentation maps on the subjects in Dataset II. **Figure 8** presents some visualized examples. As we can see from **Figure 8**, most methods presented good segmentation accuracy on Dataset II. For DeepQSMSeg, it can still not segment the DN out, which implies that it may not be able to learn the features of DN.

To better investigate the segmentation accuracy, the segmentation metrics are summarized in **Table 3**. As we can see, our proposed method achieved the best performance in terms of mean DC, mean SDC, and mean ASSD. The U-Net performed the best on HD95, while our proposed CAU-Net had a very close performance. Our proposed method achieved the highest segmentation accuracy on CN, GP, PUT, and RN, while U-Net is the best on THA and DN.

As Dataset II was acquired by using a different machine with different parameters from the training set, the overall performance of all methods degraded compared to their performance on the test set of Dataset I. Interestingly, when segmenting DN, all methods had better accuracy on Dataset II, which implies that the QSMs stemmed from STAGE had better tissue contrast on DN.

The correlations of the measured susceptibility values between the manual segmentations and the automatic segmentations are also plotted in **Figure 9**. As shown in the figure,

TABLE 3 | Numerical evaluation results on the Dataset II.

Metric		CN	GP	PUT	THA	SN	RN	DN	Average
DC	U-Net	0.7694	0.8428	0.8502	**0.7920**	0.6724	0.7685	**0.8472**	0.7918
	AU-Net	0.7511	0.8507	0.8517	0.7811	0.6949	0.7725	0.8381	0.7914
	DeepQSMSeg	0.7129	0.8199	0.8232	0.7516	**0.7544**	0.7145	0	0.6538
	Proposed	**0.7792**	**0.8519**	**0.8561**	0.7782	0.6816	**0.7750**	0.8448	**0.7953**
SDC	U-Net	0.7741	0.8729	0.8919	**0.6565**	0.8199	**0.9030**	**0.9138**	0.8331
	AU-Net	0.7565	0.8860	0.8942	0.6370	0.8343	0.8983	0.9044	0.8301
	DeepQSMSeg	0.7194	0.8491	0.8559	0.6054	**0.8942**	0.8800	0.0000	0.6863
	Proposed	**0.7847**	**0.8866**	**0.9005**	0.6343	0.8258	0.8994	0.9109	0.8346
ASSD (mm)	U-Net	0.6605	0.3863	0.3536	**0.9056**	0.5126	0.2995	**0.2683**	0.4838
	AU-Net	0.7164	**0.3482**	0.3445	0.9447	0.4704	0.3055	0.2924	0.4889
	DeepQSMSeg	0.8148	0.4584	0.4326	1.0112	**0.3216**	0.3625	∞	/
	Proposed	**0.6246**	0.3535	**0.3334**	0.9515	0.4958	**0.2992**	0.2764	**0.4763**
HD95 (mm)	U-Net	2.9219	1.9752	1.8634	**2.8873**	2.7109	**1.5878**	**1.6721**	**2.2312**
	AU-Net	3.1908	**1.9015**	1.9278	2.9552	2.5247	1.6501	1.7910	2.2773
	DeepQSMSeg	3.3003	2.3239	2.1374	3.2131	**1.9669**	1.7993	∞	/
	Proposed	**2.9194**	2.0064	**1.8305**	2.9671	2.5369	1.6748	1.7071	2.2346

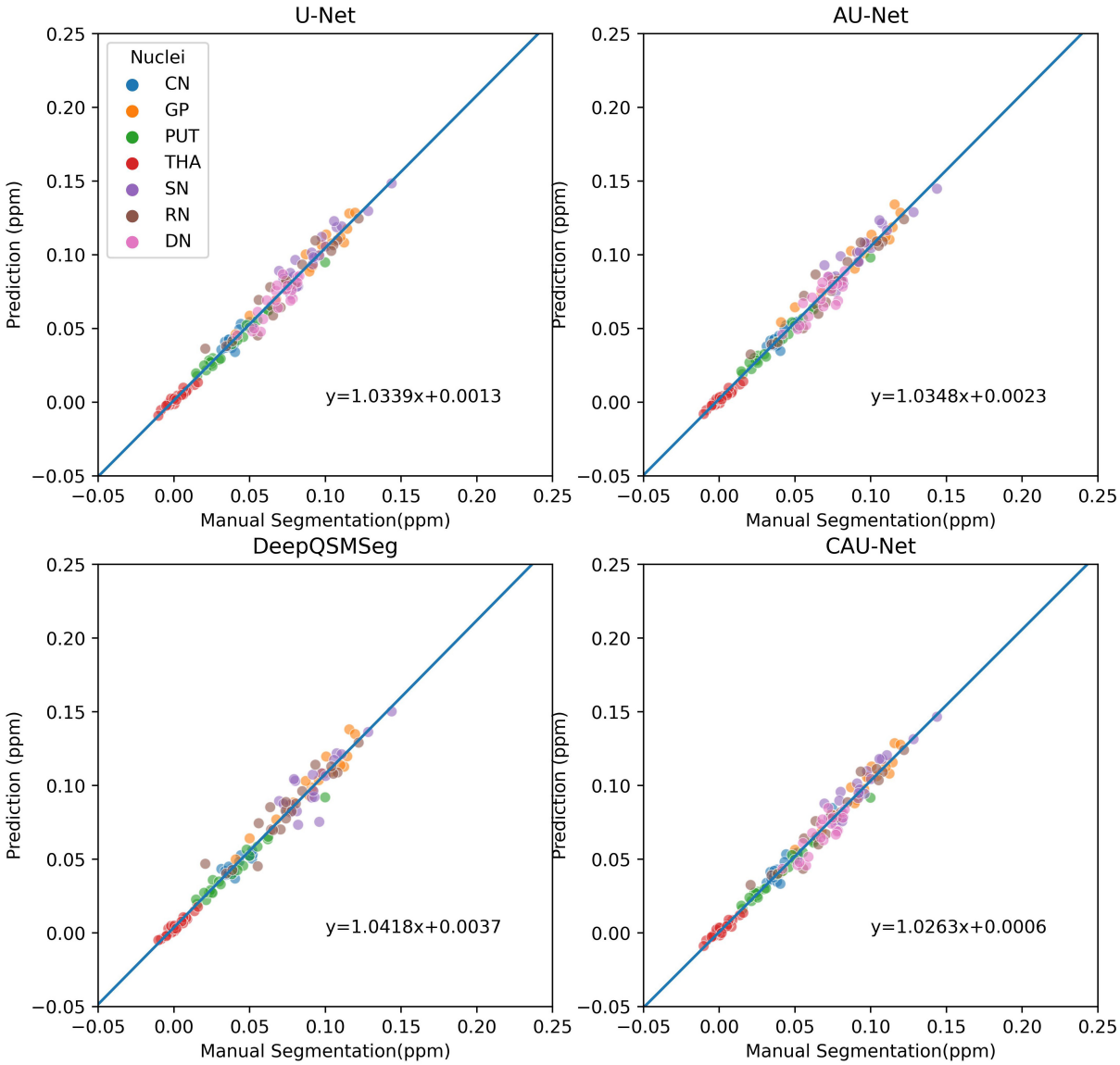


FIGURE 9 | Scatter plot of susceptibility values measured from manual segmentations and automatic segmentations on the subjects of the Dataset II. The correlation lines are also plotted. For DeepQSMSeg, we omitted the results on DN.

all segmentation methods presented a high agreement with the values measured on manual segmentations. DeepQSMSeg had the highest correlation, which, however, was calculated by omitting the DN. Our proposed method achieved the highest performance among the other three methods, which all successfully segmented all nuclei.

DISCUSSION

To further investigate the impact of various training and inference strategies on the segmentation accuracy, the segmentation performance under different training and inference settings is discussed.

Test Time Augmentation

In our proposed method, we adopted TTA to improve the segmentation accuracy during inference. To illustrate the impact of TTA on the segmentation accuracy, we evaluated the segmentation performance without using TTA as summarized in **Table 4**. To clearly illustrate the gain, the column "Delta" explicitly quantized the improvement on the average metrics of all 7 gray matter nuclei. As we can see from **Table 4**, U-Net, AU-Net, and CAU-Net had achieved prominent improvement on the DCs of all nuclei. DeepQSMSeg, on the other hand, suffered from performance loss when TTA was introduced. It implied that the DeepQSMSeg might overfit on the training set. On the other hand, our proposed method presented the most significant difference with and without TTA. It implied

that our proposed CA module might be able to filter irrelevant features from the encoder output feature maps, and had less risk in overfitting.

As **Table 4** shows, the TTA was an effective method for improving the DC. However, it is interesting to observe from **Table 4** that when TTA is adopted, the performance of U-Net and AU-Net became worse in terms of surface distance metrics, i.e., SDC, ASSD, and HD95. Our proposed CAU-Net, on the other hand, presented substantial performance improvement in all metrics when TTA was adopted. Such a phenomenon indicates that the CAU-Net could be much more robust to the input variations, and the generalization ability of the proposed CAU-Net is stronger than that of other comparative methods. The improvement in the generalization capability should be attributed to the high-pass filter nature of the CA module. By filtering out unnecessary information and only preserving the edge information on each feature map, the feature maps fed to the decoder layers were simplified by the CA module, making the decoder layers easier to utilize the spatial information.

The expense of using TTA was, however, the inference time. With TTA, as we had to make predictions on the original and augmented images, we had to take much more time for inference. For instance, in our study, as we generated 9 augmented images, the inference time with TTA would be 10 times that without TTA.

Training Strategy

This subsection would like to demonstrate the importance of properly designed training strategies. In our proposed method, we adopted data augmentation, deep supervision, and a nonuniform patch sampling strategy to train the neural network well. To demonstrate the effectiveness of training strategies, we trained the CAU-Net using different training setups. In particular, in the three experiments shown in **Table 5**, we removed the nonuniform patch sampling, deep supervision, and data augmentation as shown in **Table 1** to see their contributions to the segmentation accuracies. To make it clear, we only presented the average values of DC, SDC, ASSD, and HD95. As we can see, the segmentation accuracies reduced in all the three additional experiments, indicating that they contributed to improving the segmentation performance. The data augmentations contributed most to the DC, while the deep supervision was the most essential factor in terms of surface distance metrics.

In particular, data augmentation techniques have been shown to be one of the most essential approaches in image segmentations. It has been well known that data augmentation approaches have been beneficial for improving the classifiers' performance since the success of AlexNet. In our study, we used various data augmentation methods to improve the performance as summarized in **Table 1**. As shown in **Table 5**, after removing

TABLE 4 | Segmentation performance on the test set of Dataset I without adopting TTA during inference.

Metric		CN	GP	PUT	THA	SN	RN	DN	Average	Delta
DC	U-Net	0.8204	0.8594	0.8577	0.8593	0.7134	0.8215	0.8029	0.8192	−0.0021
	AU-Net	0.8112	0.8472	0.8533	0.852	0.7021	0.8421	0.7991	0.8153	−0.0025
	DeepQSMSeg	0.7943	0.8447	0.8348	0.8375	0.6669	0.8194	0	0.6854	+0.0237
	Proposed	0.8263	0.8669	0.8598	0.859	0.7097	0.8383	0.794	0.822	−0.0037
SDC	U-Net	0.8252	0.8897	0.8719	0.7765	0.8609	0.9342	0.8613	0.86	+0.0056
	AU-Net	0.8183	0.8753	0.87	0.7661	0.858	0.9453	0.8661	0.857	+0.0066
	DeepQSMSeg	0.77	0.8545	0.8299	0.7319	0.8098	0.9289	0	0.7036	+0.0327
	Proposed	0.8177	0.889	0.8641	0.7662	0.8562	0.9498	0.8528	0.8565	−0.0026
ASSD	U-Net	0.5045	0.3142	0.3646	0.5285	0.3736	0.2289	0.3708	0.3836	−0.0096
(mm)	AU-Net	0.519	0.3521	0.3681	0.5521	0.3872	0.1896	0.3651	0.3904	−0.0143
	DeepQSMSeg	0.6277	0.4268	0.4786	0.6453	0.4964	0.2361	∞	∞	/
	Proposed	0.5147	0.3133	0.3754	0.5514	0.3823	0.1957	0.4025	0.3908	+0.0076
HD95	U-Net	2.7117	1.5892	1.8551	2.108	2.0384	1.258	1.9874	1.9354	−0.0572
(mm)	AU-Net	2.7192	1.7431	1.8281	2.123	2.2067	1.1061	2.0545	1.9687	***−0.0898***
	DeepQSMSeg	3.1397	2.8627	2.2427	2.4159	2.4404	1.3513	∞	∞	/
	Proposed	2.6915	1.5994	1.8388	2.104	2.0481	1.1243	2.1912	1.9425	+0.0301

The Delta value indicates the difference with and without TTA.

TABLE 5 | Segmentation performance on the test set under different training strategies.

Method	DC	SDC	ASSD (mm)	HD95 (mm)
Proposed	0.8257	0.8591	0.3832	1.9124
w/o nonuniform patch sampling	0.8108 (−0.0149)	0.8455 (−0.0136)	0.4259 (+0.0427)	2.1384 (+0.2260)
w/o deep supervision	0.8048 (−0.0209)	0.8324 (−0.0267)	0.4500 (+0.0668)	2.2460 (+0.3336)
w/o data augmentation	0.7893 (−0.0364)	0.8461 (−0.0130)	0.4242 (+0.0410)	2.1782 (+0.2658)

the data augmentations from training, the DC significantly drops from 0.8257 to 0.7893. The main reason should be blamed for the small size of our dataset. When data augmentation was adopted, the methods listed in **Table 1** can generate many different versions of images, which increased the diversity of the training data and improved the representation ability of the neural network.

Deep supervision was also a critical approach that affected the performance. As **Table 5** shows, when deep supervision is absent, the DC drops to 0.0209 and the SDC drops to 0.0267. In our study, deep supervision is implemented by adding a convolution layer at each decoder stage to generate an auxiliary segmentation map. By incorporating additional classifier outputs at the middle stages, the network can be forced to learn effective representations to reduce the loss. Moreover, it also helped the deeper layers to be updated at the beginning of training, and was beneficial for improving the convergence behavior. By introducing deep supervision at the decoder layers, all the decoder layers were forced to extract spatial information from the skip connections. Combined with the edge information extracted from the CA module, much finer segmentation maps can be obtained as the decoder layers recover the feature maps to their original resolution.

The effect of the patch sampling scheme was also discussed. In our task, as it is not possible to feed the whole volume into the memory due to the limited memory size, splitting the images into patches was necessary. In our study, we chose to sample the patches with bias because the foreground voxels (i.e., the nuclei) and the background voxels are severely imbalanced. In particular, the sampling method ensured that at least two-third patches were centered at the foreground voxel during training. The nonuniform patch sampling method can also be regarded as an implicit way of data augmentation, which over-samples the foreground voxels to train the network. As we can see from **Table 5**, the segmentation performance was slightly improved by using the nonuniform patch sampling scheme. Despite that the contribution is relatively small compared to the contributions of deep supervision and data augmentation, the nonuniform patch sampling was shown to be able to further improve the performance with almost no additional computational cost.

After all, to improve the segmentation accuracy, it has been shown in our experiments that the training strategy was at least as important as developing more advanced networks. In our study, we can see that the performance gain of our proposed method came from several aspects, which are as follows: (1) the CA module that reduces the redundant information passed to the decoder layers; (2) the deep supervision's assistance in forcing the decoder layers to learn effective representations; (3) sufficient data augmentation and the bias patch sampling strategy to increase the diversity of patches; (4) TTA to utilize the ensembling of various predictions.

CONCLUSION

In our study, a deep-learning-based method was proposed for the gray matter nuclei segmentation task on T_1WI and QSM. A CA module was proposed and incorporated in the skip connections of U-Net to filter out the redundant information from the encoder feature maps. Experimental results on two test sets acquired with various parameters revealed that our proposed method could overperform all popular network structures. To investigate the origination of our proposed method, the results obtained under different training and inference strategies were also discussed, which implied that the appropriate choices of training and inference strategies were at least as important as developing more effective network structures.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Tianjin First Central Hospital Review Board and Ethics Committee. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

CC and ZL wrote this manuscript. MW and ZL implemented the programming work. CC and HW contributed to the data acquisition and data analysis. YC, SZ, and KZ contributed to the data evaluation and acquisition. WS and SX contributed to the study conception and design. ZL and SX contributed to the study conception and design, drafted and revised the article to provide important intellectual content, and agreed to be accountable for all aspects of this study. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Natural Science Foundation of China (NSFC) (grant nos. 81901728, 61871239, 81871342, and 81873888), Science and Technology Talent Cultivation Project of Tianjin Health and Family Planning Commission (grant no. RC20185), Young Talents Project of Tianjin Health Science and Technology (grant no. TJWJ2021QN011), and Tianjin Key Medical Discipline (Specialty) Construction Project.

ACKNOWLEDGMENTS

We acknowledge E. Mark Haacke from Wayne State University, who provided us the SWI and STAGE sequences and the postprocessing software to acquire the QSM images.

REFERENCES

Abdul-Rahman, H. S., Gdeisat, M. A., Burton, D. R., Lalor, M. J., Lilley, F., and Moore, C. J. (2007). Fast and robust three-dimensional best path phase unwrapping algorithm. *Appl. Opt.* 46, 6623–6635. doi: 10.1364/AO.46.006623

Beliveau, V., Nørgaard, M., Birkl, C., Seppi, K., and Scherfler, C. (2021). Automated segmentation of deep brain nuclei using convolutional neural networks and susceptibility weighted imaging. *Hum. Brain Mapp.* 42, 4809–4822. doi: 10.1002/hbm.25604

Bilgic, B., Pfefferbaum, A., Rohlfing, T., Sullivan, E. V., and Adalsteinsson, E. (2012). MRI estimates of brain iron concentration in normal aging using quantitative susceptibility mapping. *Neuroimage* 59, 2625–2635. doi: 10.1016/j.neuroimage.2011.08.077

Chai, C., Qiao, P., Zhao, B., Wang, H., Liu, G., Wu, H., et al. (2022). Brain gray matter nuclei segmentation on quantitative susceptibility mapping using dual-branch convolutional neural network. *Artif. Intell. Med.* 125:102255. doi: 10.1016/j.artmed.2022.102255

Chai, C., Wang, H., Liu, S., Chu, Z.-Q., Li, J., Qian, T., et al. (2019). Increased iron deposition of deep cerebral gray matter structures in hemodialysis patients: a longitudinal study using quantitative susceptibility mapping: brain iron overload in HD Patients. *J. Magn. Reson. Imaging* 49, 786–799. doi: 10.1002/jmri.26226

Chai, C., Yan, S., Chu, Z., Wang, T., Wang, L., Zhang, M., et al. (2015a). Quantitative measurement of brain iron deposition in patients with haemodialysis using susceptibility mapping. *Metab. Brain Dis.* 30, 563–571. doi: 10.1007/s11011-014-9608-2

Chai, C., Zhang, M., Long, M., Chu, Z., Wang, T., Wang, L., et al. (2015b). Increased brain iron deposition is a risk factor for brain atrophy in patients with haemodialysis: a combined study of quantitative susceptibility mapping and whole brain volume analysis. *Metab. Brain Dis.* 30, 1009–1016. doi: 10.1007/s11011-015-9664-2

Chang, J., Zhang, X., Chang, J., Ye, M., Huang, D., Wang, P., et al. (2018). "Brain tumor segmentation based on 3D Unet with multi-class focal loss," in *Proceedings of the 2018 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, Beijing, 1–5. doi: 10.1109/CISP-BMEI.2018.8633056

Chen, W., Liu, B., Peng, S., Sun, J., and Qiao, X. (2019). "S3D-UNet: separable 3D U-Net for brain tumor segmentation," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, eds A. Crimi, S. Bakas, H. Kuijf, F. Keyvan, M. Reyes, and T. van Walsum (Cham: Springer International Publishing), 358–368. doi: 10.1007/978-3-030-11726-9_32

Cicek, O., Abdulkadir, A., Lienkamp, S. S., Brox, T., and Ronneberger, O. (2016). 3D U-Net: learning dense volumetric segmentation from sparse annotation. *arXiv* [Preprint]. arXiv:1606.06650

Cogswell, P. M., Wiste, H. J., Senjem, M. L., Gunter, J. L., Weigand, S. D., Schwarz, C. G., et al. (2021). Associations of quantitative susceptibility mapping with Alzheimer's disease clinical and imaging markers. *Neuroimage* 224:117433. doi: 10.1016/j.neuroimage.2020.117433

Deistung, A., Schäfer, A., Schweser, F., Biedermann, U., Turner, R., and Reichenbach, J. R. (2013). Toward in vivo histology: a comparison of quantitative susceptibility mapping (QSM) with magnitude-, phase-, and R2*-imaging at ultra-high magnetic field strength. *Neuroimage* 65, 299–314. doi: 10.1016/j.neuroimage.2012.09.055

Galea, I., Durnford, A., Glazier, J., Mitchell, S., Kohli, S., Foulkes, L., et al. (2022). Iron deposition in the brain after aneurysmal subarachnoid hemorrhage. *Stroke* 53, 1633–1642. doi: 10.1161/STROKEAHA.121.036645

Gibson, E., Giganti, F., Hu, Y., Bonmati, E., Bandula, S., Gurusamy, K., et al. (2018). Automatic multi-organ segmentation on abdominal CT with dense V-networks. *IEEE Trans. Med. Imaging* 37, 1822–1834. doi: 10.1109/TMI.2018.2806309

Guan, Y., Guan, X., Xu, J., Wei, H., Xu, X., and Zhang, Y. (2021). "DeepQSMSeg: a deep learning-based sub-cortical nucleus segmentation tool for quantitative susceptibility mapping," in *Proceedings of the 2021 43rd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)*, Mexico, 3676–3679. doi: 10.1109/EMBC46164.2021.9630449

Guo, C., Szemenyei, M., Yi, Y., Wang, W., Chen, B., and Fan, C. (2021). "SA-UNet: spatial attention U-Net for retinal vessel segmentation," in *Proceedings of the 2020 25th International Conference on Pattern Recognition (ICPR)*, Milan, 1236–1242. doi: 10.1109/ICPR48806.2021.9413346

Haacke, E. M., Tang, J., Neelavalli, J., and Cheng, Y. C. N. (2010). Susceptibility mapping as a means to visualize veins and quantify oxygen saturation. *J. Magn. Reson. Imaging* 32, 663–676. doi: 10.1002/jmri.22276

He, K., Zhang, X., Ren, S., and Sun, J. (2015). "Delving deep into rectifiers: surpassing human-level performance on ImageNet classification," in *Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV)* (Santiago: IEEE), 1026–1034. doi: 10.1109/ICCV.2015.123

Ibtehaz, N., and Rahman, M. S. (2020). MultiResUNet?: rethinking the U-Net architecture for multimodal biomedical image segmentation. *Neural Netw.* 121, 74–87. doi: 10.1016/j.neunet.2019.08.025

Langkammer, C., Krebs, N., Goessler, W., Scheurer, E., Ebner, F., Yen, K., et al. (2010). Quantitative MR imaging of brain iron: a postmortem validation study. *Radiology* 257, 455–462. doi: 10.1148/radiol.10100495

Li, X., Chen, L., Kutten, K., Ceritoglu, C., Li, Y., Kang, N., et al. (2019). Multi-atlas tool for automated segmentation of brain gray matter nuclei and quantification of their magnetic susceptibility. *Neuroimage* 191, 337–349. doi: 10.1016/j.neuroimage.2019.02.016

Liu, C., Li, W., Tong, K. A., Yeom, K. W., and Kuzminski, S. (2015). Susceptibility-weighted imaging and quantitative susceptibility mapping in the brain: brain susceptibility imaging and mapping. *J. Magn. Reson. Imaging* 42, 23–41. doi: 10.1002/jmri.24768

Liu, S., Buch, S., Chen, Y., Choi, H.-S., Dai, Y., Habib, C., et al. (2017). Susceptibility-weighted imaging: current status and future directions: SWI review. *NMR Biomed.* 30:e3552. doi: 10.1002/nbm.3552

Liu, Z., Cao, C., Ding, S., Liu, Z., Han, T., and Liu, S. (2018). Towards clinical diagnosis: automated stroke lesion segmentation on multi-spectral MR image using convolutional neural network. *IEEE Access* 6, 57006–57016. doi: 10.1109/ACCESS.2018.2872939

Lowekamp, B. C., Chen, D. T., Ibáñez, L., and Blezek, D. (2013). The design of SimpleITK. *Front. Neuroinform.* 7:45. doi: 10.3389/fninf.2013.00045

Maier, O., Menze, B. H., von der Gablentz, J., Häni, L., Heinrich, M. P., Liebrand, M., et al. (2017). ISLES 2015 – a public evaluation benchmark for ischemic stroke lesion segmentation from multispectral MRI. *Med. Image Anal.* 35, 250–269. doi: 10.1016/j.media.2016.07.009

Meng, Z., Fan, Z., Zhao, Z., and Su, F. (2018). "ENS-Unet: end-to-end noise suppression U-Net for brain tumor segmentation," in *Proceedings of the 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Honolulu, HI, 5886–5889. doi: 10.1109/EMBC.2018.8513676

Menze, B. H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., et al. (2015). The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Trans. Med. Imaging* 34, 1993–2024. doi: 10.1109/TMI.2014.2377694

Miao, X., Choi, S., Tamrazi, B., Chai, Y., Vu, C., Coates, T. D., et al. (2018). Increased brain iron deposition in patients with sickle cell disease: an MRI quantitative susceptibility mapping study. *Blood* 132, 1618–1621. doi: 10.1182/blood-2018-04-840322

Minaee, S., Boykov, Y. Y., Porikli, F., Plaza, A. J., Kehtarnavaz, N., and Terzopoulos, D. (2021). Image segmentation using deep learning: a survey. *IEEE Trans. Pattern Anal. Mach. Intell.* 1. doi: 10.1109/TPAMI.2021.3059968 [Epub ahead of print].

MONAI Consortium (2021). *MONAI: Medical Open Network for AI*. Zenodo. doi: 10.5281/zenodo.5728262

Oktay, O., Schlemper, J., Folgoc, L. L., Lee, M., Heinrich, M., Misawa, K., et al. (2018). Attention U-Net: learning where to look for the pancreas. *arXiv* [Preprint]. arXiv:1804.03999

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., et al. (2019). "PyTorch: an imperative style, high-performance deep learning library," in *NIPS'19: Proceedings of the 33rd International Conference on Neural Information Processing Systems* (Red Hook, NY: Curran Associates, Inc).

Peterson, E. T., Kwon, D., Luna, B., Larsen, B., Prouty, D., De Bellis, M. D., et al. (2019). Distribution of brain iron accrual in adolescence: evidence from cross-sectional and longitudinal analysis. *Hum. Brain Mapp.* 40, 1480–1495. doi: 10.1002/hbm.24461

Pudlac, A., Burgetova, A., Dusek, P., Nytrova, P., Vaneckova, M., Horakova, D., et al. (2020). Deep gray matter iron content in neuromyelitis optica

and multiple sclerosis. *Biomed Res. Int.* 2020:6492786. doi: 10.1155/2020/6492786

Ronneberger, O., Fischer, P., and Brox, T. (2015). "U-Net: convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015 Lecture Notes in Computer Science*, eds N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi (Cham: Springer International Publishing), 234–241. doi: 10.1007/978-3-319-24574-4_28

Salami, A., Avelar-Pereira, B., Garzón, B., Sitnikov, R., and Kalpouzos, G. (2018). Functional coherence of striatal resting-state networks is modulated by striatal iron content. *Neuroimage* 183, 495–503. doi: 10.1016/j.neuroimage.2018.08.036

Schweser, F., Deistung, A., Lehr, B. W., and Reichenbach, J. R. (2011). Quantitative imaging of intrinsic magnetic tissue properties using MRI signal phase: an approach to in vivo brain iron metabolism? *Neuroimage* 54, 2789–2807. doi: 10.1016/j.neuroimage.2010.10.070

Smith, S. M. (2002). Fast robust automated brain extraction. *Hum. Brain Mapp.* 17, 143–155. doi: 10.1002/hbm.10062

Stankiewicz, J., Panter, S. S., Neema, M., Arora, A., Batt, C. E., and Bakshi, R. (2007). Iron in chronic brain disorders: imaging and neurotherapeutic implications. *Neurotherapeutics* 4, 371–386. doi: 10.1016/j.nurt.2007.05.006

Tang, R., Zhang, Q., Chen, Y., Liu, S., Haacke, E. M., Chang, B., et al. (2020). Strategically acquired gradient echo (STAGE)-derived MR angiography might be a superior alternative method to time-of-flight MR angiography in visualization of leptomeningeal collaterals. *Eur. Radiol.* 30, 5110–5119. doi: 10.1007/s00330-020-06840-7

Thomas, G. E. C., Zarkali, A., Ryten, M., Shmueli, K., Gil-Martinez, A. L., Leyland, L.-A., et al. (2021). Regional brain iron and gene expression provide insights into neurodegeneration in Parkinson's disease. *Brain* 144, 1787–1798. doi: 10.1093/brain/awab084

Valdés Hernández, M., Case, T., Chappell, F., Glatz, A., Makin, S., Doubal, F., et al. (2019). Association between striatal brain iron deposition, microbleeds and cognition 1 year after a minor ischaemic stroke. *Int. J. Mol. Sci.* 20:1293. doi: 10.3390/ijms20061293

Wang, B., Qiu, S., and He, H. (2019). "Dual encoding U-Net for retinal vessel segmentation," in *Proceedings of the Medical Image Computing and Computer Assisted Intervention – MICCAI 2019: 22nd International Conference, Shenzhen, China*, eds D. Shen, T. Liu, T. M. Peters, L. H. Staib, C. Essert, S. Zhou, et al. (Cham: Springer International Publishing), 84–92. doi: 10.1007/978-3-030-32239-7_10

Wang, S., Hu, S.-Y., Cheah, E., Wang, X., Wang, J., Chen, L., et al. (2020). U-Net using stacked dilated convolutions for medical image segmentation. *arXiv* [Preprint]. arXiv:2004.03466

Xia, S., Zheng, G., Shen, W., Liu, S., Zhang, L. J., Haacke, E. M., et al. (2015). Quantitative measurements of brain iron deposition in cirrhotic patients using susceptibility mapping. *Acta Radiol.* 56, 339–346. doi: 10.1177/0284185114525374

Yushkevich, P. A., Piven, J., Hazlett, H. C., Smith, R. G., Ho, S., Gee, J. C., et al. (2006). User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *Neuroimage* 31, 1116–1128. doi: 10.1016/j.neuroimage.2006.01.015

Zhang, X., Chai, C., Ghassaban, K., Ye, J., Huang, Y., Zhang, T., et al. (2021). Assessing brain iron and volume of subcortical nuclei in idiopathic rapid eye movement sleep behavior disorder. *Sleep* 44:zsab131. doi: 10.1093/sleep/zsab131

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Neuroscience

TYPE Original Research
PUBLISHED 01 August 2022
DOI 10.3389/fnins.2022.920765

OPEN ACCESS

EDITED BY
Yao Wu,
Children's National Hospital,
United States

REVIEWED BY
Yuan Zhou,
Institute of Psychology (CAS), China
Yong Liu,
Beijing University of Posts
and Telecommunications (BUPT),
China

*CORRESPONDENCE
Quan Zhang
tj_zhangquan1981@163.com
Meng Liang
liangmeng@tmu.edu.cn
Feng Gao
gaofeng@tju.edu.cn

SPECIALTY SECTION
This article was submitted to
Brain Imaging Methods,
a section of the journal
Frontiers in Neuroscience

RECEIVED 15 April 2022
ACCEPTED 01 July 2022
PUBLISHED 01 August 2022

CITATION
Hou A, Pang X, Zhang X, Peng Y, Li D, Wang H, Zhang Q, Liang M and Gao F (2022) Widespread aberrant functional connectivity throughout the whole brain in obstructive sleep apnea. *Front. Neurosci.* 16:920765. doi: 10.3389/fnins.2022.920765

Widespread aberrant functional connectivity throughout the whole brain in obstructive sleep apnea

Ailin Hou[1,2], Xueming Pang[2], Xi Zhang[2], Yanmin Peng[2], Dongyue Li[2], He Wang[2], Quan Zhang[3]*, Meng Liang[2]* and Feng Gao[1,4]*

[1]College of Precision Instrument and Optoelectronics Engineering, Tianjin University, Tianjin, China, [2]School of Medical Imaging and Tianjin Key Laboratory of Functional Imaging, Tianjin Medical University, Tianjin, China, [3]Department of Radiology, Characteristic Medical Center of Chinese People's Armed Police Force, Tianjin, China, [4]Tianjin Key Laboratory of Biomedical Detecting Techniques and Instruments, Tianjin University, Tianjin, China

Objective: Obstructive sleep apnea (OSA) is a sleep-related breathing disorder with high prevalence and is associated with cognitive impairment. Previous neuroimaging studies have reported abnormal brain functional connectivity (FC) in patients with OSA that might contribute to their neurocognitive impairments. However, it is unclear whether patients with OSA have a characteristic pattern of FC changes that can serve as a neuroimaging biomarker for identifying OSA.

Methods: A total of 21 patients with OSA and 21 healthy controls (HCs) were included in this study and scanned using resting-state functional magnetic resonance imaging (fMRI). The automated anatomical labeling (AAL) atlas was used to divide the cerebrum into 90 regions, and FC between each pair of regions was calculated. Univariate analyses were then performed to detect abnormal FCs in patients with OSA compared with controls, and multivariate pattern analyses (MVPAs) were applied to classify between patients with OSA and controls.

Results: The univariate comparisons did not detect any significantly altered FC. However, the MVPA showed a successful classification between patients with OSA and controls with an accuracy of 83.33% (p = 0.0001). Furthermore, the selected FCs were associated with nearly all brain regions and widely distributed in the whole brain, both within and between, many resting-state functional networks. Among these selected FCs, 3 were significantly correlated with the apnea-hypopnea index (AHI) and 2 were significantly correlated with the percentage of time with the saturation of oxygen (SaO_2) below 90% of the total sleep time (%TST < 90%).

Conclusion: There existed widespread abnormal FCs in the whole brain in patients with OSA. This aberrant FC pattern has the potential to serve as a

neurological biomarker of OSA, highlighting its importance for understanding the complex neural mechanism underlying OSA and its cognitive impairment.

KEYWORDS

resting-state functional magnetic resonance imaging, functional connectivity, multivariate pattern analyses, obstructive sleep apnea, machine learning

Introduction

Obstructive sleep apnea (OSA) is a common sleep-related breathing disorder resulting from obstruction of the upper airway, and the symptoms include snoring at night, frequent stop in breathing, and daytime sleepiness (Park et al., 2011). The major consequences of OSA contain intermittent nocturnal hypoxia and fragmented sleep (Verstraeten, 2007). About 936 million people worldwide between the ages of 30 and 69 years suffered from OSA, when the apnea-hypopnea index (AHI) $\geq$ 5/h criterion was used (Benjafield et al., 2019). OSA not only increases the risk of hypertension, cardiovascular disease and diabetes, as well as traffic accidents, but also shows an impairment of cognitive functions, such as attention, working memory, episodic memory, and executive function (Gagnon et al., 2014). Moreover, OSA has also been reported to be associated with psychological and neurological problems, such as depression, anxiety, post-traumatic stress disorder, and Alzheimer's disease (Gupta and Simpson, 2015; Vanek et al., 2020). Sleep fragmentation and intermittent nocturnal hypoxia are considered the main contributing factors to neuropsychological impairments in patients with OSA (Lim and Pack, 2014). However, the neural mechanisms are still largely unclear.

Resting-state functional magnetic resonance imaging (fMRI) provided a non-invasive and effective tool to explore the human brain. Functional connectivity (FC) was a commonly used technique for studying the neural mechanisms underlying cognitive impairment in patients with OSA. In the resting-state FC studies of OSA, researchers found abnormal FCs in patients with OSA associated with several brain regions such as insula (Zhang et al., 2015; Park et al., 2016a), hippocampus (HIP) (Song et al., 2018; Zhou et al., 2020a), amygdala (Yu et al., 2019), caudate nuclei (Song et al., 2018), and posterior cingulate cortex (PCC) (Qin et al., 2020). Besides, a fair amount of studies on OSA reported abnormal within-network and between-network FCs of resting-state brain functional networks (Khazaie et al., 2017), such as default mode network (DMN) (Zhang et al., 2013; Peng et al., 2014; Li et al., 2015, 2016b; Chen et al., 2018a), central executive network (CEN) (Zhang et al., 2013), and salience network (SN) (Yu et al., 2019). However, all these studies on the alterations of resting-state FCs in OSA were based on univariate analysis, i.e., comparing a single FC between patients and controls at a time and repeating this univariate comparison for every FC (i.e., a mass univariate analysis). Therefore, it is unclear whether OSA has a characteristic pattern of FC alterations which can serve as a neuroimaging biomarker for identifying OSA.

Multivariate pattern analysis (MVPA) is a machine learning technique that uses a pattern classifier to identify the specific spatial pattern of brain activities or connectivities in a particular experimental condition or a group of patients (Mur et al., 2009; Pereira et al., 2009). Unlike the mass univariate analysis employed in the above previous studies which only focused on one FC at a time, MVPA performs a joint analysis of all FCs in the whole brain at once and examines their spatial pattern and, thus, has greater power for detecting the differences in FCs between patients with OSA and controls. The higher sensitivity of MVPA also comes from the fact that it naturally avoids multiple comparisons problem and thus corrections for multiple comparisons are usually not needed (Liang et al., 2019). MVPA has been used successfully in detecting abnormal FC patterns and identifying neuroimaging biomarkers in other diseases, such as major depressive disorder (Zhu et al., 2020), schizophrenia (Hua et al., 2020), and social anxiety disorder (Liu et al., 2015). Zhou et al. (2020b) also used this technique based on the spatial pattern of regional homogeneity (ReHo) of resting-state neural activities to distinguish between patients with OSA and HCs. These studies have shown a promising potential of MVPA to identify the characteristic patterns of FC alterations in patients with OSA.

Therefore, in this study, we aimed to characterize the spatial patterns of resting-state FCs in OSA using MVPA and test its potential to serve as a neuroimaging biomarker to aid the diagnosis of OSA. We first performed univariate analyses to compare every FC between patients with OSA and controls, and then performed MVPA, combined with a feature selection procedure, to distinguish patients with OSA from healthy controls (HCs) using the spatial pattern of FCs. To characterize the model-selected FCs that contributed to the successful classification between patients and controls, we further examined the spatial distribution of these selected FCs and their relationship with the known resting-state

functional networks. Furthermore, the relationship between MVPA-selected FCs and disease severity of OSA was explored.

Materials and methods

Participants

This study included twenty-one male patients with moderate-to-severe OSA and twenty-one male HCs matched for handedness, education, and age. All subjects were recruited from the Sleep Laboratory of the Respiratory Department of Tianjin Medical University General Hospital. The inclusion criteria for patients with OSA included that the AHI was more than 15 times/h. The inclusion criteria for HCs included (1) the AHI $<$ 5 times/h, (2) no history of sleep breathing disorders, (3) no symptoms of nocturnal snoring confirmed by a physician, and (4) male. The exclusion criteria for both patients with OSA and HCs were as follows: (1) Other sleep disorders except OSA, (2) left-handedness, (3) history of severe hypertension, diabetes, respiratory disease, and cardiovascular disease, (4) mental diseases and other neurological conditions, (5) the score of Mini-Mental State Examination (MMSE) less than 24, (6) alcohol or illicit drug abuse or current intake of psychoactive medications, (7) body weight more than 125 kg, and (8) MRI contraindications such as claustrophobia and metallic implants or devices in the body. This study was approved by the local ethics committee and all subjects signed written informed consent.

The clinical manifestations of all patients included nocturnal snoring, irregular breathing, choking in sleep, and daytime sleepiness. None of them received drug therapy, surgery, or continuous positive pressure (CPAP) treatment. All patients underwent nocturnal polysomnography (PSG), and relevant clinical indicators were calculated based on the PSG results. According to the American Academy of Sleep Medicine (AASM) guidelines, apnea was defined as a reduction $\geq$ 90% in airflow lasting for at least 10 s during sleep and associated with persistent respiratory effort, and hypopnea was defined as a reduction $\geq$ 30% in airflow lasting for at least 10 s and accompanied by 4% or more oxygen saturation (Redline et al., 2007). The AHI was the average number of apnea and hypopnea that occurred per hour during sleep. The percentage of time with the saturation of oxygen (SaO_2) less than 90% of the total sleep time (i.e., %TST $<$ 90%) was recorded. The Epworth Sleepiness Scale (ESS) (Johns, 1991), a self-reported questionnaire assessing the severity of daytime sleepiness, was also recorded. The total score in ESS was 24. An ESS score of more than 6, 11, and 16 was defined as sleepiness, excessive sleepiness, and risky sleepiness, respectively (Kapur et al., 2017; Yu et al., 2019). Furthermore, all subjects were also assessed on MMSE, the most commonly used screening scale for cognitive impairment (Folstein et al., 1975). The maximal score of MMSE was 30. A score between 27 and 30 is considered normal, and a score $<$ 27 is considered cognitively impaired.

Data acquisition

The MR images were acquired using a 3.0 Tesla MRI scanner (Signa HDx, General Electric, Milwaukee, WI, United States) in Tianjin Medical University General Hospital. To reduce head movements and scanner noise, foam pads and earplugs were used, respectively. The resting-state fMRI data were acquired using an echo-planar imaging (EPI). Its sequence parameters were as follows: repetition times (TR) = 2,000 ms, echo time (TE) = 30 ms, flip angle (FA) = 90°, field of view (FOV) = 240 $\times$ 240 mm^2, matrix = 64 $\times$ 64, slice thickness = 3 mm, slice gap = 1 mm, and 38 axial slices. Each functional run included 180 volumes. In a single session, subjects were asked to relax without thinking about anything in particular, keep their eyes closed, and stay awake.

Functional magnetic resonance imaging data preprocessing

The fMRI data preprocessing was performed using Data Processing and Analysis of Brain Imaging (DPABI; Chinese Academy of Sciences, Beijing, China)[1] (Yan et al., 2016), which is a convenient plug-in software based on Statistical Parametric Mapping (SPM12)[2] in MATLAB platform. The first 10 volumes were discarded to eliminate the effects of the instability of the machine and the subjects' inadaptability to the environment in the very beginning of the scan. After slice-timing correction and six-dimensional head motion correction, the remaining 170 images were spatially normalized to the standard Montreal Neurological Institute (MNI) EPI template with a resampling voxel size of 3 $\times$ 3 $\times$ 3 mm^3. The effect of linear drift or trends in signal was removed. Then, several sources of spurious variance were regressed out by linear regression, such as 12 head motion parameters, global mean signal, white matter signal, cerebrospinal fluid, and the spike volumes if the frame-wise displacement (FD) exceeded 0.5 mm. A temporal band-pass filtering ($0.01 \leq f \leq 0.08$ Hz) was also performed. The head motion (the maximum displacements and maximum spin) of all participants was less than 2 mm and 2°, respectively.

Anatomical parcellation and construction of brain network

The cerebrum was segmented into 90 regions by the automated anatomical labeling (AAL) template (Tzourio-Mazoyer et al., 2002). The Pearson's correlation coefficient of the

1 http://rfmri.org/DPABI
2 http://www.fil.ion.ucl.ac.uk/spm

averaged time series between each pair of regions was calculated to define FC, and then a 90 × 90 symmetric correlation matrix was obtained for every participant. A Fisher's r-to-z transformation of the correlation coefficients was applied to improve the normality of FC values (Liu et al., 2017).

Univariate analysis

A two-sample *t*-test was used to compare every FC between OSA and HC groups, and the statistical significance for multiple comparisons was determined by three methods, namely, Bonferroni correction (corrected $p < 0.05$), a false discovery rate (FDR, $q < 0.05$), and the network-based statistic (NBS) approach (corrected $p < 0.05$ determined by 10,000 permutations, with a cluster-defining threshold of $p < 0.001$ by two-sample *t*-tests) (Zalesky et al., 2010). Besides, this univariate analysis was performed using the graph theoretical network analysis toolbox (GRETNA)[3] (Wang et al., 2015).

Multivariate pattern analysis

The MVPA was performed using the MVPANI toolbox[4] (Peng et al., 2020) to classify patients with OSA from HCs. Linear support vector machine (SVM) was used to find a hyperplane between patients with OSA and HCs which had a maximal distance to the support vectors on each side. The SVM model was trained and tested using a leave-one-participant-out cross-validation procedure. In each cross-validation, 41 participants were used to train the classifier and the remaining one participant was used to test the trained classifier. In this way, every participant was used once as a test sample, and the classification accuracy was calculated as the percentage of the correctly classified participants over all participants.

Feature selection

As the number of features (i.e., the FCs) was far more than the number of subjects, to avoid over-fitting, a feature selection based on the features' *F* scores was performed during the model training in each cross-validation step using the following procedure as implemented in the MVPANI toolbox: first, in each cross-validation step, an *F* score was calculated for each FC using an *F*-test comparing the two groups of participants (i.e., patients and controls) in the training dataset, and then all FCs were ranked according to their *F* scores; second, only the top *N* percentage of FCs were selected for building the SVM model that was trained using the training samples and then tested using the test sample; third, this feature selection procedure was repeated for all cross-validation steps for a particular percentage *N*. In this study, a series of *N* (from 10% to 100% in steps of 10%; i.e., ten percentages in total) was tested and a classification accuracy was obtained for each *N*. The final classification accuracy was determined by the highest one among the ten accuracies.

Permutation test

The statistical significance of the final classification accuracy (against the chance-level accuracy) was determined and corrected for multiple comparisons using a permutation test ($n = 10{,}000$) as follows. First, in each permutation step, exactly the same MVPA procedure as described earlier was performed, i.e., a linear SVM combined with the same feature selection procedure (i.e., feature selection based on *F* scores with 10 percentages of selected FCs from 10% to 100% in steps of 10%), except that in every cross-validation step, the class labels of the training samples were randomly shuffled; this procedure yielded 10 chance-level classification accuracies and the highest accuracy was taken as the final accuracy of this permutation step. Second, the first step was repeated 10,000 times, yielding 10,000 highest chance-level accuracies of all permutation steps with which a null distribution of chance-level accuracies was formed. Third, the 10 true classification accuracies obtained from the true labels (each corresponding to a feature selection percentage) were compared with this null distribution, resulting in a *p*-value for each true classification accuracy that was calculated as the percentage of chance-level accuracies greater than or equal to the true classification accuracy. The resultant 10 *p*-values corresponded to the family-wise-error (FWE) corrected *p*-values, and the true accuracies were considered statistically significant if $p < 0.05$.

Characterization of the multivariate pattern analysis-selected functional connectivities

To characterize the FCs contributing to the "patients vs. controls" classification, we further examined the FCs selected by the above MVPA procedure in two aspects, i.e., the spatial distribution of the selected FCs and the relationships of the selected FCs with the predefined resting-state functional networks. Here, the MVPA-selected FCs included the FCs selected in at least one cross-validation step during the feature selection that led to the highest classification accuracy.

To examine the spatial distribution of the selected FCs, we visually inspected which parts of the brain were involved in these FCs. Furthermore, we also evaluated the importance of the brain regions associated with these FCs by calculating the weighted degrees of each region. This was done for positive weights and negative weights separately by calculating the sum of all positive weights of the FCs associated with a given region and the sum of all negative weights of the FCs associated with a given region, respectively.

As the resting-state functional networks have been reported to play an important role in many sensory and cognitive functions related to OSA (Zhang et al., 2013, 2015;

3 http://www.nitrc.org/projects/gretna/

4 http://funi.tmu.edu.cn

TABLE 1 Regions belonging to each functional network.

Network	Regions with abbreviation
VN	CAL, CUN, LING, SOG, MOG, IOG, FFG
SMN	PreCG, ROL, PoCG, PCL, HES, STG
AN	SPG, ITG.R, SMA, INS, MCG, SMG, PUT, PAL
FPN	MFG, MFGorb, IFGoper, IFGtri, IPL
LS	SFGorb, OLF, REC, AMYG, CAU, THA, TPOsup, TPOmid
DMN	SFG, SFGmed, SFGmorb, IFGorb, ACG, PCG, HIP, PHG, ANG, PCUN, MTG, ITG.L

Khazaie et al., 2017; Chen et al., 2018b; Chang et al., 2020) and their disruptions have been indicated in patients with OSA (Zhang et al., 2015; Park et al., 2016b; Wu et al., 2020), we further examined the relationship of the selected FCs with the predefined resting-state functional networks. Specifically, we categorized the 90 brain regions into 7 functional networks, namely, visual network (VN), somatomotor network (SMN), dorsal attention network (DAN), ventral attention network (VAN), limbic system (LS), frontoparietal network (FPN), and DMN, according to Yeo's parcellation of the cerebral cortex (Yeo et al., 2011). In our results, we merged DAN and VAN into AN (**Table 1**). The full name of all brain regions is summarized in **Supplementary Table 1**. According to such categorization of all brain regions, the MVPA-selected FCs were divided into two sets, namely, intra-network FCs (if an FC connects two regions that belong to the same functional network) and inter-network FCs (if an FC connects two regions that belong to different functional networks). The number of intra-network FCs was standardized by the total number of all possible intra-network FCs (i.e., dividing the number of intra-network FCs by the total number of all possible intra-network FCs), and similarly, the number of inter-network FCs was standardized by the total number of all possible inter-network FCs.

Correlations between the selected functional connectivities and clinical variables

To investigate the correlation between selected the FCs and the clinical variables, Pearson correlation analyses were performed, and $p < 0.005$ was considered statistically significant.

Results

Demographic and clinical indices

There were no significant differences (two-sample t-tests, all $p > 0.05$) between patients with OSA and HCs in age, years of education, or MMSE (**Table 2**). As expected, patients with OSA had a significantly higher score for the body mass index (BMI) ($t = 3.893$, $p < 0.001$), AHI ($t = 11.762$, $p < 0.001$), %TST < 90% ($t = 3.792$, $p = 0.001$), and ESS ($t = 8.461$, $p < 0.001$).

TABLE 2 The demographic and clinical characteristics of patients with OSA and healthy controls.

	OSA patients		Healthy controls		p-value
	Mean	*SD*	Mean	*SD*	
Age (years)	44.05	7.277	40.62	11.404	0.252
Education (years)	13.48	3.092	14.76	2.914	0.173
BMI (kg/m^2)	29.52	4.231	24.95	3.173	<0.001*
MMSE	29.48	0.814	29.86	0.359	0.057
AHI	54.35	19.97	2.52	1.401	<0.001*
%TST < 90%	18.66	21.10	0.979	2.654	0.001*
ESS	14.67	7.262	1.10	1.136	<0.001*

**Significant difference between OSA and HC, $p < 0.05$. BMI, body mass index; MMSE, Mini Mental State Examination; AHI, apnea-hypopnea index;%TST < 90%, percentage of total sleep time spent at oxygen saturations less than 90%; ESS, Epworth sleepiness scale. The full name of all brain regions were summarized in Supplementary Table 1.*

Univariate comparisons of functional connectivities between patients with obstructive sleep apnea and healthy controls

The univariate analyses showed that there were no significant changes in FC between patients with OSA and HCs regardless of the method for multiple comparisons' correction ($p < 0.05$, Bonferroni corrected; $p < 0.05$, NBS corrected; $q < 0.05$, FDR corrected).

Classification between patients with obstructive sleep apnea and healthy controls

The MVPA showed successful classifications for 4 out of 10 feature selection percentages (**Figure 1A**): The classification accuracies were 83.33% when the top 10% FCs were selected ($p = 0.0001$; the corresponding specificity and sensitivity were 85.71% and 80.95%, respectively), 71.43% when the top 30%, 80%, and 100% FCs were selected ($p = 0.0076$; the corresponding specificity and sensitivity were 66.67% and 76.19%, respectively), and 69.05% for the other percentages ($p = 0.067$; the corresponding specificity and sensitivity were 71.43% and 66.67%, respectively) (**Figure 1B**). The highest classification accuracy of 83.33% was considered the final accuracy and the selected top 10% FCs (400 FCs) were further characterized. The receiver operating characteristic (ROC) curve corresponding to this highest classification accuracy (83.33%) and the

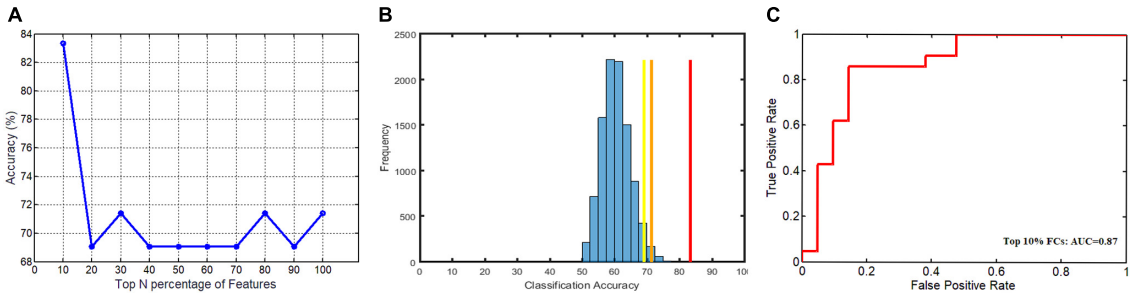


FIGURE 1
The classification accuracies of 10 feature selection percentages, the corresponding null distribution, and the receiver operating characteristic (ROC) curve for the highest classification accuracies. **(A)** The highest classification accuracy (83.33%) was obtained when the top 10% FCs were selected and two lower accuracies (71.43% and 69.05%) were obtained when different numbers of FCs were selected; **(B)** the corresponding null distribution of the highest chance-level accuracies (the blue histogram), along with the three true classification accuracies (red: 83.33%; orange: 71.43%; yellow: 69.05%); **(C)** the receiver operating characteristic (ROC) curve and the area under the curve (AUC) (0.87) corresponding to this highest classification accuracy (83.33%).

corresponding area under the curve (AUC = 0.87) are shown in Figure 1C.

Characterization of the multivariate pattern analysis-selected functional connectivities

The spatial distribution of the selected top 10% FCs (i.e., 400 FCs) in the brain is shown in Figure 2. Among the 400 selected FCs, 195 FCs showed higher weight value in patients with OSA (Figure 2A), and 205 FCs showed lower weight value in patients with OSA (Figure 2B). Moreover, the FCs with absolute weight values greater than the mean plus twice the standard deviation (i.e., absolute mean + 2SD) are shown in Figures 2C,D. To evaluate the importance of the brain regions for the classification, we also calculated the positive and negative weighted degrees of each region. A total of 16 brain regions showed significantly higher positive weighted degrees in patients with OSA than in HC (> mean + SD), including the left MFGorb, left CUN, left SMG, right PCUN, right REC, right ITG, right IOG, left REC, left IOG, left IFGorb, left FFG, right FFG, right PHG, left INS, left STG, and left CAU (Figures 3A,C), and 14 brain regions showed significantly lower negative weighted degrees in patients with OSA than in HC, including left TPOsup, right HIP, left MCG, left TPOmid, right MCG, right PHG, right REC, left PHG, right ITG, left SFGmed, left SMG, left IFGoper, right PCG, and left HEC (Figures 3B,D).

Among the 400 selected FCs, 65 FCs were intra-network FCs and 335 were inter-network FCs. The numbers of intra-network FCs were 9 in VN, 4 in SMN, 8 in AN, 4 in FPN, 30 in DMN, and 10 in LS, respectively (Figure 4, the diagonal entries; Figure 5), accounting for 9.89%, 6.06%, 7.62%, 8.89%, 11.86%, and 8.33% of intra-network FCs in these functional networks, respectively (the mean percentage of intra-network FC was 9.56% across these functional networks). The percentages of inter-network FCs between each pair of functional networks are shown in Figure 4 (the off-diagonal entries), and the mean percentage was 10.08%. To specifically look at the inter-network FCs associated with the DMN, we also showed all inter-network FCs between DMN and the other five networks in Figure 6.

Correlations between the selected functional connectivities and clinical variables

We further examined the correlations between the 400 selected FCs and the clinical variables in patients with OSA. We found that the clinical variable AHI showed negative correlations with the FC between the left CUN and the left TPOsup ($r = -0.607$, $p = 0.0035$) and with the FC between the left PHG and the left IFGoper ($r = -0.6.26$, $p = 0.0024$) and showed positive correlations with the FC between left INS and left MFGorb ($r = 0.608$, $p = 0.0035$) (Figure 7A). Moreover, the clinical variable %TST < 90% showed a positive correlation with the FCs between the right MCG and the left TPOmid ($r = 0.705$, $p = 0.00036$) and with the FC between the right PUT and ITG ($r = 0.602$, $p = 0.00386$) (Figure 7B).

Discussion

Weak alterations of resting-state functional connectivities in obstructive sleep apnea can be detected by multivariate pattern analysis

We performed both a univariate analysis and an MVPA to identify the differences in resting-state FCs between patients

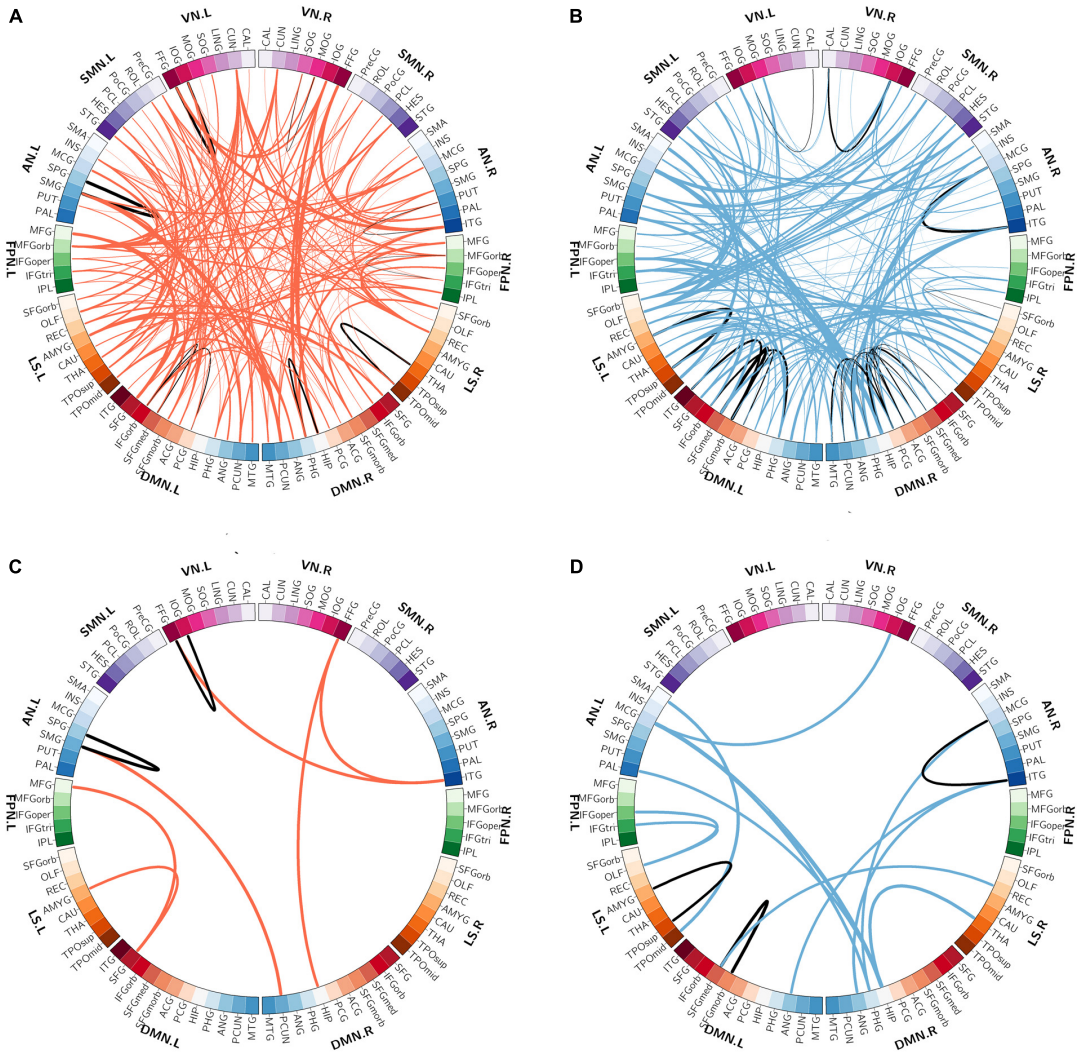


FIGURE 2
The spatial distribution of the selected functional connectivity. **(A)** The spatial distribution of the selected FCs with higher classification weights in patients with OSA compared with controls. **(B)** The spatial distribution of the selected FCs with lower classification weights in patients with OSA compared with controls. **(C)** The FCs with higher positive weight values were greater than mean + 2SD in patients with OSA. **(D)** The FCs with more negative weight values, which means the absolute weight values were more than mean + 2SD in patients with OSA. The thickness of lines represents the absolute weight values and the black lines indicated functional connectivity within one functional network.SD, standard deviation; VN, visual network; SMN, somatomotor network; AN, attention network; FPN, frontoparietal network; DMN, default mode network; LS, limbic system.

with OSA and HCs. The fact that no significant differences were identified using univariate two-sample *t*-tests, even for relatively liberal thresholds ($q < 0.05$ corrected by FDR, or $p < 0.05$ corrected by NBS), suggests that the alterations of the resting-state FCs might not be very large. However, such weak alterations of FCs in OSA can be detected as a pattern change by MVPA, confirming that the resting-state FCs were indeed altered in patients with OSA. A classification accuracy of 83.33% also suggests that the spatial pattern of resting-state FCs can successfully distinguish patients with OSA from HCs, demonstrating its potential as a neuroimaging biomarker for aiding the diagnosis of OSA.

The whole-brain resting-state functional connectivities were altered in a dispersed way in obstructive sleep apnea

Using an SVM classification algorithm combined with a feature selection procedure, we identified 400 FCs contributing to the successful classification between patients with OSA and HCs that yielded the highest classification accuracy (83.33%). Further examination of the spatial distribution of these 400 MVPA-selected FCs showed that almost all brain regions (88

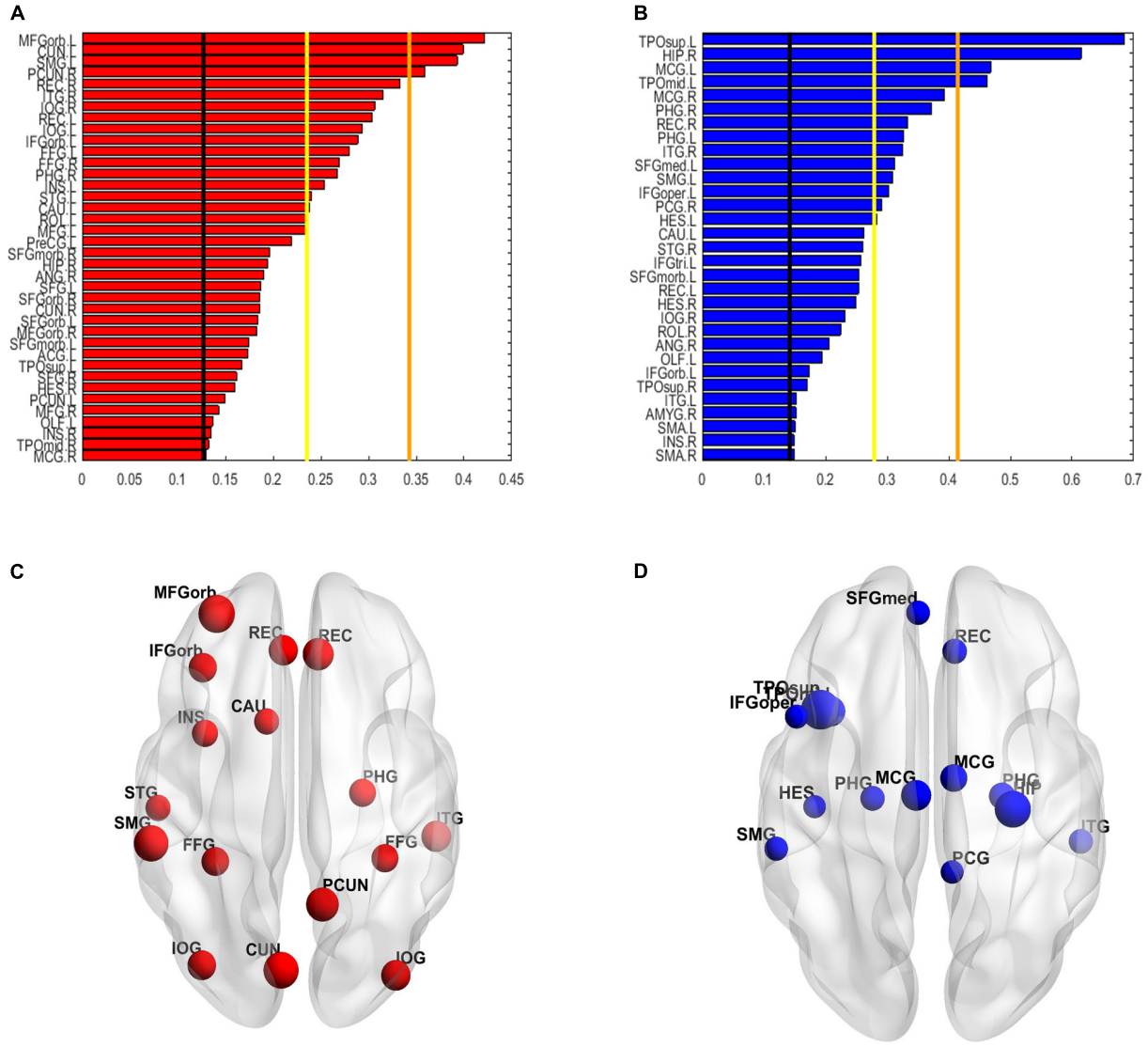


FIGURE 3
The weighted degree maps of brain regions. **(A,C)** The brain regions with positive weighted degrees greater than mean + SD. **(B,D)** The brain regions with negative weighted degrees less than mean-SD. The black line, yellow line, and orange line indicate the mean value, mean ± − SD, and mean ± − 2SD, respectively. SD, standard deviation.

out of 90) were involved in these FCs. This result suggests that OSA is likely to affect the FCs among widely distributed regions in the whole brain, rather than some local networks involving only a few particular brain regions. This is in line with previous studies investigating the FC changes in OSA. For example, Park et al. found 27 decreased FCs and 46 increased FCs in patients with OSA associated with 62 out of 90 brain regions (Park et al., 2016b). Such widely distributed FC alterations also support the previous findings that the global topological properties of the whole-brain resting-state functional network were disrupted as well in patients with OSA. For example, although a small-world topology was still preserved, the small-world property was significantly decreased (Chen et al., 2017), along with some other global topological properties such as clustering coefficient, characteristic path length, and global efficiency (Huang et al., 2019).

The relationship of the multivariate pattern analysis-selected functional connectivities with the predefined resting-state functional networks

We also characterized the relationship of the MVPA-selected FCs with the predefined resting-state functional networks and

FIGURE 4
The percentages of intra-network FCs in each functional network (the diagonal entries) and the percentages of inter-network FCs (the off-diagonal entries). The values that exceeded the average percentages were marked with red rectangular boxes. VN, visual network; SMN, somatomotor network; AN, attention network; FPN, frontoparietal network; DMN, default mode network; LS, limbic system.

found that these FCs were associated with all 7 predefined resting-state functional networks, further corroborating the finding that OSA involves widely distributed FC alterations in the whole brain. Furthermore, the quantification of these MVPA-selected FCs in terms of intra- and inter-network FCs showed that the DMN had the highest percentage of intra-network FCs among the 400 MVPA-selected FCs and also had relatively high percentages of inter-network FCs with the SMN, the AN, and the LS, suggesting the important role of the DMN-associated FC changes in OSA. This is consistent with some previous studies that have reported abnormal intra-network FCs as well as the global and local topological properties of DMN in patients with OSA compared with HCs (Zhang et al., 2013; Peng et al., 2014; Li et al., 2015, 2016a,b; Chen et al., 2018a). It is known that the DMN, including the posterior cingulate gyrus (PCG), the medial prefrontal cortex, HIP, medial temporal lobe (MTG), angular gyrus (ANG), and precuneus (PCUN) as the core regions (Li et al., 2016b), is more active during resting state but its activity is inhibited during many cognitive tasks, and the degree of inhibition even increases with the task load (Buckner et al., 2008). Our results showed positive classification weights for the FCs between the bilateral PCUN, the bilateral PCG, the right SFG, and the medial part of the right superior frontal gyrus (SFGmed), indicating that these FCs were lower in patients with OSA compared with HCs. The connectivity between the right HIP and ipsilateral parahippocampal gyrus (PHG) was an important member of the classified pattern, while Song's study found reduced FC between the right HIP and the bilateral thalamus and PHG in patients with OSA (Song et al., 2018). We also found that the intra-network FC in DMN associated with bilateral PCG was useful in OSA-HC classification. PCG, PCUN, and HIP were considered the key regions of posterior DMN (pDMN); the medial prefrontal cortex, anterior cingulate (ACG), and superior frontal gyrus belong to anterior DMN (aDMN) (Zhang et al., 2013; Chen et al., 2018a). In a previous study, Zhang et al. (2013) found that FCs of patients with OSA in aDMN were significantly decreased than HC's, while FCs in OSA were increased in pDMN. We also found that the intra-network FCs in aDMN showed negative classification weights, such as the FCs between the bilateral ACG and the bilateral medial orbital part of the superior frontal gyrus (SFGmorb), Therefore, the abnormal intra-network FC in DMN explained the functional heterogeneity of aDMN and pDMN. Chen et al. (2018a) reported abnormal FCs within the DMN and decreased network topological properties such as the clustering coefficient and the local efficiency of the DMN.

Moreover, some previous studies have reported abnormal FC between DMN and other brain regions. Zhang et al. (2015) found that the FCs between key nodes of the DMN (the bilateral ACG, right PCG, bilateral SFG, and bilateral medial prefrontal cortex) and the AN (the right INS) were significantly decreased in patients with OSA. Song et al. (2018) reported that the FCs between the nodes in DMN (HIP and ANG) and the nodes in LS (THA and CAU) were significantly abnormal in patients with OSA. As the DMN has been suggested to play an important role in many cognitive functions such as regulating emotion, consciousness, memory, and introspection, our present findings, together with the previous results, suggest that the disrupted FCs associated with the DMN in OSA may underlie the cognitive impairments observed in patients with OSA.

Some papers have also found visual dysfunction in patients with OSA. Giora et al. (2017) found that the reaction time in a visual task for patients with OSA was significantly longer than HCs. Moghimi et al. (2013) detected that the nerve fiber indicator was significantly reduced in patients with OSA, and patients with OSA had a higher prevalence rate of glaucoma and ocular hypertension. The calcarine cortex (CAL) is a core region of the visual recognition network (Tao et al., 2013) and was reported to be associated with the shifting of attention to the intended visual target and the modulation of visual input through attention. Liu et al. (2017) found the voxel mirrored homotopic connectivity (VMHC) in bilateral CAL, and VMHC value in CAL was positively correlated with AHI (Yamagishi et al., 2005). Zhang et al. (2013) found that the right cuneus (CUN) exhibited reduced gray matter volume (GMV) in patients with OSA that imply the visual attention deficit of OSA. However, there are few studies on functional disconnection associated with VN in patients with OSA. In the current OSA-HC classification FC pattern, the percentage of intra-network FC in VN was higher than the percentage of intra-network FC in the whole brain, but the percentage of inter-network FC between VN and the other network was lower than the percentage of inter-network in the whole brain.

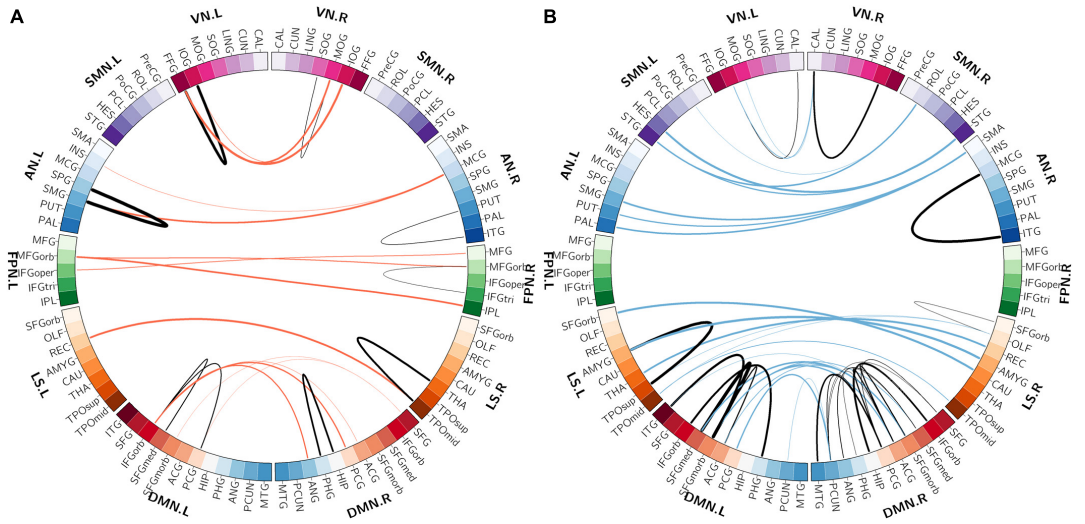


FIGURE 5
The distributions of intra-network FCs. **(A)** The spatial distribution of the selected intra-network FCs with higher classification weights in patients with OSA compared with controls. **(B)** The spatial distribution of the selected intra-network FCs with lower classification weights in patients with OSA compared with controls. The thickness of lines represents the absolute weight values and the black lines indicate functional connectivity within one functional network. VN, visual network; SMN, somatomotor network; AN, attention network; FPN, frontoparietal network; DMN, default mode network; LS, limbic system.

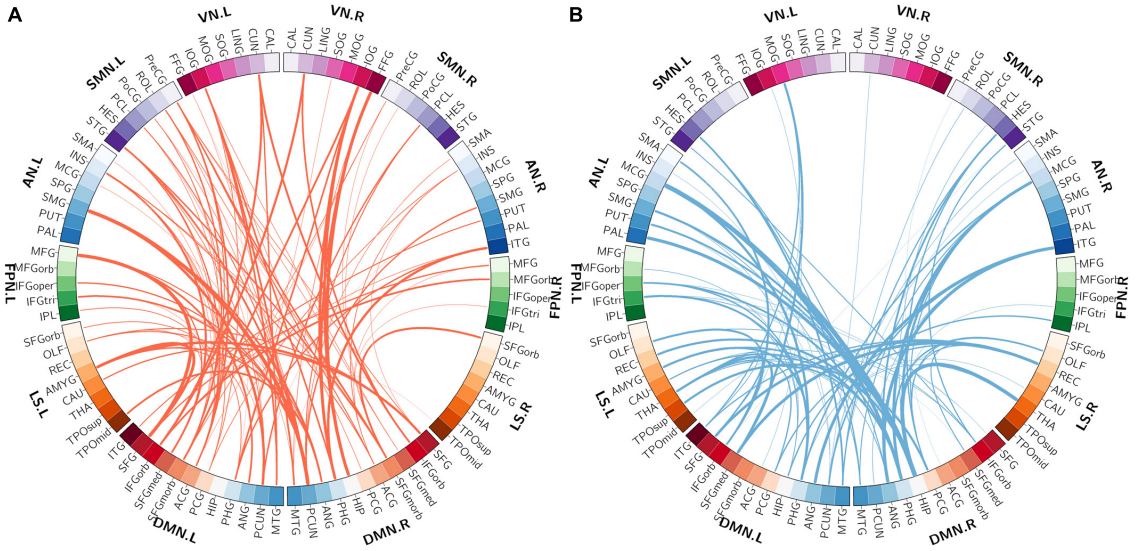


FIGURE 6
The distribution of inter-network FCs between the DMN and the other five functional networks. **(A)** The spatial distribution of the selected inter-network FCs between DMN and the other networks with higher classification weights in patients with OSA compared with controls. **(B)** The spatial distribution of the selected inter-network FCs between DMN and the other networks with lower classification weights in patients with OSA compared with controls. The thickness of lines represents the absolute weight values. VN, visual network; SMN, somatomotor network; AN, attention network; FPN, frontoparietal network; DMN, default mode network; LS, limbic system.

Furthermore, the FC between left CAL and left IOG, FC between right CAL and bilateral IOG, and left MOG play an important role in differentiating patients with OSA and HCs, and the classification weight of these connectivities was negative (HC-OSA).

Correlations between the functional connectivities and the disease severity

We found that some MVPA-selected FCs were significantly correlated with the AHI and the%TST < 90%. AHI is the main

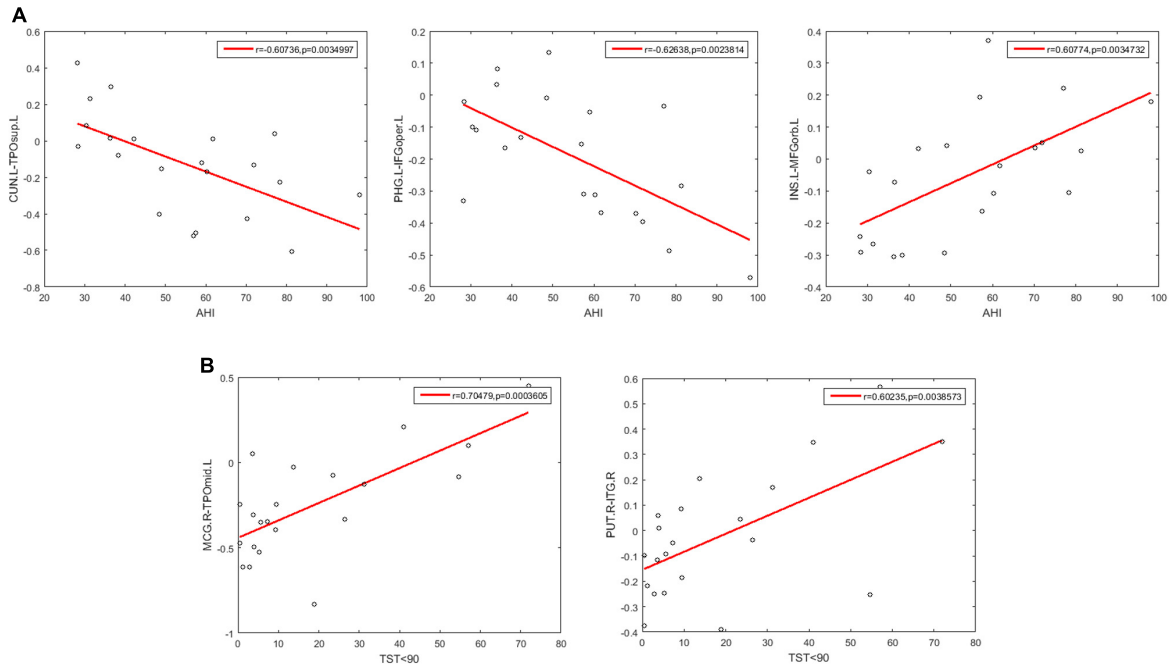


FIGURE 7
Scatter plots showing significant correlations between the FCs and the clinical variables in patients with OSA. **(A)** The correlations between the FCs and AHI. **(B)** The correlations between the FCs and % TST < 90%. AHI, apnea-hypopnea index; %TST < 90%, percentage of total sleep time spent at oxygen saturations less than 90%; CUN.L, left Cuneus; TPOsup.L, left Temporal pole: superior temporal gyrus; PHG.L, left Parahippocampal gyrus; IFGoper.L, left Inferior frontal gyrus, opercular part; INS.L, left Insula; MFGorb.L, left Middle frontal gyrus, orbital part; MOG.R, right Middle occipital gyrus; TPOmid.L, left Temporal pole: middle temporal gyrus; PUT.R, right Lenticular nucleus, putamen; ITG.R, right Inferior temporal gyrus.

indicator of the severity of OSA. In this study, the FCs between the left PHG (DMN) and left IFGoper (FPN) and between the left CUN (VN) and left TOPsup (LS) showed negative correlations with AHI, while the FC between the left INS (AN) and left MFGorb (FPN) showed a positive correlation of AHI. Although this has not been reported in previous studies, our result suggests these FCs might be indicative of the AHI in OSA.

The FCs between the right MCG (AN) and the left TPOmid (LS), and between the right PUT (AN) and the right ITG (AN) showed a positive correlation with % TST < 90%. It is noticeable that both FCs were associated with the functional network AN (attention network). Attention is a primary cognitive function, involving selective attention, sustained attention, and attention distribution, which will further affect other cognitive functions (Olaithe et al., 2018), and some previous studies have reported attentional impairments in all three aspects of attention in patients with OSA (Verstraeten et al., 2004; Vanek et al., 2020). Even though the treatment of continuous positive airway pressure (CPAP) could improve alertness and attention (Verstraeten and Cluydts, 2004), it did not seem to be able to restore the quality of attention to normal levels in patients with OSA (Lau et al., 2010). Our results provide evidence for the neural mechanisms of attention impairment in patients with OSA, which may be related to the disrupted FCs in the AN due to hypoxemia during sleep, and such attention deficits in patients with OSA may be more resistant to treatment.

Limitations

There are several limitations in this study. First, the sample size in this study was relatively small and only recruited male subjects. The statistical significance of the correlations between the FCs and the clinical variables was not corrected for multiple comparisons also due to the small sample size. Therefore, large sample dataset and female patients with OSA should be included in future studies to confirm our results. Second, we only used FC to distinguish patients with OSA from HCs. Whether merging different imaging measures could improve the classification accuracy in distinguishing patients with OSA from HCs needs to be further studied.

Conclusion

The findings in this study revealed that the resting-state FCs were altered in OSA and the disrupted FCs were widely distributed and involved almost all resting-state functional

networks in the whole brain of patients with OSA. The successful classification between patients with OSA and HCs obtained using machine learning techniques also indicates that the altered resting-state FCs are indicative of the severity of the disease and have the potential to serve as a neuroimaging biomarker of OSA.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by the Medical Research Ethics Committee of Tianjin Medical University General Hospital. The patients/participants provided their written informed consent to participate in this study.

Author contributions

AH and ML: study design and article writing. QZ: acquisition of data. AH, XP, and HW: analysis and interpretation of data. XZ, YP, and DL: technical guidance. ML, QZ, and FG: manuscript review and editing. ML and FG: supervision. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by the National Natural Science Foundation of China (Grant Nos. 81871393, 62075156, 81971694, and 81401401) and the Natural Science Foundation of Tianjin (Grant No. 16JCQNJC10900).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer YL declared a past co-authorship with the author ML to the handling editor.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2022.920765/full#supplementary-material

References

Benjafield, A. V., Ayas, N. T., Eastwood, P. R., Heinzer, R., Ip, M. S. M., Morrell, M. J., et al. (2019). Estimation of the global prevalence and burden of obstructive sleep apnoea: a literature-based analysis. *Lancet. Respir. Med.* 7, 687–698. doi: 10.1016/S2213-2600(19)30198-5

Buckner, R. L., Andrews-Hanna, J. R., and Schacter, D. L. (2008). The brain's default network: anatomy, function, and relevance to disease. *Ann. N. Y. Acad. Sci.* 1124, 1–38. doi: 10.1196/annals.1440.011

Chang, Y. T., Chen, Y. C., Chen, Y. L., Hsu, S. W., Yang, F. Y., Lee, C. C., et al. (2020). Functional connectivity in default mode network correlates with severity of hypoxemia in obstructive sleep apnea. *Brain Behav.* 10:e01889. doi: 10.1002/brb3.1889

Chen, L., Fan, X., Li, H., Ye, C., Yu, H., Gong, H., et al. (2018a). Topological Reorganization of the Default Mode Network in Severe Male Obstructive Sleep Apnea. *Front. Neurol.* 9:363. doi: 10.3389/fneur.2018.00363

Chen, L. T., Fan, X. L., Li, H. J., Ye, C. L., Yu, H. H., Xin, H. Z., et al. (2018b). Aberrant brain functional connectome in patients with obstructive sleep apnea. *Neuropsychiatr. Dis. Treat.* 14, 1059–1070. doi: 10.2147/NDT.S161085

Chen, L. T., Fan, X. L., Li, H. J., Nie, S., Gong, H. H., Zhang, W., et al. (2017). Disrupted small-world brain functional network topology in male patients with severe obstructive sleep apnea revealed by resting-state fMRI. *Neuropsychiatr. Dis. Treat.* 13, 1471–1482. doi: 10.2147/NDT.S135426

Folstein, M. F., Folstein, S. E., and McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* 12, 189–198. doi: 10.1016/0022-3956(75)90026-6

Gagnon, K., Baril, A. A., Gagnon, J. F., Fortin, M., Decary, A., Lafond, C., et al. (2014). Cognitive impairment in obstructive sleep apnea. *Pathol. Biol.* 62, 233–240. doi: 10.1016/j.patbio.2014.05.015

Giora, E., Galbiati, A., Marelli, S., Zucconi, M., and Ferini-Strambi, L. (2017). Evidence of perceptive impairment in OSA patients investigated by means of a visual search task. *Cortex* 95, 136–142. doi: 10.1016/j.cortex.2017.08.004

Gupta, M. A., and Simpson, F. C. (2015). Obstructive sleep apnea and psychiatric disorders: a systematic review. *J. Clin. Sleep Med.* 11, 165–175. doi: 10.5664/jcsm.4466

Hua, M., Peng, Y., Zhou, Y., Qin, W., Yu, C., and Liang, M. (2020). Disrupted pathways from limbic areas to thalamus in schizophrenia highlighted by whole-brain resting-state effective connectivity analysis. *Prog. Neuropsychopharmacol. Biol. Psychiat.* 99:109837. doi: 10.1016/j.pnpbp.2019.109837

Huang, Y., Liu, Y., Zhao, D., Liu, B., Zhang, H., Huang, Z., et al. (2019). Small-world properties of the whole-brain functional networks in patients with

obstructive sleep apnea-hypopnea syndrome. *Sleep Med.* 62, 53–58. doi: 10.1016/j.sleep.2018.08.037

Johns, M. W. (1991). A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep* 14, 540–545. doi: 10.1093/sleep/14.6.540

Kapur, V. K., Auckley, D. H., Chowdhuri, S., Kuhlmann, D. C., Mehra, R., Ramar, K., et al. (2017). Clinical practice guideline for diagnostic testing for adult obstructive sleep apnea: an American Academy of sleep medicine clinical practice guideline. *J. Clin. Sleep Med.* 13, 479–504. doi: 10.5664/jcsm.6506

Khazaie, H., Veronese, M., Noori, K., Emamian, F., Zarei, M., Ashkan, K., et al. (2017). Functional reorganization in obstructive sleep apnoea and insomnia: a systematic review of the resting-state fMRI. *Neurosci. Biobehav. Rev.* 77, 219–231. doi: 10.1016/j.neubiorev.2017.03.013

Lau, E. Y., Eskes, G. A., Morrison, D. L., Rajda, M., and Spurr, K. F. (2010). Executive function in patients with obstructive sleep apnea treated with continuous positive airway pressure. *J. Int. Neuropsychol. Soc.* 16, 1077–1088.

Li, H., Li, L., Shao, Y., Gong, H., Zhang, W., Zeng, X., et al. (2016a). Abnormal Intrinsic Functional Hubs in Severe Male Obstructive Sleep Apnea: evidence from a Voxel-Wise Degree Centrality Analysis. *PLoS One* 11:e0164031. doi: 10.1371/journal.pone.0164031

Li, H. J., Nie, X., Gong, H. H., Zhang, W., Nie, S., and Peng, D. C. (2016b). Abnormal resting-state functional connectivity within the default mode network subregions in male patients with obstructive sleep apnea. *Neuropsychiatr. Dis. Treat.* 12, 203–212. doi: 10.2147/NDT.S97449

Li, H. J., Dai, X. J., Gong, H. H., Nie, X., Zhang, W., and Peng, D. C. (2015). Aberrant spontaneous low-frequency brain activity in male patients with severe obstructive sleep apnea revealed by resting-state functional MRI. *Neuropsychiatr. Dis. Treat.* 11, 207–214. doi: 10.2147/NDT.S73730

Liang, M., Su, Q., Mouraux, A., and Iannetti, G. D. (2019). Spatial patterns of brain activity preferentially reflecting transient pain and stimulus intensity. *Cereb. Cortex* 29, 2211–2227. doi: 10.1093/cercor/bhz026

Lim, D. C., and Pack, A. I. (2014). Obstructive sleep apnea and cognitive impairment: addressing the blood-brain barrier. *Sleep Med. Rev.* 18, 35–48. doi: 10.1016/j.smrv.2012.12.003

Liu, F., Guo, W., Fouche, J. P., Wang, Y., Wang, W., Ding, J., et al. (2015). Multivariate classification of social anxiety disorder using whole brain functional connectivity. *Brain Struct. Funct.* 220, 101–115. doi: 10.1007/s00429-013-0641-4

Liu, F., Wang, Y. F., Li, M. L., Wang, W. Q., Li, R., Zhang, Z. Q., et al. (2017). Dynamic functional network connectivity in idiopathic generalized epilepsy with generalized tonic-clonic seizure. *Hum. Brain Mapp.* 38, 957–973. doi: 10.1002/hbm.23430

Moghimi, S., Ahmadraji, A., Sotoodeh, H., Sadeghniat, K., Maghsoudipour, M., Fakhraie, G., et al. (2013). Retinal nerve fiber thickness is reduced in sleep apnea syndrome. *Sleep Med.* 14, 53–57. doi: 10.1016/j.sleep.2012.07.004

Mur, M., Bandettini, P. A., and Kriegeskorte, N. (2009). Revealing representational content with pattern-information fMRIan introductory guide. *Soc. Cogn. Affect. Neurosci.* 4, 101–109. doi: 10.1093/scan/nsn044

Olaithe, M., Bucks, R. S., Hillman, D. R., and Eastwood, P. R. (2018). Cognitive deficits in obstructive sleep apnea: insights from a meta-review and comparison with deficits observed in COPD, insomnia, and sleep deprivation. *Sleep Med. Rev.* 38, 39–49. doi: 10.1016/j.smrv.2017.03.005

Park, B., Palomares, J. A., Woo, M. A., Kang, D. W., Macey, P. M., Yan-Go, F. L., et al. (2016a). Aberrant insular functional network integrity in patients with obstructive sleep apnea. *Sleep* 39, 989–1000. doi: 10.5665/sleep.5738

Park, B., Palomares, J. A., Woo, M. A., Kang, D. W., Macey, P. M., Yan-Go, F. L., et al. (2016b). Disrupted functional brain network organization in patients with obstructive sleep apnea. *Brain Behav.* 6:e00441. doi: 10.1002/brb3.441

Park, J. G., Ramar, K., and Olson, E. J. (2011). Updates on definition, consequences, and management of obstructive sleep apnea. *Mayo. Clin. Proc.* 86, 549–554. doi: 10.4065/mcp.2010.0810

Peng, D. C., Dai, X. J., Gong, H. H., Li, H. J., Nie, X., and Zhang, W. (2014). Altered intrinsic regional brain activity in male patients with severe obstructive sleep apnea: a resting-state functional magnetic resonance imaging study. *Neuropsychiatr. Dis. Treat.* 10, 1819–1826. doi: 10.2147/NDT.S67805

Peng, Y., Zhang, X., Li, Y., Su, Q., Wang, S., Liu, F., et al. (2020). MVPANI: a toolkit with friendly graphical user interface for multivariate pattern analysis of neuroimaging data. *Front. Neurosci.* 14:545. doi: 10.3389/fnins.2020.00545

Pereira, F., Mitchell, T., and Botvinick, M. (2009). Machine learning classifiers and fMRI: a tutorial overview. *Neuroimage* 45, S199–S209. doi: 10.1016/j.neuroimage.2008.11.007

Qin, Z., Kang, D., Feng, X., Kong, D., Wang, F., and Bao, H. (2020). Resting-state functional magnetic resonance imaging of high altitude patients with obstructive sleep apnoea hypopnoea syndrome. *Sci. Rep.* 10:15546. doi: 10.1038/s41598-020-72339-2

Redline, S., Budhiraja, R., Kapur, V., Marcus, C. L., Mateika, J. H., Mehra, R., et al. (2007). The scoring of respiratory events in sleep: reliability and validity. *J. Clin. Sleep Med.* 3, 169–200.

Song, X., Roy, B., Kang, D. W., Aysola, R. S., Macey, P. M., Woo, M. A., et al. (2018). Altered resting-state hippocampal and caudate functional networks in patients with obstructive sleep apnea. *Brain Behav.* 8:e00994. doi: 10.1002/brb3.994

Tao, H., Guo, S., Ge, T., Kendrick, K. M., Xue, Z., Liu, Z., et al. (2013). Depression uncouples brain hate circuit. *Mol. Psychiat.* 18, 101–111. doi: 10.1038/mp.2011.127

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15, 273–289. doi: 10.1006/nimg.2001.0978

Vanek, J., Prasko, J., Genzor, S., Ociskova, M., Kantor, K., Holubova, M., et al. (2020). Obstructive sleep apnea, depression and cognitive impairment. *Sleep Med.* 72, 50–58. doi: 10.1016/j.sleep.2020.03.017

Verstraeten, E. (2007). Neurocognitive effects of obstructive sleep apnea syndrome. *Curr. Neurol. Neurosci. Rep.* 7, 161–166.

Verstraeten, E., and Cluydts, R. (2004). Executive control of attention in sleep apnea patients: theoretical concepts and methodological considerations. *Sleep Med. Rev.* 8, 257–267. doi: 10.1016/j.smrv.2004.01.001

Verstraeten, E., Cluydts, R., Pevernagie, D., and Hoffmann, G. (2004). Executive function in sleep apnea: controlling for attentional capacity in assessing executive attention. *Sleep* 27, 685–693.

Wang, J., Wang, X., Xia, M., Liao, X., Evans, A., and He, Y. (2015). Corrigendum: GRETNA: a graph theoretical network analysis toolbox for imaging connectomics. *Front. Hum. Neurosci.* 9:458. doi: 10.3389/fnhum.2015.00458

Wu, Y., Zhao, W., Chen, X., Wan, X., and Lei, X. (2020). Aberrant awake spontaneous brain activity in obstructive sleep apnea: a review focused on resting-state EEG and resting-state fMRI. *Front. Neurol.* 11:768. doi: 10.3389/fneur.2020.00768

Yamagishi, N., Goda, N., Callan, D. E., Anderson, S. J., and Kawato, M. (2005). Attentional shifts towards an expected visual target alter the level of alpha-band oscillatory activity in the human calcarine cortex. *Brain Res. Cogn. Brain Res.* 25, 799–809. doi: 10.1016/j.cogbrainres.2005.09.006

Yan, C. G., Wang, X. D., Zuo, X. N., and Zang, Y. F. (2016). DPABI: data processing & analysis for (resting-state) brain imaging. *Neuroinformatics* 14, 339–351. doi: 10.1007/s12021-016-9299-4

Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., et al. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J. Neurophysiol.* 106, 1125–1165. doi: 10.1152/jn.00338.2011

Yu, H., Chen, L., Li, H., Xin, H., Zhang, J., Wei, Z., et al. (2019). Abnormal resting-state functional connectivity of amygdala subregions in patients with obstructive sleep apnea. *Neuropsychiatr. Dis. Treat.* 15, 977–987. doi: 10.2147/NDT.S191441

Zalesky, A., Fornito, A., and Bullmore, E. T. (2010). Network-based statistic: identifying differences in brain networks. *Neuroimage* 53, 1197–1207. doi: 10.1016/j.neuroimage.2010.06.041

Zhang, Q., Qin, W., He, X., Li, Q., Chen, B., Zhang, Y., et al. (2015). Functional disconnection of the right anterior insula in obstructive sleep apnea. *Sleep Med.* 16, 1062–1070. doi: 10.1016/j.sleep.2015.04.018

Zhang, Q., Wang, D., Qin, W., Li, Q., Chen, B., Zhang, Y., et al. (2013). Altered resting-state brain activity in obstructive sleep apnea. *Sleep* 36, 651–659B. doi: 10.5665/sleep.2620

Zhou, L., Liu, G., Luo, H., Li, H., Peng, Y., Zong, D., et al. (2020a). Aberrant Hippocampal Network Connectivity Is Associated With Neurocognitive Dysfunction in Patients With Moderate and Severe Obstructive Sleep Apnea. *Front. Neurol.* 11:580408. doi: 10.3389/fneur.2020.580408

Zhou, L., Shan, X., Peng, Y., Liu, G., Guo, W., Luo, H., et al. (2020b). Reduced regional homogeneity and neurocognitive impairment in patients with moderate-to-severe obstructive sleep apnea. *Sleep Med.* 75, 418–427. doi: 10.1016/j.sleep.2020.09.009

Zhu, X., Yuan, F., Zhou, G., Nie, J., Wang, D., Hu, P., et al. (2020). Cross-network interaction for diagnosis of major depressive disorder based on resting state functional connectivity. *Brain Imag. Behav.* 15, 1279–1289. doi: 10.1007/s11682-020-00326-2

frontiers | Frontiers in Neuroscience

TYPE Original Research
PUBLISHED 09 August 2022
DOI 10.3389/fnins.2022.976229

OPEN ACCESS

EDITED BY
Yao Wu,
Children's National Hospital,
United States

REVIEWED BY
Jun Shi,
Shanghai University, China
Mingli Zhang,
McGill University Health Centre,
Canada

*CORRESPONDENCE
Yande Ren
8198458ryd@qdu.edu.cn

SPECIALTY SECTION
This article was submitted to
Brain Imaging Methods,
a section of the journal
Frontiers in Neuroscience

RECEIVED 23 June 2022
ACCEPTED 15 July 2022
PUBLISHED 09 August 2022

CITATION
Zhao F, Pan H, Li N, Chen X, Zhang H, Mao N and Ren Y (2022) High-order brain functional network for electroencephalography-based diagnosis of major depressive disorder.
Front. Neurosci. 16:976229.
doi: 10.3389/fnins.2022.976229

High-order brain functional network for electroencephalography-based diagnosis of major depressive disorder

Feng Zhao[1], Hongxin Pan[1], Na Li[1], Xiaobo Chen[1], Haicheng Zhang[2], Ning Mao[2] and Yande Ren[3]*

[1]School of Computer Science and Technology, Shandong Technology and Business University, Yantai, China, [2]Department of Radiology, Yantai Yuhuangding Hospital, Yantai, China, [3]Department of Radiology, The Affiliated Hospital of Qingdao University, Qingdao, China

Brain functional network (BFN) based on electroencephalography (EEG) has been widely used to diagnose brain diseases, such as major depressive disorder (MDD). However, most existing BFNs only consider the correlation between two channels, ignoring the high-level interaction among multiple channels that contain more rich information for diagnosing brain diseases. In such a sense, the BFN is called low-order BFN (LO-BFN). In order to fully explore the high-level interactive information among multiple channels of the EEG signals, a scheme for constructing a high-order BFN (HO-BFN) based on the "correlation's correlation" strategy is proposed in this paper. Specifically, the entire EEG time series is firstly divided into multiple epochs by sliding window. For each epoch, the short-term correlation between channels is calculated to construct a LO-BFN. The correlation time series of all channel pairs are formulated by these LO-BFNs obtained from all epochs to describe the dynamic change of short-term correlation along the time. To construct HO-BFN, we cluster all correlation time series to avoid the problems caused by high dimensionality, and the correlation of the average correlation time series from different clusters is calculated to reflect the high-order correlation among multiple channels. Experimental results demonstrate the efficiency of the proposed HO-BFN in MDD identification, and its integration with the LO-BFN can further improve the recognition rate.

KEYWORDS

electroencephalography, brain functional networks, major depressive disorder, high-order brain functional network, disease classification

Introduction

Major depressive disorder (MDD) is a kind of common brain disease that is characterized by persistent and significant low mood, slow thinking, and cognitive function impairment (Holma et al., 2014; LeMoult and Gotlib, 2019; Liu et al., 2020). In the statistics of the World Health Organization (WHO), MDD has become the second largest serious disease in the world (Melchior et al., 2013) and has brought a heavy burden on patients and their families (Zhang et al., 2019). According to medical researches, the accurate early identification of MDD is important, and it can not only effectively relieve the pain of the patients, but also directly reduce the tragedy of suicide (Cipriani et al., 2018; Grunebaum et al., 2018). However, early neuroimaging-based MDD diagnosis is very challenging, because the changes of brain functional connectivity (FC) are considerably complicated. Electroencephalography (EEG) of high temporal resolution (Babajani and Soltanian-Zadeh, 2006) can well describe the temporal evolution of complex FC during brain activity and thereby becomes the best choice for MDD research.

Brain functional network (BFN) constructed based on EEG has been widely used in the diagnosis of brain diseases (Wang et al., 2015; Li et al., 2020). Since brain activity is dynamic in nature, some studies have shown that the dynamic change of FC over the whole scanning time may be the intrinsic feature of brain function (Damaraju et al., 2014; Cohen and D'Esposito, 2016; Kudela et al., 2017). Many studies try to describe the dynamic changes of FC between channels by using sliding windows to construct BFN and the relationship between these dynamic changes and brain diseases (Wee et al., 2016; Guo et al., 2017; Sun et al., 2019; Zhang et al., 2021). Sun et al. (2019) constructed BFN based on EEG signals by sliding windows, confirming that MDD had abnormal cognitive processing. Zhang et al. (2021) used sliding windows to construct BFN based on EEG signals, and the results showed that the brain regions of MDD patients were significantly altered.

Although the aforementioned EEG-based BFN helps us to understand the brain activities of the MDD patients, most of them only reflect the low-order FC (LO-FC) between two channels (as shown in **Figure 1A**), ignoring the fact that high-order FC (HO-FC) among channels could also be changed for MDD patients (as shown in **Figure 1B**). For the ease of description, we call BFN based on conventional LO-FC as low-order BFN (LO-BFN). In essence, brain activity is complex, and the HO-FC usually contains more abstract information than the LO-FC, and it helps to reveal high-level and more complex interaction information (Chen et al., 2016). Therefore, it is of clinical significance to investigate effective methods of constructing a high-order BFN (HO-BFN) that better reflects the complex interaction among multiple channels and simulates the mechanisms of the deep brain, providing rich discriminative information for the diagnosis of mental disorders (Plis et al., 2014; Chen et al., 2016; Zhang et al., 2016).

In this work, we propose a novel method to construct HO-BFN for MDD classification. Specifically, the entire EEG signals of a subject are divided into multiple overlapping time series by sliding windows, and the correlation between two channels within one window is computed as the LO-FC. The network constructed by the LO-FC is called the LO-BFN, reflecting the dynamic change of FC throughout the whole scanning time. Significantly different from the LO-BFN, each vertex of the HO-BFN represents one pair of channels, and each edge represents the correlation between the channel pairs. In this way, the HO-BFN involves more channels and can reveal high-level and more interaction among brain regions.

It is noted that the scale of the HO-BFN is very huge, and it may suffer from the dilemmas of high dimensionality and small size samples. The main reason lies in the fact that the scale of HO-BFN will increase exponentially when the number of EEG channels increases. To address the issue, we introduce hierarchical clustering (Yu et al., 2015) to construct HO-BFN. In other words, similar LO-FC time series are clustered into one group, and the average LO-FC time series are computed for each group. After that, the HO-BFN is constructed based on the correlation between the two groups. As a result, the HO-BFN constructed by hierarchical clustering can not only reduce computing time and memory requirements, but also reflect the HO-FC among multiple channel pairs (more than four channels) and capture more useful and complex information.

In summary, the main contributions of the paper line are twofold: (1) A HO-FC representation strategy is proposed to capture high-order interactions among multiple channels of EEG signals. In fact, the HO-BFN is used to characterize the complex interactions among brain regions, and it has been applied in fMRI and achieved good results (Zhao et al., 2018). However, to the best of our knowledge, few studies have used EEG-based HO-BFN to reveal the complex interactions among EEG channels. (2) The HO-BFN is constructed in both time and frequency domains based on the "correlation's correlation" strategy. Specifically, we first compute the correlation between two channels to obtain the LO-BFN, and then the HO-BFN is subsequently derived by computing the correlation between each pair of channels from the LO-BFN. Then, we further apply HO-BFN to computer-aided diagnosis for MDD. The experimental results show that HO-BFN provides complementary identification information to the LO-BFN and that combining HO-BFN and LO-BFN can further improve the accuracy of MDD diagnosis.

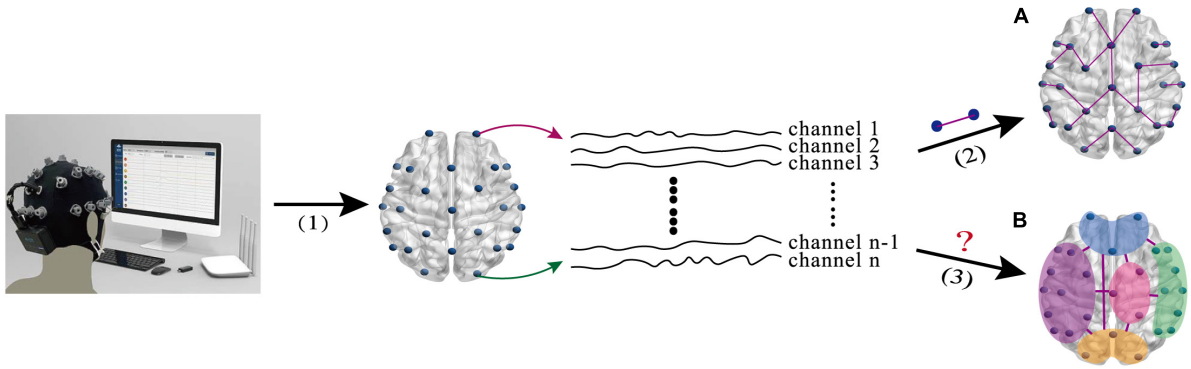


FIGURE 1
The intuitive explanation of brain network research based on EEG signals. **(A)** The low-order FC (LO-FC) between two channels. **(B)** High-order FC (HO-FC) among channels. (1) EEG signal acquisition(1) EEG signal acquisition, (2) calculation of correlation between two channels, and (3) calculation of correlation among multiple channels.

Materials and data preprocessing

The EEG data used in this study came from the publicly available Multi-modal Open Dataset for Mental-disorder Analysis (MODMA) dataset (Cai et al., 2020). It included 24 patients with MDD (12 male and 12 female) and 29 normal controls (NC) (20 male and 9 female). The MDD group was 16–52 years old, and the NC group was 19–51 years old; all the subjects were right-handed, and their education level was primary school or above. In addition, MDD patients had a health Questionnaire 9-item (PHQ-9) (Spitzer et al., 1999) score greater than or equal to 5 and had not received psychotropic medication for 2 weeks. **Table 1** shows the statistics of the subjects.

In data acquisition, 128-channel HydroCel Geodesic Sensor Net and Net Station acquisition software were used to record EEG signals for five minutes. Taking Cz as the reference, the sampling rate was 250 Hz. In order to reduce the interference of EEG data, the subjects were required to close their eyes and keep awake to avoid any unnecessary eye movement, saccade, and blink. The collected EEG data were filtered by 0.1–40 Hz and inhibited by 48–52 Hz to eliminate the data interference caused by baseline drift and electrical interference.

TABLE 1 Demographic information of the subjects.

	MDD	NC	*p*-value
Gender (M/F)	12/12	20/9	0.1600^a
Age (mean ± SD)	30.9 ± 21.1	30.9 ± 20.1	0.9880^b
PHQ-9 (mean ± SD)	18.3 ± 7.3	2.6 ± 2.6	0.0000^b
GAD-7 (mean ± SD)	13.4 ± 11.4	2.1 ± 4.9	0.0000^b

MDD, major depression disorder; NC, normal control; M, male; F, female; PHQ-9, Patient Health Questionnaire-9item; GAD-7, generalized anxiety disorder-7. p^a: Statistical significance level was calculated by χ^2-test; p^b: Statistical significance level was obtained by two-sample, two-tailed *t*-test.

The processed data was then re-referenced against REST (Yao, 2001). Finally, after the above steps, some high-power content was contained in the remaining data points and some EEG epochs were removed by the Artifact Subspace Reconstruction (ASR) plugin (Chang et al., 2018; Pion-Tonachini et al., 2018). In this study, theta (4–8 Hz), alpha (8–13 Hz), and beta (13–40 Hz) bands calculated by fast Fourier transform were selected in the frequency domain, which had been proved to be much distinct in the identification of depression (Nyström et al., 1986; Knott et al., 2001; Jaworska et al., 2012).

Methods

Figure 2 shows the overall pipeline of our method, which includes six steps: (1) constructing LO-BFN by sliding window; (2) clustering all low-order correlation time series; (3) constructing HO-BFN by calculating the correlation between clusters; (4) selecting and extracting features from each constructed BFN; (5) constructing support vector machines (SVMs) based on the selected features in both LO-BFN and HO-BFN; and (6) fusing the decision scores of multiple SVMs to predict whether each subject is MDD or NC.

Construction of low-order BFN

In order to construct LO-BFN, we first use the sliding window to divide the entire EEG signals into $H = [(M - W)/s] + 1$ overlapping windows, where M is the image volume during the entire scan period, and W and s are the window width and step size of the sliding window, respectively. Then, we calculate the correlation between $x_i^l(h)$ and $x_j^l(h)$, where $x_i^l(h)$ and $x_j^l(h)$ denote the i-th and j-th channels under the h-th window of the l-th subject, respectively.

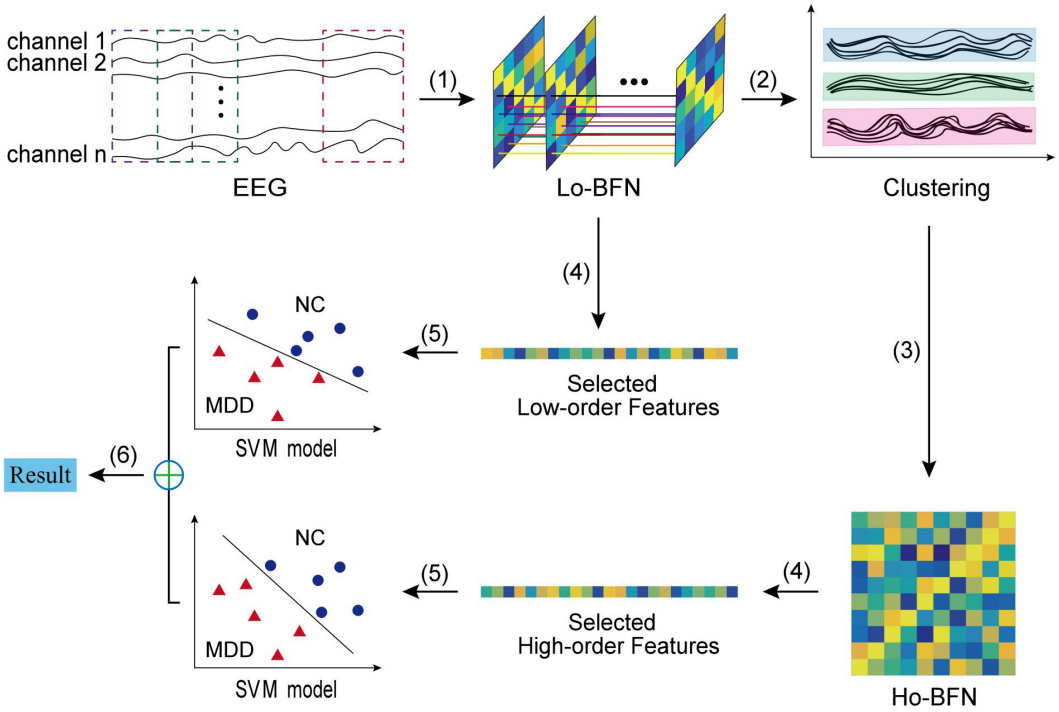


FIGURE 2
The flowchart of the proposed BFN classification framework, including six main steps: (1) constructing LO-BFN; (2) clustering the time series; (3) constructing HO-BFN; (4) feature selection; (5) constructing SVM model; and (6) classification fusion.

In the frequency domain, we use the phase lag index (PLI) (Stam et al., 2007; Peraza et al., 2012) to calculate the channel correlation, which is robust to volume conduction artifacts. The PLI is denoted as:

$$C_{ij}^{l} = \left| \frac{1}{N} \sum_{n=1}^{N} sign\left(\varphi_i^l(t_n) - \varphi_j^l(t_n) \right) \right| \quad (1)$$

where N is the sample number, *sign* is the sign function, and $\varphi_i^l(t_n) - \varphi_j^l(t_n)$ is the phase synchronization of channels x_i^l and x_j^l at time t_n . Among them, $\varphi^l(t_n)$ can be obtained by analyzing the signal based on Hilbert transform (Bruns, 2004).

On the other hand, in the time domain, we calculate the channel correlation by using the Pearson correlation coefficient (PCC) (Eslami and Saeed, 2018) as follows:

$$C_{ij}^{l} = corr(x_i^l, x_j^l) \quad (2)$$

Therefore, for the l-th subject, the h-th subnetwork of LO-BFN is constructed as $G_L^l(h) = \left(\left\{ x_i^l(h) \right\}, \left\{ C_{ij}^l(h) \right\} \right)$ (as shown in **Figure 3A**), where $\left\{ x_i^l(h) \right\}$ is vertices and $\left\{ C_{ij}^l(h) \right\}$ is the weights of the edges connecting the i-th and j-th nodes. Then, we construct H subnetworks to form LO-BFN $G_L^l = \left[G_L^l(1), G_L^l(2), \ldots, G_L^l(H) \right]$ for each subject, which describes the change in FC strength of all channel pairs over time.

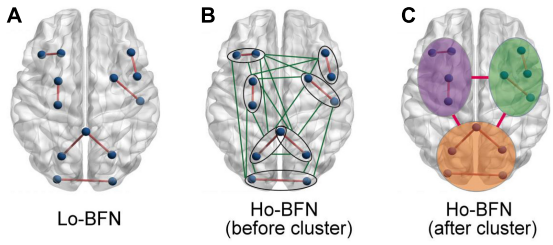


FIGURE 3
Schematic diagram of different BFNs. **(A)** Schematic diagram of the brain region of LO-BFN. **(B)** Schematic diagram of brain regions of HO-BFN before clustering. **(C)** Schematic diagram of brain regions of HO-BFN after clustering.

Construction of high-order BFN

In order to capture high-level FCs, we adopt the strategy of "correlation's correlation". That is, based on the LO-BFN, the PCC is used to calculate the correlation between the LO-FC time series of the l-th subject, which is called HO-FC, denoted as:

$$H_{ij,pq}^{l} = corr(C_{ij}^{l}, C_{pq}^{l}) \quad (3)$$

where C_{ij}^l is the LO-FC time series between the i-th and the j-th channels of the l-th subject and C_{pq}^l is the LO-FC time

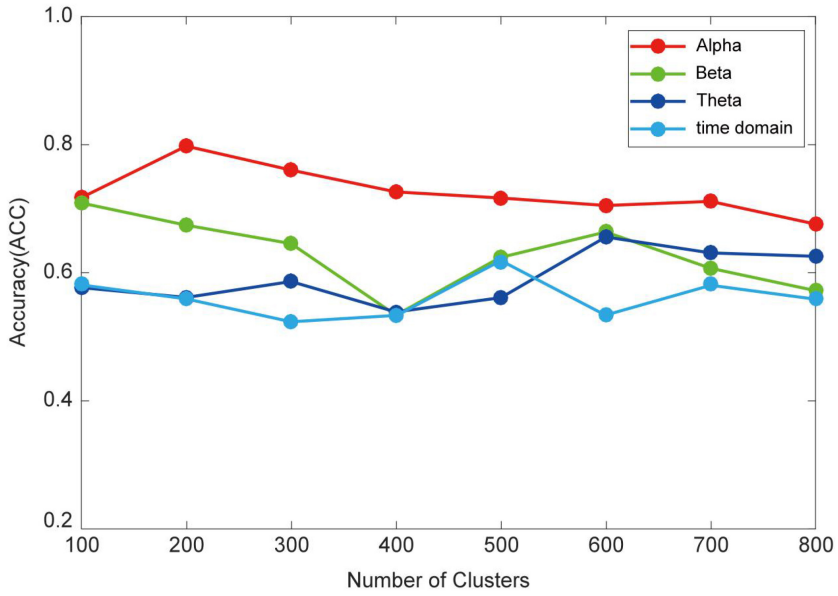


FIGURE 4
Recognition accuracy of HO-BFN with different number of clusters.

TABLE 2 MDD classification using different BFNs.

Network	ACC (%)	TPR (%)	TNR (%)	PPV	NPV	F1 (%)
Alpha-LO	60.67	59.17	60.83	64.81	54.86	61.78
Alpha-HO	79.78	80.83	78.33	82.25	77.08	81.48
Alpha-Fu	83.98	84.34	82.67	85.80	81.39	85.01
Beta-LO	64.52	66.17	62.17	68.17	59.96	67.07
Beta-HO	70.90	71.50	69.83	74.39	66.68	72.71
Beta-Fu	76.63	77.83	74.00	78.82	73.24	78.25
Theta-LO	69.63	71.50	68.50	73.43	66.29	72.23
Theta-HO	65.58	64.50	67.17	70.46	60.68	67.13
Theta-Fu	77.65	80.83	74.00	79.44	76.04	80.08
Frequency domain-Fu	**86.62**	**89.17**	**82.67**	**86.48**	**86.19**	**87.77**
Time domain-LO	59.16	60.50	58.17	63.75	54.56	61.94
Time domain-HO	61.23	62.67	60.33	65.85	56.85	64.11
Time domain-Fu	65.79	69.67	60.33	68.17	61.92	68.78

LO = LO-BFN; HO = HO-BFN; Fu, the fusion of LO-BFN and HO-BFN. For example, alpha-LO means the LO-BFN in the alpha band, and note that frequency domain-Fu means the fusion of all BFNs of three bands in the frequency domain. Values highlighted in bold indicate the best results.

series between the p-th and the q-th channels of the l-th subject. Therefore, $H^{l}_{ij,pq}$ can represent the HO-FC among the four channels of the l-th subject at most, that is, the correlation between the FC between the i-th and j-th channels and the FC between the p-th and q-th channels. Physiologically, difference in FC among different channels in MDD patients and healthy individual subjects can be used to identify MDD.

If we have 128 channels in our study, the dimensionality of the constructed LO-BFN is 128 × 128. Thus, a large-scale HO-BFN will be constructed (as shown in Figure 3B) in Eq. 3; that is, the dimensionality is $(128 \times 128)^2$, and the constructed network contains at least thousands of nodes and millions of edges. It is a critical problem that the dimensionality is too large, and it will introduce high computational complexity for the subsequent feature extraction and selection procedures. Besides, the generalization performance of the HO-BFN learning system may also degrade.

To this end, we will reduce the network dimensionality by clustering the LO-FC time series. Specifically, the LO-FC time series of the subjects are clustered into different clusters to find the potential interaction patterns. Then, the HO-FC between the respective average LO-FC time series in clusters

TABLE 3 Brain regions corresponding to channels of interest.

Brain area	Channels
Frontal (F)	E2, E3, E4, E5, E9, E10, E11, E12, E15, E16, E18, E19, E22, E23, E24, E26, E27, E123, E124
Left temporal (LT)	E28, E33, E34, E35, E39, E40, E41, E45, E46, E47, E50, E51, E52, E58
Central (C)	E6, E7, E13, E20, E29, E30, E31, E36, E37, E42, E53, E54, E55, E79, E80, E86, E87, E93, E104, E105, E106, E111, E112, E118
Right temporal (RT)	E92, E96, E97, E98, E101, E102, E103, E108, E109, E110, E115, E116, E117, E122
Posterior (P)	E59, E60, E61, E62, E65, E66, E67, E70, E71, E72, E75 E76, E77, E78, E83, E84, E85, E90, E91,

The channels in Table 3 are channels of interest, while the other channels are marginal and do not belong to the brain regions classified above.

is calculated. Compared with the previous large-scale HO-BFN, the dimensionality of the HO-BFN constructed in this way is greatly reduced (as shown in Figure 3C). This method not only preserves important interactive information, but also avoids the problems of high computational complexity and low generalization performance.

In order to ensure that the clustering results in different subjects are consistent, the FC matrices of all subjects are first accumulated together, so that to connect the time series of LO-FC in the same channel pairs in all subjects into a long vector. That is, the long vector connected by the time series $\{C_{ij}^l\}_{1<=l<=R}$ of the LO-FC between the i-th and j-th channels of all subjects is $C_{ij} = \left[C_{ij}^1, C_{ij}^2, \cdots, C_{ij}^R\right]$, where R is the number of subjects.

After the C_{ij} is obtained, we divide it into k clusters by hierarchical clustering. The cluster centers are calculated by averaging all the FC long vectors in the k-th cluster of the l-th subject as the following Eq. 4:

$$\tilde{C}_k^l = \frac{\sum_{C_{ij}\in\omega_k} C_{ij}^l}{|\omega_k|} \tag{4}$$

where $|\omega_k|$ denotes the total number of FC long vectors in cluster k. Finally, the PCC between the two clusters $\tilde{C}_{k1}^l$ and $\tilde{C}_{k2}^l$ of the l-th subject is calculated as follows:

$$\tilde{H}_{k1,k2}^l = corr(\tilde{C}_{k1}^l, \tilde{C}_{k2}^l) \tag{5}$$

Finally, we obtain a small-scale HO-BFN $G_{\tilde{H}}^l = \left(\left\{\tilde{C}_{k1}^l\right\}, \left\{\tilde{H}_{k1,k2}^l\right\}\right)$, taking $\left\{\tilde{C}_{k1}^l\right\}$ as vertices and $\left\{\tilde{H}_{k1,k2}^l\right\}$ as the weights of edges.

Feature extraction, selection, classification, and fusion

Both LO-BFN and HO-BFN of the l-th subject, i.e., G_L^l and $G_{\tilde{H}}^l$, are used for the subsequent classification. Due to the possible phase mismatch of all FC matrices of LO-BFN in each subject, the dynamic characteristics of different subjects do not completely correspond. Therefore, to avoid this situation, we calculate the average FC matrix of each subject's LO-BFN as the low-order feature. For the ease of computation, we vectorize the averaged LO-FC matrix of the l-th subject into f_L^l, which is called the low-order feature vector. Similarly, we vectorize the FC matrix $\tilde{H}^l$ of HO-BFN of the l-th subject to $f_{\tilde{H}}^l$ as the high-order feature vector.

Both low-order feature vectors and high-order feature vectors have a large number of features, introducing irrelevant or redundant information for subsequent MDD classification. Therefore, we use t-test and Least Absolute Shrinkage and Selection Operator (LASSO) (Tibshirani, 2011) methods to select features for high-order and low-order feature vectors, which can effectively remove redundant features. Specifically, we first perform t-test on both the low-order feature vector f_L and the high-order feature vector $f_{\tilde{H}}$ on the training set and select the features that are significantly different from f_L and $f_{\tilde{H}}$ as preliminary features, denoted $\bar{f}_L$ and $\bar{f}_{\tilde{H}}$, respectively.

Then, we use LASSO to further remove redundant features and select the features most related to MDD. Let I^l be the labels for the l-th subject. Specifically, if the l-th subject is MDD, $I^l = -1$; if the l-th subject is NC, $I^l = +1$; and α is set to be the weight vector for feature selection. The objective of LASSO is defined as:

$$\min_{\alpha} \frac{1}{2}\sum_{l=1}^{N}\left|\left|I^l\left\langle -\bar{f}^l,\ \alpha\right\rangle\right|\right|_2^2 + \lambda||\alpha||_1 \tag{6}$$

where $\langle\cdot,\cdot\rangle$ is the inner product operator and λ is the regularization parameter. For simplicity, let $\bar{\bar{f}}_L$ and $\bar{\bar{f}}_{\tilde{H}}$ denote selected the final feature sets from the feature vectors $\bar{f}_L$ and $\bar{f}_{\tilde{H}}$, respectively.

Finally, the MDD is classified by using SVM with a linear kernel in this paper. Herein, we train two linear SVM classifiers by using both $\bar{\bar{f}}_L$ and $\bar{\bar{f}}_{\tilde{H}}$ features, respectively. The final results are obtained by fusing the decision scores of two SVM classifiers by linear combination. Among them, we set the weight $\beta \in [0.1, 0.2, \ldots, 0.9]$ for each SVM. The weights of the F classifiers to be fused are set to be $\beta_1, \beta_2, \ldots, \beta_F$ and satisfy $\beta_1 + \beta_2 + \ldots + \beta_F = 1$.

Experiments results

To evaluate the effectiveness of our proposed method, we analyze the impact of clustering parameters on HO-BFN. The classification ability of LO-BFN, HO-BFN, and their fusion for MDD is evaluated by six different indicators, i.e., accuracy (ACC), sensitivity or true positive rate (TPR), specificity or true negative rate (TNR), precision or positive predictive value (PPV), negative predictive value (NPV), and F1 score. Finally,

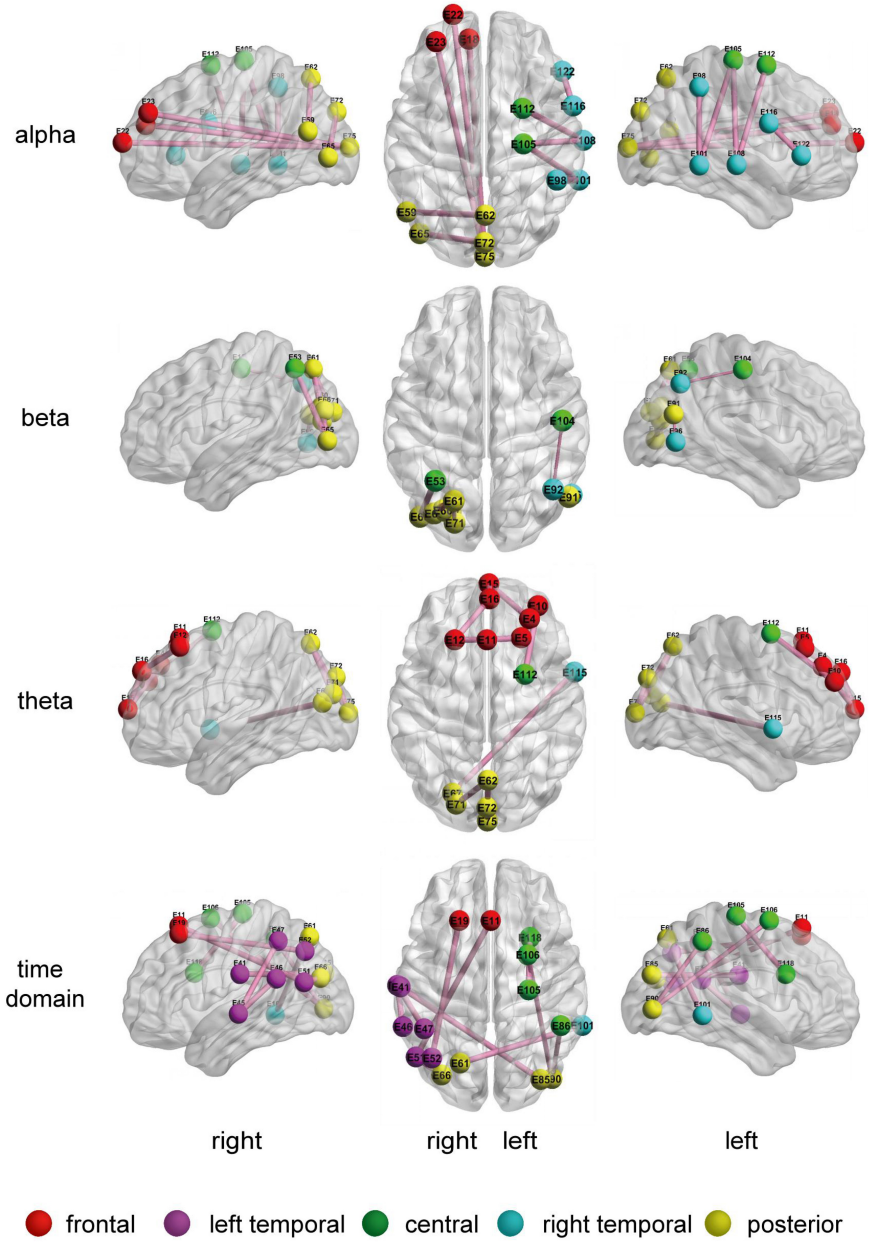


FIGURE 5
The 10 most frequently selected connection diagrams for different LO-BFNs.

we conduct feature analysis to understand the role of each channel in MDD diagnosis.

Furthermore, we adopt a nested ten-fold cross-validation (CV) strategy consisting of two nested loops to evaluate the effectiveness of our proposed method. In the outer loop, all data are randomly divided into ten subsets of roughly the same size, where one subset is selected as the testing set, while the other nine subsets are selected as the training set. In the inner loop, the data of the training set are merged and redivided into ten subsets of similar size, nine of which are used to adjust the hyper-parameters and one for model evaluation. We report the average accuracy of classification results across the ten-fold CV. Then, in order to avoid any possible bias in fold selection, this procedure is repeated 10 times, with a different random partitioning of samples each time. Finally, the average accuracy of 10 repetitions is reported. Since the performance of our method also depends on hyper-parameters, such as W and s in sliding window, k in clustering, p in t-test, λ in LASSO, and c in SVM model. The optimal hyper-parameters can be determined when the average classification accuracy

TABLE 4 The brain regions selected from different HO-BFNs.

Cluster	Alpha	Beta	Theta	Time domain
Cluster 1	LT, LC, LP-RP	LF-RF, RC, RT	LT, LC	LC, LP, RT
Cluster 2	LF-RF, LT-RT, LC-RC, RP	LF, LT, RC, RP	LF-RF, LC-RC, RT	LF-RF, LT-RT, LC-RC
Cluster 3	LF-RF	RC, RT, RP	LF-RF, LC	LF-RF, LT, LC
Cluster 4	LF-RF, LT, LP-RP	LF-RF, LT-RT, LC-RC, LP-RP	LF, LC	LF-RF, LT-RT, LC-RC

reaches its optimum. In our experiment, we fix the size of the sliding window, i.e., $W = 10000$, $s = 1000$, and determine the optimal values of other parameters within the following range: $k \in [100, 200, \ldots, 800]$, $p \in [0.01, 0.02, \ldots, 0.05]$, $\lambda \in [0.1, 0.2, \ldots, 0.9]$, and $c \in [2^{-4}, 2^{-3}, \ldots, 2^{4}]$.

The clustering effectiveness on high-order BFN

To reduce the computational complexity of the HO-BFN, the hierarchical clustering is employed to construct HO-BFNs. The hyper-parameter k in hierarchical clustering indicates the cluster number, and it has a crucial influence on the constructed HO-BFN, and further affects the final classification results. In the experiment, we optimize HO-BFN by adjusting the clustering number k. Figure 4 shows the classification results when k takes different numbers.

As can be seen from Figure 4 that for the clustering parameter k, when the time domain takes the value of 500, alpha takes the value of 200, beta takes the value of 100, and theta takes the value of 600, the HO-BFN generates a relatively satisfactory classification result. The ACC is greatly affected by k; i.e., classification performance is very sensitive to clustering parameters. Different HO-BFNs have different performances, indicating that different HO-BFNs contain different levels of MDD diagnosis information. Therefore, we can conclude that it is necessary to select k carefully toward a better understanding of dynamics in brains.

Comparison of major depressive disorder diagnosis using different brain functional networks

In this subsection, we further study how HO-BFN contributes to MDD diagnosis. We construct and analyze HO-BFN from the time domain and alpha, beta, and theta bands in the frequency domain, respectively. In the experiment, we train and test the classifiers of LO-BFN and HO-BFN, respectively, and determine the parameter combination that can produce the best ACC.

Due to the complex connections in the brain, it is difficult to fully capture the relationship between different brain regions through a single type of BFN. In order to further improve the classification performance, we adopt the linear fusion of the SVM integrated decision score to combine LO-BFN and HO-BFN (Zhao et al., 2020) and analyze their fusion performance. In addition, we also believe that different BFNs constructed by the three bands in the frequency domain can reflect FC between channels from different views, which are complementary. Therefore, we further linearly fuse the decision scores of BFNs in three bands. Table 2 shows the classification performance of different BFNs, and the best classification results are highlighted in bold.

From Table 2, we can draw the following conclusions: (1) BFNs constructed in different ways have different performance, implying that each BFN provides meaningful and various information for MDD identification; (2) fusing LO-BFN and HO-BFN is better than that of a single network, and the ACC is relatively increased in 4%, indicating that LO-BFN and HO-BFN are complementary to each other in classifying MDD; (3) the performance of each brain network in the frequency domain is obviously better than that of the BFNs in the time domain, indicating that it is more effective to extract features in the frequency domain.

The most discriminative features for major depressive disorder diagnosis

To identify the most discriminative features in MDD diagnosis, we select ten frequently selected LO-FC features and two frequently selected HO-FC features based on the t-test and LASSO regression ten-fold cross-validations of ten times. The higher selection frequency of FC indicates stronger reliability and discriminative ability.

Similar to previous studies (Bian et al., 2014), we divide the brain into five regions, i.e., frontal (F), left temporal (LT), central (C), right temporal (RT), and posterior (P). The frontal region is further divided into left frontal (LF) and right frontal (RF), the central region is further divided into left central (LC) and right central (RC), and the posterior is further divided into left posterior (LP) and right posterior (RP). Table 3 shows the details of these brain regions.

Figure 5 shows the 10 most discriminative FC feature maps for the different LO-BFNs. The node color indicates the brain region the channel belongs to, the connection line represents the correlation between two channels, and the line width indicates the frequency. The thicker line represents the higher frequency. It can be seen that the frequently selected FC features in LO-BFN often appear in the LF, RC, RT, and LP regions in the alpha band; left central-right central (LC-RC), RT and left posterior-right posterior (LP-RP) regions in the beta band; the left frontal-right frontal (LF-RF), RC, RT, and LP regions in the theta band,

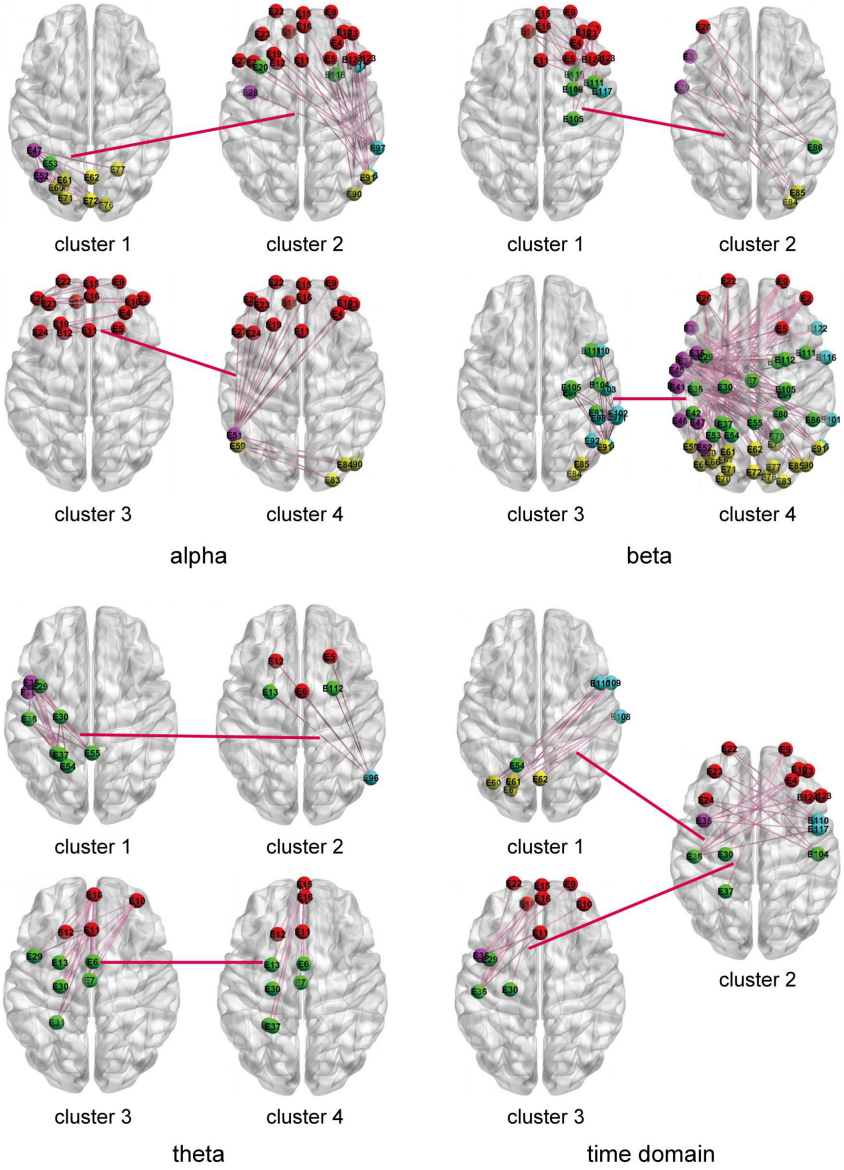


FIGURE 6
The two most frequently selected connection diagrams for different HO-BFNs. Each brain diagram represents the FC of interest channels across the brain region in a cluster.

and the LF, RC, left temporal-right temporal (LT-RT) and left posterior-right posterior (LP-RP) regions in the time domain.

The HO-BFN is constructed based on clustering. Each cluster is used as the vertex of the network, and the HO-FC is used as the edge of the network. For HO-BFN, we select the two most discriminative cluster pairs and describe the FC feature diagram between these two cluster pairs in Figure 6, where each brain map represents a cluster, the connection between brain maps represents the connection between clusters, that is, HO-FC, and the connections in the brain map represent the FC between channels belonging to this cluster, that is, LO-FC. Table 4 shows the brain regions involved in the most discriminative clusters that are selected from different HO-BFNs. From Figure 6 and Table 4, we can observe that the most frequently selected features in HO-BFN are mainly distributed in LF-RF, LT-RT, and LC-RC regions of the theta band, other brain regions except RT in the time domain, as well as the whole brain region of the alpha and beta band.

Discussion

In this paper, we propose a novel method to construct HO-BFN based on the "correlation's correlation" strategy, which

can simultaneously capture low-order features reflecting FC information between any two channels and high-order features reflecting FC information among multiple channels, so as to better simulate the mechanisms of the deep brain and provide more discriminative information for the diagnosis of mental disorders. We believe that different BFNs can mine disease-disturbed BFN variation information from different aspects, which has better performance in MDD classification experiments. According to the experimental results, we will discuss HO-BFNs in more detail.

In the experiment to explore the influence of the clustering parameter on the classification accuracy of HO-BFN, we find that the classification result changes with the change of the clustering parameter. When the clustering parameter is too small or too large, the accuracy of diagnosis will gradually decrease. This can be understood from two aspects: (1) when the clustering parameter is too small, the LO-FC time series with different dynamic changes may be divided into the same cluster, which will reduce the similarity of each cluster, so that the HO-BFN constructed by the mean sequence of each class as the vertex is unreliable; and (2) when the clustering parameter is too large, the LO-FC time series with similar dynamic changes may be divided into different clusters, which will increase the number of features extracted from HO-BFN, thus resulting in more redundant features and causing the decrease in accuracy and generalization ability. Thus, choosing the suitable number of clusters for HO-BFN is the key to achieving MDD classification and improving the classification accuracy.

According to the experimental results in **Table 2**, we find that the proposed BFNs in the frequency domain are usually more discriminative than in the time domain, and the alpha and theta bands in the frequency domain are more discriminative than the beta band, indicating that the brain functional structure of MDD patients has undergone significant changes in these two bands. Therefore, we believe that the BFNs of alpha and theta play an important role in the pathogenesis of MDD. Several previous studies have also reported the same or similar conclusion: abnormal brain function in MDD patients occurs at certain frequency bands. For example, Fingelkurts et al. (2010) showed that the FC in the alpha and theta bands of EEG in MDD was impaired. Hosseinifard et al. (2013) found that there were differences in alpha bands between the MDD group and the NC group.

To further prove the effectiveness of BFNs in MDD disease diagnosis, we trace the BFNs to which the features of the classifier used for training belong. The experimental results in **Figures 5**, **6** show that: (1) the channel pairs selected by LO-BFN hardly overlap with those of HO-BFN, indicating that the FC features extracted from LO-BFN and HO-BFN are complementary to each other. (2) RC and RT are significantly different in all LO-BFNs, and they are related to the regulation of attention, long-term memory, and emotion. In previous studies, Fan et al. (2013) found that abnormal RT superior gyrus activity could be a potential marker of suicidal tendencies in MDD patients. Sun et al. (2019) observed differences in MDD patients in the F, T, and C of the theta band and in the T and C of the alpha band. Zhang et al. (2021) found significant modifications in brain synchrony of LF, T, and RT in MDD patients. Our observations are generally consistent with these studies (Fan et al., 2013; Sun et al., 2019; Zhang et al., 2021).

However, this study has some limitations. Firstly, either the low-order or the high-order FC is based on the correlation instead of the inherent causality. The relatively small sample size and unbalanced data may also affect the result of the analysis. Although most conclusions obtained by our method are generally consistent with the previous relevant studies, the experimental results may have a potential bias due to the heterogeneity of the experiment (Hassler and Thadewald, 2010). In the future work, we will focus on the influence of nonsensical and biased correlation and investigate more advanced models to calculate correlation, such as causality inference, based on larger datasets to obtain more accurate and sufficient information about the brain changes of MDD patients. Secondly, the sliding window algorithm should set the step size and window width, but we fixed the window width and step length in this study. Further studies will investigate the influence of different parameter settings. Finally, we linearly combine the decision scores of the LO-BFN and HO-BFN at the decision-making level, and this linear combination may not fully mine the complementary information, thereby affecting the classification accuracy. Therefore, in our future work, we will further increase the accuracy of MDD diagnosis by using more advanced information fusion strategies.

Conclusion

In this paper, we propose a framework for constructing HO-BFN based on the "correlation's correlation" strategy and capture the high-order correlations across different channels for MDD diagnosis. We use hierarchical clustering to reduce the computational complexity of the HO-BFN. Experimental results demonstrate that: (1) the proposed HO-BFN can provide discriminative information for the MDD identification. (2) Fusing high-order and low-order BFNs can significantly improve the recognition rate of MDD patients. (3) The most discriminative brain regions are associated with the regulation of attention, and these findings are consistent with the daily behavior of MDD patients.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: http://modma.lzu.edu.cn/data/index/.

Author contributions

FZ: conceptualization, methodology, writing – review, and editing. HP: conceptualization, software, writing – original draft, methodology, formal analysis, investigation, and validation. NL: validation. XC, HZ, NM, and YR: writing – review, and editing. All authors contributed to the article and approved the submitted version.

Funding

FZ was supported in part by the National Natural Science Foundation of China (62176140, 82001775, 61772319, 61873177, 61972235, 61976125, and 61976124) and Doctoral Scientific Research Foundation of Shandong Technology and Business University (No. BS202016).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Babajani, A., and Soltanian-Zadeh, H. (2006). Integrated MEG/EEG and fMRI model based on neural masses. *IEEE Trans. Biomed. Eng.* 53, 1794–1801. doi: 10.1109/TBME.2006.873748

Bian, Z., Li, Q., Wang, L., Lu, C., Yin, S., and Li, X. (2014). Relative power and coherence of EEG series are related to amnestic mild cognitive impairment in diabetes. *Front. Aging Neurosci.* 6:11. doi: 10.3389/fnagi.2014.00011

Bruns, A. (2004). Fourier-, Hilbert- and wavelet-based signal analysis: Are they really different approaches? *J. Neurosci. Methods* 137, 321–332. doi: 10.1016/j.jneumeth.2004.03.002

Cai, H., Gao, Y., Sun, S., Li, N., Tian, F., Xiao, H., et al. (2020). *MODMA Dataset: A Multi-Modal Open Dataset for Mental-Disorder Analysis*. Available online at: http://modma.lzu.edu.cn/data/index/ (accessed February 18, 2021).

Chang, C. Y., Hsu, S. H., Pion-Tonachini, L., and Jung, T. P. (2018). Evaluation of artifact subspace reconstruction for automatic EEG artifact removal. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* 2018, 1242–1245.

Chen, X., Zhang, H., Gao, Y., Wee, C. Y., Li, G., and Shen, D. (2016). High-order resting-state functional connectivity network for MCI classification. *Hum. Brain Mapp.* 37, 3282–3296. doi: 10.1002/hbm.23240

Cipriani, A., Furukawa, T. A., Salanti, G., Chaimani, A., Atkinson, L. Z., Ogawa, Y., et al. (2018). Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: A systematic review and network meta-analysis. *Lancet (London, England)* 391, 1357–1366.

Cohen, J. R., and D'Esposito, M. (2016). The Segregation and integration of distinct brain networks and their relationship to cognition. *J. Neurosci.* 36, 12083–12094.

Damaraju, E., Allen, E. A., Belger, A., Ford, J. M., McEwen, S., Mathalon, D. H., et al. (2014). Dynamic functional connectivity analysis reveals transient states of dysconnectivity in schizophrenia. *NeuroImage Clin.* 5, 298–308.

Eslami, T., and Saeed, F. (2018). Fast-GPU-PCC: A GPU-based technique to compute pairwise Pearson's correlation coefficients for time series data-fMRI study. *High Throughput* 7:11. doi: 10.3390/ht7020011

Fan, T., Wu, X., Yao, L., and Dong, J. (2013). Abnormal baseline brain activity in suicidal and non-suicidal patients with major depressive disorder. *Neurosci. Lett.* 534, 35–40. doi: 10.1016/j.neulet.2012.11.032

Fingelkurts, A. A., Fingelkurts, A. A., Rytsl, H., Suominen, K., Isomets, E., and Khknen, S. (2010). Impaired functional connectivity at EEG alpha and theta frequency bands in major depression. *Hum. Brain Mapp.* 28, 247–261.

Grunebaum, M. F., Galfalvy, H. C., Choo, T. H., Keilp, J. G., Moitra, V. K., Parris, M. S., et al. (2018). Ketamine for rapid reduction of suicidal thoughts in major depression: A midazolam-controlled randomized clinical trial. *Am. J. Psychiatry* 175, 327–335. doi: 10.1176/appi.ajp.2017.17060647

Guo, H., Liu, L., Chen, J., Xu, Y., and Jie, X. (2017). Alzheimer classification using a minimum spanning tree of high-order functional network on fMRI dataset. *Front. Neurosci.* 11:639. doi: 10.3389/fnins.2017.00639

Hassler, U., and Thadewald, T. (2010). Nonsensical and biased correlation due to pooling heterogeneous samples. *J. R. Stat. Soc.* 52, 367–379.

Holma, K. M., Haukka, J., Suominen, K., Valtonen, H. M., Mantere, O., Melartin, T. K., et al. (2014). Differences in incidence of suicide attempts between bipolar I and II disorders and major depressive disorder. *Bipolar Disord.* 16, 652–661.

Hosseinifard, B., Moradi, M. H., and Rostami, R. (2013). Classifying depression patients and normal subjects using machine learning techniques and nonlinear features from EEG signal. *Comput. Methods Programs Biomed.* 109, 339–345.

Jaworska, N., Blier, P., Fusee, W., and Knott, V. (2012). α Power, α asymmetry and anterior cingulate cortex activity in depressed males and females. *J. Psychiatric Res.* 46, 1483–1491. doi: 10.1016/j.jpsychires.2012.08.003

Knott, V., Mahoney, C., Kennedy, S., and Evans, K. (2001). EEG power, frequency, asymmetry and coherence in male depression. *Psychiatry Res.* 106, 123–140.

Kudela, M., Harezlak, J., and Lindquist, M. A. (2017). Assessing uncertainty in dynamic functional connectivity. *NeuroImage* 149, 165–177.

LeMoult, J., and Gotlib, I. H. (2019). Depression: A cognitive perspective. *Clin. Psychol. Rev.* 69, 51–66.

Li, G., Jiang, Y., Jiao, W., Xu, W., Huang, S., Gao, Z., et al. (2020). The maximum eigenvalue of the brain functional network adjacency matrix: Meaning and application in mental fatigue evaluation. *Brain Sci.* 10:92. doi: 10.3390/brainsci10020092

Liu, Q., He, H., Yang, J., Feng, X., Zhao, F., and Lyu, J. (2020). Changes in the global burden of depression from 1990 to 2017: Findings from the Global Burden of Disease study. *J. Psychiatric Res.* 126, 134–140. doi: 10.1016/j.jpsychires.2019.08.002

Melchior, M., Chastang, J. F., Head, J., Goldberg, M., Zins, M., Nabi, H., et al. (2013). Socioeconomic position predicts long-term depression trajectory: A 13-year follow-up of the GAZEL cohort study. *Mol. Psychiatry* 18, 112–121. doi: 10.1038/mp.2011.116

Nyström, C., Matousek, M., and Hällström, T. (1986). Relationships between EEG and clinical characteristics in major depressive disorder. *Acta Psychiatr. Scand.* 73, 390–394.

Peraza, L. R., Asghar, A. U., Green, G., and Halliday, D. M. (2012). Volume conduction effects in brain network inference from electroencephalographic recordings using phase lag index. *J. Neurosci. Methods* 207, 189–199. doi: 10.1016/j.jneumeth.2012.04.007

Pion-Tonachini, L., Hsu, S. H., Chang, C. Y., Jung, T. P., and Makeig, S. (2018). Online automatic artifact rejection using the real-time EEG source-mapping toolbox (REST). *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* 2018, 106–109. doi: 10.1109/EMBC.2018.8512191

Plis, S. M., Sui, J., Lane, T., Roy, S., Clark, V. P., Potluru, V. K., et al. (2014). High-order interactions observed in multi-task intrinsic networks are dominant indicators of aberrant brain function in schizophrenia. *NeuroImage* 102(Pt 1), 35–48. doi: 10.1016/j.neuroimage.2013.07.041

Spitzer, R. L., Kroenke, K., and Williams, J. B. (1999). Validation and utility of a self-report version of PRIME-MD: The PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. *JAMA* 282, 1737–1744.

Stam, C. J., Nolte, G., and Daffertshofer, A. (2007). Phase lag index: Assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources. *Hum. Brain Mapp.* 28, 1178–1193. doi: 10.1002/hbm.20346

Sun, S., Li, X., Zhu, J., Wang, Y., La, R., Zhang, X., et al. (2019). Graph theory analysis of functional connectivity in major depression disorder with high-density resting state EEG data. *IEEE Trans. Neural Syst. Rehabil. Eng.* 27, 429–439. doi: 10.1109/TNSRE.2019.2894423

Tibshirani, R. (2011). Regression shrinkage and selection via the lasso: A retrospective. *J. R. Stat. Soc. Ser. B (Statistical Methodology)* 73, 267–288.

Wang, N., Zhang, L., and Liu, G. (2015). EEG-based research on brain functional networks in cognition. *Biomed. Mater. Eng.* 26(Suppl. 1), S1107–S1114.

Wee, C. Y., Yang, S., Yap, P. T., and Shen, D. (2016). Sparse temporally dynamic resting-state functional connectivity networks for early MCI identification. *Brain Imaging Behav.* 10, 342–356. doi: 10.1007/s11682-015-9408-2

Yao, D. (2001). A method to standardize a reference of scalp EEG recordings to a point at infinity. *Physiol. Meas.* 22, 693–711.

Yu, M., Hillebrand, A., Tewarie, P., Meier, J., van Dijk, B., Van Mieghem, P., et al. (2015). Hierarchical clustering in minimum spanning trees. *Chaos (Woodbury, N.Y.)* 25:023107.

Zhang, B., Yan, G., Yang, Z., Su, Y., Wang, J., and Lei, T. (2021). Brain functional networks based on resting-state EEG data for major depressive disorder analysis and classification. *IEEE Trans. Neural Syst. Rehabil. Eng.* 29, 215–229. doi: 10.1109/TNSRE.2020.3043426

Zhang, H., Chen, X., Shi, F., Li, G., Kim, M., Giannakopoulos, P., et al. (2016). Topographical information-based high-order functional connectivity and its application in abnormality detection for mild cognitive impairment. *J. Alzheimers Dis.* 54, 1095–1112. doi: 10.3233/JAD-160092

Zhang, X., Shen, J., Din, Z. U., Liu, J., Wang, G., and Hu, B. (2019). Multimodal depression detection: Fusion of electroencephalography and paralinguistic behaviors using a novel strategy for classifier ensemble. *IEEE J. Biomed. Health Inform.* 23, 2265–2275. doi: 10.1109/JBHI.2019.2938247

Zhao, F., Chen, Z., Rekik, I., Lee, S. W., and Shen, D. (2020). Diagnosis of autism spectrum disorder using central-moment features from low- and high-order dynamic resting-state functional connectivity networks. *Front. Neurosci.* 14:258. doi: 10.3389/fnins.2020.00258

Zhao, F., Zhang, H., Rekik, I., An, Z., and Shen, D. (2018). Diagnosis of autism spectrum disorders using multi-level high-order functional networks derived from resting-state functional MRI. *Front. Hum. Neurosci.* 12:184. doi: 10.3389/fnhum.2018.00184

frontiers | Frontiers in Neuroscience

TYPE Original Research
PUBLISHED 09 September 2022
DOI 10.3389/fnins.2022.944585

OPEN ACCESS

EDITED BY
Yao Wu,
Children's National Hospital,
United States

REVIEWED BY
Zhang Tao,
Xihua University, China
Wei Zhao,
Hunan Normal University, China

*CORRESPONDENCE
Chuanjun Zhuo
chuanjunzhuotjmh@163.com
Shaohong Zou
zoushaohong_xj@163.com
Hongjun Tian
thj-home@163.com

SPECIALTY SECTION
This article was submitted to Brain Imaging Methods, a section of the journal Frontiers in Neuroscience

RECEIVED 15 May 2022
ACCEPTED 04 August 2022
PUBLISHED 09 September 2022

CITATION
Chen J, Zhang X, Qu Y, Peng Y, Song Y, Zhuo C, Zou S and Tian H (2022) Exploring neurometabolic alterations in bipolar disorder with suicidal ideation based on proton magnetic resonance spectroscopy and machine learning technology. *Front. Neurosci.* 16:944585. doi: 10.3389/fnins.2022.944585

Exploring neurometabolic alterations in bipolar disorder with suicidal ideation based on proton magnetic resonance spectroscopy and machine learning technology

Jiayue Chen[1,2,3], Xinxin Zhang[4], Yuan Qu[5], Yanmin Peng[6,7], Yingchao Song[6,7], Chuanjun Zhuo[1,2,3,8,9,10]*, Shaohong Zou[11]* and Hongjun Tian[1,2]*

[1]Department of Psychiatry, Tianjin Fourth Center Hospital, The Fourth Central Clinical College, Tianjin Medical University, Tianjin, China, [2]Department of Key Laboratory of Real Time Imaging of Brian Circuits in Psychiatry and Neurology (RTIBNP_Lab), Tianjin Fourth Center Hospital, The Fourth Central Clinical College, Tianjin Medical University, Tianjin, China, [3]Department of Psychiatry, School of Basic Medical Science, Tianjin Medical University, Tianjin, China, [4]Department of Medical Imaging, Tianjin Children's Hospital, Tianjin, China, [5]Department of Radiology, People's Hospital of Xinjiang Uygur Autonomous Region, Ürümqi, China, [6]School of Medical Imaging, Tianjin Medical University, Tianjin, China, [7]Tianjin Key Laboratory of Functional Imaging, Tianjin Medical University, Tianjin, China, [8]Psychiatric-Neuroimaging-Genetics and Comorbidity Laboratory (PNGC_Lab), Tianjin Anding Hospital, Mental Health Teaching Hospital of Tianjin Medical University, Tianjin, China, [9]Department of Psychiatry, First Affiliated Hospital of Zhengzhou University, Zhengzhou, China, [10]Department of Psychiatry, First Hospital/First Clinical Medical College of Shanxi Medical University, Taiyuan, China, [11]Department of Clinical Psychology, People's Hospital of Xinjiang Uygur Autonomous Region, Ürümqi, China

Bipolar disorder (BD) is associated with a high risk of suicide. We used proton magnetic resonance spectroscopy (^{1}H-MRS) to detect biochemical metabolite ratios in the bilateral prefrontal white matter (PWM) and hippocampus in 32 BD patients with suicidal ideation (SI) and 18 BD patients without SI, identified potential brain biochemical differences and used abnormal metabolite ratios to predict the severity of suicide risk based on the support vector machine (SVM) algorithm. Furthermore, we analyzed the correlations between biochemical metabolites and clinical variables in BD patients with SI. There were three main findings: (1) the highest classification accuracy of 88% and an area under the curve of 0.9 were achieved in distinguishing BD patients with and without SI, with N-acetyl aspartate (NAA)/creatine (Cr), myo-inositol (mI)/Cr values in the bilateral PWM, NAA/Cr and choline (Cho)/Cr values in the left hippocampus, and Cho/Cr values in the right hippocampus being the features contributing the most; (2) the above seven features could be used to predict Self-rating Idea of Suicide Scale scores ($r = 0.4261$, $p = 0.0302$); and (3) the level of neuronal function in the left hippocampus may be related to the duration of illness, the level of membrane phospholipid catabolism in the left hippocampus may be related to the severity of depression, and the level of inositol metabolism in the left PWM may be related to the age of onset in BD patients with SI. Our results showed that the combination of multiple brain biochemical metabolites

could better predict the risk and severity of suicide in patients with BD and that there was a significant correlation between biochemical metabolic values and clinical variables in BD patients with SI.

KEYWORDS

bipolar disorder, suicidal ideation, proton magnetic resonance spectroscopy, machine learning, multivariate pattern analysis, support vector machine

Introduction

Bipolar disorder (BD) is a lifelong severe mental disorder characterized by alternating high and low emotions. Approximately 2%−3% of individuals in the world are troubled by bipolar and related disorders. The World Health Organization (WHO) reported that BD is expected to rise to sixth place on the list of global burdens of disease by 2030. Some studies have shown that BD is the disease associated with the highest risk of suicide among all major mental disorders (Merikangas et al., 2011), and the associated suicide rate is 20 to 30 times higher than that in the common population (Plans et al., 2019; Carvalho et al., 2020). In terms of suicide risk, suicidal ideation (SI) is considered to be one of the important predictors of future suicidal behavior (Nock et al., 2008). Therefore, paying close attention to SI can reduce the risk of early suicide in patients with BD. However, the current assessment of suicide risk is still mainly based on many sociodemographic characteristics and clinical risk factors, which usually have poor predictive accuracy. In addition, due to the influence of patients' subjectivity and shame, almost 80% of patients who attempted suicide do not explain their suicidal thoughts to doctors (Gosnell et al., 2019). Consequently, strengthening the search for specific neurobiological markers will help to predict future suicide risk in BD patients with SI.

At present, the neurobiological mechanisms underlying SI in patients with BD are unknown. Magnetic resonance imaging (MRI) technology has been widely used to understand the pathophysiological mechanisms underlying mental diseases, and such research has made corresponding progress in revealing the neurobiological changes in brain regions of suicidal patients with BD. For example, the graph theory analysis of global brain functional connectivity in resting-state functional MRI showed that the distribution of intrinsic connectivity in the bilateral ventromedial prefrontal cortex of patients with BD who attempted suicide was significantly lower than that in patients without attempted suicide and was related to the severity of SI (Sankar et al., 2022). Relevant studies have found that in patients with a current or prior diagnosis of depression or BD, the intensity of SI was associated with weaker connections of the limbic network with the hippocampus, default mode network, dorsal attention network, and executive control network (Chin Fatt et al., 2021). In studies evaluating low-frequency fluctuations (ALFF) and gray matter volume in the prefrontal cortex, it was shown that the ALFF values in the medial prefrontal cortex, ventral prefrontal cortex, and dorsolateral prefrontal cortex in BD patients with suicide attempts were significantly higher than those in patients without suicide attempts (Zhao et al., 2021). In gray matter and white matter-related studies, bilateral hippocampal gray matter volume and right ventral frontal white matter fractional anisotropy were found to decrease in BD patients with suicide attempts (Fan et al., 2019). From the above research, we can speculate that the structural and functional abnormalities in the prefrontal lobe and hippocampus may be related to SI in patients with BD.

Proton magnetic resonance spectroscopy (^{1}H-MRS) is a non-invasive and non-radioactive technique used to study the levels of biochemical metabolites in the brain, including N-acetyl aspartate (NAA), choline (Cho), myo-inositol (mI), and creatine (Cr), which can provide relevant information about neuronal integrity and neurotransmitter levels. Some studies have shown that the NAA/Cr values in the left prefrontal white matter (PWM) in patients with BD II (Zhong et al., 2014) and bilateral PWM in depressed patients with BD, compared with healthy controls, were decreased (Lai et al., 2019), furthermore, NAA/Cr+phosphocreatine and NAA/Cho values in the bilateral hippocampus were significantly decreased in patients with first-episode BD I (Atmaca et al., 2006). Glutamic acid and glutamine complex (Glx) levels and NAA/Glx in the anterior cingulate cortex could distinguish depression patients with and without SI (Lewis et al., 2020). However, there are few studies on ^{1}H-MRS in BD patients with suicide, and the results have either not been significant or not been very consistent (Rocha et al., 2015; Zhong et al., 2018).

In recent years, multivariate pattern analysis (MVPA) has become an effective analysis method that can often detect differences in neuroimaging data that cannot be detected by traditional univariate statistical methods by combining information from many features (Nielsen et al., 2020). Support vector machine (SVM) has been one of the most widely used machine learning algorithms to identify neurobiological markers of various neuropsychiatric disorders, which has high sensitivity and specificity in distinguishing patients with

BD from those with other neuropsychiatric disorders or healthy subjects, as well as predicting clinical outcomes of neuropsychiatric disorders (Orrù et al., 2012; Liu et al., 2015, 2017a; Librenza-Garcia et al., 2017; Orban et al., 2018; Schwarz et al., 2019; Yang et al., 2019; Ji et al., 2020; Li et al., 2020). However, to date, the research on ^{1}H-MRS of BD using the SVM algorithm is very limited, and only one study described the potential of ^{1}H-MRS in predicting the first emotional episode of young BD offspring by using SVM (Zhang et al., 2021). Research on ^{1}H-MRS in BD patients with SI is nonexistent.

Therefore, in this study, we used ^{1}H-MRS to detect the biochemical metabolite ratios in the bilateral PWM and hippocampus in BD patients with and without SI, identified potential brain biochemical differences in BD patients with SI, and used these abnormal metabolite ratios to predict the severity of suicide risk based on the SVM algorithm. In addition, we examined the relationships between biochemical metabolite ratios and clinical variables. In short, the purpose of this study was to explore changes in brain biochemical metabolites in BD patients with SI, to identify high-accuracy neuroimaging predictors that may be used to evaluate the risk of suicide, and then to provide early identification and assessment of the risk or tendency of suicidal behavior in patients with BD in future.

Methods

Participants and clinical assessments

We recruited 50 patients with BD from the Department of Clinical Psychology of the People's Hospital of Xinjiang Uygur Autonomous Region in China. There was a restriction on the age of all participants, who ranged from 18 to 55 years. The diagnosis of BD was confirmed by two experienced psychiatrists according to the structured clinical interview criteria of the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR). The evaluation of the clinical status was conducted utilizing the 17-item Hamilton Depression Scale (HAMD) (Hamilton, 1960) and the Young Mania Rating Scale (YMRS) (Young et al., 1978). All patients had a YMRS total score $<$ 7 and a 17-item HAMD total score $>$17. Moreover, an inclusion criterion for all participants was that they were right-handed. The exclusion criteria included the following: (1) other mental disorders meeting DSM-IV-TR criteria; (2) a history of organic brain disorder and serious physical illness; (3) a history of epilepsy, severe brain injury, or coma caused by other reasons for more than 5 min; (4) alcohol or drug abuse; (5) pregnancy or lactation; and (6) contraindications for magnetic resonance scanning, such as metal implants or claustrophobia.

We used the Self-rating Idea of Suicide Scale (SIOSS) to assess whether the patients had SI. There were 26 items on the SIOSS that included four factors: despair factor, optimism factor, sleep factor, and concealment factor. All items were scored with a "yes" or "no" answer. The assessment criteria for SI were the sum of the total scores of the despair factor, optimism factor, and sleep factor $\geq$12 and the concealment score $<$ 4. The higher the score was, the stronger the SI. The scale is simple and easy for patients to understand, and it has good validity and reliability (Xia et al., 2002). According to SIOSS scores, there were 32 patients with SI and 18 patients without SI. In addition, our study evaluates the severity of anxiety symptoms in all patients with BD by utilizing the 14-item Hamilton Anxiety Scale (HAMA).

Image data acquisition

The MRI scanning was conducted on all subjects using an Ingenia 3.0T MRI scanner (Ingenia, Philips Healthcare, Netherlands) with a 15-channel phased-array head coil. The following precautions were taken before scanning: (1) coffee, tobacco, and alcohol were banned in 24 h before the MRI scanning, and the body temperature need to be kept at a normal level to rule out feverish patients; (2) the doctor told the subjects not to move during the MRI scanning; and (3) the subjects were required to rest quietly for half an hour before the examination. In addition, earplugs and headphones were used to reduce noise, and head movement was reduced by using foam pads.

One 3DT1FSPGR sequence that covered the whole brain was used for the anatomical localization of MRS. T2-weighted images (T2WI) [repetition time (TR) = 3,000 ms, echo time (TE) = 95 ms], T1-weighted images (T1W-IR) [TR = 2,000 ms, TE = 20 ms, inversion time (TI) = 800 ms], and T2 fluid-attenuated inversion recovery (T2FLAIR) (TR = 8,000 ms, TE = 270 ms, TI = 2,000 ms) were routinely conducted. The voxel size = 0.8 $\times$ 0.8 $\times$ 6 mm^3, there were 18 slices, with a field of view (fov)=24 cm. The scanning planes of three sequences were duplicated by each other. Before ^{1}H-MRS sequence localization, it was necessary to judge whether there were organic lesions, and the anatomical positions of the bilateral PWM and hippocampus were determined by experienced radiologists. In three dimensions, none of the volume of interest (VOI) involved sulci or cerebrospinal fluid. The size of a single voxel was 20 $\times$ 20 $\times$ 20 mm^3. The scanning parameters were as follows: TE = 35 ms; TR = 2,000 ms; and average (superposition) number of signals (NSA) = 96. To prevent the disturbance of spectral line quality by some disadvantageous factors surrounding the spectrum voxel, the saturation band could be manually placed near the VOI, which was essential to ensure that the air, bone, fat, or blood vessels surrounding the VOI were adequately suppressed. During scanning, the chemical shift selective saturation method was used to optimize the water suppression to ensure that the water suppression rate was $>$99% and the full width at half maximum of the water peak was

<10 Hz. The acquisition time of the ^{1}H-MRS sequence was 10 min (Chen et al., 2021).

Data processing

The spectral view software of the Philips 3.0T workstation (ISP 7.0, Philips Healthcare, the Netherlands) was used for ^{1}H-MRS data postprocessing. The specific data processing steps were as follows: residual water peak removal, signal attenuation correction, spectral line interpolation, Fourier transform, spectral line phase correction, baseline level adjustment, the selection of peak frequency position, line width setting, and peak Gaussian fitting. The NAA/Cr, Cho/Cr, and mI/Cr ratios in the bilateral PWM and hippocampus were used to analyze the changes in brain biochemical metabolism, and an experienced radiologist evaluated the quality of the spectrum and analyzed the data.

Univariate statistical analysis

The univariate data were analyzed by SPSS 24.0 software, and the significance level of the two tails was set at $p < 0.05$. When the data were continuous variables that conformed to a normal distribution, the *t*-test was conducted, and the data are expressed as the mean ± standard deviation. When the continuous variable did not conform to a normal distribution, the non-parametric test was selected, and the data are expressed as the median (upper quartile, lower quartile). The data of discontinuous variables were analyzed by the chi-square test. Then, the results with statistically significant differences between the two groups were further corrected by the Bonferroni correction method for multiple comparisons. In determining whether the concentration of brain biochemical metabolites in BD patients with SI was related to clinical variables, Pearson correlation analysis was used for the two groups of variables conforming to the normal distribution, and Spearman correlation analysis was used for those not conforming to the normal distribution. In addition, we plotted the receiver operating characteristic (ROC) curves and calculated corresponding areas under the curve (AUC) for the abnormal metabolic values of BD patients with and without SI.

MVPA

This analysis was conducted using SVM (i.e., support vector classification, SVC) from the MVPANI toolbox (Peng et al., 2020). We first used the leave-one-out cross-validation (LOOCV) method, leaving only one subject for the test set each time, and the remaining 49 subjects for the training set, which resulted in the model being trained 50 times and tested 50 times. Although the calculation method using LOOCV is complicated, it has a high utilization rate of samples, which is suitable for small sample research. Each feature was standardized before cross-validation: the values in each row (each row represents a sample) were normalized by transforming all values in each sample to z scores with a mean of 0 and a standard deviation (SD) of 1 using the following equation:

$$z_{ij} = \frac{x_{ij} - mean\,(x_{i.})}{SD\,(x_{i.})}\ for\ the\ i^{th}\ row$$

We adopted the feature selection method of F-score in this study. The specific process was shown as follows: in the process of each LOOCV, all features were arranged in order from high to low according to the size of F-values of an F-test in BD patients with and without SI in the training dataset. Then, we selected the number of the feature with the highest F-values of 10% (the number of the feature was one) to train a new classifier in the training dataset and the performance of the classifier was tested using the reserved test dataset, so the classification accuracy of this LOOCV was generated. According to the LOOCV process, 50 models and 50 feature sets were generated, and the features contained in each feature set in the 50 feature sets were not all the same. In addition, the features in each feature set have corresponding weight values. Finally, we could calculate the average classification accuracy, the average weight values of the features, and the frequency of each feature of 50 models. Next, increase the number of feature by 10% each time, and repeat the above feature selection process until all features were selected, and the number of features was 2, 3, 4, 6, 7, 8, 9, 10, 12. Therefore, 10 feature sets with different feature numbers and 10 average classification accuracies were finally obtained. The statistical significance of classification accuracy was determined by the permutation test, and the significance threshold was $p < 0.05$. In this study, we conducted 10,000 random permutation tests, obtained 10,000 random classification accuracies, and then used these 10,000 random classification accuracies to construct the null distribution. The *p*-value was the percentage greater than or equal to the actual classification accuracy ($p = 0.0001$, i.e., 1/10,000). In the training set of feature selection with the F-score, because we performed 10 independent MVPA analyses, the above *p*-values calculated from the permutation tests needed to be further corrected by the Bonferroni correction method for multiple comparisons. In addition to obtaining classification accuracy, we calculated the receiver operating characteristic (ROC) curves and the corresponding areas under the curve (AUCs) for each classification.

Furthermore, SVM (i.e., support vector regression, SVR) was applied to predict SIOSS scores in patients with BD. We used the biochemical metabolic values corresponding to the highest classification accuracy from the above feature selection process as the feature and SIOSS scores in patients with BD as the regression target of the regression analysis. We also made use of the LOOCV method to divide all subjects into the training set

and test set, and each feature was standardized in the process of cross-validation (the method was the same as before). Finally, we calculated the Pearson correlation coefficient between predicted SIOSS scores and observed SIOSS scores. Similarly, we used the permutation test to determine the statistical significance of the Pearson correlation coefficient using the followed specific methods. The predicted SIOSS scores of patients with BD were randomly disrupted. With the disrupted SIOSS scores, we recalculated the Pearson correlation coefficients between the predicted SIOSS scores and the observed SIOSS scores. We repeated the above process 10,000 times and compared the actual Pearson correlation coefficient with the null distribution based on the null distribution composed of 10,000 Pearson correlation coefficients. The *p*-value was the proportion of the Pearson correlation coefficient obtained by random permutation tests greater than or equal to the actual value in 10,000 random permutation tests. The level of statistical significance was $p < 0.05$.

Results

Demographics

Thirty-two BD patients with SI (11 men, 21 women; 34.69 ± 10.51 years old; age range: 18–54 years) and 18 BD patients

TABLE 1 Demographic and clinical data of BD patients with and without SI.

	BD with SI ($n = 32$)	BD without SI ($n = 18$)	z/t/χ2	*p*
Gender (male/female)	11/21	9/9	1.172	0.279
Positive family history (yes/no)	8/24	1/17	1.781	0.182
Education (years)	16 (15, 16)	16 (15, 16)	−0.494	0.621[b]
Duration of illness (month)	48 (24, 81)	60 (20.5, 63)	−0.457	0.647[b]
Age (years)	34.69 ± 10.51	32.89 ± 10.97	0.572	0.570[a]
Age of onset (years)	29.88 ± 10.11	28.61 ± 9.64	0.432	0.668[a]
17-item HAMD score	18.5 (18, 21)	19 (18, 19.25)	−0.492	0.623[b]
14-item HAMA score	13.97 ± 4.04	14.67 ± 3.80	−0.598	0.553[a]
SIOSS score	18 (15, 20)	10.5 (6, 11)	−5.856	0.000[b,*]

BD, bipolar disorder; SI, suicidal ideation; HAMD, Hamilton Depression Scale; HAMA, Hamilton Anxiety Scale; SIOSS, Self-rating Idea of Suicide Scale; [a]Independent samples t-test; [b]Mann–Whitney U test; [*]$p < 0.05$ significant.

TABLE 2 Comparisons of proton magnetic resonance spectroscopy in the bilateral PWM and hippocampus of BD patients with and without SI.

	BD with SI ($n = 32$)	BD without SI ($n = 18$)	z/t	*p*	p^c
Left PWM					
NAA/Cr	2.02 (1.78, 2.07)	1.78 (1.37, 2.06)	−1.598	0.110[b]	
Cho/Cr	0.95 ± 0.33	0.98 ± 0.39	−0.247	0.806[a]	
mI/Cr	0.52 (0.41, 0.61)	0.53 (0.45, 1.11)	−1.275	0.202[b]	
Right PWM					
NAA/Cr	1.79 (1.54, 1.91)	1.63 (1.27, 1.71)	−2.024	0.043[b,*]	0.516
Cho/Cr	1.11 (0.77, 1.17)	0.98 (0.69, 1.15)	−0.689	0.491[b]	
mI/Cr	0.97 (0.55, 2.06)	0.5 (0.35, 0.68)	−2.615	0.009[b,*]	0.108
Left hippocampus					
NAA/Cr	1.80 (1.52, 1.88)	2.09 (1.67, 2.31)	−1.827	0.068[b]	
Cho/Cr	0.90 (0.74, 0.90)	1.33 (1.04, 1.44)	−3.369	0.001[b,*]	0.012*
mI/Cr	0.75 (0.61, 0.86)	0.78 (0.47, 0.86)	−0.214	0.830[b]	
Right hippocampus					
NAA/Cr	1.86 (1.55, 1.86)	1.78 (1.40, 1.96)	−0.898	0.369[b]	
Cho/Cr	1.06 (0.78, 1.06)	1.12 (0.78, 1.18)	−1.204	0.229[b]	
mI/Cr	0.71 (0.44, 0.71)	0.65 (0.57, 0.70)	−0.857	0.391[b]	

BD, bipolar disorder; SI, suicidal ideation; PWM, prefrontal whiter matter; NAA, N-acetylaspartate.
Cho, choline; mI, myo-inositol; Cr, creatine; [a]Independent samples t-test; [b]Mann–Whitney U test.
[c]Bonferroni correction; [*]$p < 0.05$ significant.

without SI (nine men, nine women; 32.89 ± 10.97 years old; the age range: 18–54 years) underwent demographics and clinical evaluation in this study. Table 1 shows the demographic and clinical variables of all participants in the study. There were no significant differences in age, gender, educational level, family history, duration of illness, and age of onset between BD patients with and without SI (all $p > 0.05$). In terms of clinical symptoms, the SIOSS score in BD patients with SI was higher than that in BD patients without SI ($p < 0.05$). No significant differences were found in the 17-item HAMD and 14-item HAMA scores between the groups (all $p > 0.05$).

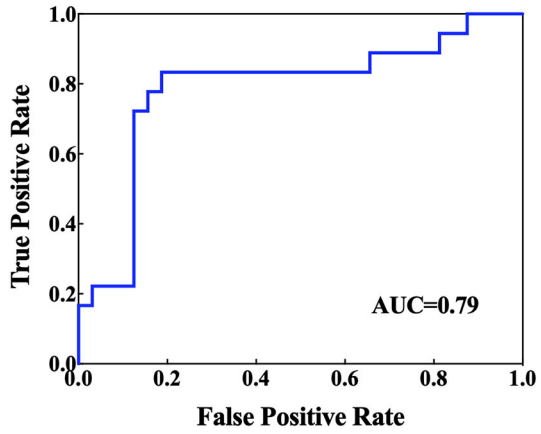


FIGURE 1
Receiver operating characteristic (ROC) curves and corresponding areas under the curve (AUCs) for the brain biochemical metabolites that showed significant differences in the univariate statistical analysis.

Comparisons of ^{1}H-MRS in the bilateral PWM and hippocampus of BD patients with and without SI

Table 2 presents the comparative results of the NAA/Cr, Cho/Cr, and mI/Cr ratios in the bilateral PWM and hippocampus of BD patients with and without SI. After using the Bonferroni correction method, the Cho/Cr ratios in the left hippocampus were significantly lower in BD patients with SI than in BD patients without SI ($p < 0.05$). However, there were no significant differences in the Cho/Cr ratios in the right hippocampus, NAA/Cr and mI/Cr ratios in the bilateral hippocampus, and NAA/Cr, Cho/Cr, and mI/Cr ratios in the bilateral PWM between the two groups (all $p > 0.05$). Furthermore, the AUC of the Cho/Cr values in the left hippocampus was 0.79 (Figure 1).

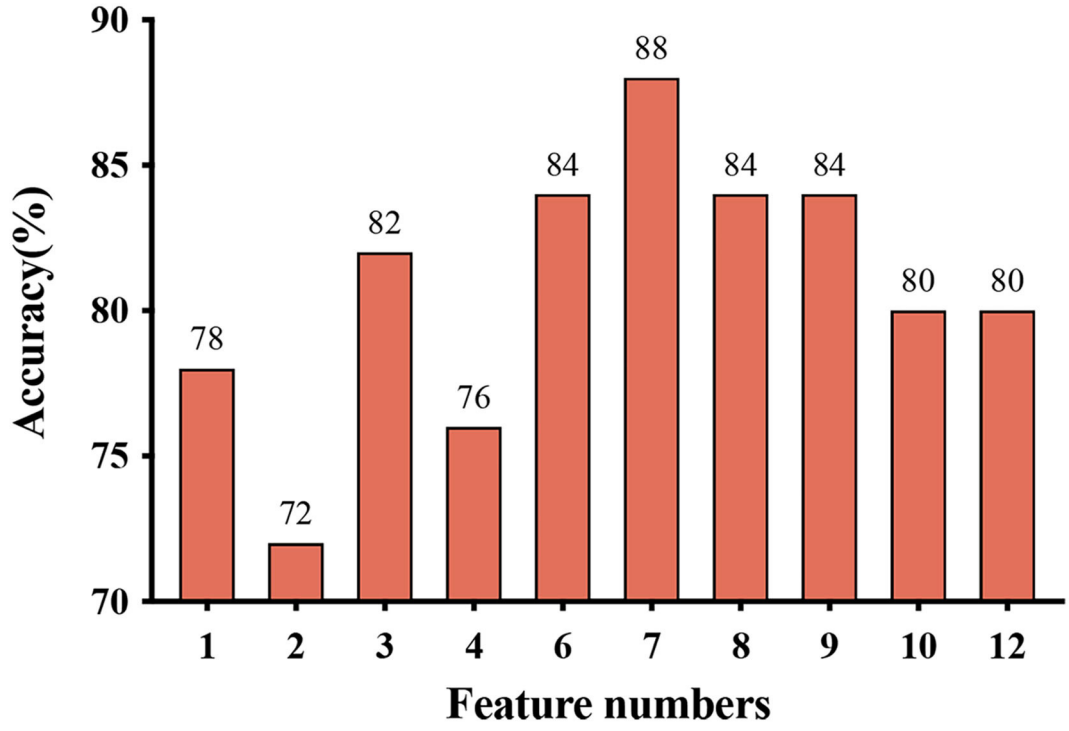


FIGURE 2
Feature numbers and classification accuracy for each classifier when using feature selection.

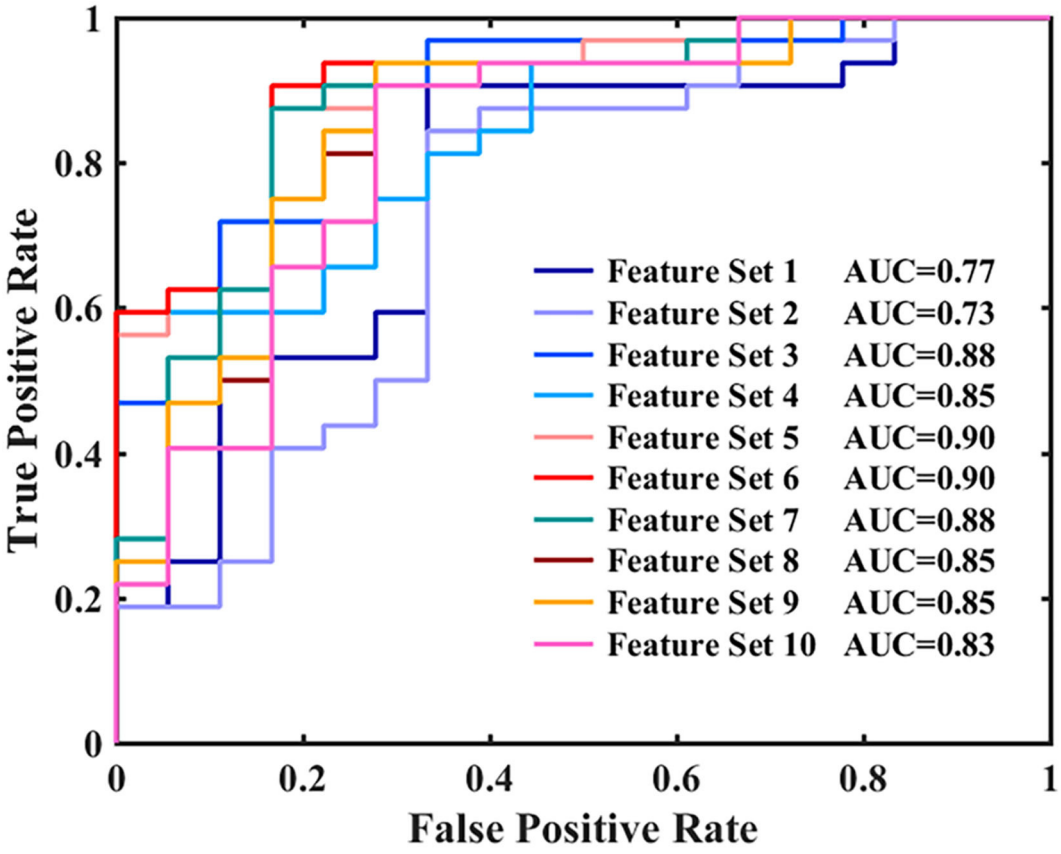


FIGURE 3
Receiver operating characteristic (ROC) curves and areas under the curves (AUCs) for each classifier when using feature selection.

MVPA of biochemical metabolite ratios in BD patients with and without SI

Figures 2, 3, and Table 3 show the number of features, classification accuracy, sensitivity, specificity, AUC, and *p*-value obtained by the permutation test when the F-score was used for feature selection in each classifier. When the classification accuracy was 78%, the AUC was 0.77 and the number of the feature was one, the corresponding features were Cho/Cr ratios in the left hippocampus. When the classification accuracy was 72%, the AUC was 0.73 and the number of the feature was two, the corresponding features were mI/Cr ratios in the right PWM and Cho/Cr ratios in the left hippocampus. When the classification accuracy was 82%, the AUC was 0.88 and the number of the feature was three, the corresponding features were mI/Cr ratios in the bilateral PWM and Cho/Cr ratios in the left hippocampus. When the classification accuracy was 76%, the AUC was 0.85 and the number of the feature was four, the corresponding features were NAA/Cr ratios in the right PWM, mI/Cr ratios in the bilateral PWM, and Cho/Cr ratios in the left hippocampus. When the classification accuracy was 84%, the AUC was 0.90 and the number of the feature was six, the corresponding features were NAA/Cr ratios in the bilateral PWM, mI/Cr ratios in the bilateral PWM, and NAA/Cr ratios and Cho/Cr ratios in the left hippocampus. When the classification accuracy was 88%, the AUC was 0.90 and the number of the feature was seven, the corresponding features were NAA/Cr and mI/Cr ratios in the bilateral PWM, NAA/Cr and Cho/Cr ratios in the left hippocampus, and Cho/Cr ratios in the right hippocampus. When the classification accuracy was 84%, the AUC was 0.88 and the number of the feature was eight, the corresponding features were NAA/Cr and mI/Cr ratios in the bilateral PWM, Cho/Cr ratios in the left PWM, NAA/Cr ratios in the left hippocampus, and Cho/Cr ratios in the bilateral hippocampus. When the classification accuracy was 84%, the AUC was 0.85 and the number of the feature was nine, the corresponding features were NAA/Cr and mI/Cr ratios in the bilateral PWM, NAA/Cr and Cho/Cr ratios in the bilateral hippocampus, and Cho/Cr ratios in the left PWM. When the classification accuracy was 80%, the AUC was 0.85 and the number of the feature was 10, the corresponding features were NAA/Cr and mI/Cr ratios in the bilateral PWM, Cho/Cr ratios

TABLE 3 Results for each classifier corresponding to feature selection.

Feature set	Feature numbers	Sensitivity	Specificity	Accuracy (%)	AUC	p^a value
Feature set 1	1	0.91	0.56	78	0.77	<0.001*
Feature set 2	2	0.88	0.44	72	0.73	0.277
Feature set 3	3	0.88	0.72	82	0.88	0.008*
Feature set 4	4	0.88	0.56	76	0.85	0.117
Feature set 5	6	0.94	0.67	84	0.90	0.002*
Feature set 6	7	0.94	0.78	88	0.90	<0.001*
Feature set 7	8	0.91	0.72	84	0.88	0.002*
Feature set 8	9	0.91	0.72	84	0.85	0.002*
Feature set 9	10	0.84	0.72	80	0.85	0.021*
Feature set 10	12	0.84	0.72	80	0.83	0.016*

AUC, areas under the curve; [a]Bonferroni correction; *$p < 0.05$ significant.

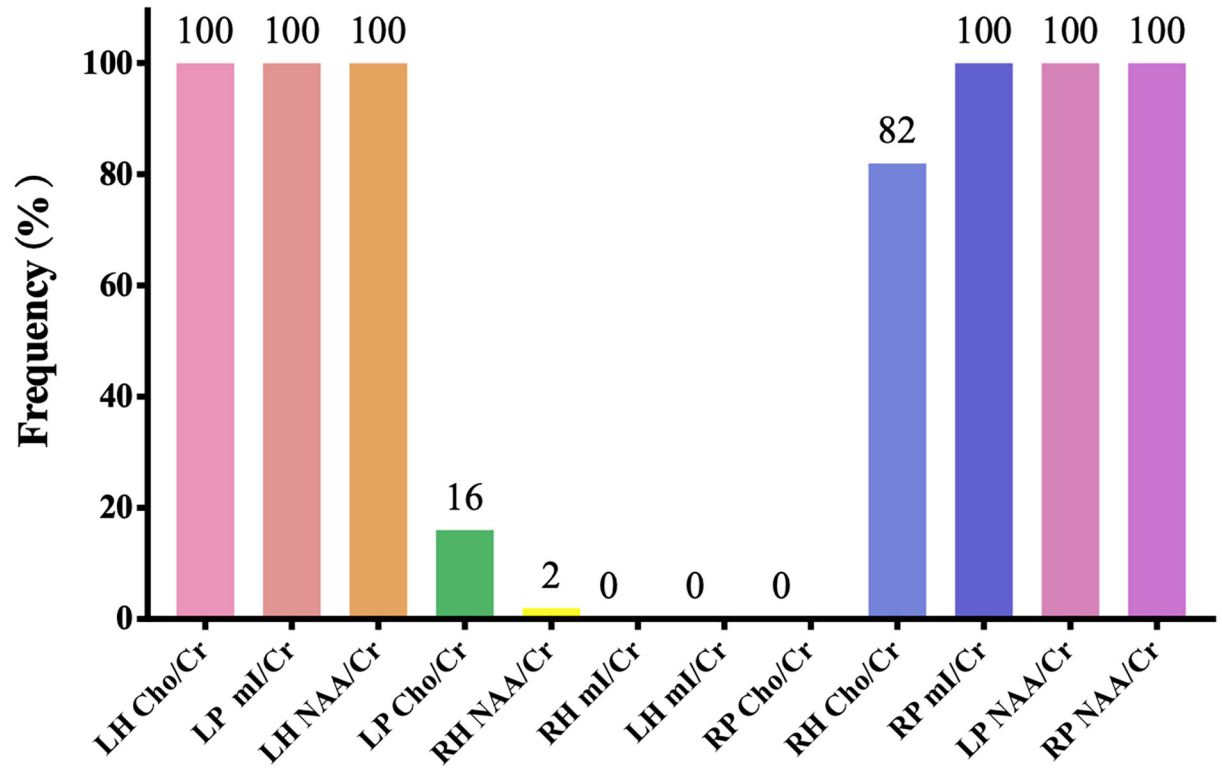


FIGURE 4
The frequency of each feature corresponding to the highest classification accuracy (LH, left hippocampus; RH, right hippocampus; LP, left prefrontal whiter matter; RP, right prefrontal whiter matter; NAA, N-acetylaspartate; Cho, choline; mI, myo-inositol; Cr, creatine).

in the left PWM, NAA/Cr and Cho/Cr ratios in the bilateral hippocampus, and mI/Cr ratios in the left hippocampus. When the classification accuracy was 80%, the AUC was 0.83 and the number of the feature was 12, the corresponding features were NAA/Cr, Cho/Cr, and mI/Cr ratios in the bilateral PWM and hippocampus. According to the above results, we found that the sixth feature set (the number of the feature was seven) had the highest classification accuracy of 88% and the AUC was 0.9. Figures 4, 5 show the frequency and average weight values of each feature corresponding to the highest classification accuracy in the above feature selection process. We concluded that the NAA/Cr ratios in the bilateral PWM, the mI/Cr ratios in the right PWM, and the Cho/Cr ratios in the right hippocampus were higher and the NAA/Cr and Cho/Cr ratios in the left

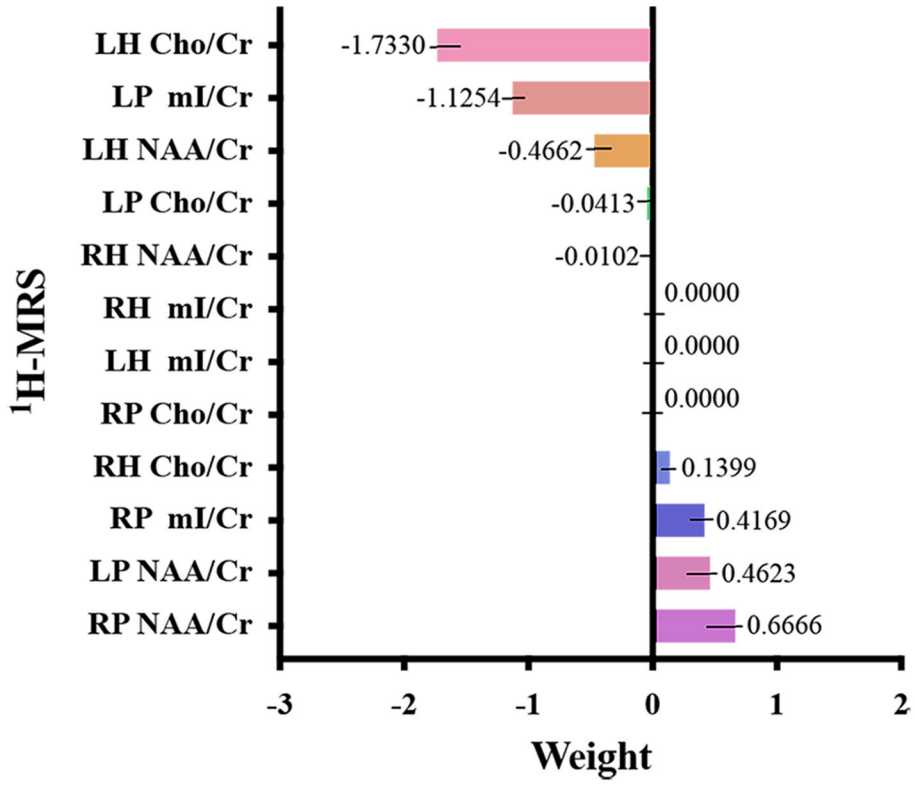


FIGURE 5
Weight value of each feature corresponding to the highest classification accuracy (LH, left hippocampus; RH, right hippocampus; LP, left prefrontal whiter matter; RP, right prefrontal whiter matter; NAA, N-acetylaspartate; Cho, choline; mI, myo-inositol; Cr, creatine).

hippocampus and the mI/Cr ratios in the left PWM were lower in BD patients with SI than in BD patients without SI. Moreover, we found that the above seven features could be used to predict the severity of suicide risk ($r = 0.4261$, $p = 0.0302$) (Figure 6).

Correlations between biochemical metabolite ratios and clinical variables in BD patients with SI

Table 4 summarizes the correlations between biochemical metabolite ratios and age of onset, duration of illness, 14-item HAMA score, and 17-item HAMD score in BD patients with SI. Notably, the NAA/Cr ratios in BD patients with SI were positively correlated with the duration of illness ($r = 0.354$, $p < 0.05$), and the Cho/Cr ratios in BD patients with SI were positively correlated with 17-item HAMD scores in the left hippocampus ($r = 0.372$, $p < 0.05$). In addition, the mI/Cr ratios in the left PWM showed a positive correlation with age of onset ($r = 0.372$, $p < 0.05$). However, there were no correlations between the other biochemical metabolite ratios and clinical variables in BD patients with SI (all $p > 0.05$).

Discussion

In the univariate statistical analysis, we found that the Cho/Cr values in the left hippocampus of BD patients with SI were decreased, suggesting that BD patients with SI had a decrease in membrane phospholipid metabolism in the hippocampus. In addition, the AUC of Cho/Cr values in the left hippocampus was relatively high (0.79), indicating that it has good diagnostic accuracy for BD patients with SI. To the best of our knowledge, there have been no studies on ^{1}H-MRS in the hippocampus of BD patients with SI. Previous studies have suggested that functional disorders of the frontal limbic network play an essential role in the mechanisms associated with suicide among patients with BD (Giakoumatos et al., 2013). The frontal limbic network is composed of the frontal lobes, cingulate gyrus, and subcortical brain structures, such as the hippocampus, amygdala, and other brain regions (Phillips et al., 2003, 2008). Among them, the hippocampus has excellent effects on encoding and recalling the emotional meaning of events, which may affect the emotional response and regulation process (Richard-Devantoy et al., 2015). Studies have indicated that the gray matter volume in the hippocampus and orbitofrontal cortex of BD patients with suicide attempts, compared with BD

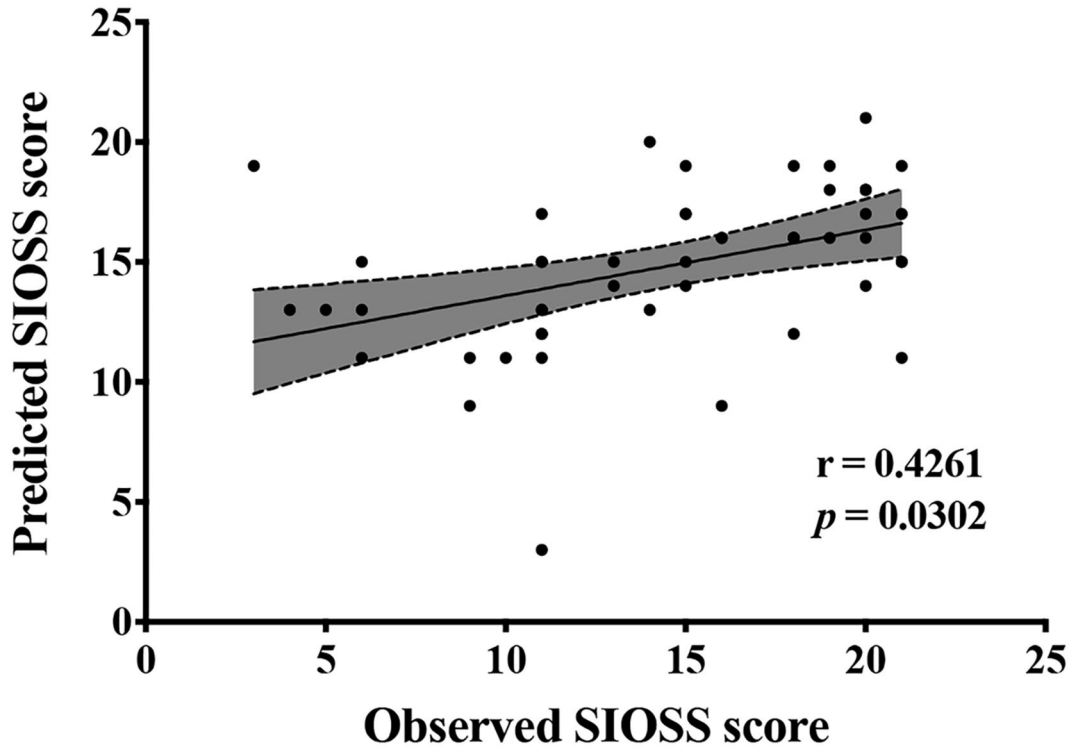


FIGURE 6
Biochemical metabolic values corresponding to the highest classification accuracy in the feature selection predict Self-rating Idea of Suicide Scale (SIOSS) scores.

patients without suicide attempts, was significantly decreased (Johnston et al., 2017), and the amplitude of ALFF values in the right hippocampus was increased in severely depressed patients with a history of suicide attempts (Lan et al., 2019). Thus, the hippocampus also plays a major role in the mechanism of suicide in patients with BD. Cho is one of the components of membrane phospholipid metabolism that is closely related to membrane phospholipid decomposition and synthesis, reflects the metabolic level of membrane phospholipids, and is a marker of membrane integrity (Strakowski et al., 2005). In addition, Cho is more abundant in astrocytes and oligodendrocytes, which reflects changes in glial metabolism and function (Atmaca and Yildirim, 2012). A recent study found that cerebral vimentin-immunoreactive astrocytes showed a widespread reduction in depressive disorder patients who died of suicide, suggesting that dysfunction in astrocytes was associated with suicide and depression (O'Leary et al., 2021), and the results of our study provide new evidence for this.

In the MVPA, our results found that the highest classification accuracy obtained by putting all features into the classifier was 88% and the AUC was 0.90, which was higher than the diagnostic prediction ability of the Cho/Cr values in the left hippocampus in univariate statistical analysis (AUC = 0.79). This suggested that the use of multiple biochemical metabolic values more effectively identified SI in patients with BD. Based on the number of features, the frequency of feature occurrence, and the weight values obtained by feature selection, our results suggested that in addition to the metabolic values identified in the univariate statistical analysis, we further found that the NAA/Cr values in the bilateral PWM, the mI/Cr values in the right PWM, and the Cho/Cr values in the right hippocampus were higher in BD patients with SI than in those without SI and the NAA/Cr values in the left hippocampus and the mI/Cr values in the left PWM were lower in those with SI than in those without SI by using the MVPA. Many neuroimaging studies have reported that the dysfunction of some neural circuits is related to SI and behavior (Ding et al., 2017; Johnston et al., 2017; Brown et al., 2020). The abnormal function of the frontal lobes could lead to disorders of emotional executive control and emotional pain processing (van Heeringen et al., 2010). The results of some studies have shown that dysfunction in the PWM regions was most common in BD patients with a history of suicide attempts (Hozer and Houenou, 2016). BD patients with a history of suicide attempts had decreased white matter in the caudal frontal lobe and the left orbitofrontal cortex (Mahon et al., 2012; Johnston et al., 2017). Therefore, abnormal white matter in the prefrontal lobes may be related to the pathogenesis of suicide in patients with BD. NAA reflects the vitality of neurons and can

TABLE 4 Correlations between biochemical metabolite ratios and clinical variables in BD patients with SI.

	Age of onset		Duration of illness		14-item HAMA		17-item HAMD	
	r	*p*	*r*	*p*	*r*	*p*	*r*	*p*
Left PWM								
NAA/Cr	−0.050	0.787	0.084	0.646	−0.206	0.258	−0.335	0.061
Cho/Cr	−0.173	0.343	−0.235	0.195	0.171	0.349	0.042	0.819
mI/Cr	0.372	0.036*	0.111	0.544	−0.180	0.323	0.037	0.841
Right PWM								
NAA/Cr	0.024	0.895	0.247	0.173	0.002	0.991	0.094	0.610
Cho/Cr	−0.329	0.066	−0.216	0.236	0.142	0.438	0.190	0.299
mI/Cr	−0.188	0.303	0.066	0.719	0.248	0.171	0.179	0.328
Left Hippocampus								
NAA/Cr	−0.030	0.871	0.354	0.047*	0.128	0.486	0.116	0.528
Cho/Cr	0.008	0.964	−0.318	0.076	0.072	0.695	0.372	0.036*
mI/Cr	0.060	0.744	−0.260	0.151	0.092	0.616	0.195	0.286
Right hippocampus								
NAA/Cr	0.108	0.555	0.050	0.787	0.126	0.492	−0.097	0.597
Cho/Cr	0.309	0.086	0.191	0.296	−0.053	0.774	−0.020	0.916
mI/Cr	0.048	0.792	−0.085	0.642	−0.146	0.424	−0.215	0.238

BD, bipolar disorder; SI, suicidal ideation; PWM, prefrontal whiter matter; NAA, N-acetyl aspartate; Cho, choline; mI, myo-inositol; Cr, creatine; HAMA, Hamilton Anxiety Scale; HAMD, Hamilton Depression Scale; $^{*}p < 0.05$ significant.

10.3389/fnins.2022.944585

maintain the osmotic pressure of neurons cell (Baslow, 2003). NAA content is closely related to the function of neurons and is a marker of neuron density and survival. Simultaneously, NAA is related to the function of mitochondria and myelin formation and is an energy marker of neuronal mitochondria. The results of our study suggested that BD patients with SI had enhanced prefrontal neuronal function or mitochondrial dysfunction. At present, there are many controversies about the study of NAA in the prefrontal lobes of patients with BD. Some studies suggested that the NAA values in the prefrontal lobes were decreased, while others thought that the NAA values were increased or at a normal level (Zhong et al., 2014, 2018; Moon et al., 2015; Liu et al., 2017b; Patel et al., 2018). The reasons for the inconsistent results may be related to magnetic resonance technical parameters, disease status, the influence of drugs, and other factors. For example, studies have reported that NAA values were increased in the prefrontal cortex after lithium treatment (Hajek et al., 2012). A follow-up study of BD patients treated with lithium for 4 weeks also confirmed that lithium could induce an increase in NAA levels (Moore et al., 2000). More high-quality samples need to be evaluated for further discussion in future.

Myo-inositol exists in glial cells and is a marker of glial cells. It is involved in the circulation of inositol phosphate, the regulation of neuronal permeability, and the catabolism of phospholipids in the cell membrane. It has a nutritional and protective effect on neurons (Dager et al., 2008). In the MVPA, our study found that the mI/Cr values in the right PWM were increased and the mI/Cr values in the left prefrontal matter were decreased in BD patients with SI. According to previous studies, we learned that when the synthesis of inositol phosphate is blocked, mI content will increase, and the excitatory transmitter induced by inositol phosphate will decrease, thus resulting in depression (Wu et al., 2001). In addition, a plasma metabolomic study found that the peak of inositol was higher in depressive disorder patients with suicide attempts than in depressive disorder patients without suicide attempts (Zhou, 2013). Based on the above results, it could be concluded that the increase in inositol in depressed patients with BD might be related to suicide. However, there were also inconsistent findings. Shimon et al. (1997) measured the content of the inositol and its synthetic enzyme, inositol monophosphatase in postmortem brain samples of suicide victims, patients with BD, and normal controls. The results showed that the levels of inositol in the frontal cortex of the suicide victims and patients with BD were significantly lower than those of the normal controls. We found that the changes in inositol levels in the left and right PWM were inconsistent in BD patients with SI, which might be related to the asymmetry of frontal lobe function. Therefore, our results indicated that the abnormal level of inositol metabolism or the disorder of glial cell function in the PWM might be the pathophysiological mechanism in BD patients with SI.

The above research results showed that, first, the metabolic values with differences in univariate statistical analysis coincided with the results of the MVPA, and the frequency of the occurrence of the feature included in the highest classification accuracy of 88% was consistent with the size of features average weight values, indicating that our model has good stability and high feature sensitivity. Second, the results of this study also fully confirmed the advantage of MVPA, i.e., it could find brain biochemical metabolic differences that could not be detected by univariate statistical analyses. Through the MVPA results, we found that the neuron function in the PWM and hippocampus, inositol metabolism level in the PWM, and membrane phospholipid catabolism level in the hippocampus were functionally disrupted. These differences in brain biochemical metabolism values might also be neurobiological mechanisms associated with BD with SI. Moreover, we further found that the biochemical metabolic values corresponding to the highest classification accuracy in the feature selection could predict SIOSS scores in patients with BD. At present, many studies have focused on univariate statistical analyses to explore the correlations between imaging characteristics and clinical variables in BD patients with SI. A few studies have recently used the multiple regression model and LOOCV method to predict suicide scale scores with dynamic ALFF values in BD and depression patients with SI (Li et al., 2019; Gong et al., 2020). However, we are not aware of any study that has reported using biochemical metabolic values to predict the severity of SI based on MVPA. In the current study, we found that biochemical metabolic values could successfully predict the severity of SI in patients with BD, which further suggested that biochemical metabolic values may be a more powerful predictive neural marker for SI among patients with BD. In summary, according to our research results, we firmly believe that the combination of multiple biochemical metabolites can help to identify BD patients with SI and predict the severity of SI, thereby reducing suicide mortality, and the model had sufficiently high enough accuracy, sensitivity, and specificity.

In the correlation analysis between brain biochemical metabolites and clinical variables in BD patients with SI, we found that the NAA/Cr values in the left hippocampus were positively correlated with the duration of illness, suggesting that the longer the duration of illness was, the more active the neuronal function in the left hippocampus among BD patients with SI. Some studies have shown that suicide attempts by patients with BD are related to longer duration of illness (Lijffijt et al., 2014). The NAA/Cr values in the hippocampus of patients with BD I were positively correlated with the duration of illness (Cui et al., 2009), which was consistent with the results of our study. However, some studies have found that the hippocampal NAA/Cr values in first-episode patients with BD are not associated with the duration of illness (Atmaca et al., 2006). The inconsistencies in the results of the above studies

might be related to the selected samples and the influence of medications. In future, first-episode patients with BD need to be selected for further evaluation. Moreover, our study also found that the mI/Cr values in the left PWM were positively correlated with the age of onset. Some studies have suggested that there was a significant correlation between suicidal behavior and age of onset in patients with BD (Song et al., 2012). Singh et al. (2013) showed that mI/Cr values were related to age and increased with age. Therefore, our results suggested that the level of inositol metabolism in the left PWM of BD patients with SI may increase with the age of onset. In addition, our results showed that the Cho/Cr values in the left hippocampus were positively correlated with HAMD scores, suggesting that the more severe the degree of depression in BD patients with SI was, the stronger the level of membrane phospholipid catabolism in the left hippocampus.

To the best of our knowledge, this study is the first to explore the differences in brain biochemical metabolites in BD patients with SI and use abnormal metabolite ratios to predict the severity of suicide risk based on the SVM algorithm. However, our research also has some limitations. First, our sample size was limited, and there were no data from multiple centers to verify each other, so the generalization ability of the model needs to be further confirmed. Second, the samples were unbalanced, which may affect the robustness of the model. In future, we need to expand the sample size and maintain the sample balance for further research and discussion. In addition, this study focused only on ^{1}H-MRS, which could be combined with other imaging techniques to find high-accuracy suicide predictors for patients with BD at multiple levels. Finally, this study was a cross-sectional study, which was unable to predict whether BD patients with SI will commit suicide, and therefore, longitudinal research will be needed to further elaborate on the future suicide risk of patients with BD.

Conclusion

In conclusion, our study showed that the combination of multiple brain biochemical metabolites could better predict the risk and severity of suicide in patients with BD. Moreover, the abnormal levels of inositol metabolism in the PWM, neuron function in the PWM and hippocampus, and membrane phospholipid catabolism level in the hippocampus may be neuropathophysiological mechanisms underlying SI among patients with BD. In addition, in BD patients with SI, the level of neuronal function in the left hippocampus may be related to the duration of illness, the level of membrane phospholipid catabolism in the left hippocampus may be related to the severity of depression, and the level of inositol metabolism in the left PWM may be related to the age of onset. In future, we can expand the sample size and conduct multicenter prospective research combined with multimodal data, including genetics, electroencephalography, and different imaging methods, to further elucidate suicide prediction indicators in patients with BD to provide a basis for the early detection of suicide behavior in patients with BD.

Data availability statement

The data used to support the findings of this study are available from the corresponding author upon request.

Ethics statement

The studies involving human participants were reviewed and approved by Ethics Committee of the People's Hospital of Xinjiang Uygur Autonomous Region. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

JC, XZ, CZ, SZ, and HT conceived and designed this study. JC, XZ, YP, and YS conducted the data analysis. YQ performed the magnetic resonance imaging scanning. JC and SZ collected the data. JC, CZ, SZ, and HT wrote and critically reviewed the first draft of the manuscript. All authors contributed to the writing, critical review, and approval of the final manuscript.

Funding

This work was supported by the National Key Research and Development Program of China (2018YFC1314301), the Natural Science Foundation of Xinjiang Uygur Autonomous Region (2017D01C109), and the Tianshan innovation team project of Xinjiang Uygur Autonomous Region (2022D14011).

Acknowledgments

We would like to thank all participants for their contributions to this study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Atmaca, M., and Yildirim, H. (2012). Altered neurochemical ingredient of hippocampus in patients with bipolar depression. *Depress. Res. Treat*. 2012, 485249. doi: 10.1155/2012/485249

Atmaca, M., Yildirim, H., Ozdemir, H., Poyraz, A. K., Tezcan, E., and Ogur, E. (2006). Hippocampal ^{1}H MRS in first-episode bipolar I patients. *Prog. Neuropsychopharmacol. Biol. Psychiatry*. 30, 1235–1239. doi: 10.1016/j.pnpbp.2006.03.032

Baslow, M. H. (2003). N-acetylaspartate in the vertebrate brain: metabolism and function. *Neurochem. Res*. 28, 941–953. doi: 10.1023/A:1023250721185

Brown, V. M., Wilson, J., Hallquist, M. N., Szanto, K., and Dombrovski, A. Y. (2020). Ventromedial prefrontal value signals and functional connectivity during decision-making in suicidal behavior and impulsivity. *Neuropsychopharmacology*. 45, 1034–1041. doi: 10.1038/s41386-020-0632-0

Carvalho, A. F., Firth, J., and Vieta, E. (2020). Bipolar disorder. *N. Engl. J. Med*. 383, 58–66. doi: 10.1056/NEJMra1906193

Chen, J., Zou, S., Qu, Y., Zhang, C., Zhang, Y., Tang, X., et al. (2021). Neurometabolic alterations in bipolar disorder with anxiety symptoms: a proton magnetic resonance spectroscopy study of the prefrontal whiter matter. *Psychiatry Res*. 299,113859. doi: 10.1016/j.psychres.2021.113859

Chin Fatt, C. R., Jha, M. K., Minhajuddin, A., Mayes, T., Ballard, E. D., and Trivedi, M. H. (2021). Dysfunction of default mode network is associated with active suicidal ideation in youths and young adults with depression: findings from the T-RAD study. *J. Psychiatr. Res*. 142, 258–262. doi: 10.1016/j.jpsychires.2021.07.047

Cui, L., Li, T., Jiang, L., Deng, W., Huang, C., Chen, Z., et al. (2009). Comparative study of proton magnetic resonance spectroscopy on the hippocampus between patients with paranoid schizophrenia and patients with bipolar mania. *Chin J Nerv Ment Dis*. 35, 355–358. doi: 10.3760/cma.j.issn.1674-6554.2009.11.012

Dager, S. R., Corrigan, N. M., Richards, T. L., and Posse, S. (2008). Research applications of magnetic resonance spectroscopy to investigate psychiatric disorders. *Top. Magn. Reson. Imaging*. 19, 81–96. doi: 10.1097/RMR.0b013e318181e0be

Ding, Y., Pereira, F., Hoehne, A., Beaulieu, M. M., Lepage, M., Turecki, G., et al. (2017). Altered brain processing of decision-making in healthy first-degree biological relatives of suicide completers. *Mol. Psychiatry*. 22, 1149–1154. doi: 10.1038/mp.2016.221

Fan, S., Lippard, E. T. C., Sankar, A., Wallace, A., Johnston, J. A. Y., Wang, F., et al. (2019). Gray and white matter differences in adolescents and young adults with prior suicide attempts across bipolar and major depressive disorders. *J. Affect. Disord*. 245, 1089–1097. doi: 10.1016/j.jad.2018.11.095

Giakoumatos, C. I., Tandon, N., Shah, J., Mathew, I. T., Brady, R. O., Clementz, B. A., et al. (2013). Are structural brain abnormalities associated with suicidal behavior in patients with psychotic disorders? *J. Psychiatr. Res*. 47, 1389–1395. doi: 10.1016/j.jpsychires.2013.06.011

Gong, J., Chen, G., Zhou, M., Jia, Y., Zhong, S., Chen, F., et al. (2020). Characteristics of temporal dynamics of intrinsic brain activity in unmedicated bipolar disorder with suicidality. *Aust. N. Z. J. Psychiatry*. 54, 1115–1124. doi: 10.1177/0004867420948960

Gosnell, S. N., Fowler, J. C., and Salas, R. (2019). Classifying suicidal behavior with resting-state functional connectivity and structural neuroimaging. *Acta Psychiatr. Scand*. 140, 20–29. doi: 10.1111/acps.13029

Hajek, T., Bauer, M., Pfennig, A., Cullis, J., Ploch, J., O'Donovan, C., et al. (2012). Large positive effect of lithium on prefrontal cortex N-acetylaspartate in patients with bipolar disorder: 2-centre study. *J. Psychiatry Neurosci*. 37, 185–192. doi: 10.1503/jpn.110097

Hamilton, M. (1960). A rating scale for depression. *J. Neurol. Neurosurg. Psychiatry*. 23, 56–62. doi: 10.1136/jnnp.23.1.56

Hozer, F., and Houenou, J. (2016). Can neuroimaging disentangle bipolar disorder? *J A?ect.Disord*. 195, 199–214. doi: 10.1016/j.jad.2016.01.039

Ji, L., Meda, S. A., Tamminga, C. A., Clementz, B. A., Keshavan, M. S., Sweeney, J. A., et al. (2020). Characterizing functional regional homogeneity (ReHo) as a B-SNIP psychosis biomarker using traditional and machine learning approaches. *Schizophr. Res*. 215, 430–438. doi: 10.1016/j.schres.2019.07.015

Johnston, J. A. Y., Wang, F., Liu, J., Blond, B. N., Wallace, A., Liu, J., et al. (2017). Multimodal neuroimaging of frontolimbic structure and function associated with suicide attempts in adolescents and young adults with bipolar disorder. *Am. J. Psychiatry*. 174, 667–675. doi: 10.1176/appi.ajp.2016.15050652

Lai, S., Zhong, S., Shan, Y., Wang, Y., Chen, G., Luo, X., et al. (2019). Altered biochemical metabolism and its lateralization in the cortico-striato-cerebellar circuit of unmedicated bipolar II depression. *J. Affect. Disord*. 259, 82–90. doi: 10.1016/j.jad.2019.07.021

Lan, M. J., Rizk, M. M., Pantazatos, S. P., Rubin-Falcone, H., Miller, J. M., Sublette, M. E., et al. (2019). Resting-state amplitude of low-frequency fluctuation is associated with suicidal ideation. *Depress. Anxiety*. 36, 433–441. doi: 10.1002/da.22888

Lewis, C. P., Port, J. D., Blacker, C. J., Sonmez, A. I., Seewoo, B. J., Leffler, J. M., et al. (2020). Altered anterior cingulate glutamatergic metabolism in depressed adolescents with current suicidal ideation. *Transl. Psychiatry*. 10, 119. doi: 10.1038/s41398-020-0792-z

Li, H., Cui, L., Cao, L., Zhang, Y., Liu, Y., Deng, W., et al. (2020). Identification of bipolar disorder using a combination of multimodality magnetic resonance imaging and machine learning techniques. *BMC Psychiatry* 20, 488. doi: 10.1186/s12888-020-02886-5

Li, J., Duan, X., Cui, Q., Chen, H., and Liao, W. (2019). More than just statics: temporal dynamics of intrinsic brain activity predicts the suicidal ideation in depressed patients. *Psychol. Med*. 49, 852–860. doi: 10.1017/S0033291718001502

Librenza-Garcia, D., Kotzian, B. J., Yang, J., Mwangi, B., Cao, B., Pereira Lima, L. N., et al. (2017). The impact of machine learning techniques in the study of bipolar disorder: a systematic review. *Neurosci. Biobehav. Rev*. 80, 538–554. doi: 10.1016/j.neubiorev.2017.07.004

Lijffijt, M., Rourke, E. D., Swann, A. C., Zunta-Soares, G. B., and Soares, J. C. (2014). Illness-course modulates suicidality-related prefrontal gray matter reduction in women with bipolar disorder. *Acta Psychiatr. Scand*. 130, 374–387. doi: 10.1111/acps.12314

Liu, F., Guo, W., Fouche, J. P., Wang, Y., Wang, W., Ding,. J., et al. (2015). Multivariate classification of social anxiety disorder using whole brain functional connectivity. *Brain Struct. Funct*. 220, 101–115. doi: 10.1007/s00429-013-0641-4

Liu, F., Wang, Y., Li, M., Wang, W., Li, R., Zhang, Z., et al. (2017a). Dynamic functional network connectivity in idiopathic generalized epilepsy with generalized tonic-clonic seizure. *Hum. Brain Mapp*. 38, 957–973. doi: 10.1002/hbm.23430

Liu, T., Wang, Y., Zhong, S., Wang, B., Liao, X., Lai, S., et al. (2017b). A comparison of neurometabolites between remitted bipolar disorder and depressed bipolar disorder: a proton magnetic resonance spectroscopy study. *J. Affect. Disord*. 211, 153–161. doi: 10.1016/j.jad.2017.01.009

Mahon, K., Burdick, K. E., Wu, J., Ardekani, B. A., and Szeszko, P. R. (2012). Relationship between suicidality and impulsivity in bipolar I disorder: a diffusion tensor imaging study. *Bipolar Disord*. 14, 80–89. doi: 10.1111/j.1399-5618.2012.00984.x

Merikangas, K. R., Jin, R., He, J. P., Kessler, R. C., Lee, S., Sampson, N. A., et al. (2011). Prevalence and correlates of bipolar spectrum disorder in the world mental health survey initiative. *Arch Gener Psych*. 68, 241–251. doi: 10.1001/archgenpsychiatry.2011.12

Moon, C. M., Kang, H. K., and Jeong, G. W. (2015). Metabolic change in the right dorsolateral prefrontal cortex and its correlation with symptom severity in patients with generalized anxiety disorder: Proton MR spectroscopy at 3 Tesla. *Psychiatry Clin. Neurosci*. 69, 422–430. doi: 10.1111/pcn.12279

Moore, G. J., Bebchuk, J. M., Hasanat, K., Chen, G., Seraji-Bozorgzad, N., Wilds, I. B., et al. (2000). Lithium increases N-acetyl-aspartate in the human brain: in

vivo evidence in support of bcl-2's neurotrophic effects? *Biol. Psychiatry*. 48, 1–8. doi: 10.1016/S0006-3223(00)00252-3

Nielsen, A. N., Barch, D. M., Petersen, S. E., Schlaggar, B. L., and Greene, D. J. (2020). Machine learning with neuroimaging: evaluating its applications in psychiatry. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* 5, 791–798. doi: 10.1016/j.bpsc.2019.11.007

Nock, M. K., Borges, G., Bromet, E. J., Alonso, J., Angermeyer, M., Beautrais, A., et al. (2008). Cross-national prevalence and risk factors for suicidal ideation, plans and attempts. *Br. J. Psychiatry*. 192, 98–105. doi: 10.1192/bjp.bp.107.040113

O'Leary, L. A., Belliveau, C., Davoli, M. A., Ma, J. C., Tanti, A., Turecki, G., et al. (2021). Widespread decrease of cerebral vimentin-immunoreactive astrocytes in depressed suicides. *Front. Psychiatry*. 12, 640963. doi: 10.3389/fpsyt.2021.640963

Orban, P., Dansereau, C., Desbois, L., Mongeau-Pérusse, V., Giguère, C. É., Nguyen, H., et al. (2018). Multisite generalizability of schizophrenia diagnosis classification based on functional brain connectivity. *Schizophr. Res*. 192, 167–171. doi: 10.1016/j.schres.2017.05.027

Orrù, G., Pettersson-Yeo, W., Marquand, A. F., Sartori, G., and Mechelli, A. (2012). Using support vector machine to identify imaging biomarkers of neurological and psychiatric disease: a critical review. *Neurosci. Biobehav. Rev*. 36, 1140–1152. doi: 10.1016/j.neubiorev.2012.01.004

Patel, N. C., Cecil, K. M., Strakowski, S. M., Adler, C. M., and DelBello, M. P. (2018). Neurochemical alterations in adolescent bipolar depression: a proton magnetic resonance spectroscopy pilot study of the prefrontal cortex. *J. Child Adolesc. Psychopharmacol*. 18, 623–627. doi: 10.1089/cap.2007.151

Peng, Y., Zhang, X., Li, Y., Su, Q., Wang, S., Liu, F., et al. (2020). MVPANI: a toolkit with friendly graphical user interface for multivariate pattern analysis of neuroimaging data. *Front. Neurosci*. 14, 545. doi: 10.3389/fnins.2020.00545

Phillips, M. L., Drevets, W. C., Rauch, S. L., and Lane, R. (2003). Neurobiology of emotion perception II: implications for major psychiatric disorders. *Biol. Psychiatry*. 54, 515–528. doi: 10.1016/S0006-3223(03)00171-9

Phillips, M. L., Ladouceur, C. D., and Drevets, W. C. (2008). A neural model of voluntary and automatic emotion regulation: implications for understanding the pathophysiology and neurodevelopment of bipolar disorder. *Mol. Psychiatry*. 13, 833–857. doi: 10.1038/mp.2008.65

Plans, L., Barrot, C., Nieto, E., Rios, J., Schulze, T. G., Papiol, S., et al. (2019). Association between completed suicide and bipolar disorder: a systematic review of the literature. *J Affect Disord.* 242, 111–122. doi: 10.1016/j.jad.2018.08.054

Richard-Devantoy, S., Berlim, M. T., and Jollant, F. (2015). Suicidal behaviour and memory: a systematic review and meta-analysis. *World J. Biol. Psychiatry*. 16, 544–566. doi: 10.3109/15622975.2014.925584

Rocha, M. V., Nery-Fernandes, F., Guimarães, J. L., Quarantini Lde, C., de Oliveira, I. R., et al. (2015). Normal metabolic levels in prefrontal cortex in euthymic bipolar I patients with and without suicide attempts. *Neural. Plast*. 2015, 165180. doi: 10.1155/2015/165180

Sankar, A., Scheinost, D., Goldman, D. A., Drachman, R., Colic, L., Villa, L. M., et al. (2022). Graph theory analysis of whole brain functional connectivity to assess disturbances associated with suicide attempts in bipolar disorder. *Transl. Psychiatry*. 12, 7. doi: 10.1038/s41398-021-01767-z

Schwarz, E., Doan, N. T., Pergola, G., Westlye, L. T., Kaufmann, T., Wolfers, T., et al. (2019). Reproducible grey matter patterns index a multivariate, global alteration of brain structure in schizophrenia and bipolar disorder. *Transl. Psychiatry*. 9, 12. doi: 10.1038/s41398-018-0225-4

Shimon, H., Agam, G., Belmaker, R. H., Hyde, T. M., and Kleinman, J. E. (1997). Reduced frontal cortex inositol levels in postmortem brain of suicide victims and patients with bipolar disorder. *Am. J. Psychiatry*. 154, 1148–1150. doi: 10.1176/ajp.154.8.1148

Singh, M. K., Jo, B., Adleman, N. E., Howe, M., Bararpour, L., Kelley, R. G., et al. (2013). Prospective neurochemical characterization of child offspring of parents with bipolar disorder. *Psychiatry Res*. 214, 153–160. doi: 10.1016/j.pscychresns.2013.05.005

Song, J. Y., Yu, H. Y., Kim, S. H., Hwang, S. S., Cho, H. S., Kim, Y. S., et al. (2012). Assessment of risk factors related to suicide attempts in patients with bipolar disorder. *J. Nerv. Ment. Dis.* 200, 978–984. doi: 10.1097/NMD.0b013e3182718a07

Strakowski, S. M., Delbello, M. P., and Adler, C. M. (2005). The functional neuroanatomy of bipolar disorder: a review of neuroimaging findings. *Mol. Psychiatry* 10, 105–116. doi: 10.1038/sj.mp.4001585

van Heeringen, K., Van den Abbeele, D., Vervaet, M., Soenen, L., and Audenaert, K. (2010). The functional neuroanatomy of mental pain in depression. *Psychiatry Res*. 181, 141–144. doi: 10.1016/j.pscychresns.2009.07.011

Wu, R., Silverstone, P., and Lin, Y. (2001). No significant change of brain myo-inositol is observed in bipolar affective disorder after sodium valproate medication by *in vivo* proton MR spectroscopy. *Proc. Intl. Soc. Mag. Reson. Med*. 9, 1012.

Xia, C., Wang, D., Wu, S., and Ye, J. (2002). Preliminary development of the self-rating idea of suicide scale. *J. Clin. Psychiatry*. 12, 100–102.

Yang, J., Pu, W., Ouyang, X., Tao, H., Chen, X., Huang, X., et al. (2019). Abnormal connectivity within anterior cortical midline structures in bipolar disorder: evidence from integrated MRI and functional MRI. *Front. Psychiatry*. 10, 788. doi: 10.3389/fpsyt.2019.00788

Young, R. C., Biggs, J. T., Ziegler, V. E., and Meyer, D. A. (1978). A rating scale for mania: reliability, validity and sensitivity. *Br. J. Psychiatry*. 133, 429–435. doi: 10.1192/bjp.133.5.429

Zhang, W., Nery, F. G., Tallman, M. J., Patino, L. R., Adler, C. M., Strawn, J. R., et al. (2021). Individual prediction of symptomatic converters in youth offspring of bipolar parents using proton magnetic resonance spectroscopy. *Eur. Child Adolesc. Psychiatry*. 30, 55–64. doi: 10.1007/s00787-020-01483-x

Zhao, Y., Wang, L., Edmiston, E. K., Womer, F. Y., Jiang, X., Wu, F., et al. (2021). Alterations in gray matter volumes and intrinsic activity in the prefrontal cortex are associated with suicide attempts in patients with bipolar disorder. *Psychiatry Res. Neuroimaging* 307, 111229. doi: 10.1016/j.pscychresns.2020.111229

Zhong, S., Wang, Y., Lai, S., Liu, T., Liao, X., Chen, G., et al. (2018). Associations between executive function impairment and biochemical abnormalities in bipolar disorder with suicidal ideation. *J. Affect Disord*. 241, 282–290. doi: 10.1016/j.jad.2018.08.031

Zhong, S., Wang, Y., Zhao, G., Xiang, Q., Ling, X., Liu, S., et al. (2014). Similarities of biochemical abnormalities between major depressive disorder and bipolar depression: a proton magnetic resonance spectroscopy study. *J. Affect Disord*. 168, 380–386. doi: 10.1016/j.jad.2014.07.024

Zhou, C. (2013). *The metabonomics of plasma of major depressive disorder with suicide attempts* (dissertation). Chongqing: Chongqing Medical University.

frontiers | Frontiers in Neuroimaging

TYPE Original Research
PUBLISHED 22 September 2022
DOI 10.3389/fnimg.2022.917806

OPEN ACCESS

EDITED BY
Yao Wu,
Children's National Hospital,
United States

REVIEWED BY
Nusrat Sharmin,
Military Institute of Science and Technology (MIST), Bangladesh
Silvia Basaia,
San Raffaele Hospital (IRCCS), Italy

*CORRESPONDENCE
Guillaume Theaud
guillaume.theaud@usherbrooke.ca

†Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in the analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

SPECIALTY SECTION
This article was submitted to Brain Imaging Methods, a section of the journal Frontiers in Neuroimaging

RECEIVED 11 April 2022
ACCEPTED 09 June 2022
PUBLISHED 22 September 2022

DORIS: A diffusion MRI-based 10 tissue class deep learning segmentation algorithm tailored to improve anatomically-constrained tractography

Guillaume Theaud[1,2]*, Manon Edde[1], Matthieu Dumont[2], Clément Zotti[2], Mauro Zucchelli[3], Samuel Deslauriers-Gauthier[3], Rachid Deriche[3], Pierre-Marc Jodoin[2,4] and Maxime Descoteaux[1,2] for the Alzheimer's Disease Neuroimaging Initiative†

[1]Sherbrooke Connectivity Imaging Laboratory (SCIL), Université de Sherbrooke, Sherbrooke, QC, Canada, [2]Imeka Solutions Inc., Sherbrooke, QC, Canada, [3]Athena Project-Team, Inria Sophia Antipolis-Méditerranée, Université Côte D'Azur, Nice, France, [4]Videos & Images Theory and Analytics Laboratory (VITAL), Université de Sherbrooke, Sherbrooke, QC, Canada

Modern tractography algorithms such as anatomically-constrained tractography (ACT) are based on segmentation maps of white matter (WM), gray matter (GM), and cerebrospinal fluid (CSF). These maps are generally estimated from a T1-weighted (T1w) image and then registered in diffusion weighted images (DWI) space. Registration of T1w to diffusion space and partial volume estimation are challenging and rarely voxel-perfect. Diffusion-based segmentation would, thus, potentially allow not to have higher quality anatomical priors injected in the tractography process. On the other hand, even if FA-based tractography is possible without T1 registration, the literature shows that this technique suffers from multiple issues such as holes in the tracking mask and a high proportion of generated broken and anatomically implausible streamlines. Therefore, there is an important need for a tissue segmentation algorithm that works directly in the native diffusion space. We propose **DORIS**, a DWI-based deep learning segmentation algorithm. DORIS outputs 10 different tissue classes including WM, GM, CSF, ventricles, and 6 other subcortical structures (putamen, pallidum, hippocampus, caudate, amygdala, and thalamus). DORIS was trained and validated on a wide range of subjects, including 1,000 individuals from 22 to 90 years old from clinical and research DWI acquisitions, from 5 public databases. In the absence of a "true" ground truth in diffusion space, DORIS used a silver standard strategy from Freesurfer output registered onto the DWI. This strategy is extensively evaluated and discussed in the current study. Segmentation maps provided by

DORIS are quantitatively compared to Freesurfer and FSL-fast and the impacts on tractography are evaluated. Overall, we show that DORIS is fast, accurate, and reproducible and that DORIS-based tractograms produce bundles with a longer mean length and fewer anatomically implausible streamlines.

KEYWORDS

diffusion magnetic resonance imaging, tractography, anatomical constraints, image segmentation, machine learning

1. Introduction

Diffusion MRI is often used to explore structural brain connectivity using tractography. Tractography algorithms use orientation fields (Pierpaoli and Basser, 1996; Kreher et al., 2005; Peled et al., 2006; Tournier et al., 2007; Descoteaux, 2008; Jeurissen et al., 2014) to reconstruct the main white matter pathways of the brain. Moreover, tractography algorithms are mostly based on tracking masks to guide the algorithm where it is allowed to go and where it should stop. The easiest way to obtain a tracking mask is by thresholding the DTI fractional anisotropy (FA) map (Côté et al., 2013; Farquharson et al., 2013; Chamberland et al., 2014; Jeurissen et al., 2019; Vanderweyen et al., 2020), and this is usually recommended for pathological brains (Theaud et al., 2019; Vanderweyen et al., 2020). However, this thresholding does not support 3-way crossing areas resulting in holes in the tracking mask and can introduce biases in further analyses such as along tract-profiling and tractometry (Bells et al., 2011; Cousineau et al., 2017). In addition, since FA-based tracking cannot enforce anatomical constraints, it is known to suffer from broken streamlines that terminate prematurely in the white matter (WM). FA-based tracking also suffers from anatomically implausible streamlines that wrongly go through gray matter (GM) and cerebrospinal fluid (CSF) voxels (Girard et al., 2014). This is why nowadays, it is often recommended to use anatomical constraints to ensure streamlines reach GM, subcortical structures, or exit the brainstem, and not terminate in CSF voxels (Smith et al., 2012; Girard et al., 2014). This gave birth to the family of *"anatomically-constrained tractography (ACT)"* algorithms based on more precise masks of WM, GM, and CSF (Smith et al., 2012; Girard et al., 2014; Aydogan and Shi, 2020). More recently, surface-enhanced tractography (SET) (St-Onge et al., 2018, 2021) was proposed to further improve ACT. SET is also based on the same paradigm as ACT but adds another constraint using the cortical surface mesh to initialize tracking and enforce the ending streamline segment orthogonal to the cortex.

Adding a layer of anatomical prior has obvious benefits but also comes with its computational challenges, as an error in this prior often leads to inadequate or suboptimal tractography. Hence, in principle, all algorithms based on ACT or SET require voxel-perfect tissue segmentation in diffusion space to work optimally. Currently, tissue segmentation algorithms such as FSL-fast (Zhang et al., 2001), Atropos (Avants et al., 2009), or Freesurfer (Fischl, 2012) are based on T1- or T2-weighted images and thus, always require a registration step to bring the segmented tissue maps into diffusion space. This registration step is not perfect and is sensitive to preprocessing steps such as brain extraction (Chen et al., 2019). It is also a step that can take multiple hours (from 1 to 10 h), depending on the tool used. Additionally, just like registration algorithms, structural segmentation algorithms, despite spectacular improvements over the past couple of years, are not voxel-perfect wither everywhere in the brain. Segmentation errors often occur in partial volume voxels between tissues (e.g., WM-GM partial volume) and in nuclei extraction. Partial volume effects, nuclei not segmented, or registration errors can thus have a negative impact on tractography algorithms. For example, tractography can be incorrectly allowed to end streamlines in partial volume near the ventricles due to segmentation errors and/or registration inaccuracies in that area. Moreover, the absence of nuclei boundaries allows tracking to go through them without meeting a stopping criterion. Furthermore, the registration issues impact endpoints of streamlines and forbid streamlines to traverse narrow WM corridors such as around the insula, external capsule, and other deep structures. As a result, further analysis such as tractometry can be biased and accumulate undesired errors along streamlines.

Diffusion MRI segmentation Diffusion-based segmentation algorithms address registration issues that are not voxel-perfect, which can have a negative impact on tractography results. Moreover, segmenting in diffusion MRI space permits to obtain tissue maps faster in the whole pipeline removing the dependency on T1w preprocessing and T1w registration. For these reasons, diffusion-based segmentation algorithms have started to appear as promising tools in the literature. Recently, Zhang et al. (2021) developed a deep learning algorithm to segment dMRI into WM/GM/CSF classes. The model was trained, validated, and tested on 190 young subjects (under 40 years old) from 5 databases (Glasser et al., 2013; Casey et al., 2018; Tong et al., 2020; Garza-Villarreal et al., 2021). In addition, it required a multi-shell dMRI

acquisition to extract kurtosis features (Jensen and Helpern, 2010; Steven et al., 2014), which could be considered a limitation as such a technique cannot work on single-shell clinical dMRI acquisitions. On the other hand, Little and Beaulieu (2021) created an algorithm to segment the GM ribbon based on a combination of fractional anisotropy from DTI and the powdered average DWI from a single-shell diffusion acquisition. Their method was designed to quantify DTI measures under the cortex in a young and healthy cohort of 15 adults. Hence, these two segmentation algorithms were not meant to enhance tractography. Moreover, these two algorithms were developed only on subjects under 40 years of age and could, thus, be misadapted for elderly subjects with larger ventricles, brain atrophy, and white matter hyperintensities. Furthermore, both the Little and Beaulieu (2021) and Zhang et al. (2021) methods did not fully manage deep nuclei, which are very important in tractography to retrieve bundles that connect cortical GM to nuclei such as the optic radiations, among others. Other DWI-based segmentation algorithms permit to extract a 3-class WM/GM/CSF segmentation (Li et al., 2006; Liu et al., 2007; Saygin et al., 2011; Ye et al., 2012; Yap et al., 2015; Zhang et al., 2015; Visser et al., 2016; Battistella et al., 2017; Ciritsis et al., 2018; Nie et al., 2018; Cheng et al., 2020; Wang et al., 2020) but, as the two methods previously presented, they are not adapted to enhance the tractography.

Contributions of our study To address the aforementioned issues, we present DORIS: a novel diffusion MRI-based 10 tissue class deep learning segmentation algorithm tailored to improve anatomically-constrained tractography. DORIS is based on a DenseUnet model (Kaku et al., 2019), a convolutional neural network composed of dense blocks in the encoder and decoder path. DORIS is trained and validated on 1,000 subjects from 22 to 90 years old with single and multi-shell acquisitions from 5 databases. In the absence of a "true" ground truth in diffusion space, DORIS uses a silver standard strategy from Freesurfer (Fischl, 2012) output registered onto the DWI. Even if not a perfect ground truth or gold standard, this strategy will be extensively evaluated and discussed later. To train our model, we test several different potential diffusion-based features as input. We test 5 DTI- and HARDI-derived measures as input channels and also test 3 other input channel variants including rotation invariant features (Zucchelli et al., 2020), DTI, and spherical harmonic measures. These 4 input channel variants highlight the possibility of using single and/or multi-shell acquisitions and easily adding DORIS in a dMRI processing pipeline. DORIS predicts a total of 10 tissue class labels: the WM, the GM, the ventricles, the CSF around the brain, and 6 subcortical regions (putamen, pallidum, hippocampus, caudate, amygdala, and thalamus). The goal of DORIS is to segment WM, GM, ventricles, and nuclei voxel maps from diffusion MRI data *only* so that they can be used to perform anatomically-constraint tractography. In summary, this study presents: (i) a large training and validation set covering a wide range of DWI acquisitions and ages, (ii) extensive testing of optimal diffusion measures that drive the DenseUnet algorithm, (iii) a quantitative evaluation of DORIS against Freesurfer and FSL-fast, (iv) speed acceleration compared to well-known segmentation algorithms, and (v) qualitative and quantitative anatomically-constrained tractography benefits.

2. Methods

2.1. Datasets

Training and validation datasets come from the Alzheimer's Disease Neuroimaging Initiative (ADNI), the Parkinson's Progression Markers Initiative (PPMI), the Human Connectome Project (HCP-1200), and the UK Biobank databases.

Alzheimer's Disease Neuroimaging Initiative was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). Data were also obtained from the Parkinson's Progression Markers Initiative (PPMI) database (www.ppmi-info.org/access-data-specimens/download-data). HCP-1200 data is described in Glasser et al. (2013) and information about the UK Biobank database can be found at www.ukbiobank.ac.uk. As seen in Table 1, these databases cover a wide range of ages (22–90 years old). A total number of 1,000 subjects were used for training and validation sets, 750 subjects for training and 250 for validation.

Testing data comes from the Penthera3T (Paquette et al., 2019) database. Penthera3T is a test-retest database containing scans of 12 subjects from 24 to 30 years old. Each subject was scanned 6 times (two sessions of three scans per session) for a total of 72 acquisitions. We used the 72 acquisition to evaluate DORIS (testing dataset) and be able to evaluate the reproducibility of the segmentation in test-retest. Moreover, some subjects from ADNI and HCP will be used to gauge the performance of DORIS and the impact it has on the tractography.

2.2. Data processing

Data processing was mostly done using Nextflow pipelines (Di Tommaso et al., 2017; Kurtzer et al., 2017). Train, validation, and test sets were preprocessed using the TractoFlow pipeline (Theaud et al., 2020) with its default parameters. *B*-values below 1,200 s/mm^2 were used for DTI metrics while fiber orientation

TABLE 1 Datasets (training, validation, or testing set), age distribution, number of subjects, *b*-value, number of directions and resolution in mm^3 isotropic (excepted ADNI) for ADNI, PPMI, HCP, UKBiobank (UKB), and Penthera3T (P3T).

Database	Usage	Age range	Nb subjects used	*b*-value(s)	Nb dirs	Resolution
ADNI	Train/valid/testing*	60–90	250	1,000	41	1.3 x 1.3 x 2.7
PPMI	Train/valid	35–80	250	1,000	64	2
HCP	Train/valid/testing*	22–35	250	1,000, 2,000, 3,000	90, 90, 90	1.25
UKB	Train/valid	40–69	250	1,000, 2,000	50, 50	2
P3T	Testing	24–30	72**	300, 1,000, 2,000	8, 32, 60	2

* For illustration and example purpose.
** 12 subjects were scanned six times (two sessions with three scans per session).

distribution function (fODF) metrics were computed with *b*-values above 700 s/mm^2. The fODF was generated using a spherical harmonics order of 8, the same fiber response function (Descoteaux, 2008) for all the subjects (15, 4, 4) x 10^{-4} mm^2/s and with all the shells that required the previously presented requirement. Freesurfer (Fischl, 2012) was used on the raw T1 weighted images. Further details on Freesurfer are provided in Section 2.5.

The data processing pipeline is illustrated in Figure 1A. Data generated by TractoFlow (Theaud et al., 2020) and Freesurfer (Fischl, 2012) were manually verified using a quality control (QC) step with dmriqc flow https://github.com/scilus/dmriqc_flow. Full data processing took 15,000 CPU h, 2,000 GPU h, and 4 Tb of storage. Manual quality control took the authors 50 h of work.

2.3. Model

DenseUNet is the deep neural network upon which DORIS is built. This model's architecture was presented by Kaku et al. (2019) and had a similar objective to our study by proposing a segmentation algorithm of brain tissues in several labels based on T1w. Moreover, in this study, they compared the DenseUNet with a classical U-Net that was, in their specific case, less accurate than the DenseUNet. Due to excellent segmentation capabilities and the similarity of Kaku et al. (2019) with our segmentation objective, the DenseUNet was selected to segment diffusion MRI due to its. This model selection choice will be further discussed in the Section 4.

For DORIS, we replaced ReLU with LeakyReLU and, due to memory constraints, image patches were used instead of the full image. In addition, we used exponential logarithmic loss (Wong et al., 2018) developed for unbalanced labels and small structures. This loss is defined as:

$$L_{Exp} = w_{Dice}L_{Dice} + w_{CE}L_{CE}, \quad (1)$$

with w_{Dice} and w_{CE} being weights for the Dice loss L_{Dice} and the cross entropy loss L_{CE}. The learning rate was chosen to be 0.0001. Moreover, we used data augmentation (rotation, scaling, shearing, and axis flip) on the training set (value ranges for rotation, scaling and shearing are available in Supplementary Materials). Training of the Dense-UNet went on for 163 epochs before it was stopped by an early stopping criterion.

2.4. Hyperparameter optimization

We optimized 5 hyperparameters listed in Figure 1B. This includes four different input channels i.e.,:

1. DWI ($b = 1{,}000$ s/mm^2) fitted with spherical harmonics (SH) of order 4 with b0 concatenated (to add the non-diffusion T2w contrast). This image is a simplified representation of the DWI.
2. The 6 eigenvalues of the tensor matrix from DTI fit. The eigenvalues are the rotation invariant features of the diffusion tensor.
3. The rotation invariant features (from $b = 1{,}000$ s/mm^2 DWI) (Zucchelli et al., 2020) of the spherical harmonics. These are the 4th-order HARDI shape representation equivalent of the DTI eigenvalues.
4. Four DTI- and HARDI-derived measures include maximum and total apparent fiber density (AFD max and AFD Total), the DTI axial diffusivity (AD), mean diffusivity (MD), and radial diffusivity (RD). As pointed out by Chamberland et al. (2019), these 4 diffusion measures are good representative features that maximize the variance of the dMRI data using principal component analyses (PCA).

These images were selected because of the straightforwardness of computing them with public open-access software and because they can be extracted from any single-shell DWI acquisition, which makes them more suitable for future translation to a wide range of applications.

We also optimized the number of feature maps in the first layer (30, 35, and 40), the patch size (32 x 32 x 32, 64 x 64 x 64, and 128 x 128 x 128), the number of convolution layers (2, 3, and 4), and the batch size (1, 2, 3, and 4). These hyperparameters

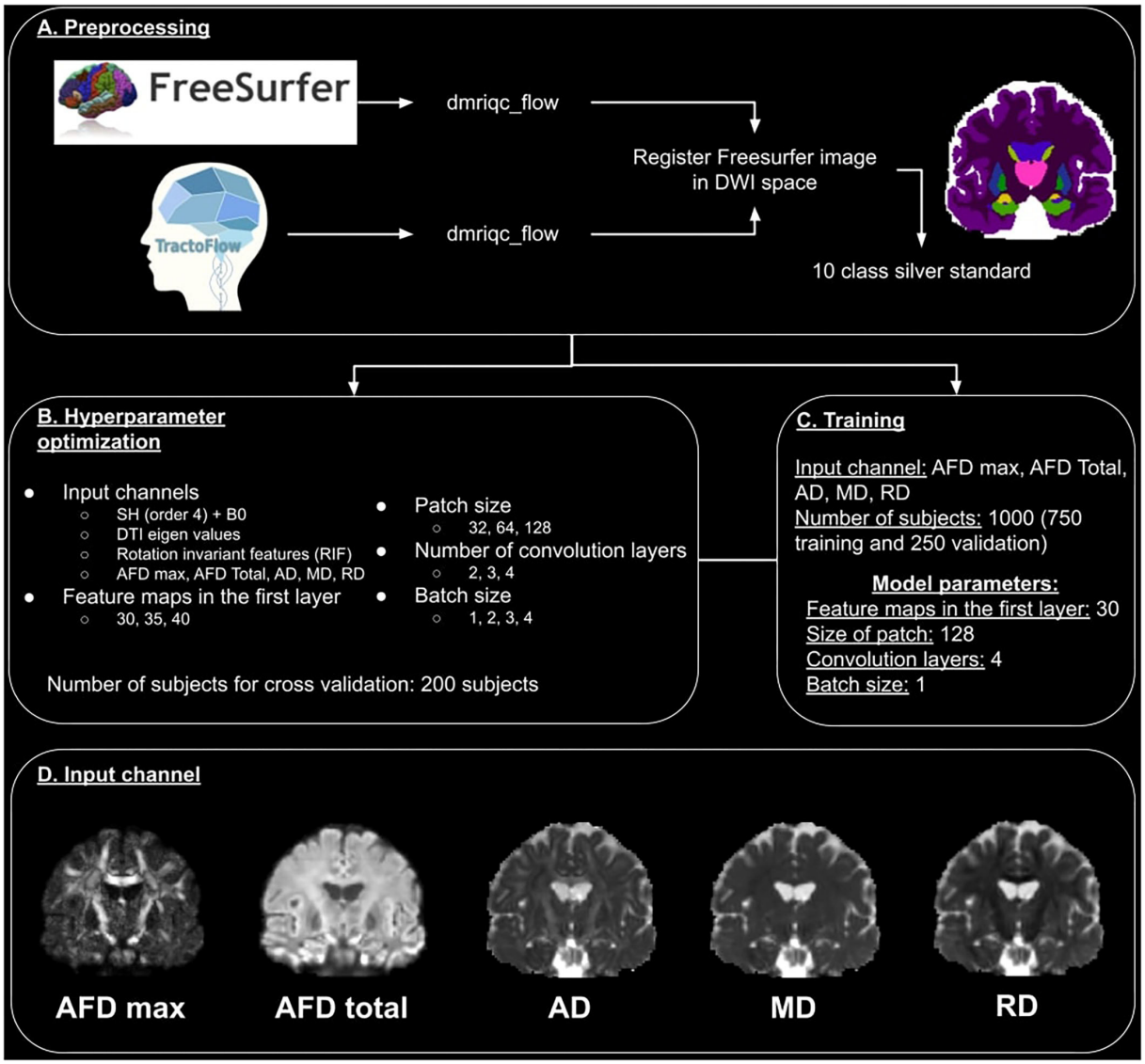


FIGURE 1
Overview of the DORIS processing pipeline. **(A)** Represents the preprocessing pipeline used for training, validation, and test datasets. **(B)** Illustrates the hyperparameter optimization and input channel selection. **(C)** Describes the training process. **(D)** Illustrates the 5 images used as an input channel.

were optimized with data of 200 subjects from the dataset previously described.

The hyperparameter search revealed that the best input channel to use is the 4D image made of AFD max, AFD Total, AD, MD, and RD images (Chamberland et al., 2019). A representative illustration of these 5 images is shown in Figure 1D. The best number of feature maps in the first layer is 30, the best patch size is 128, the best number of convolution layers is 4 and the best batch size is 1. The summary of this optimization is shown in Figure 1C.

2.5. Silver standard

Since the notion of "true" ground truth (or gold standard) is unavailable in diffusion space, we used Freesurfer (Fischl, 2012) to create our "reference maps" in diffusion space, or what we prefer to call our **silver standard**. As illustrated in Figure 2A, the native T1-weighted images were first processed by Freesurfer 6.0 to generate wmparc and aparc+aseg images, which were then registered in diffusion space using an ANTs registration operation (the exact command is available in the

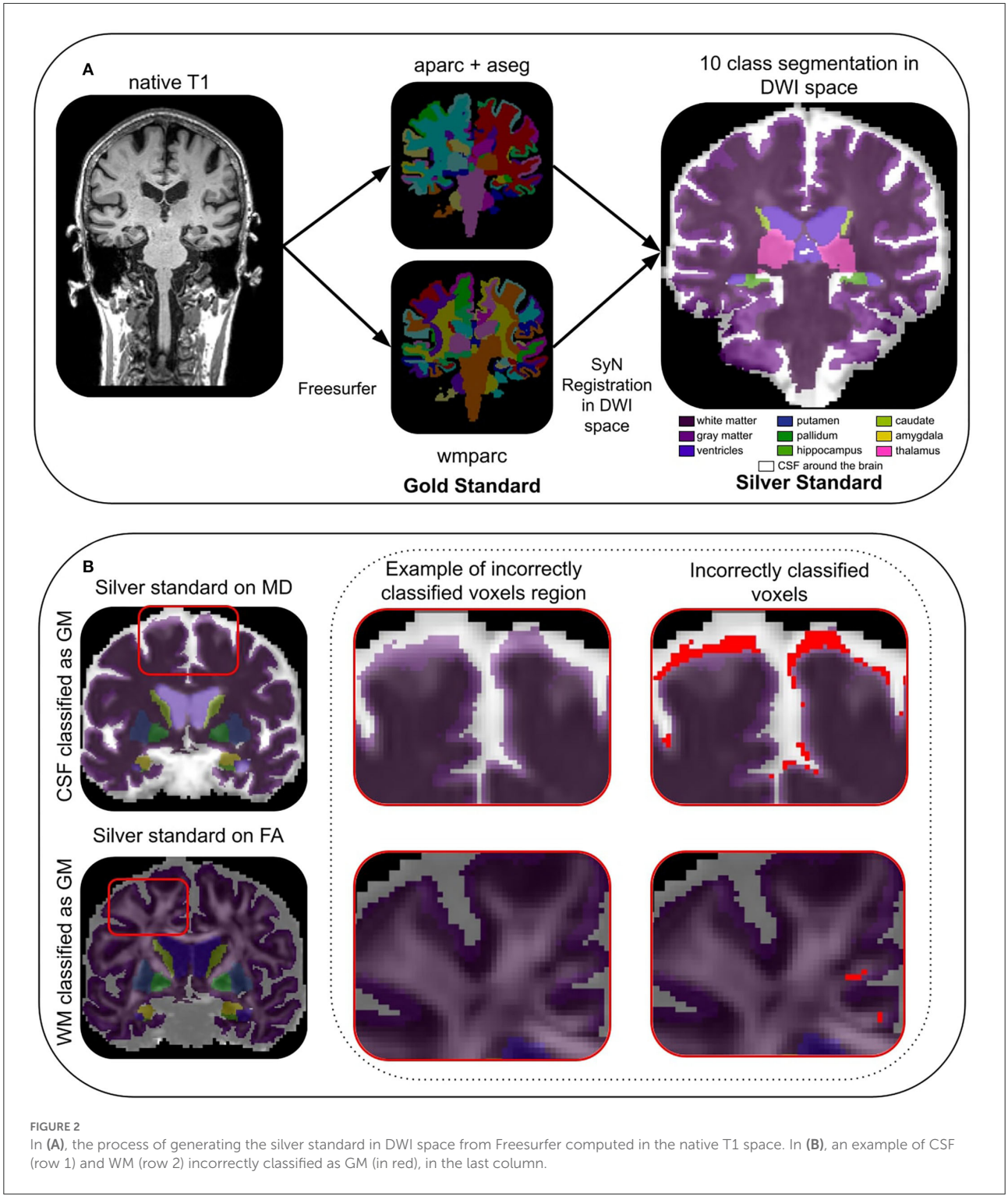


FIGURE 2
In **(A)**, the process of generating the silver standard in DWI space from Freesurfer computed in the native T1 space. In **(B)**, an example of CSF (row 1) and WM (row 2) incorrectly classified as GM (in red), in the last column.

Supplementary Materials) (Avants et al., 2009). The registration used the nearest neighbor interpolation and an affine + non-linear SyN warping transformations from the Tractoflow output. Next, we concatenated Freesurfer labels from these registered wmparc and aparc+aseg images to create our **silver standard** reference maps with the following 10 labels: (1) white matter, (2) gray matter, (3) ventricles, (4) putamen, (5) pallidum, (6) hippocampus, (7) caudate, (8) amygdala, (9) thalamus, and (10) CSF around the brain, as seen in Figure 2.

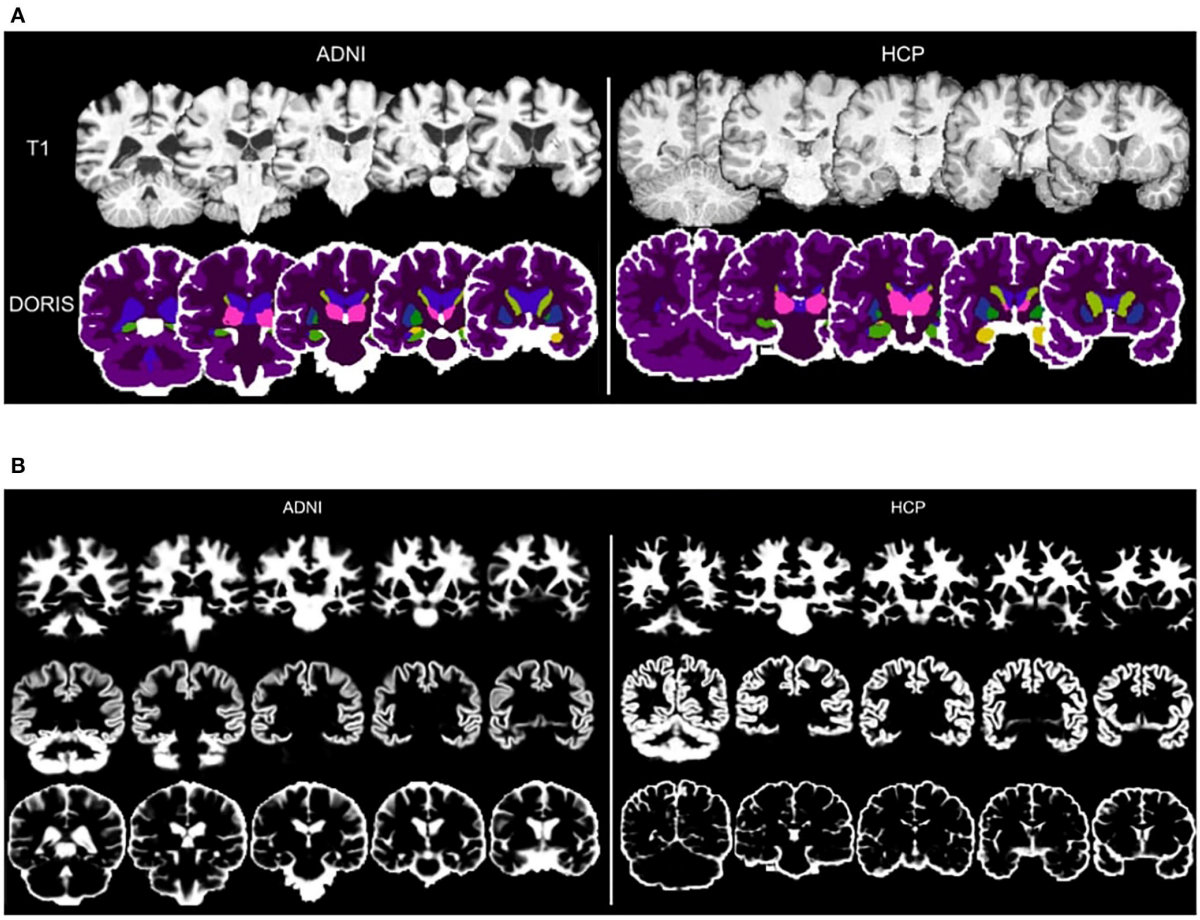


FIGURE 3
Binary and probability segmentation maps obtained from DORIS on ADNI and HCP subject. **(A)** DORIS binary segmentation on a ADNI and HCP subject. T1 image illustrate differences between an aging and young brain. **(B)** DORIS probability segmentation. of WM, GM and CSF (combining ventricules and CSF around the brain) on a ADNI and HCP subject. The mosaic containing the nuclei is available in Supplementary Materials.

2.5.1. Incorrectly classified voxels

Due to the registration step, our silver standard suffers from issues that we previously described, i.e., some voxels are wrongly aligned with the dMRI data, which leads to **incorrectly classified voxels**. This is expected considering the nature of our silver standard construction methodology. Some of these incorrectly classified voxels can be automatically detected using measures in diffusion space such as MD and FA from DTI maps. Based on a theoretical CSF mean diffusivity of $3 \times 10^{-3} mm^2/s$ reported in the literature (Koo et al., 2009; Pasternak et al., 2009; Zhang et al., 2012), a "safe" CSF mask (Dumont et al., 2019) containing the ventricles and the part of the constrained CSF between the skull and the brain is extracted. To extract this "safe" CSF mask, MD voxels higher than $2 \times 10^{-3} mm^2/s$ (Groen et al., 2011) were selected. To quantify the number of CSF voxels classified as GM voxels, we intersect the "safe" CSF mask with the GM region of DORIS as well as that of the silver standard. As shown in Figure 2B, the number of voxels in the two intersections corresponds to the number of CSF voxels incorrectly classified as GM. For the "safe" WM mask (Dumont et al., 2019), we considered voxels with a FA value above 0.3 as WM (Chamberland et al., 2014). Then, we intersect the GM label with the "safe" WM mask. The number of voxels in the two intersections corresponds to the number of WM voxels incorrectly classified as GM. Finally, note that the CSF around the brain label will only be used to compute the incorrectly classified voxels and not used in the other analysis. Indeed, CSF around the brain is not useful for tractography purposes.

Incorrectly classified voxels will permit us to address the limitation of our silver standard and explore solutions, with DORIS in native diffusion space, that could potentially have fewer incorrectly classified voxels than in the silver standard itself. The goal of DORIS is to generate fewer incorrectly classified voxels than the silver standard due to the large number of subjects in the training set and to diffusion measures inputted into the model. This issue will be extensively discussed in the Section 4.

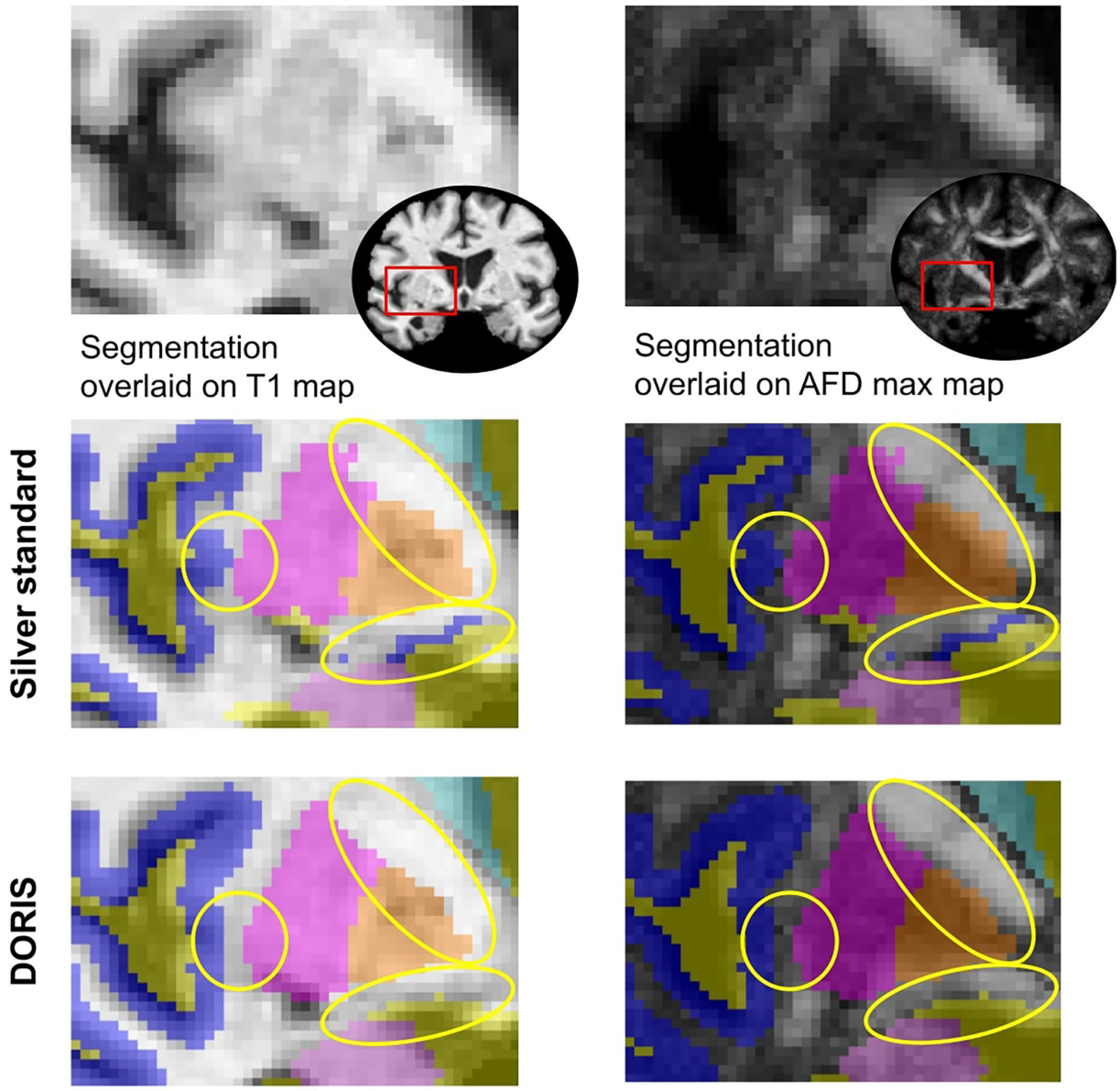


FIGURE 4
Silver standard and DORIS segmentation near the insula overlaid on the T1 and AFD max map. Yellow circles show where DORIS has a better segmentation than the silver standard.

2.6. Evaluation

The DORIS evaluation consists of 7 steps:

1. Qualitative results of DORIS on ADNI and HCP subjects,
2. A computation time comparison between DORIS and state-of-the-art algorithms (Fastsurfer; Henschel et al., 2020, FSL-fast; Zhang et al., 2001, and Freesurfer; Fischl, 2012),
3. A comparison between a manual segmentation, DORIS, and the silver standard in a small region-of-interest (ROI),
4. The number of incorrectly classified voxels generated by DORIS compared to the silver standard,
5. A comparison between the volume of labels from DORIS and the silver standard to ensure the reproducibility in test-retest,
6. A Dice score (Dice, 1945) between DORIS and the silver standard,
7. Quantitative and qualitative analyses about the impact of DORIS on tractography.

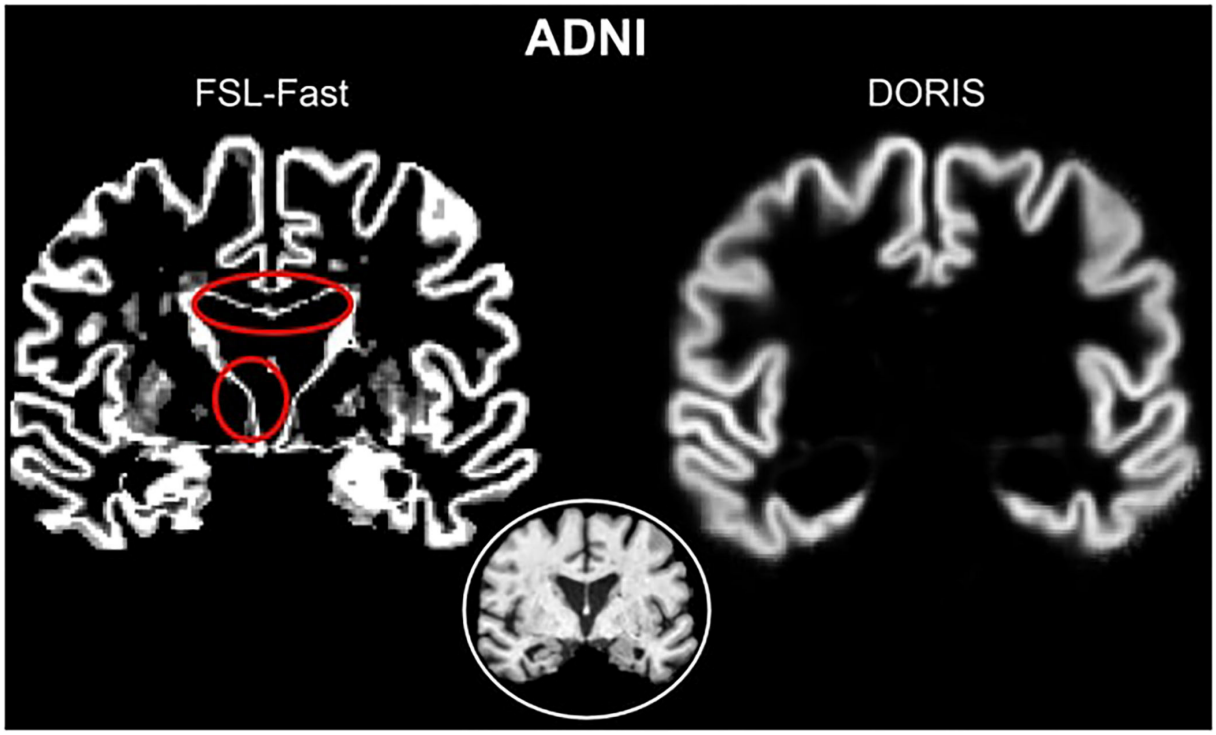


FIGURE 5
Red circles show WM-CSF partial volume incorrectly classified as GM in FSL-Fast. However, DORIS does not classify these partial volume voxels as GM.

2.6.1. Evaluation using a manual segmentation

As introduced previously, our silver standard is imperfect. Thus, to compare DORIS with the silver standard, an ROI located near the insula (illustrated in Figure 6) was manually segmented by an expert. We targeted that region due to its complexity of being segmented. Indeed, this area is composed of small white matter corridors between the insula, the putamen, the thalamus, and the caudate nuclei. To facilitate the manual segmentation process, nuclei were not separated from the GM cortex, i.e., they are both considered the same label for this evaluation experiment. To have a better analysis of DORIS, the manual segmentation will be used as a reference. For DORIS and the silver standard, the same labels were extracted from the ROI to make a qualitative analysis. Finally, quantitative analysis will be done by computing the Dice score between DORIS and the manual segmentation; and between the silver standard and the manual segmentation.

2.6.2. Evaluation using Penthera3T

Incorrectly classified voxel To determine if one of the segmentation generates significantly more outliers than the other, a two-sided *t*-test was also performed between incorrectly classified voxels generated by DORIS and the silver standard. As for the volume test, the significance level was fixed at $p < 0.001$.

Volume For DORIS and the silver standard, we compute the volume in mm^3 of each of the 9 labels (excluding the CSF around the brain). To validate if the volume is significantly different, a related two-sided *t*-test is performed between sessions for DORIS and the silver standard. The volume is determined as significantly different in the case where $p < 0.001$ (using scipy). Moreover, the volume for each label is averaged across the full dataset to obtain a global mean volume for each label.

Dice For each label and each acquisition, we compute the Dice score (Dice, 1945) between DORIS and our silver standard. Then, we average the Dice scores of each subject to obtain one average Dice score per label across all the testing datasets. We also compute 2 additional WM map DWI based as a comparison based on classical FA threshold higher than 0.15 and WM signal fraction from multi-shell multi-tissue (MSMT) fODF (Jeurissen et al., 2014) higher than 0.1.

2.6.3. Impact of DORIS on tractography

To evaluate the impact of DORIS on tractography, we performed a particle filtering tractography (PFT) (Girard et al., 2014), which is part of the ACT family of algorithms (Smith

et al., 2012). The PFT algorithm is known to be more restricted by probabilistic tissue maps. Acquisitions used for this qualitative analysis come from ADNI and Penthera3T database to cover a large age range of dMRI quality and anatomical difficulty (less WM in ADNI, enlarged ventricles, thinner cortex, presence of white matter hyperintensities). To compare our DORIS-based tractograms, we also compute a standard tractogram using FSL-fast maps, as done in Tractoflow (Theaud et al., 2020). Tracking parameters were the same for both techniques: a probabilistic tractography using a step size of 0.5mm and a maximal angle between two steps of 20 degrees. The fODF used in both tracking was computed using a spherical harmonic order of 8. We launched 5 seeds per voxel from the WM mask (obtained from DORIS or FSL-fast).

To extract white matter bundles, we ran RecobundleX (RBx) (Garyfallidis et al., 2015; Rheault, 2020) on both DORIS and FSL-Fast Tractoflow tractograms. From RBx, 6 bundles were used for quantitative and qualitative analyses: (i) Superior longitudinal fasciculus (SLF), (ii) whole corpus callosum (CC), (iii) Inferior fronto-occipital fasciculus (IFOF), (iv) fornix (FX), (v) anterior, and (vi) posterior commissure (AC/PC). Based on Maier-Hein et al. (2017), we considered the SLF and CC as "easy to track" bundles, the IFOF as "hard to track" bundles, and the FX and AC/PC as "very hard to track" bundles. For these 6 bundles, the number of streamlines and the mean bundle length are reported. Finally, the number of streamlines per bundle is divided by the total number of streamlines and expressed as a percentage ($ratio = (number\ of\ streamlines\ in\ the\ bundle\ /\ number\ of\ streamlines\ in\ the\ tractogram) \times 100$).

3. Results

3.1. DORIS segmentation on ADNI and HCP

Figure 3A shows a qualitative example of DORIS (bottom row) performance on ADNI and HCP subjects. Figure 3B shows, for the same subjects, the probabilistic maps obtained by DORIS. For the ADNI subject, DORIS correctly classified the WM even with the presence of aging lesions (white matter hyperintensities). Moreover, DORIS separates the GM cortex from the nuclei (see Supplementary Materials for nuclei figure). As shown in Figure 4, DORIS is better than the silver standard to segment the nuclei and tiny white matter corridors.

Then, as illustrated in Figure 5 with an ADNI subject, DORIS does not classify WM-CSF partial volume (around the ventricles) as GM. This misclassification of WM-CSF partial volume is visible in FSL-Fast segmentation. To easily compare DORIS to FSL-Fast and the silver standard, figures with probabilistic maps obtained by FSL-Fast and silver standard segmentation on these two subjects are available in Supplementary Materials.

TABLE 2 Computation time per subject for DORIS, FastSurfer, FSL-Fast, and FreeSurfer.

Algorithm	Computation time	SyN registration time
DORIS	48 s	N/A
FastSurfer	42 s	45 m
FSL-Fast	2 m: 12 s	45 m
Freesurfer	10 h	45 m

In the last column, SyN registration computation time are reported to bring the T1 image into DWI space is described.

3.2. Computation time

We evaluated the computation time of DORIS compared to three well-known segmentation algorithms based on T1. Table 2 shows the computation time for each algorithm. DORIS has a similar computation time to FastSurfer (48 and 42 s, respectively), it is faster than FSL-fast (approximately 2 min) and orders of magnitude faster than the *recon-all* command from Freesurfer (10 h). It is important to note that even if some algorithms (FastSurfer or FSL-Fast) have a computation time comparable to DORIS, they need an extra 45 min to bring their segmentation map into the diffusion space.

3.3. Manual segmentation

Figure 6B shows the manual outline of the silver standard (A) and DORIS (B) overlaid on top of the manual segmentation maps. The outlines of DORIS are in better agreement with the manual segmentation than the silver standard, especially along the GM and CSF edges. The mean Dice score of DORIS with respect to the manual segmentation is 0.82, which reveals a good agreement. For the WM, GM, and CSF, DORIS Dice scores are respectively 0.87, 0.70, and 0.89. As for the silver standard, its Dice scores are lower with 0.71 for the mean Dice score and 0.78, 0.64, and 0.72 for the WM, GM, and CSF. Overall, this experience shows a Dice score increase, between the silver standard and DORIS, of 11.5% for the WM, 9.4% for the GM, and 23.6% for the CSF.

3.4. Penthera3T

3.4.1. Incorrectly classified voxels

CSF voxels classified as GM First, we present the number of CSF voxels incorrectly classified as GM. As shown in Figure 7A, while DORIS generates 11,459 incorrectly classified voxels, the silver standard generates 3 times as many incorrectly classified voxels: 30,999. This corresponds to a 171% significant increase of incorrectly classified voxels for the silver standard compared to DORIS (related t-tests $p = 1.13 \times 10^{-31}$).

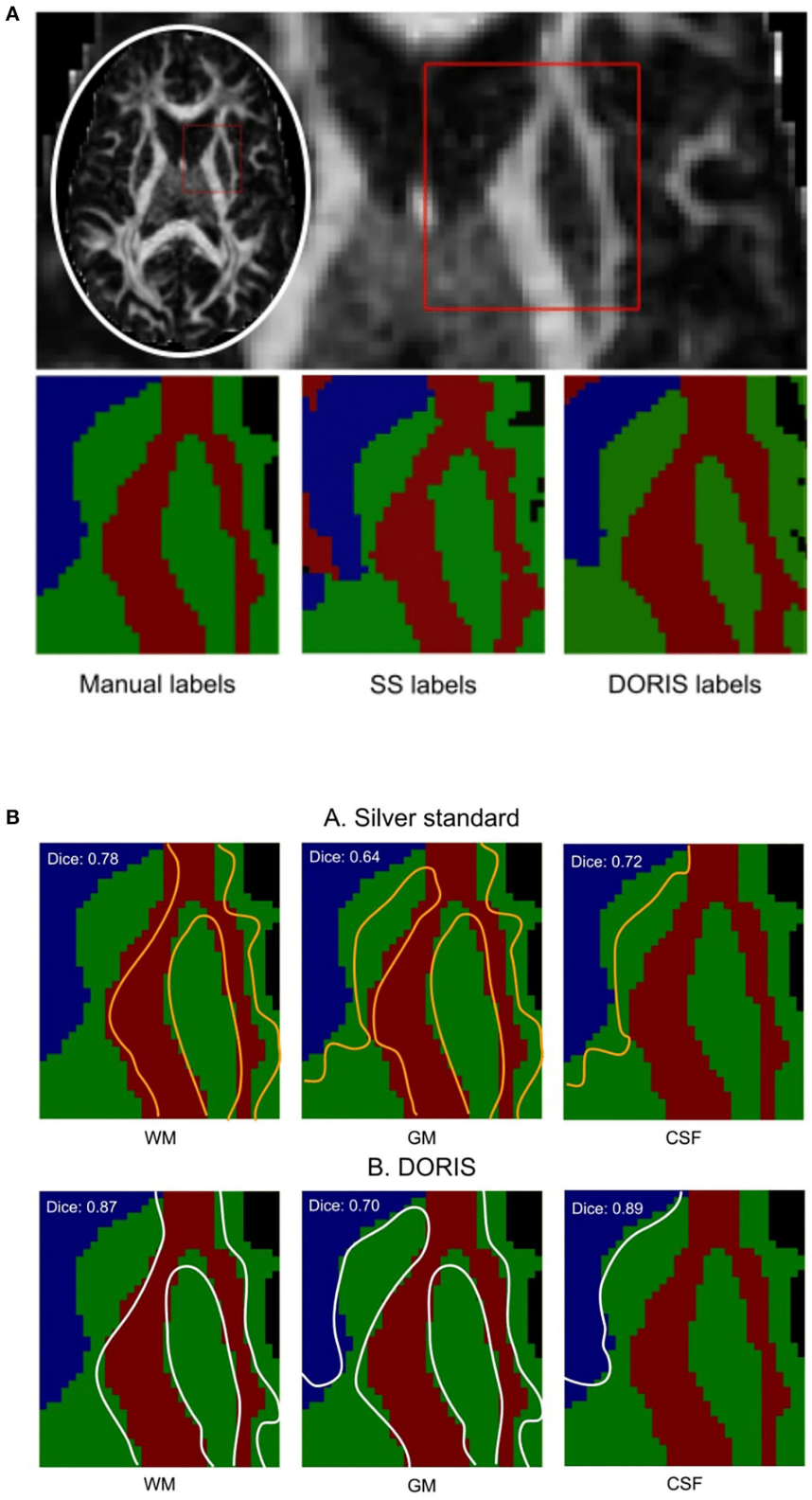


FIGURE 6
Manual segmentation qualitative and quantitative analyses. Based on the Dice score and segmentation alignment, DORIS is more in agreement with the manual segmentation than the silver standard. **(A)** Region of interest outlined in red square for qualitative and quantitative analyses of the three segmentation results (Manual, Silver Standard (SS) and DORIS). **(B)** Comparison between manual segmentation and silver standard **(A)**, and DORIS **(B)**. Tissue edges of the silver standard (orange) and DORIS (white) are overlaid on the manual segmentation for qualitative assessment of the segmentation quality.

FIGURE 7
The number of incorrectly classified WM and CSF voxels. DORIS predicts significantly fewer CSF and WM voxels incorrectly classified as GM than the silver standard. **(A)** CSF voxels incorrectly classified as GM. Significance level of $p < 0.001$ is reached and indicates significantly more incorrectly classified voxels in the silver standard than DORIS. **(B)** WM voxels incorrectly classified as GM. Significance level of $p < 0.001$ is reached and indicates significantly more incorrectly classified voxels in the silver standard than DORIS.

WM voxels classified as GM Figure 7B shows the number of WM voxels incorrectly classified as GM. DORIS gets 8,422 incorrectly classified voxels against twice as many for the silver standard. This corresponds to a significant increase of 98% (related t-tests $p = 2.12 \times 10^{-37}$).

3.4.2. Volume

At first, we compare the volume of each region between DORIS and the silver standard using the testing dataset, ignoring the test-retest for now. For the WM and GM regions, Figure 8A shows a certain volume difference between DORIS and the silver standard. The average WM volume for DORIS is 5,49,175 mm^3 (std: 65,178 mm^3) against 5,09,096 mm^3 (std: 55,475 mm^3) for the silver standard; an increase of 7.9%. As for the GM, DORIS gets and average volume of 6,36,705 mm^3 (std: 53,742 mm^3) against 6,50,521 mm^3 (std: 50,136 mm^3) for the silver standard volume an increase of 2.2%. The WM and GM volume are significantly different between DORIS and the silver standard. Related t-test results for WM and GM volumes between DORIS and the silver standard are available in Supplementary Materials.

Figure 8B shows volumes for the smallest regions (ventricles, putamen, pallidum, hippocampus, caudate amygdala, and thalamus labels) produced by DORIS and the silver standard. The percentage of volume differences varied from 0.5% for the hippocampus to 33.5% for the pallidum (a volume table is available in the Supplementary Materials). Except for the hippocampus and amygdala, related t-tests revealed significantly different volumes in all the smallest labels between DORIS and the silver standard. Related t-test scores are available in Supplementary Materials.

DORIS and Silver Standard in test-retest dataset As seen in Figures 9A,B, there is no statistically significant variation in the DORIS computed volumes in test-retest between the two sessions across all subjects (smallest p-value across labels: $p = 0.107$). As seen in Figures 9C,D, similarly to DORIS, no statistically significant differences are observed for the silver standard volumes (smallest p-value across labels: $p = 0.06$). Full related t-test results for DORIS and the silver standard are available in Supplementary Materials.

3.4.3. Dice

Dice scores for each region are illustrated in Figure 10. The mean Dice score across all the regions is 0.72. The lower Dice score is 0.63 for the caudate nuclei and the highest is 0.81 for the WM. FA-based WM had a Dice score of 0.55 compared to the silver standard WM mask, whereas multi-shell multi-tissue based WM had a better Dice score (0.58) than the silver standard. These two methods serve as a reference for the DORIS WM Dice performance. Overall, FA-based thresholding and multi-shell multi-tissue spherical deconvolution from MRtrix (Supplementary Materials) give a generally lower Dice score than DORIS with the silver standard in the whole WM.

3.5. Impact of DORIS on tractography

Figure 11 shows qualitative results of PFT tracking using DORIS-based segmentation maps. These qualitative results underline that the DORIS-based tractogram does not go through the nuclei and enables the exploration of the WM under the gyri, a well-documented hard-to-track region (Maier-Hein et al.,

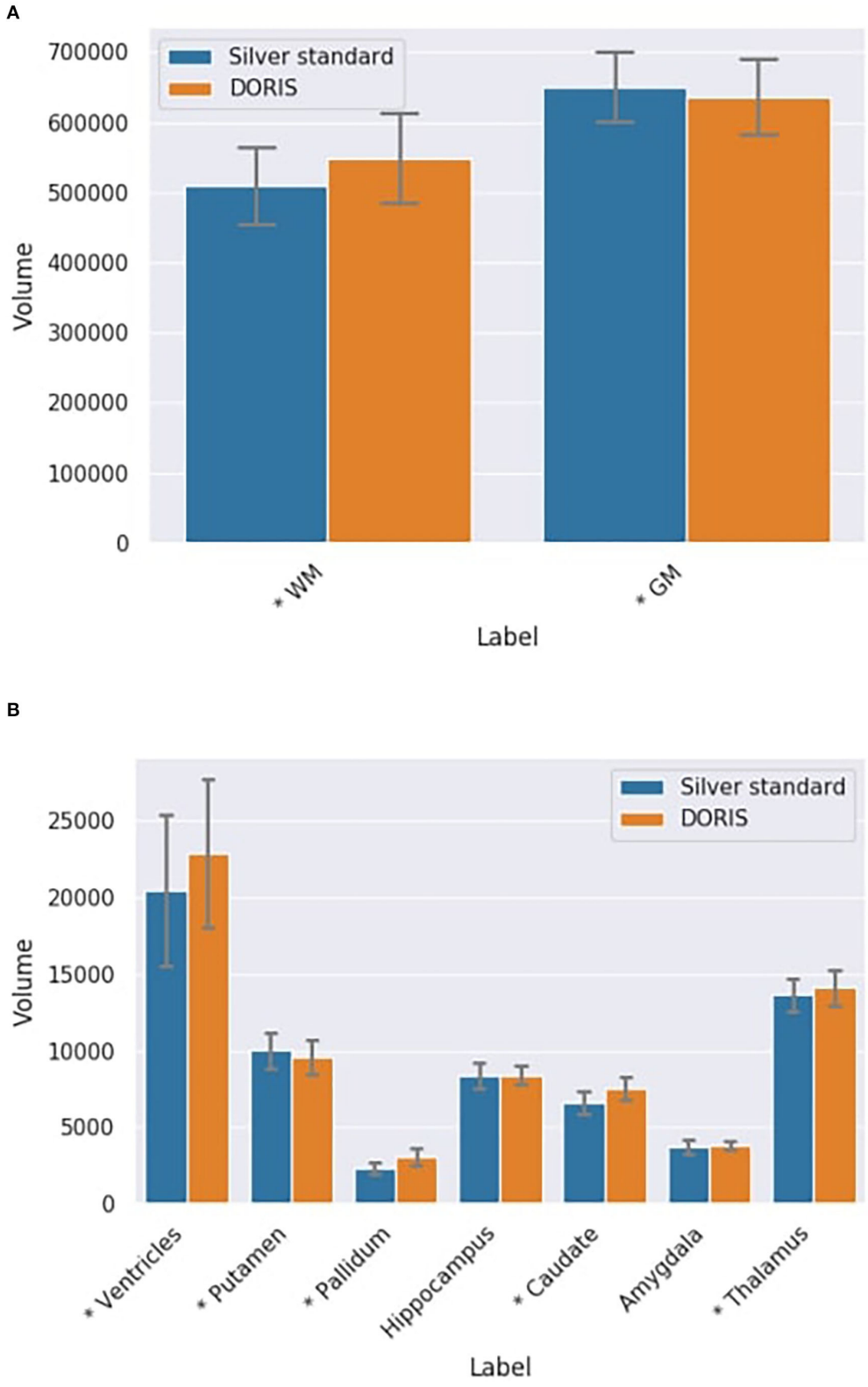


FIGURE 8
Label volumes in mm^3 for DORIS and the silver standard. DORIS and silver standard have label volumes significantly different except for the hippocampus and the amygdala. **(A)** WM and GM volumes in mm^3 (*$p < 0.001$) for DORIS and the silver standard. **(B)** Small label volumes in mm^3 (*$p < 0.001$) for DORIS and the silver standard.

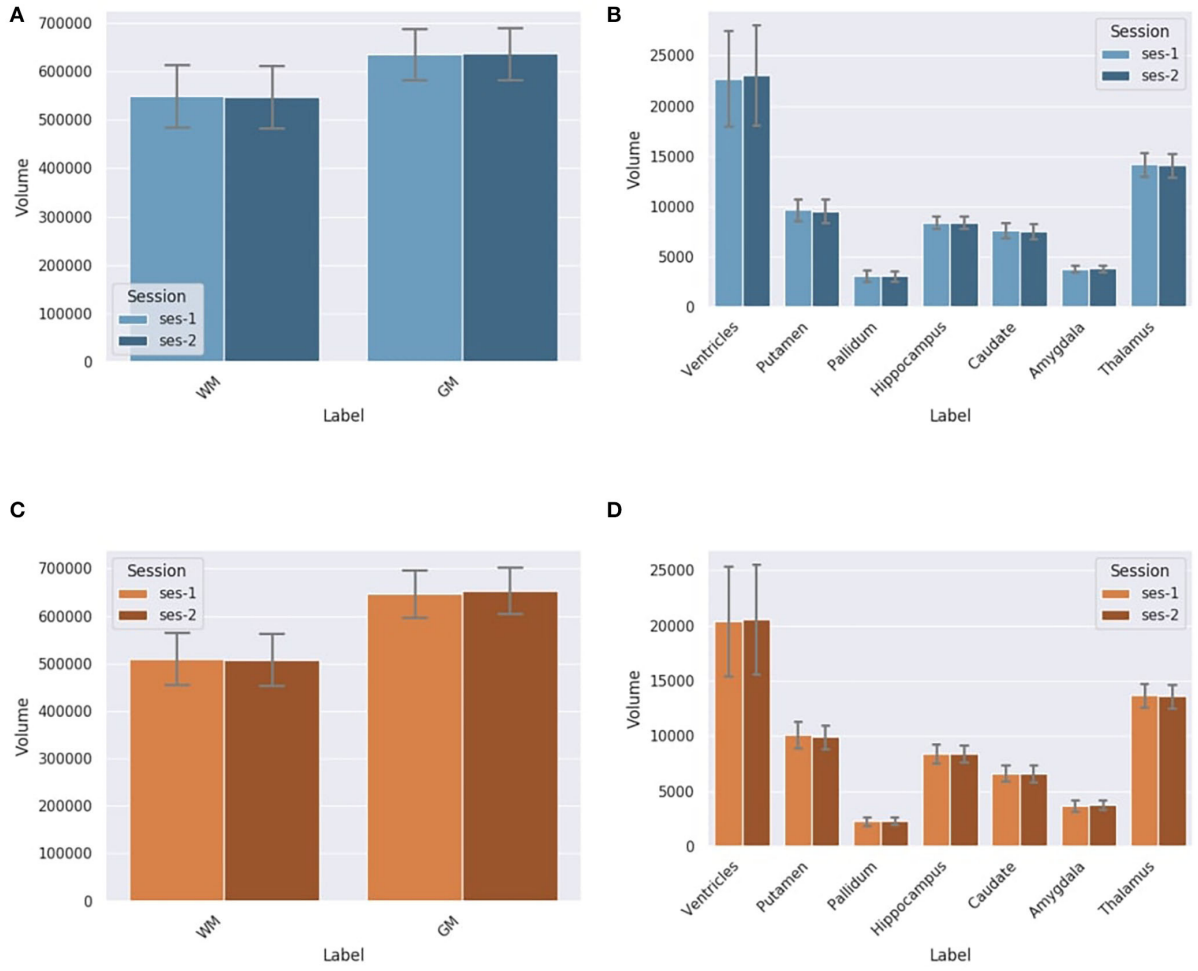


FIGURE 9
Label volumes in mm^3 for DORIS and the silver standard (SS) in test-retest for session 1 and session 2. Label volumes in test-retest are not significantly different for DORIS and the silver standard segmentation. **(A)** WM and GM volumes in mm^3 for DORIS in test-retest for session 1 and session 2. **(B)** Small label volumes in mm^3 for DORIS in test-retest for session 1 and session 2. **(C)** WM and GM volumes in mm^3 for the SS in test-retest for session 1 and session 2. **(D)** Small label volumes in mm^3 for the SS in test-retest for session 1 and session 2.

2017; Mandonnet et al., 2018; St-Onge et al., 2018, 2021). Figure 12 also shows the 6 bundles (SLF, whole CC, IFOF, FX, and AC/PC) extracted from Penthera and ADNI subjects. The 6 bundles were retrieved in the Penthera3T subject. However, the FX and the AC/PC were not retrieved in the ADNI subject, which is not surprising due to the brain atrophy and thinning of WM corridors. For the Penthera3T subject (Figures 12A, 13), the FSL-Fast-based tractogram has no streamline in the FX and got only a fraction of the AC/PC. On the other hand, the DORIS-based tractogram contains all 6 bundles. As for the number of streamlines, while the FSL-fast Tractoflow tractogram has slightly more streamlines in the SLF and IFOF, the DORIS-based CC, FX, and AD/PC are in much better shape. For the ratio and the mean length, DORIS outperforms FSL-Fast for the 6 bundles. As for the ADNI subject (Figure 12B), the number of streamlines, the ratio, and the mean bundle length are higher for the 3 bundles (SLF, whole CC, and IFOF) with the DORIS segmentation map. Figure 13 shows, in the ADNI subject, that parts of the whole CC are not extracted due to the presence of WM aging lesions.

4. Discussion

DORIS presents a good computation time and outperforms some of the state-of-the-art algorithms doing the segmentation at the same time as FastSurfer, in less than 1 min per subject. Moreover, DORIS has a good accuracy compared to a manual segmentation and exhibits good reproducibility performances in test-retest experiments. DORIS also has fewer incorrectly classified voxels than our silver standard extracted from Freesurfer registered into diffusion space.

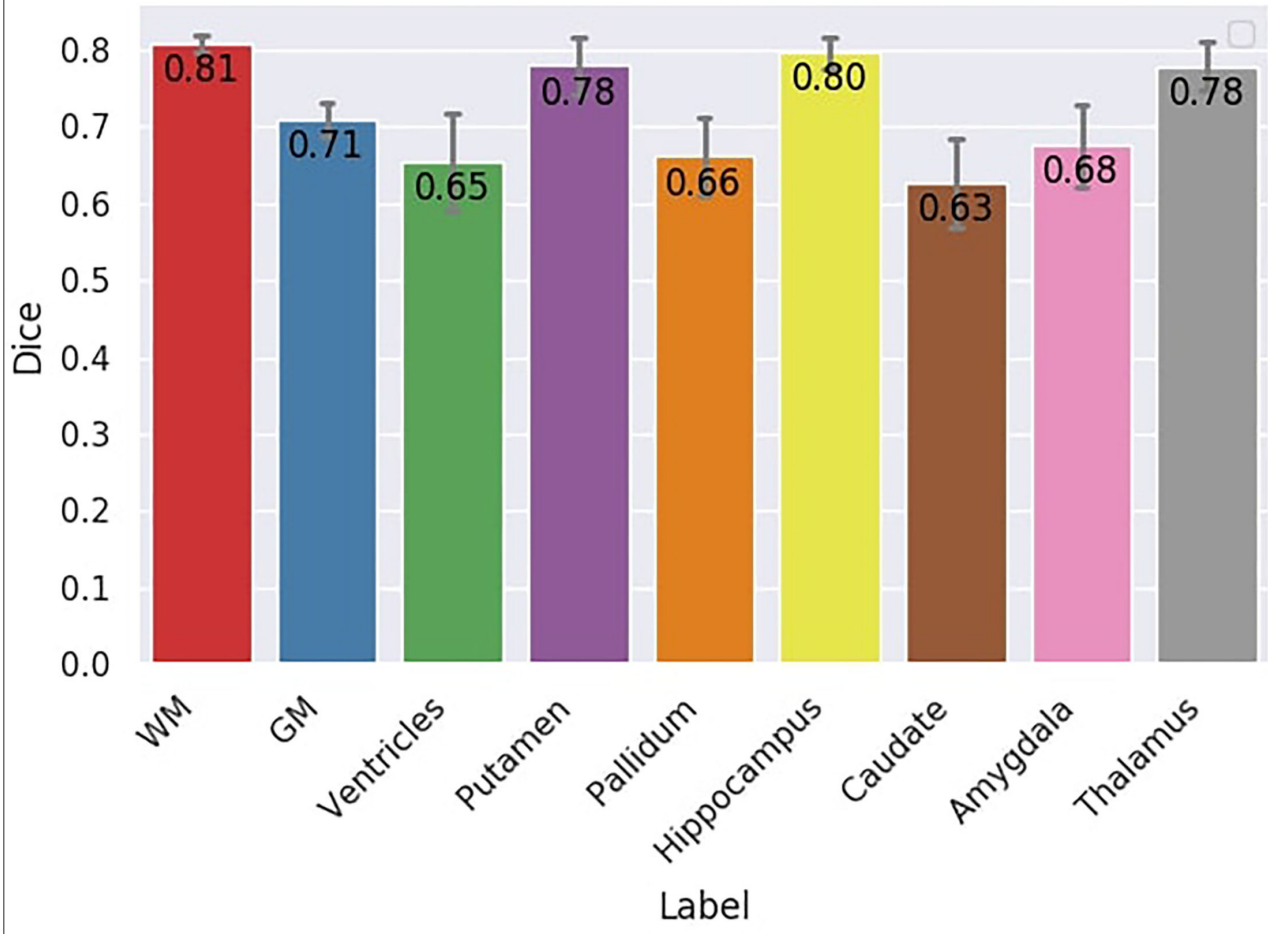


FIGURE 10
Dice score between DORIS and the silver standard. The Dice score between DORIS and the silver standard varied between 0.63 and 0.81.

Finally, empirical results show that DORIS is a good candidate to improve anatomically constrained tractography. With preliminary tractography results, DORIS seems to produce improved anatomically constrained tractograms and, thus, permits the extraction of hard-to-track bundles in tiny corridors like the fornix and the anterior and posterior commissures.

4.1. DORIS and the silver standard

Dice scores from the *Results* section show that DORIS performs well and is more accurate than the silver standard compared to the manual segmentation. In addition, manual segmentation analyses show a good overall performance of DORIS, with Dice scores between 0.70 and 0.89. This confirms that the registration step introduces errors that affect the silver standard (as shown in Section 3.2).

With DORIS, the number of incorrectly classified voxels is significantly less than the silver standard. In addition, region volumes across the full testing dataset varied between DORIS and the silver standard. DORIS produces a larger WM volume than Freesurfer. This volume difference is an advantage for tractography algorithms that have access to more voxels to traverse in hard-to-track regions. Even if volumes of tissue labels are different between the silver standard and DORIS, DORIS can be considered more accurate than the silver standard due to its fewer incorrectly classified voxels. Moreover, as shown in Figure 4, the nuclei shapes were better with DORIS than with our silver standard. This better definition helped the tractography reconstruction of some WM bundles, such as the IFOF, that go near these nuclei.

Overall, results suggest that DORIS outperforms the silver standard. This can be explained by the location of silver standard errors that are different between subjects. Silver standard errors are not the same among the subjects even if it is the same region in the brain (illustrated in Supplementary Materials). Hence, given our large training set of 750 subjects, we hypothesize

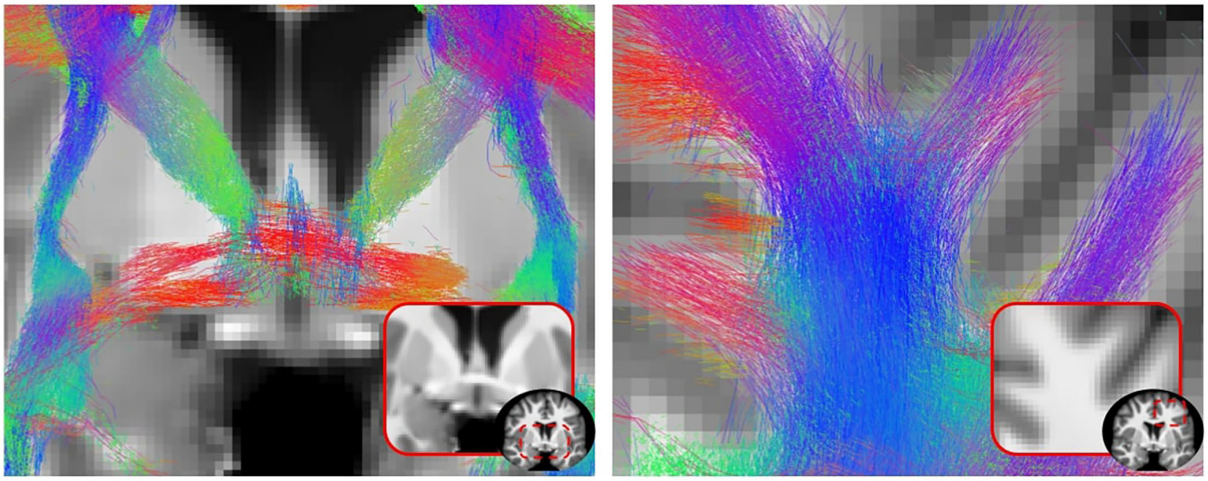


FIGURE 11
2D view of a particle filtering tractography (PFT) performed using DORIS segmentation on a Penthera3T subject. PFT DORIS-based enables to not go through the nuclei with a good gyrus coverage.

that DORIS was, therefore, able to generalize and make fewer segmentation errors than the silver standard.

4.2. Computation time

As shown in Table 2, the registration step adds 45 min of processing time for the state-of-the-art algorithms, a burden that DORIS does not suffer from. However, in a tractography context, the T1 weighted image must also be preprocessed (e.g., denoised, resampled) and then registered in diffusion space (Theaud et al., 2020). Adding other T1 preprocessing steps adds another 1 or 2 h of computation time, depending on the software used. Using DORIS, these 1–2 h of the processing could be optional if the goal is to obtain a whole brain tractogram. However, T1 processing and registration in DWI space remain useful for qualitative or volume analyses.

4.3. Limitations

One major limitation of DORIS is the silver standard used for training and validation. As mentioned in the introduction, a proper ground truth in diffusion space does not exist in the community. Even if manual segmentation has proven useful to validate the DORIS segmentation, it is impossible to do this on an entire brain and multiple subjects. This would require countless hours of one or more neuroanatomists who would have to work on 2 mm resolution images switching between $b = 0$ or FA or T1 images (or others) to perform the manual segmentation. Such a gold standard is, thus, unlikely to ever exist. On the other hand, DWI-based segmentation algorithms do exist but none of them have been developed to enhance tractography and all of them produce a maximum of 3 tissue classes (WM, GM, and CSF) (Li et al., 2006; Liu et al., 2007; Yap et al., 2015; Zhang et al., 2015, 2021; Visser et al., 2016; Ciritsis et al., 2018; Nie et al., 2018; Cheng et al., 2020; Wang et al., 2020; Little and Beaulieu, 2021). In the MRtrix3 software (Tournier et al., 2019), WM, GM, and CSF masks can be extracted from signal fractions obtained from multi-shell multi-tissue fODF (Jeurissen et al., 2014). However, as demonstrated in the results, this method is limited and has not been evaluated on tractography yet. Also, WM, GM, and CSF masks can be extracted with a single-shell 3-tissue fODF version developed by Dhollander and Connelly (2016) called the SS3T-CSD method. However, SS3T-CSD signal fraction maps have not been confronted with ACT. Jeurissen and Szczepankiewicz (2021) and Karan et al. (2021) recently showed that having a tensor-value DWI using linear and spherical encoding helps WM, GM, and CSF signal fraction estimation. This method requires a very specific multi-dimensional b-tensor encoding acquisition scheme and is not easily applicable to actual well-known databases such as HCP, ADNI, and UKBiobank. For all the above reasons, we chose to use our silver standard, based on T1 segmentation registered in DWI space, to develop DORIS. Another limitation is the input channel selection that was only based on the Dice score. As we showed in the results, the Dice score did not fully reflect the segmentation quality and a visual human-based segmentation could be worth investigating in the future. However, as mentioned before, this human-based visual inspection, on a large number of subjects and parameters tested

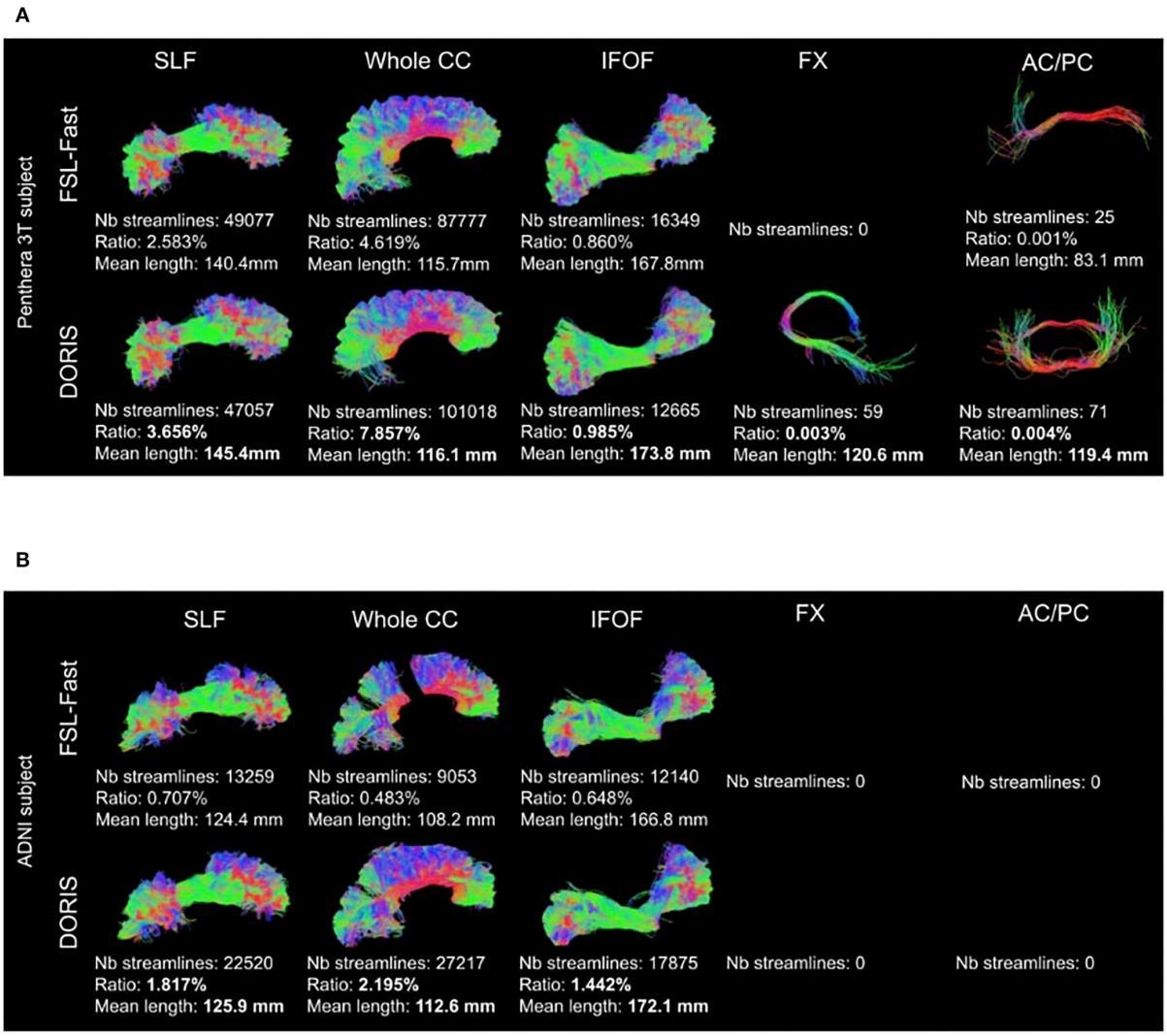


FIGURE 12
Bundles extracted from **(A)** Penthera3T and **(B)** ADNI subjects. For each bundle, we reported the number of streamlines, the ratio, and the mean length of the bundle. For both Penthera3T and ADNI subjects, DORIS based bundles have a better ratio and mean length. For Penthera 3T, DORIS based tractogram reconstructe the Fornux and the commisures correctly. For the ADNI subject, DORIS based tractogram was not impacted by the presence of aging lesions. **(A)** Superior longitudinal fasciculus (SLF), whole corpus callosum (CC), inferior longitudinal occipital fasciculus (IFOF), Fornix (FX), anterior and posterior commissure (AC/PC) extracted from a Penthera3T subject tractogram based on FSL-Fast (top row) and DORIS (bottom row). **(B)** Superior longitudinal fasciculus (SLF), whole corpus callosum (CC), inferior longitudinal occipital fasciculus (IFOF), Fornix (FX), anterior and posterior commissure (AC/PC) extracted from an ADNI subject tractogram based on FSL-Fast (top row) and DORIS (bottom row).

during the hyperparameter optimization, would most likely be infeasible in practice.

4.4. Impact on tractography

We showed that tractograms based on the DORIS segmentation had longer streamlines than the FSL-Fast Tractoflow equivalent. This difference can be explained by the larger WM volume predicted by DORIS and fewer broken streamlines caused by errors in the tissue classes. Indeed, due to partial volume between the CSF and the WM classified as GM by FSL-Fast, PFT tracking can terminate streamlines in this partial volume, even if this ending point is anatomically incorrect. This advantage of less broken fiber generated by DORIS is also highlighted by the percentage of streamlines in the bundle. This percentage is always higher using DORIS than the FSL-Fast Tractoflow equivalence.

DORIS also seems to be robust to aging white matter hyperintensities, as shown in the corpus callosum example of the

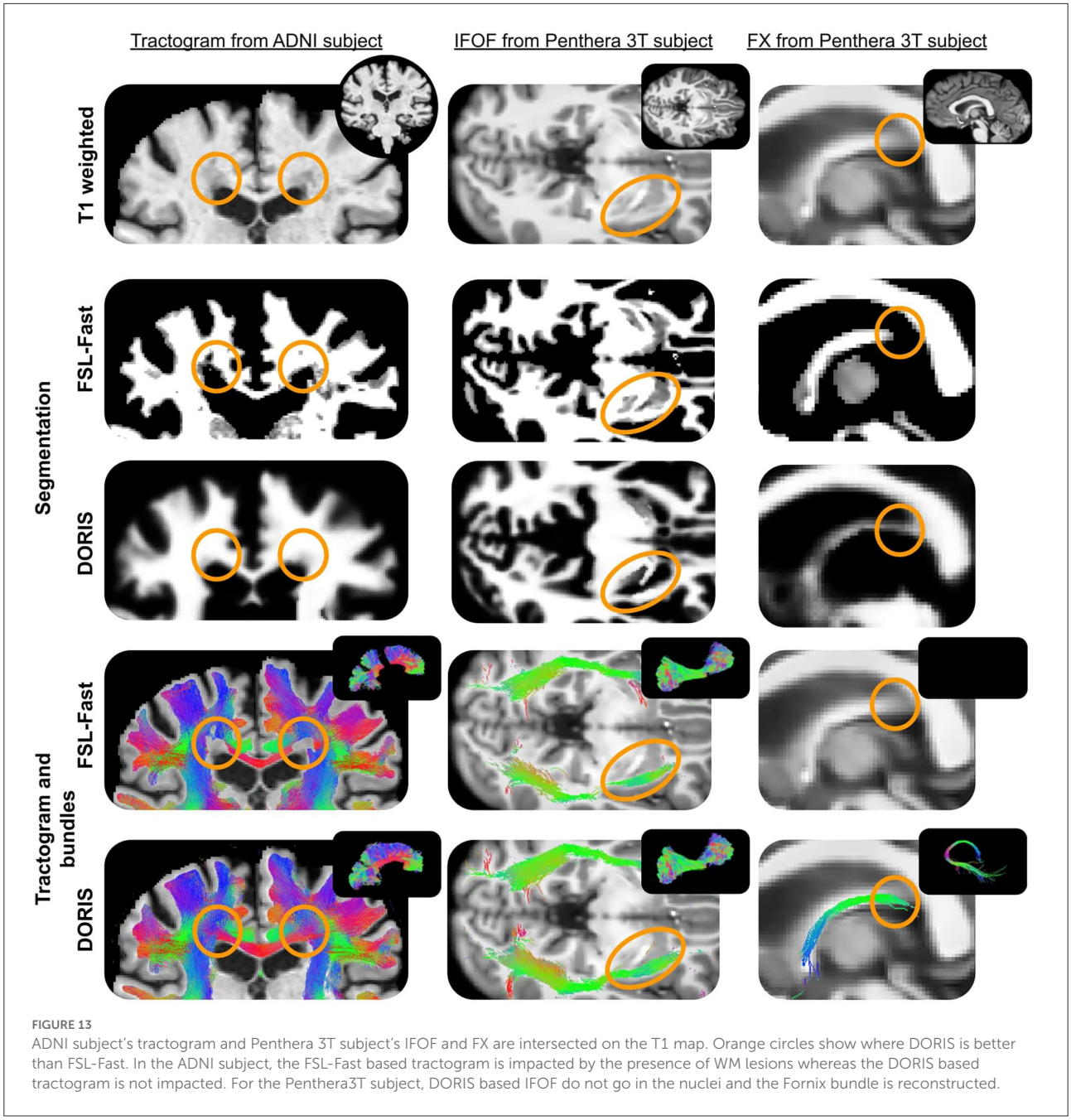


FIGURE 13
ADNI subject's tractogram and Penthera 3T subject's IFOF and FX are intersected on the T1 map. Orange circles show where DORIS is better than FSL-Fast. In the ADNI subject, the FSL-Fast based tractogram is impacted by the presence of WM lesions whereas the DORIS based tractogram is not impacted. For the Penthera3T subject, DORIS based IFOF do not go in the nuclei and the Fornix bundle is reconstructed.

ADNI subject. Indeed, the ADNI subject clearly shows a hole in the whole CC based on FSL-Fast. However, using DORIS, the whole CC is not cut and is biased by aging lesions. This result is possible due to the large age range of the training set of DORIS.

4.5. Future study

Preliminary results of DORIS-based segmentation used in conjunction with tractography highlight the potential of a native diffusion space segmentation algorithm embedded in the tractography process. More analyses will be made in the near future to quantify the impact of DORIS on connectomics and tractometry.

Another improvement that is possible is to explore different models beyond the denseUnet. Indeed, in this study, we showed that the denseUnet model performs well to do the 10-class segmentation task, which can then improve the bundle reconstruction from tractography using the segmentation output. However, a follow-up study is needed to find the best model to do this task. Due to the big training set we presented, the smallest model architecture as an Unet (Ronneberger et al.,

2015) or a fully convolutional network (Long et al., 2015) could work as well as the denseUnet. This study could consist of doing hyperparameter optimizations for each model and then, a quantitative evaluation specifically on the impact of the segmentation on anatomically-constrained tractography reconstructions.

Finally, a new class could be added to segment white matter lesions from aging subjects or subjects with anomalies. This lesion class could permit to do tractometry (Cousineau et al., 2017) analyses directly under the lesion, around it, and in the normal appearing white matter.

5. Conclusion

DORIS is the first algorithm to precisely segment voxels of 10 brain tissue classes purely based on diffusion MRI measures. DORIS can work on both single- and multi-shell diffusion MRI acquisition. This study shows the importance of using a big learning set (750 subjects), with a wide age range (from 22 to 90 years old) and variable image quality, to generalize and not bias the learned model. DORIS is fast, accurate, reproducible, and enhances anatomically-constrained tractography producing longer and less anatomically implausible streamlines.

Data availability statement

Five datasets have been used in this paper, namely Penthera3T, ADNI, PPMI, HCP, and the UKBiobank. Penthera3T, ADNI, PPMI and HCP are publicly accessible and can be found here: https://zenodo.org/record/2602049#.YymIeNVBzJU, https://adni.loni.usc.edu/data-samples/access-data/, https://www.ppmi-info.org/access-data-specimens/download-data and https://www.humanconnectome.org/study/hcp-young-adult/document/1200-subjects-data-release. As for UKBiobank, restrictions apply to the dataset as it cannot be readily available because it is only released by contract. Requests to access the datasets should be directed to https://www.ukbiobank.ac.uk/enable-your-research/apply-for-access.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent from the patients/participants or patients/participants' legal guardian/next of kin was not required to participate in this study in accordance with the national legislation and the institutional requirements.

Author contributions

GT and MDe contributed to conception, design of the study, and wrote the first draft of the manuscript. GT and MDu processed the database. ME contributed to the statistical analysis. MZ, SD-G, and CZ contributed to developing part of the code used in the manuscript. P-MJ, MDe, and ME revised the first versions of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

Funding

This study has been supported by the French government, through the 3IA Côte d'Azur Investments in the Future project managed by the National Research Agency (ANR) with the reference number ANR-19-P3IA-0002. This study has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (ERC Advanced Grant Agreement No 694665: CoBCoM – Computational Brain Connectivity Mapping).

Acknowledgments

Thanks to the Neuroinformatics Chair of the Sherbrooke University who helped push forward neuroscience research. Thanks to Mauro Zucchelli for the rotation invariant feature scripts, Victor Paul-Émile Nozais, and Graham Little for various discussions about DORIS and Philippe Poulin for advice about this research. Thanks to Imeka's GPU#2 for its perseverance. A special thanks to MITACS for funding this work through the Mitacs Accelerate program. This research has been conducted using data from UK Biobank, a major biomedical database. This research uses UK Biobank under the 51659 ID number. Data were provided in part by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University. Data collection and sharing for this project were funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery

Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.;NeuroRx Research; Neurotrack Technologies;Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics.The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California. Data used in the preparation of this article were obtained from the Parkinson's Progression Markers Initiative (PPMI) database (www.ppmi-info.org/data). Forup -to-date information on the study, visit www.ppmi-info.org. PPMI – a public-private partnership - is funded by the Michael J. Fox Foundation for Parkinson's Research and funding partners, including (list the full names of all of the PPMI funding partners found at www.ppmi-info.org/fundingpartners). A special thanks to MITACS for funding this work through the Mitacs Accelerate program.

Conflict of interest

Authors GT, MDu, CZ, P-MJ, and MDe are employed by the company Imeka Solutions Inc.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnimg.2022.917806/full#supplementary-material

References

Avants, B. B., Tustison, N., and Song, G. (2009). Advanced normalization tools (ants). *Insight J.* 2, 1–35. doi: 10.54294/uvnhin

Aydogan, D. B., and Shi, Y. (2020). Parallel transport tractography. *IEEE Trans. Med. Imaging* 40, 635–647. doi: 10.1109/TMI.2020.3034038

Battistella, G., Najdenovska, E., Maeder, P., Ghazaleh, N., Daducci, A., Thiran, J.-P., et al. (2017). Robust thalamic nuclei segmentation method based on local diffusion magnetic resonance properties. *Brain Struct. Function* 222, 2203–2216. doi: 10.1007/s00429-016-1336-4

Bells, S., Cercignani, M., Deoni, S., Assaf, Y., Pasternak, O., Evans, C., et al. (2011). "Tractometry-comprehensive multi-modal quantitative assessment of white matter along specific tracts," in *Proceedings of ISMRM, Vol. 678*. Montréal, QC, 1.

Casey, B., Cannonier, T., Conley, M. I., Cohen, A. O., Barch, D. M., Heitzeg, M. M., et al. (2018). The adolescent brain cognitive development (abcd) study: imaging acquisition across 21 sites. *Dev. Cogn. Neurosci.* 32, 43–54. doi: 10.1016/j.dcn.2018.03.001

Chamberland, M., Raven, E. P., Genc, S., Duffy, K., Descoteaux, M., Parker, G. D., et al. (2019). Dimensionality reduction of diffusion mri measures for improved tractometry of the human brain. *Neuroimage* 200, 89–100. doi: 10.1016/j.neuroimage.2019.06.020

Chamberland, M., Whittingstall, K., Fortin, D., Mathieu, D., and Descoteaux, M. (2014). Real-time multi-peak tractography for instantaneous connectivity display. *Front. Neuroinform.* 8, 59. doi: 10.3389/fninf.2014.00059

Chen, D. Q., Dell'cqua, F., Rokem, A., Garyfallidis, E., Hayes, D. J., Zhong, J., et al. (2019). Diffusion weighted image co-registration: investigation of best practices. *BioRxiv* 864108. doi: 10.1101/864108

Cheng, H., Newman, S., Afzali, M., Fadnavis, S. S., and Garyfallidis, E. (2020). Segmentation of the brain using direction-averaged signal of dwi images. *Magn. Reson. Imaging* 69, 1–7. doi: 10.1016/j.mri.2020.02.010

Ciritsis, A., Boss, A., and Rossi, C. (2018). Automated pixel-wise brain tissue segmentation of diffusion-weighted images via machine learning. *NMR Biomed.* 31, e3931. doi: 10.1002/nbm.3931

Côté, M.-A., Girard, G., Boré, A., Garyfallidis, E., Houde, J.-C., and Descoteaux, M. (2013). Tractometer: towards validation of tractography pipelines. *Med. Image Anal.* 17, 844–857. doi: 10.1016/j.media.2013.03.009

Cousineau, M., Jodoin, P.-M., Garyfallidis, E., Côté, M.-A., Morency, F. C., Rozanski, V., et al. (2017). A test-retest study on parkinson's ppmi dataset yields statistically significant white matter fascicles. *Neuroimage Clin.* 16, 222–233. doi: 10.1016/j.nicl.2017.07.020

Descoteaux, M. (2008). *High angular resolution diffusion MRI: from local estimation to segmentation and tractography* (Ph.D. thesis). Université Nice Sophia Antipolis.

Dhollander, T., and Connelly, A. (2016). "A novel iterative approach to reap the benefits of multi-tissue csd from just single-shell (+ b= 0) diffusion mri data," in *Proceedings of ISMRM, Vol. 24*. Singapore, 3010.

Di Tommaso, P., Chatzou, M., Floden, E. W., Barja, P. P., Palumbo, E., and Notredame, C. (2017). Nextflow enables reproducible computational workflows. *Nat. Biotechnol.* 35, 316–319. doi: 10.1038/nbt.3820

Dice, L. R. (1945). Measures of the amount of ecologic association between species. *Ecology* 26, 297–302. doi: 10.2307/1932409

Dumont, M., Roy, M., Jodoin, P.-M., Morency, F. C., Houde, J.-C., Xie, Z., et al. (2019). Free water in white matter differentiates mci and ad from control subjects. *Front. Aging Neurosci.* 11, 270. doi: 10.3389/fnagi.2019.00270

Farquharson, S., Tournier, J.-D., Calamante, F., Fabinyi, G., Schneider-Kolsky, M., Jackson, G. D., et al. (2013). White matter fiber tractography: why we need to move beyond dti. *J. Neurosurg.* 118, 1367–1377. doi: 10.3171/2013.2.JNS121294

Fischl, B. (2012). Freesurfer. *Neuroimage* 62, 774–781. doi: 10.1016/j.neuroimage.2012.01.021

Garyfallidis, E., Ocegueda, O., Wassermann, D., and Descoteaux, M. (2015). Robust and efficient linear registration of white-matter fascicles in the space of streamlines. *Neuroimage* 117, 124–140. doi: 10.1016/j.neuroimage.2015.05.016

Garza-Villarreal, E. A., Alcala-Lozano, R., Fernandez-Lozano, S., Morelos-Santana, E., Dávalos, A., Villica na, V., et al. (2021). Clinical and functional connectivity outcomes of 5-hz repetitive transcranial magnetic stimulation as an add-on treatment in cocaine use disorder: A double-blind randomized controlled trial. *Biol. Psychiatry* 6, 745–757. doi: 10.1016/j.bpsc.2021.01.003

Girard, G., Whittingstall, K., Deriche, R., and Descoteaux, M. (2014). Towards quantitative connectivity analysis: reducing tractography biases. *Neuroimage* 98, 266–278. doi: 10.1016/j.neuroimage.2014.04.074

Glasser, M. F., Sotiropoulos, S. N., Wilson, J. A., Coalson, T. S., Fischl, B., Andersson, J. L., et al. (2013). The minimal preprocessing pipelines for the human connectome project. *Neuroimage* 80, 105–124. doi: 10.1016/j.neuroimage.2013.04.127

Groen, W. B., Buitelaar, J. K., Van Der Gaag, R. J., and Zwiers, M. P. (2011). Pervasive microstructural abnormalities in autism: a dti study. *J. Psychiatry Neurosci.* 36, 32. doi: 10.1503/jpn.090100

Henschel, L., Conjeti, S., Estrada, S., Diers, K., Fischl, B., and Reuter, M. (2020). Fastsurfer-a fast and accurate deep learning based neuroimaging pipeline. *Neuroimage* 219, 117012. doi: 10.1016/j.neuroimage.2020.117012

Jensen, J. H., and Helpern, J. A. (2010). Mri quantification of non-gaussian water diffusion by kurtosis analysis. *NMR Biomed.* 23, 698–710. doi: 10.1002/nbm.1518

Jeurissen, B., Descoteaux, M., Mori, S., and Leemans, A. (2019). Diffusion mri fiber tractography of the brain. *NMR Biomed.* 32, e3785. doi: 10.1002/nbm.3785

Jeurissen, B., and Szczepankiewicz, F. (2021). Multi-tissue spherical deconvolution of tensor-valued diffusion mri. *Neuroimage* 245, 118717. doi: 10.1016/j.neuroimage.2021.118717

Jeurissen, B., Tournier, J.-D., Dhollander, T., Connelly, A., and Sijbers, J. (2014). Multi-tissue constrained spherical deconvolution for improved analysis of multi-shell diffusion mri data. *Neuroimage* 103, 411–426. doi: 10.1016/j.neuroimage.2014.07.061

Kaku, A., Hegde, C. V., Huang, J., Chung, S., Wang, X., Young, M., et al. (2019). Darts: Denseunet-based automatic rapid tool for brain segmentation. *arXiv preprint arXiv:1911.05567*. doi: 10.48550/arXiv.1911.05567

Karan, P., Reymbaut, A., Gilbert, G., and Descoteaux, M. (2021). Enabling constrained spherical deconvolution and diffusional variance decomposition with tensor-valued diffusion mri. *bioRxiv*. doi: 10.1101/2021.04.07.438845

Koo, B.-B., Hua, N., Choi, C.-H., Ronen, I., Lee, J.-M., and Kim, D.-S. (2009). A framework to analyze partial volume effect on gray matter mean diffusivity measurements. *Neuroimage* 44, 136–144. doi: 10.1016/j.neuroimage.2008.07.064

Kreher, B., Schneider, J., Mader, I., Martin, E., Hennig, J., and Il'Yasov, K. (2005). Multitensor approach for analysis and tracking of complex fiber configurations. *Magn. Reson. Med.* 54, 1216–1225. doi: 10.1002/mrm.20670

Kurtzer, G. M., Sochat, V., and Bauer, M. W. (2017). Singularity: Scientific containers for mobility of compute. *PLoS ONE* 12, e0177459. doi: 10.1371/journal.pone.0177459

Li, H., Liu, T., Young, G., Guo, L., and Wong, S. T. (2006). "Brain tissue segmentation based on dwi/dti data," in *3rd IEEE International Symposium on Biomedical Imaging: Nano to Macro, 2006* (Arlington, VA: IEEE), 57–60.

Little, G., and Beaulieu, C. (2021). Automated cerebral cortex segmentation based solely on diffusion tensor imaging for investigating cortical anisotropy. *Neuroimage* 237, 118105. doi: 10.1016/j.neuroimage.2021.118105

Liu, T., Li, H., Wong, K., Tarokh, A., Guo, L., and Wong, S. T. (2007). Brain tissue segmentation based on dti data. *Neuroimage* 38, 114–123. doi: 10.1016/j.neuroimage.2007.07.002

Long, J., Shelhamer, E., and Darrell, T. (2015). "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Boston, MA: IEEE), 3431–3440.

Maier-Hein, K. H., Neher, P. F., Houde, J.-C., Côté, M.-A., Garyfallidis, E., Zhong, J., et al. (2017). The challenge of mapping the human connectome based on diffusion tractography. *Nat. Commun.* 8, 1–13. doi: 10.1038/s41467-017-01285-x

Mandonnet, E., Sarubbo, S., and Petit, L. (2018). The nomenclature of human white matter association pathways: proposal for a systematic taxonomic anatomical classification. *Front. Neuroanat.* 12, 94. doi: 10.3389/fnana.2018.00094

Nie, D., Wang, L., Adeli, E., Lao, C., Lin, W., and Shen, D. (2018). 3-d fully convolutional networks for multimodal isointense infant brain image segmentation. *IEEE Trans. Cybern.* 49, 1123–1136. doi: 10.1109/TCYB.2018.2797905

Paquette, M., Gilbert, G., and Descoteaux, M. (2019). *Penthera 3T*. Zenodo. doi: 10.5281/zenodo.2602049

Pasternak, O., Sochen, N., Gur, Y., Intrator, N., and Assaf, Y. (2009). Free water elimination and mapping from diffusion mri. *Magn. Reson. Med.* 62, 717–730. doi: 10.1002/mrm.22055

Peled, S., Friman, O., Jolesz, F., and Westin, C.-F. (2006). Geometrically constrained two-tensor model for crossing tracts in dwi. *Magn. Reson. Imaging* 24, 1263–1270. doi: 10.1016/j.mri.2006.07.009

Pierpaoli, C., and Basser, P. J. (1996). Toward a quantitative assessment of diffusion anisotropy. *Magn. Reson. Med.* 36, 893–906. doi: 10.1002/mrm.1910360612

Rheault, F. (2020). *Analyse et reconstruction de faisceaux de la matière blanche* (Ph.D. thesis). Université de Sherbrooke.

Ronneberger, O., Fischer, P., and Brox, T. (2015). "U-net: convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (Springer), 234–241.

Saygin, Z. M., Osher, D. E., Augustinack, J., Fischl, B., and Gabrieli, J. D. (2011). Connectivity-based segmentation of human amygdala nuclei using probabilistic tractography. *Neuroimage* 56, 1353–1361. doi: 10.1016/j.neuroimage.2011.03.006

Smith, R. E., Tournier, J.-D., Calamante, F., and Connelly, A. (2012). Anatomically-constrained tractography: improved diffusion mri streamlines tractography through effective use of anatomical information. *Neuroimage* 62, 1924–1938. doi: 10.1016/j.neuroimage.2012.06.005

Steven, A. J., Zhuo, J., and Melhem, E. R. (2014). Diffusion kurtosis imaging: an emerging technique for evaluating the microstructural environment of the brain. *Am. J. Roentgenol.* 202, W26-W33. doi: 10.2214/AJR.13.11365

St-Onge, E., Al-Sharif, N., Girard, G., Theaud, G., and Descoteaux, M. (2021). Cortical surfaces integration with tractography for structural connectivity analysis. *Brain Connect.* 11, 505–517. doi: 10.1089/brain.2020.0930

St-Onge, E., Daducci, A., Girard, G., and Descoteaux, M. (2018). Surface-enhanced tractography (set). *Neuroimage* 169, 524–539. doi: 10.1016/j.neuroimage.2017.12.036

Theaud, G., Fortin, D., Morency, F., and Descoteaux, M. (2019). "Brain tumors: a challenge for tracking algorithms," in *27th Scientific Meeting of the International Society for Magnetic Resonance in Medicine*, Montréal, QC.

Theaud, G., Houde, J.-C., Boré, A., Rheault, F., Morency, F., and Descoteaux, M. (2020). Tractoflow: a robust, efficient and reproducible diffusion mri pipeline leveraging nextflow singularity. *Neuroimage* 218, 116889. doi: 10.1016/j.neuroimage.2020.116889

Tong, Q., He, H., Gong, T., Li, C., Liang, P., Qian, T., et al. (2020). Multicenter dataset of multi-shell diffusion mri in healthy traveling adults with identical settings. *Scientific Data* 7, 1–7. doi: 10.1038/s41597-020-0493-8

Tournier, J.-D., Calamante, F., and Connelly, A. (2007). Robust determination of the fibre orientation distribution in diffusion mri: non-negativity constrained super-resolved spherical deconvolution. *Neuroimage* 35, 1459–1472. doi: 10.1016/j.neuroimage.2007.02.016

Tournier, J.-D., Smith, R., Raffelt, D., Tabbara, R., Dhollander, T., Pietsch, M., et al. (2019). Mrtrix3: a fast, flexible and open software framework for medical image processing and visualisation. *Neuroimage* 202, 116137. doi: 10.1016/j.neuroimage.2019.116137

Vanderweyen, D. C., Theaud, G., Sidhu, J., Rheault, F., Sarubbo, S., Descoteaux, M., et al. (2020). The role of diffusion tractography in refining glial tumor resection. *Brain Struct. Function* 225, 1413–1436. doi: 10.1007/s00429-020-02056-z

Visser, E., Keuken, M. C., Douaud, G., Gaura, V., Bachoud-Levi, A.-C., Remy, P., et al. (2016). Automatic segmentation of the striatum and globus pallidus using mist: Multimodal image segmentation tool. *Neuroimage* 125, 479–497. doi: 10.1016/j.neuroimage.2015.10.013

Wang, J., Cheng, H., and Newman, S. D. (2020). Sparse representation of dwi images for fully automated brain tissue segmentation. *J. Neurosci. Methods* 343, 108828. doi: 10.1016/j.jneumeth.2020.108828

Wong, K. C., Moradi, M., Tang, H., and Syeda-Mahmood, T. (2018). "3D segmentation with exponential logarithmic loss for highly unbalanced object sizes," in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (Granada: Springer), 612–619.

Yap, P.-T., Zhang, Y., and Shen, D. (2015). "Brain tissue segmentation based on diffusion mri using l 0 sparse-group representation classification," in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (Munich: Springer), 132–139.

Ye, C., Bogovic, J. A., Bazin, P.-L., Prince, J. L., and Ying, S. H. (2012). "Fully automatic segmentation of the dentate nucleus using diffusion weighted images," in *2012 9th IEEE International Symposium on Biomedical Imaging (ISBI)* (Barcelona: IEEE), 1128-1131.

Zhang, F., Breger, A., Cho, K. I. K., Ning, L., Westin, C.-F., O'Donnell, L. J., et al. (2021). Deep learning based segmentation of brain tissue from diffusion mri. *Neuroimage* 233, 117934. doi: 10.1016/j.neuroimage.2021.117934

Zhang, H., Schneider, T., Wheeler-Kingshott, C. A., and Alexander, D. C. (2012). Noddi: practical in vivo neurite orientation dispersion and density imaging of the human brain. *Neuroimage* 61, 1000–1016. doi: 10.1016/j.neuroimage.2012.03.072

Zhang, W., Li, R., Deng, H., Wang, L., Lin, W., Ji, S., et al. (2015). Deep convolutional neural networks for multi-modality isointense infant brain image segmentation. *Neuroimage* 108, 214–224. doi: 10.1016/j.neuroimage.2014.12.061

Zhang, Y., Brady, M., and Smith, S. (2001). Segmentation of brain mr images through a hidden markov random field model and the expectation-maximization algorithm. *IEEE Trans. Med. Imaging* 20, 45–57. doi: 10.1109/42.906424

Zucchelli, M., Deslauriers-Gauthier, S., and Deriche, R. (2020). A computational framework for generating rotation invariant features and its application in diffusion mri. *Med. Image Anal.* 60, 101597. doi: 10.1016/j.media.2019.101597

CITATION

Theaud G, Edde M, Dumont M, Zotti C, Zucchelli M, Deslauriers-Gauthier S, Deriche R, Jodoin P-M and Descoteaux M (2022) DORIS: A diffusion MRI-based 10 tissue class deep learning segmentation algorithm tailored to improve anatomically-constrained tractography. *Front. Neuroimaging* 1:917806. doi: 10.3389/fnimg.2022.917806

frontiers | Frontiers in Human Neuroscience

TYPE Original Research
PUBLISHED 30 September 2022
DOI 10.3389/fnhum.2022.944543

OPEN ACCESS

EDITED BY
Feng Liu,
Tianjin Medical University General Hospital, China

REVIEWED BY
Liang Zhan,
University of Pittsburgh, United States
Xin Zhang,
South China University of Technology, China
Xingwei An,
Tianjin University, China

*CORRESPONDENCE
Shu Zhang
shu.zhang@nwpu.edu.cn

SPECIALTY SECTION
This article was submitted to Brain Imaging and Stimulation, a section of the journal Frontiers in Human Neuroscience

RECEIVED 15 May 2022
ACCEPTED 13 September 2022
PUBLISHED 30 September 2022

CITATION
Yu S, Shi E, Wang R, Zhao S, Liu T, Jiang X and Zhang S (2022) A hybrid learning framework for fine-grained interpretation of brain spatiotemporal patterns during naturalistic functional magnetic resonance imaging. *Front. Hum. Neurosci.* 16:944543. doi: 10.3389/fnhum.2022.944543

A hybrid learning framework for fine-grained interpretation of brain spatiotemporal patterns during naturalistic functional magnetic resonance imaging

Sigang Yu[1], Enze Shi[1], Ruoyang Wang[1], Shijie Zhao[2], Tianming Liu[3], Xi Jiang[4] and Shu Zhang[1]*

[1]Center for Brain and Brain-Inspired Computing Research, Department of Computer Science, Northwestern Polytechnical University, Xi'an, China, [2]School of Automation, Northwestern Polytechnical University, Xi'an, China, [3]Cortical Architecture Imaging and Discovery Lab, Department of Computer Science and Bioimaging Research Center, The University of Georgia, Athens, GA, United States, [4]The Clinical Hospital of Chengdu Brain Science Institute, MOE Key Lab for Neuroinformation, School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu, China

Naturalistic stimuli, including movie, music, and speech, have been increasingly applied in the research of neuroimaging. Relative to a resting-state or single-task state, naturalistic stimuli can evoke more intense brain activities and have been proved to possess higher test–retest reliability, suggesting greater potential to study adaptive human brain function. In the current research, naturalistic functional magnetic resonance imaging (N-fMRI) has been a powerful tool to record brain states under naturalistic stimuli, and many efforts have been devoted to study the high-level semantic features from spatial or temporal representations *via* N-fMRI. However, integrating both spatial and temporal characteristics of brain activities for better interpreting the patterns under naturalistic stimuli is still underexplored. In this work, a novel hybrid learning framework that comprehensively investigates both the spatial (*via* Predictive Model) and the temporal [*via* convolutional neural network (CNN) model] characteristics of the brain is proposed. Specifically, to focus on certain relevant regions from the whole brain, regions of significance (ROS), which contain common spatial activation characteristics across individuals, are selected *via* the Predictive Model. Further, voxels of significance (VOS), whose signals contain significant temporal characteristics under naturalistic stimuli, are interpreted *via* one-dimensional CNN (1D-CNN) model. In this article, our proposed framework is applied onto the N-fMRI data during naturalistic classical/pop/speech audios stimuli. The promising performance is achieved *via* the Predictive Model to differentiate the different audio categories. Especially for distinguishing the classic and speech audios, the accuracy of classification is up to 92%. Moreover, spatial ROS and VOS are effectively obtained. Besides, temporal characteristics of the high-level semantic features are investigated on the frequency domain *via* convolution kernels of 1D-CNN model, and we effectively bridge the "semantic gap"

between high-level semantic features of N-fMRI and low-level acoustic features of naturalistic audios in the frequency domain. Our results provide novel insights on characterizing spatiotemporal patterns of brain activities *via* N-fMRI and effectively explore the high-level semantic features under naturalistic stimuli, which will further benefit the understanding of the brain working mechanism and the advance of naturalistic stimuli clinical application.

KEYWORDS

naturalistic stimuli, spatiotemporal, fMRI, convolutional neural network, Predictive Model

Introduction

Naturalistic stimuli, including movie, music, and speech, are close to real-life experience for human and have been increasingly applied in the research field of neuroimaging (Vanderwal et al., 2019; Wang et al., 2020; Saarimäki, 2021). Relative to a resting-state or single-task state, naturalistic stimuli can evoke more intense brain activities and have been proved to possess higher test–retest reliability, suggesting greater potential to study adaptive human brain function (Martinez-Garcia et al., 2012; Sonkusare et al., 2019; Simony and Chang, 2020). Especially, naturalistic audios, containing rich dynamic auditory stimuli, have been widely adopted as stimulus materials in brain function analysis. To study the functional characteristics of brain activities under naturalistic auditory stimuli, brain imaging, e.g., naturalistic functional magnetic resonance imaging (N-fMRI), has shown great potential to record the brain states and bring more explanations for the brain working mechanism (Saarimäki, 2021).

In recent years, researchers have presented many interesting findings with N-fMRI, including brain function and data reliability. For example, Lahnakoski et al. (2012) explored the detailed relationship between superior temporal sulcus (STS) and social features during watching movie clips with preselected social signals. Wang et al. (2017) proved that the reliability of connectivity and graph theoretical measures of brain networks is significantly improved during naturalistic stimuli over resting-state. Shain et al. (2020) found that human mechanisms generate predictions about upcoming words of a naturalistic sentence mainly by cognitive processes. Although these researches brought a lot of new views, they had a limited contribution to bridge the "semantic gap" (Ozcelik et al., 2022; Raposo et al., 2022). To be specific, "semantic gap" between high-level semantic features and low-level acoustic features is still large, where the former is features contained in N-fMRI with the high-level perception of human, and the latter is features merely extracted from the audios according to dynamics, rhythm, timber, pitch, and tonal (Jiang et al., 2012; Zhao et al., 2014; Lad and Patel, 2021).

In order to interpret the brain conditions and obtain the significant high-level semantic features from fMRI, researchers have made great efforts in terms of spatial and temporal analysis of brain activities. For the spatial analysis, Jiang et al. (2012) developed a computational framework to model the brain imaging space (BIS) high-level features from fMRI and achieved well classification accuracy to differentiate music/speech audios. Zhao et al. (2014) adopted brain network components to decode biologically plausible auditory saliency and effectively decoded the auditory saliency features. Hu et al. (2015) explored the brain regions and functional interactions during semantics categorization based on sparse multinomial logistic regression (SMLR) algorithm. Çelik et al. (2019) proposed a spatially informed voxel-wise modeling (SPIN-VM) technique and achieved well sensitivity in the assessment of fine-grained cortical representations. Zhang et al. (2021) developed a voxel-based state space modeling method and achieved a better understanding of high-dimensional brain activity elicited by complex, open-ended naturalistic tasks. For the temporal analysis, Chen and Hu (2018) applied recurrent neural network (RNN) model based on GRU to capture the sequential information in fMRI data, which can extract GRU patterns and identify subjects. Yan et al. (2019) proposed a multiscale RNN model, using time courses of fMRI independent components directly and achieving promising performance on classification task. Wang et al. (2020) explored functional brain networks underlying auditory saliency *via* a multivariate brain decoding approach. Moreover, much evidence has been proposed that integration of both brain spatial and temporal representations can benefit interpreting the brain state and improve the characterization of the patterns of different brain states. Zhang Y. et al. (2019) proposed a two-stage deep belief network (DBN)–based blind source separation (BSS) method and used it to explore functional brain networks in naturalistic fMRI data. Ren et al. (2021) proposed a volumetric neural architecture search and deep belief network (NAS-DBN) framework to model the N-fMRI volume images, which uncover the hierarchical temporal responses and spatial distributions at multiple scales under naturalistic stimuli. These works extracted

both spatial and temporal features from fMRI by various frameworks and achieved satisfying performance, especially for the deep learning–based framework. However, due to the limited size and considerable individual difference of N-fMRI, an efficient and effective framework, which can overcome the issue of the small dataset and retain reasonable classification and prediction performance, is largely needed. Therefore, we aim to study the fine-grained interpretation of brain spatiotemporal patterns on the voxel level, which can benefit to bridge the "semantic gap."

To design a suitable framework, choosing appropriate learning algorithms is the key issue for differentiating the brain states and identifying meaningful biomarkers. With the rapid development of deep learning and prediction approaches, they have shown the great power of explanation and generalization in human neuroscience (Rosenberg et al., 2018; Wen et al., 2018). Traditionally, studies on fMRI mainly focused on finding correlations between brain and stimulus. The generated models can work well on trained individual data, but they are difficult to be effective on brand-new individuals. To solve this problem, in recent years, Predictive Models have been proposed to build a generalized neuroimaging model, which is designed to predict individual observation and generalize to new individual data (Scheinost et al., 2019). Besides, convolutional neural network (CNN) has been widely used in neuroimaging, especially for the fMRI, for it can extract the local features which contribute to classification, prediction, and identification (Anwar et al., 2018). Therefore, in this work, we propose a novel hybrid learning framework that comprehensively studies both brain spatial (by Predictive Model) and temporal (by CNN model) characteristics *via* N-fMRI. Specifically, the whole brain functional spatial patterns under stimuli of three different audio categories (classic/pop/speech) are obtained *via* N-fMRI and serve as input to Predictive Model to achieve audio classification and significant spatial feature identification. Then, regions of significance (ROS) were generated based on the widely used AAL90 atlas (Rolls et al., 2020), and the significant spatial features identified by Predictive Model. Next, voxel-level signals in the ROS were extracted and fed to the one-dimensional CNN (1D-CNN) model to explore the voxels that consistently contribute to audio classification from the temporal perspective and investigate the characteristics of their temporal representations.

Our experimental results show that the proposed framework can achieve promising performance in naturalistic audio classification. Especially for distinguishing the classic and speech audios, the accuracy of classification is up to 92%. Based on the proposed framework, we can effectively characterize the spatiotemporal features of brain functional activity under N-fMRI. Brain regions in the temporary lobe and other regions related to audio are successfully identified, and signals of voxels in these regions are interpreted on spatiotemporal features. In addition, through qualitative temporal analysis of brain high-level semantic features and low-level acoustic features, we alleviate the semantic gaps from the view of frequency domain.

The key characteristics of this work can be summarized as three perspectives. First, an effective and hybrid learning computational framework, which integrates both brain spatial and temporal analysis, is proposed to study the brain states and high-level semantic features from N-fMRI. Second, ROS with spatial activation features across individuals and voxels of significance (VOS) with temporal characteristics were identified at a finer spatiotemporal scale. Third, "semantic gap" between high-level semantic features and low-level acoustic features is attempted to alleviate *via* our proposed framework.

The remainder of this article is organized as follows: Section "Materials and methods" briefly introduces the overall methods, including data acquisition, high-level feature extraction, Predictive Model, and 1D-CNN model. Section "Results" provides the experimental results, and section "Conclusion" concludes this article.

Materials and methods

Overview

Aiming at the fine-grained interpretation of spatiotemporal patterns of brain activities, workflow of the proposed framework is shown in **Figure 1**, which mainly includes two stages. At Stage 1, significant spatial features of brain activities corresponding to different naturalistic audio categories are identified *via* machine learning (ML) based Predictive Model, with the input of brain activation patterns extracted from N-fMRI data. Then, ROS are selected based on the AAL atlas (Rolls et al., 2020) and significant spatial features identified by Predictive Model, thus we focus on certain relevant regions instead of the whole brain. At Stage 2, voxel-level signals in the ROS are extracted and fed to 1D-CNN model to explore the voxels that consistently contribute to audio classification from the temporal perspective and investigate the characteristics of their temporal representations. Furthermore, the relationship between high-level semantic features and low-level acoustic features is explored.

Data preprocessing and pattern extraction

Naturalistic audio data description

Three typical categories of music/speech: classical music (CLA), pop music (POP), and speech (SPE) are adopted as the stimulus materials in the dataset. For each category, 7 representative excerpts are selected for fMRI scanning. The excerpts are taken from legal copies of MP3 compressed audio files to ensure a variety of different recording qualities. Each

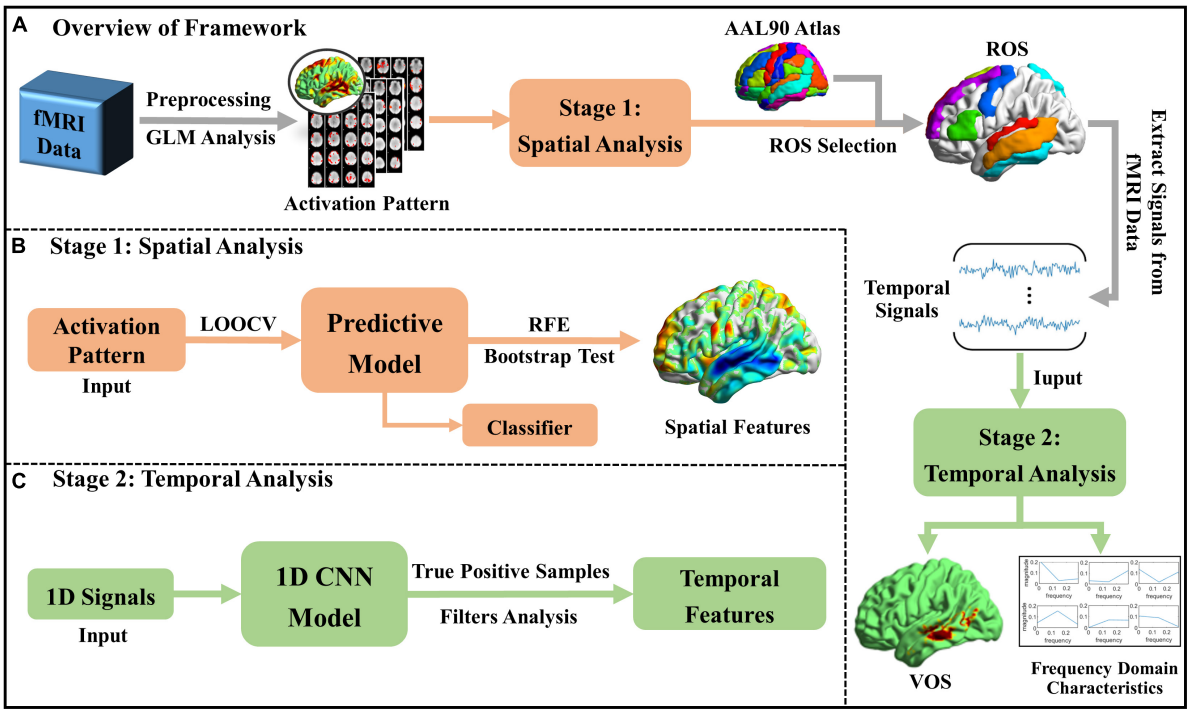


FIGURE 1
Workflow of the proposed framework. **(A)** The overview of the whole framework; **(B)** spatial analysis at Stage 1; and **(C)** temporal analysis at Stage 2.

excerpt is 90 s, totally making 3 $s \times 7$ s $\times$ 90 $s = 31.5$ min of audio data for each subject. Before the N-fMRI collection, we randomly compose the audio excerpts into two parts, each of which is about 15 min long. Between the two parts, participants could have a break with flexible duration.

Naturalistic functional magnetic resonance imaging data description

Collection of N-fMRI was conducted at the University of Georgia (UGA) under UGA Institutional Review Board (IRB) approval. Since fMRI scanning is costly and time-consuming, seven young healthy college students were recruited for this study and scanned in a GE 3T Sigma MRI system (GE Healthcare, Milwaukee, WI, USA) using an 8-channel head coil. The TR of this scan is set to be 1.5 s, and each excerpt collects 60 volumes of fMRI. Other parameters are as follows: 64 $\times$ 64 matrix size, 4 mm slice thickness, 220 mm Field of View (FOV), 30 axis slices, TE $=$ 25 ms, and ASSET $=$ 2. Note that for each subject and each excerpt, we obtained the 4D N-fMRI data with the size of (60, 91, 109, 91), where 60 is the number of volumes, and (91, 109, 91) is the size of one volume image.

Naturalistic functional magnetic resonance imaging data preprocessing

The preprocessing pipeline included motion correction, slice-timing correction, smoothing, registration, and normalization (Churchill et al., 2017) using FMRIB Software Library (FSL) (Jenkinson et al., 2012). To perform group-wise analysis on participants with a biological difference, all N-fMRI data are registered from individual space into the standard Montreal Neurological Institute (MNI) 152 standard space (Evans et al., 2012) by the FMRIB's Linear Image Registration Tool (FLIRT) (Jenkinson and Smith, 2001).

Activation pattern extraction

To extract brain activation patterns, we adopted fMRI Expert Analysis Tool (FEAT) to conduct first-level General Linear Modeling (GLM) analysis by modeling task design corresponding to each 90 s naturalistic stimuli of each excerpt of an audio category (Woolrich et al., 2001). The preprocessed 4D N-fMRI data are adopted as input, and the 3D activation pattern with a size of (91, 109, 91) is obtained by GLM analysis. In total, there are 49 brain patterns (7 subjects $\times$ 7 excerpts) for one category of audios, which would serve as input to the Predictive Model.

Spatial analysis with Predictive Model

To perform spatial analysis across individuals, Predictive Model, which is designed to predict individual observation and generalize to new individual data, is needed. Recently,

Kohoutová et al. (2020) proposed an ML-based framework that consists of model-, feature-, and biology-level assessments to provide complementary results that support the interpretability of Predictive Model. Motivated by this framework, we adopt Predictive Model to analyze the N-fMRI data from a spatial perspective, striving to obtain the common significant spatial features across individuals. To focus on certain relevant regions from the whole brain, we select ROS based on AAL90 atlas (Rolls et al., 2020) and the significant spatial features identified by Predictive Model, so that we can better interpret those significant spatial features in the further analysis.

Procedure of Predictive Model

The major procedure of the Predictive Model for N-fMRI analysis includes extraction of brain activation patterns (input features of Predictive Model), selection of appropriate learning algorithms, training and classification, verification of classification performance, and identification of features related to classification. To establish an effective Predictive model for brain state differentiation during naturalistic stimuli, three key points are necessary to be determined: (1) Input Features; (2) Learning Algorithm; and (3) Training Strategy.

Input features

Activation patterns of the whole brain are adopted as the key features of input to the Predictive Model. In order to reduce the size of fMRI activation patterns, we adopted the commonly used MNI 152 T1 brain mask to extract the voxels within brain space in this work. MNI 152 T1 brain mask contains voxels with values of either "1" or "0." Voxels with value "1" are located inside the brain, and voxels with value "0" are outside the brain. Since the fMRI data we adopted have been registered to the MNI 152 standard space, targeted voxels could be located and extracted by the mask. As a result, we reduced the voxel number of an activation pattern from about 910 thousand ($91 \times 109 \times 91$, whole space of fMRI image) to about 220 thousand (whole space of brain areas).

Learning algorithm

For the learning algorithms, multiple regression, LASSO regression, support vector machine (SVM), and support vector regression (SVR) are potential algorithms for classification and prediction (Ray, 2019). Among them, SVM is one of the most popular machine learning algorithms in current neuroimaging literature and has been proven to show promising performances on a small dataset (Scheinost et al., 2019). Considering the relatively small N-fMRI data in this work, SVM has a great advantage over other algorithms to obtain promising performance. Therefore, we adopted SVM to build the computing kernel in the Predictive Model. To fit the requirements of SVM, each sample with 220 thousand voxels was flattened into a one-dimensional vector. In this model, we selected the widely used linear kernel, which maps low-dimensional non-linear data to higher-dimensional space.

Training strategy

For the training strategy, we evaluated the performance of the Predictive Model on N-fMRI data by measuring the prediction accuracy. Here, we adopted Leave-One-Out-Cross-Validation (LOOCV), a commonly used method in machine learning, as the training strategy. For each binary classification task, the total number of available patterns is 98, corresponding to the 98 music excerpts (7 subjects $\times$ 7 excerpts $\times$ 2 categories). Depending on the LOOCV strategy, we treated each of the patterns as a sample, and samples of each participant, in turn, served as a testing set and samples of the rest participants as the training set. This process was totally repeated 7 (number of participants) times to reduce the size of the features with different combinations of training and testing data.

Moreover, to quantitatively illustrate the classification performance of the Predictive Model, we adopt the widely used statistical tool, receiver operating characteristic (ROC) curve (Obuchowski and Bullen, 2018), to describe the accuracy of the Predictive Model. As shown in Eq. 1, Sensitivity (S_N) and Specificity (S_P) are a conditional probability of correctly identifying the true samples and false samples, respectively. The learned classifier with a low false positive rate and a high true positive rate suggests promising classification performance:

$$\begin{cases} S_N = \frac{TP}{TP+FN} \\ S_P = \frac{TN}{TN+FP} \end{cases} \quad (1)$$

where TP, FN, FP, and TN are the true positive samples, false negative samples, false positive samples, and true negative samples in predicted outcomes, respectively.

Comparison with baseline technologies

To adequately compare the classification performance, we set the comparison experiments from two perspectives: learning algorithms and feature extraction methods. For different learning algorithms, we chose methods of Support Vector Regression (SVR) and Principal Components Regression (PCR) as the baseline comparison. For different feature extraction methods, we chose the acoustic features (according to dynamics, rhythm, timber, pitch, and tonal) of the naturalistic audios and brain connectivity of N-fMRI as the baseline comparison, which was presented by Jiang et al. (2012).

Spatial features identification

To identify the significant spatial features from the classifier, we choose two commonly used methods: bootstrap tests (Erlikhman and Caplovitz, 2017) and recursive feature elimination (RFE) (Craddock et al., 2009), to identify features that are most related to the pattern classification. Bootstrap tests identify features that make stable contributions to prediction across participants. Features with stable weights in the classifier will be identified as significant features. RFE fits a model and removes the weakest features until the specified number of features is reached. During the iterative training procedure of

Predictive Model, features corresponding to its lowest weights are eliminated from the training dataset until the optimal number of features is left. Theoretically, both bootstrap tests and RFE can identify significant spatial features in the Predictive Model.

Regions of significance selection

Considering that the significant spatial features identified by Predictive Model are distributed over multiple brain regions, in this section, we selected regions with the most significant spatial features as ROS for future temporal analysis based on AAL atlas (Rolls et al., 2020). Specifically, for each AAL region, we calculated the ratio of the voxel number of spatial features to the voxel number of all voxels, which are defined as r_i in Eq. 2, and then we selected ROS with a larger r_i from the whole brain:

$$r_i = \frac{n}{m} \tag{2}$$

where i is the index of AAL brain region, r is the feature voxel proportion of a region, n is the voxel number of the significant spatial features in a region, and m is the total voxel number of a region.

Temporal analysis with one-dimensional convolutional neural network

Although Predictive Model can conduct audio classification and provide related significant spatial features, the full representations of brain patterns are still to be explored. Based on the selected ROS, we further analyzed the temporal features of N-fMRI in this section. Considering the advantage of the CNN model to automatically extract the local features from the input samples, we further introduced 1D-CNN model for temporal analysis into the framework. In the 1D-CNN model, N-fMRI signals of voxels in ROS are extracted as the input with label of audio categorizes, the local temporal features that related to audio categorizes are effectively characterized.

One-dimensional convolutional neural network model architecture and experiment design

Architecture of one-dimensional convolutional neural network

In the previous research on resting state fMRI (rs-fMRI), four convolutional layers were adopted in 1D-CNN model and achieved promising classification performance (Zhang S. et al., 2019). Considering that signals in N-fMRI have richer features than those in rs-fMRI and single-task fMRI (Saarimäki, 2021), we set up five convolutional layers to train the model and extract hidden features from the input signals with a length of 60, as shown in **Figure 2**. Each of the last four convolutional layers is followed by a max-pooling layer, by which the output feature maps will be reduced to half the size of the input after passing through. At the end of 1D-CNN model is the dense layer (Softmax), whose input is the feature maps of input signals and output is the predictive result. In addition, architecture parameters are empirically set as follows: the number of the convolution filters is 40, 32, 24, 16, and 8 in each convolutional layer, respectively. The size of all kernels is 7, which is suitable for analyzing the features of signals with a length of 60. To keep the size of output feature maps the same as the input feature maps in each convolutional layer, we set the same padding and stride equal to 1. Moreover, the optimizer is SGD, activation function is Relu, batchsize is 64, and learning rate is 1×10^{-4}. The loss function is selected as "binary cross entropy."

Experiment design

In this section, we conducted experiments using N-fMRI data from CLA and SPE tasks as examples and used 98 excerpts (7 subjects × 7 excerpts × 2 categories) of available N-fMRI data. We applied each ROS as a mask to extract voxel temporal signals from all the 98 excerpts of the preprocessed N-fMRI data and thus we possessed 98 excerpts of signals for each voxel in the studied ROS. Further, we adopted LOOCV training strategy in the 1D-CNN model. Specifically, for voxels of the studied ROS, we successively selected the signals of one excerpt as test samples and signals of other 97 excerpts as training samples. In this stage, we conducted the experiment on each of ROS separately, i.e., the number of separate parallel experiment is equal to the number of ROS. We finally obtain 98 sets of predicted results for voxels in each of the ROS.

Voxels of significance identification

In the 1D-CNN model, voxel signals with significant temporal characteristics could be classified with a higher classification accuracy, which can contribute to audio classification. Based on the LOOCV strategy described in section "One-dimensional convolutional neural network model architecture and experiment design," there will be 98 runs for each voxel in each of the ROS. Here, we counted the true positive predicted results among the 98 sets of results for each voxel. If the ratio of number of true positive results to the 98 sets of results reaches a certain proportion (60% was chosen empirically) for a voxel, we consider that signals within this voxel are significantly related to the audio category and defined the voxel as VOS. Thus, VOS are identified in this stage. By studying the area that consists of VOS, we can explore how brain function works under naturalistic auditory stimuli at a finer spatial scale.

Exploration of relationship between high-level semantic features and low-level acoustic features *via* one-dimensional convolutional neural network model

"Semantic gap" between high-level semantic features (obtained from N-fMRI data) and low-level acoustic features

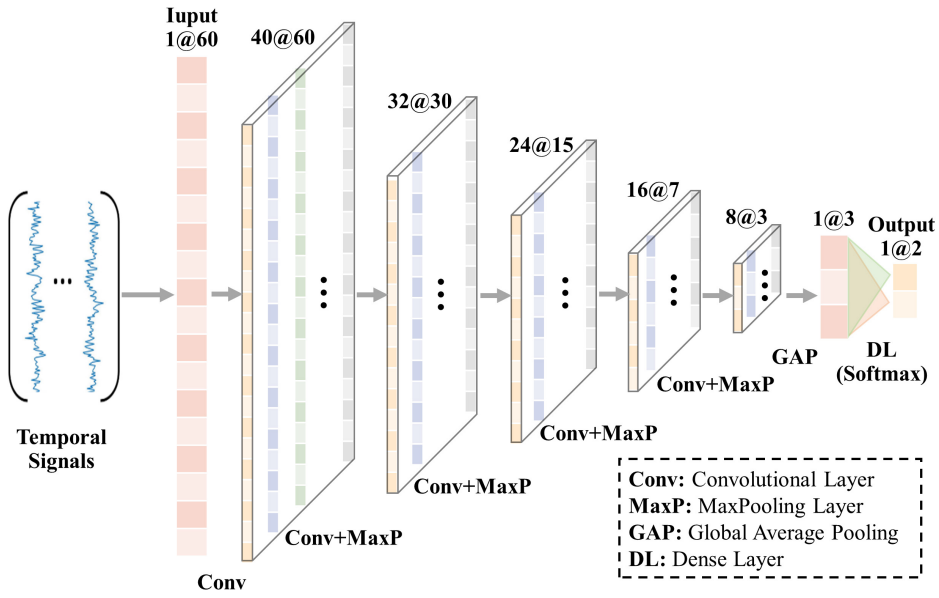


FIGURE 2
Architecture of 1D-CNN model in our framework.

(obtained from the audios) is a key problem in the N-fMRI study. Researchers have discovered more and more frequency-specific biological interpretations from fMRI (Yuen et al., 2019). In this section, the proposed framework may bring new insights into alleviating the "semantic gap" by analyzing frequency domain features of N-fMRI and audios.

In the 1D-CNN model, the convolution kernels are constantly trained and optimized in the process of convolution with input signals. The significant features of voxel signals that contribute to the classification of audio categories are embedded in those kernels (Huang et al., 2018). In this article, we selected the convolution kernels in the last convolutional layer to explore which category they are related to. To be specific, feature maps generated by convolution kernels in the last convolutional layer connect to the dense layer, thus each of these convolution kernels can be mapped to a pair of weights, which are related to the prediction result of audio categories (Lin et al., 2013). Therefore, we extracted the values of convolution kernels in the last convolutional layer and the weights in the dense layer, and then established a one-to-one relationship between kernels and categories (Liu et al., 2019; Zhang S. et al., 2019). By the analysis of convolution kernels, we can generate the temporal features of N-fMRI related to different audio categories in each region separately.

To identify and interpret the difference in temporal features of fMRI, the learned convolution kernels could be transferred into the frequency domain to explore the frequency characteristics of fMRI (Liu et al., 2019; Zhang S. et al., 2019; Jiang et al., 2020). Since the size of the convolution kernels at the temporal domain is 7, the frequency domain includes three points, i.e., about half size of the kernel, as shown in **Figure 3**. To efficiently and comprehensively study the shape of the features from the frequency domain, six typical types are observed and defined as L, L', V, V', Γ, and Γ', which are shown in **Figure 3**. Moreover, the six shapes are mathematically defined in Eq. 3:

$$\begin{cases} L: X(1) > mean(X), X(2) < mean(X), X(3) < mean(X) \\ L': X(1) > mean(X), X(2) < mean(X), X(3) < mean(X) \\ V: X(1) > mean(X), X(2) < mean(X), X(3) < mean(X) \\ V': X(1) > mean(X), X(2) < mean(X), X(3) < mean(X) \\ \Gamma: X(1) > mean(X), X(2) < mean(X), X(3) < mean(X) \\ \Gamma': X(1) > mean(X), X(2) < mean(X), X(3) < mean(X) \end{cases} \tag{3}$$

where $X(n)$ represents the nth value in the X-shape frequency domain array.

For the low-level acoustic features, we calculated two typical acoustic features: mel frequency cepstral coefficients (MFCCs) (Grama and Rusu, 2017) and Spectral Centroid (Prasetio et al., 2021).

Mel frequency cepstral coefficients are coefficients that collectively make up a Mel-Frequency Cepstrum (MFC), which have been widely used in automatic speech and speaker recognition. Specifically, the human ears in listening act like filters, which are better at identifying small changes in audio at lower frequencies (blow 1000 Hz) but not good at higher frequencies (higher than 1000 Hz). Mel-scale is a scale that relates the human perceived frequency to the actually measured frequency f. The formula to convert frequency f to Mel-scale $Mel(f)$ is illustrated in Eq. 4 (Gupta et al., 2013). MFCCs are

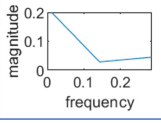
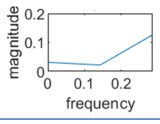
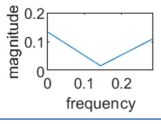
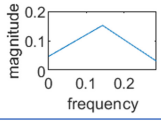
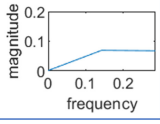
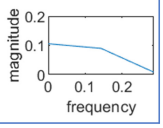
Shape	L	L'	V	V'	Γ	Γ'
Magnitude	Mainly at low frequency	Mainly at high frequency	Mainly at the low and high frequency	Mainly at the medium frequency	Mainly at medium and high frequency	Mainly at low and medium frequency
Example						

FIGURE 3
The six typical shapes of frequency domain features of convolution kernels.

commonly derived as follows: (1) Take the Fourier transform of the audio signal; (2) Map the powers of the spectrum obtained above onto the Mel-scale; (3) Take the logs of the powers at each of the Mel frequencies; (4) Take the discrete cosine transform of the list of Mel log powers; and (5) The MFCCs are the amplitudes of the resulting spectrum:

$$Mel(f) = 2595 \times lg(1 + \frac{f}{700}) \quad (4)$$

where f is the actual measured frequency, $Mel(f)$ is the Mel-scale, 2,595 and 700 are the commonly used constants in Mel-scale formula.

Spectral Centroid is one of the important physical parameters describing the properties of timber, which indicates where the centroid of the spectrum is located (Prasetio et al., 2021). Generally, the audios with dark and deep quality tend to have more low-frequency components and relatively low Spectral Centroid, while the audios with bright and cheerful quality mostly concentrate on high frequency and relatively high Spectral Centroid. It is calculated from the Fourier transform frequency and amplitude information, as defined in Eq. 5:

$$Centroid = \frac{\sum_{n=0}^{N-1} f(n)\, x(n)}{\sum_{n=0}^{N-1} x(n)} \quad (5)$$

Where $x(n)$ represents the weighted frequency value or magnitude of bin n, and $f(n)$ represents the center frequency of n.

Results

In our experiment, the Predictive Model is configured based on CanlabCore toolbox, which is available at https://github.com/canlab/CanlabCore. The 1D-CNN model is built based on Keras (a deep learning application programming interface), which runs on top of the machine learning platform TensorFlow 2.6.0 (Géron, 2019). The computing environment is a server with NVIDIA Geforce GTX 3090 with 24 GB GPU.

Characterizing spatial patterns *via* Predictive Model

Visualization of input features for Predictive Model

As the key features fed into the Predictive Model, activation patterns obtained by GLM analysis are visualized in **Figure 4**. As shown in **Figure 4**, the activation regions of POP and SPE are relatively similar, especially in the Temporal lobe which contains significant activations. However, the activation patterns of CLA are significantly different from that of POP and SPE, which include parts of the Frontal lobe and Parietal lobe. In terms of audio components, CLA audios only have instrumental sounds without vocals, whereas POP and SPE both have vocals. It may reveal the attention mechanism of Temporal lobe, which focuses more on the vocals than other sounds.

Classification performance of Predictive Model

We visualized the ROC curves for the classification results of the Predictive Model for classification performance assessment. **Figure 5** shows the ROC curve of three binary classification tasks, in which the accuracy of CLA/POP, POP/SPE, and CLA/SPE is 81, 62, and 92%, respectively. The results support

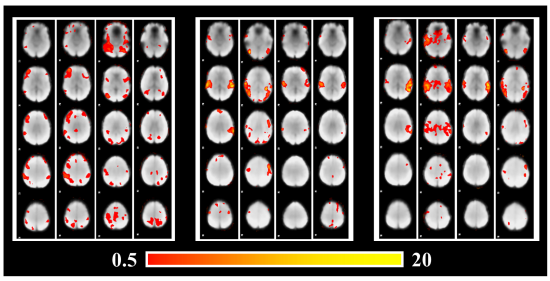


FIGURE 4
Brain activation patterns under stimuli of different audio categories.

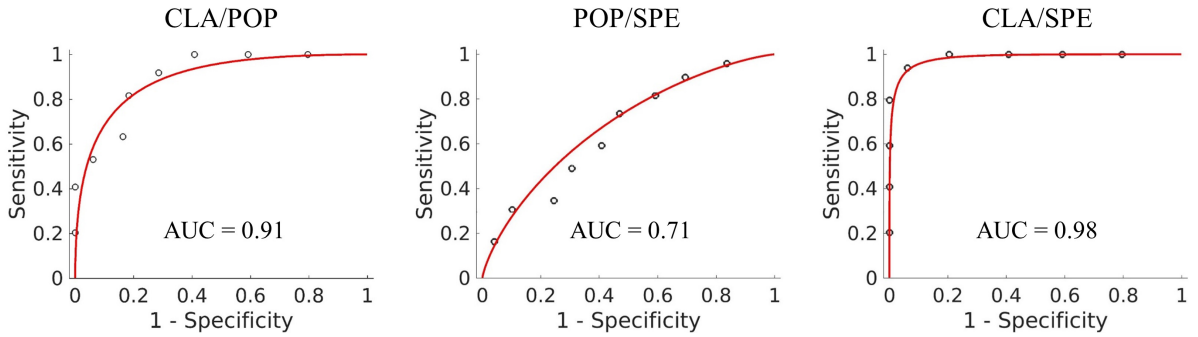


FIGURE 5
ROC curve of three binary classification tasks.

TABLE 1 Comparison of different learning algorithms and feature extraction methods.

Comparison	Algorithm/method	Accuracy		
		CLA/POP	POP/SPE	CLA/SPE
This article	Whole brain activation/SVM	**83 ± 3.8%**	**65 ± 4.8%**	91 ± 2.9%
Different learning algorithms	Whole brain activation/SVR	73 ± 4.5%	56 ± 5.0%	**96 ± 2.0%**
	Whole brain activation/PCR	71 ± 4.6%	55 ± 5.0%	94 ± 2.4%
Different feature extraction methods	Functional connectivity matrix/SVM	70.5%	63.5%	75%
Jiang et al., 2012	Acoustic features/SVM	52.5%	49.5%	63.5%

our justification in section "Visualization of input features for Predictive Model" that activation patterns of POP and SPE are relatively tough to be distinguished and activation patterns of CLA are easier to be identified. Besides, the sensitivity and specificity are 80 and 82% for CLA/POP task, 73 and 51% for POP/SPE task, and 96 and 88% for CLA/SPE task, respectively. In POP/SPE task, the specificity of 51% indicates that the probability of misjudging negative samples is high, which is the main reason for low classification accuracy. Moreover, in order to further explore the difference among three audio categories, we provide the results of three classifications, for details please refer to the **Supplementary material**.

Besides, the results of the baseline comparison are provided in **Table 1**. When comparing with different learning algorithms, SVM performs best in CLA/POP and POP/SPE classification task and generates well performance in CLA/SPE classification task. When comparing with different feature extraction methods, features of whole brain activation perform better than both functional connectivity and acoustic features. Overall, SVM performs the most robustness classification performance and features of whole brain activation generate the best classification performance among these feature extraction methods.

Exploration of the significant spatial features

Significant spatial features identified from the bootstrap tests and RFE procedure were visualized in **Figure 6** *via* the BrainNet Viewer (Xia et al., 2013). As shown in **Figure 6**, the positive activation (color "red") represents the meaningful features of the former category, the negative activation (color "blue") represents the meaningful features of the latter category. It is worth noting that both the positive and negative features contribute to the classification.

Figure 6 demonstrates that the spatial maps of bootstrap tests and RFE results are similar, consistently demonstrating the effectiveness of the Predictive Model for spatial feature identification. Although the results of bootstrap tests cover more brain areas, the activation features from RFE provide a stronger contrast between two categories with less area. Moreover, for the classification of CLA/POP and CLA/SPE, activation features are more concentrated on specific locations, while positive and negative activation areas stay away from each other. However, it is quite clear that positive and negative activation areas overlap for the classification of POP/SPE, indicating that even at the semantic level, POP and SPE are still hard to be distinguished.

Identification of regions of significance

Based on the obtained bootstrap tests and RFE results, brain regions with the most positive and negative activation features are counted and then visualized in **Figure 7** *via* the BrainNet Viewer (Xia et al., 2013). From **Figure 7**, we found that some brain areas are related to certain classical categories, like "superior frontal gyrus medial," as shown in red circle. We also found that some areas are related to more than one category.

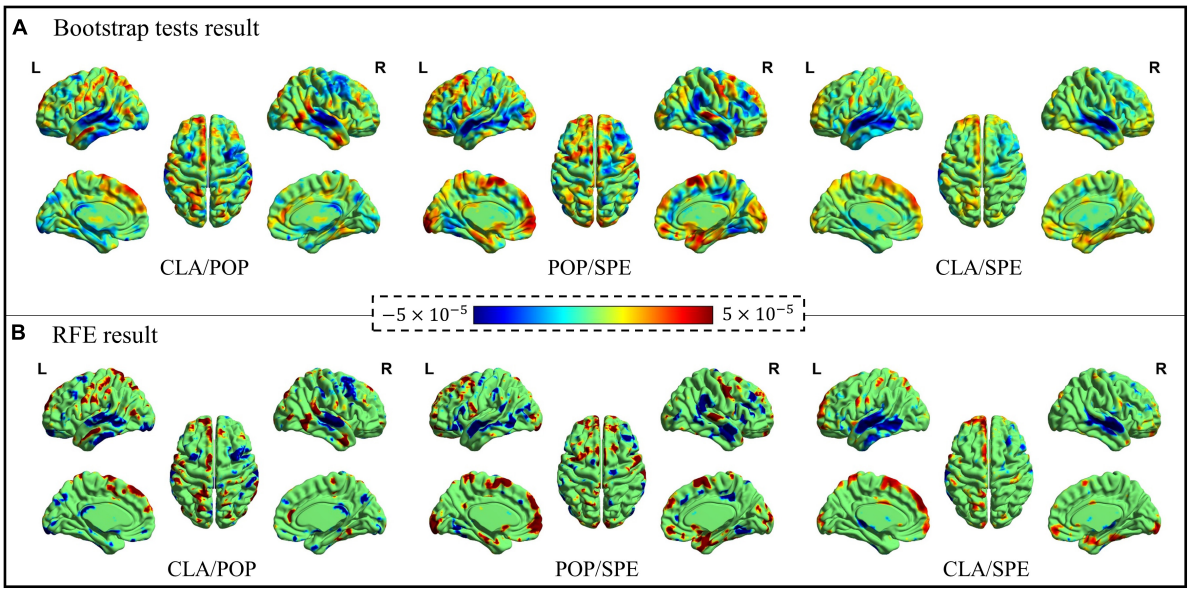


FIGURE 6
Bootstrap tests and RFE results of significant spatial features. **(A)** Significant features learned by Bootstrap tests, and **(B)** significant features learned by RFE.

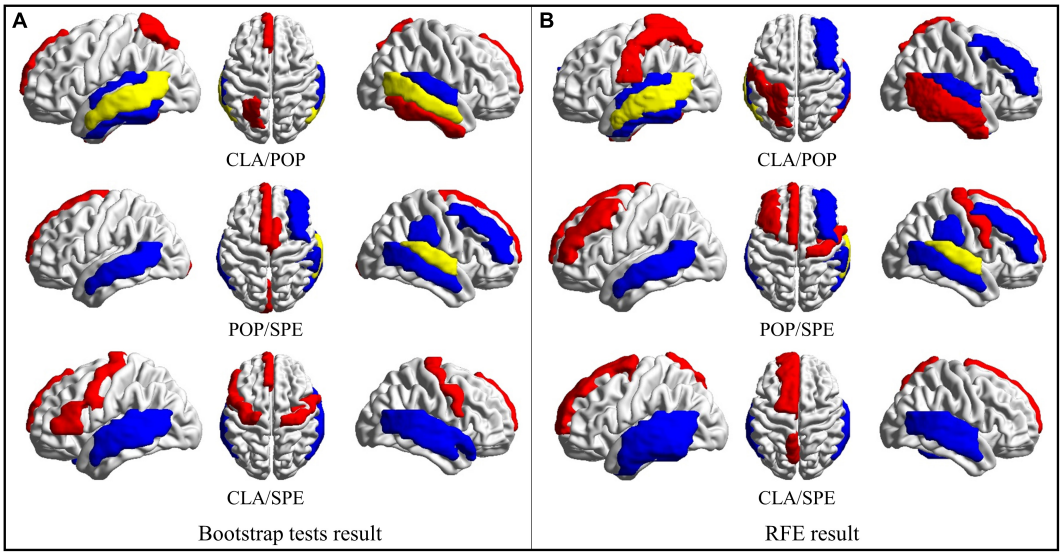


FIGURE 7
Regions of brain where activation is significant. **(A)** ROS selected by Bootstrap test results; **(B)** ROS selected by RFE results. (Red represents areas with positive activation, blue represents areas with negative activation, and yellow represents areas with both positive and negative activation).

For example, both pop and speech categories are related to the "superior temporal gyrus," which is the yellow region shown in POP/SPE result in Figure 7. Both classical and speech categories cause "middle temporal gyrus" to activate, which is the yellow region shown in CLA/POP result in Figure 7. These results reveal an interesting phenomenon that some certain brain regions are consistently related to the specific category of audios, while some are merely activated by a certain category of audios.

In order to further study the spatial and temporal characteristics of ROS, CLA/SPE categories, which provide the best classification performance in Predictive Model, are selected as the test bed for further analysis. For each audio category and each significant feature identification method, we selected

TABLE 2 Regions with most spatial features identified by Bootstrap tests and RFE.

Audio category	Positive (CLA)		Negative (SPE)	
Methods	**Bootstrap**	**RFE**	**Bootstrap**	**RFE**
Division	SFGmed.L	SFGmed.L	MTG.L	MTG.L
	IFGtriang.L	SMA.L	STG.R	STG.R
	PreCG.L	PCUN.L	MTG.R	MTG.R
	PHG.R	SFGdor.L	STG.L	STG.L
	PreCG.R	PHG.R	TPOsup.R	ITG.L
	FFG.R	FFG.R	ITG.L	CAU.R
Intersection	SFGmed.L		MTG.L	
	PHG.R		STG.R	
	FFG.R		MTG.R	
			STG.L	
			ITG.L	

the top 6 regions with the most significant features (total 24 regions). Table 2 shows the regions with the most significant features selected by Bootstrap tests and RFE in Predictive Model. From Table 2 we can see that regions selected by Bootstrap tests and RFE both have difference and intersection. To identify significant spatial features as much as possible and further select comprehensive brain regions for the temporal analysis in the following section, we combined the regions (total 16 regions) learned by the two methods and visualized the 16 ROS associated with CLA/SPE, as shown in Figure 8. To simplify the expression, we refer to each region in the format of the abbreviations in the following sections. For the full name and abbreviations of regions involved in this article, please refer to the Supplementary material.

Characterizing spatiotemporal patterns *via* one-dimensional convolutional neural network

Effectiveness of one-dimensional convolutional neural network model *via* model evaluation experiment

To verify the validity of 1D-CNN model, MTG. L brain region was randomly chosen to perform the validation

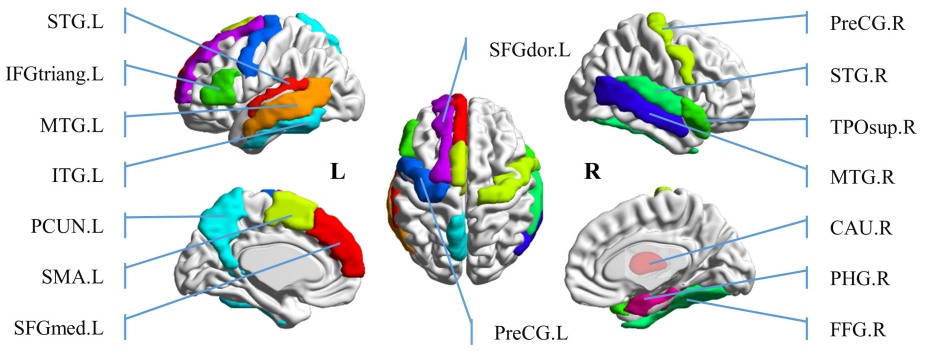


FIGURE 8
Visualization of top 16 brain regions containing the most significant spatial features.

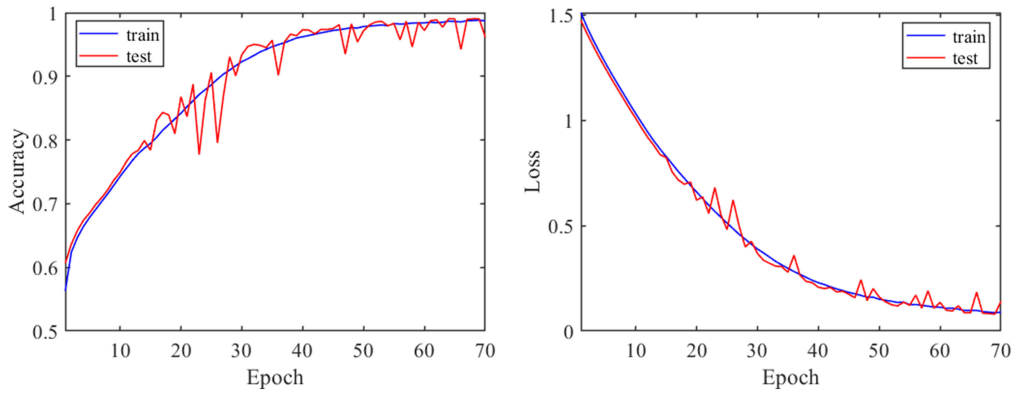


FIGURE 9
Performance of 1D-CNN model *via* model evaluation experiment.

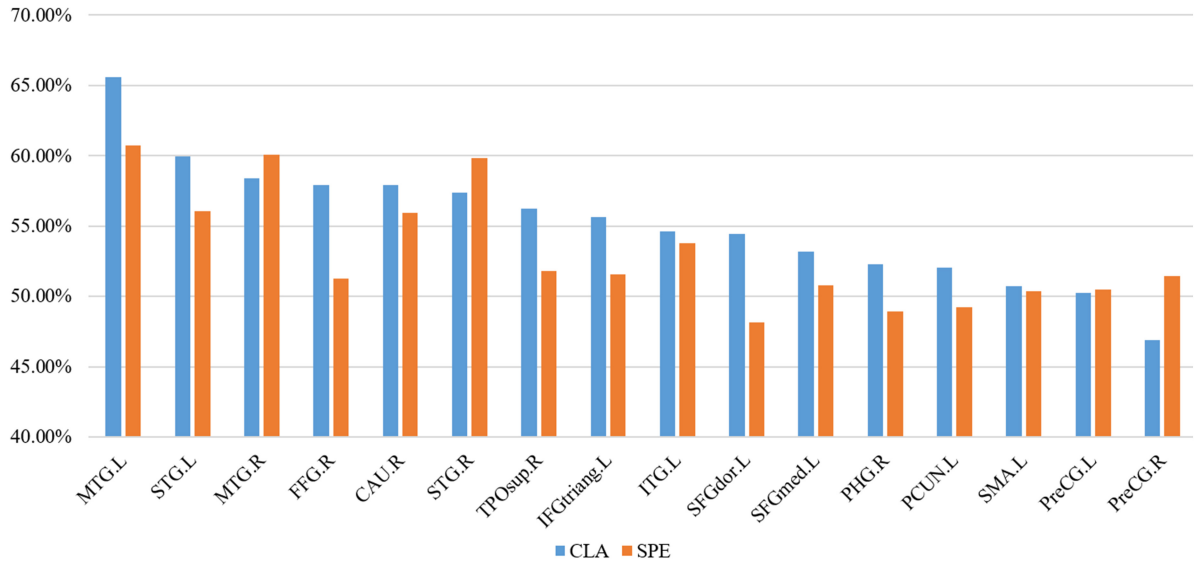


FIGURE 10
The proportion of VOS in each ROS.

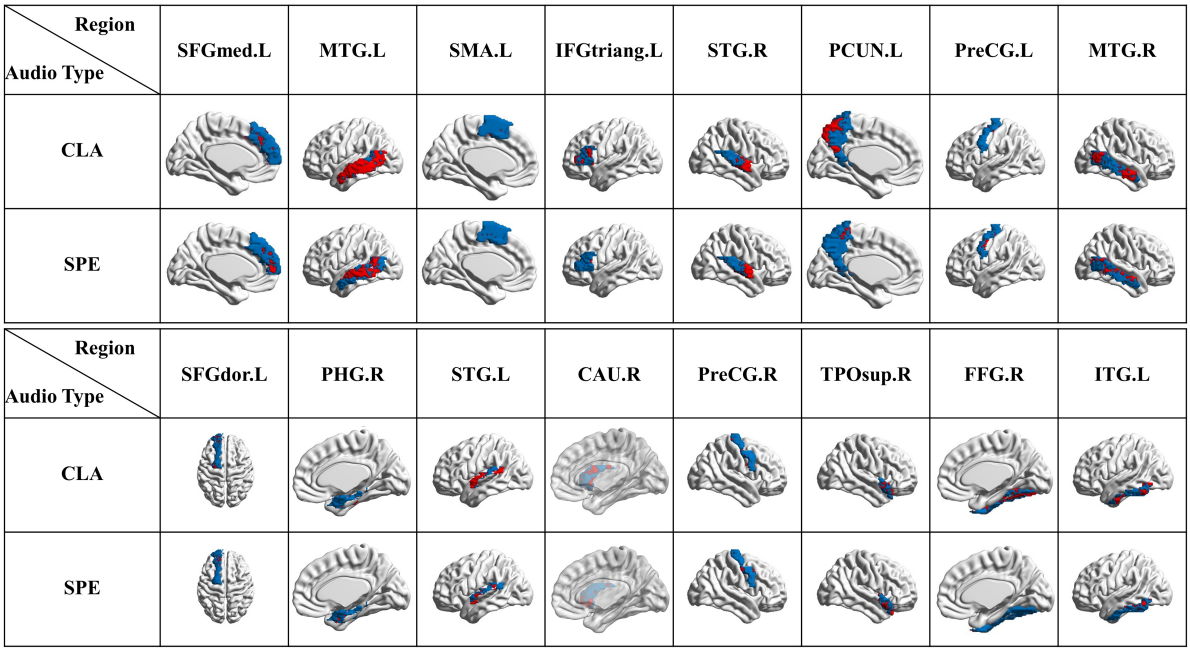


FIGURE 11
Visualization of VOS in each ROS (blue represents each single brain region, and red represents the collection of VOS in each region).

experiment. N-FMRI signals from MTG. L brain region of all participants during CLA and SPE tasks were extracted as the dataset, which was then divided into the training set and testing set followed by the ratio of 4:1. Both the training and testing classification performances are shown in **Figure 9**. It can be seen that with the increase in epoch, the accuracy of training and testing both increased up to about 97%, and the corresponding loss value decreased and converged from 1.5 to less than 0.1. The performance indicates that the proposed 1D-CNN model has satisfying performance in classifying CLA and SPE signals.

Interpreting the voxels of significance

For the 16 selected ROS, we calculated the proportion of VOS number to the total voxel number in both CLA and SPE

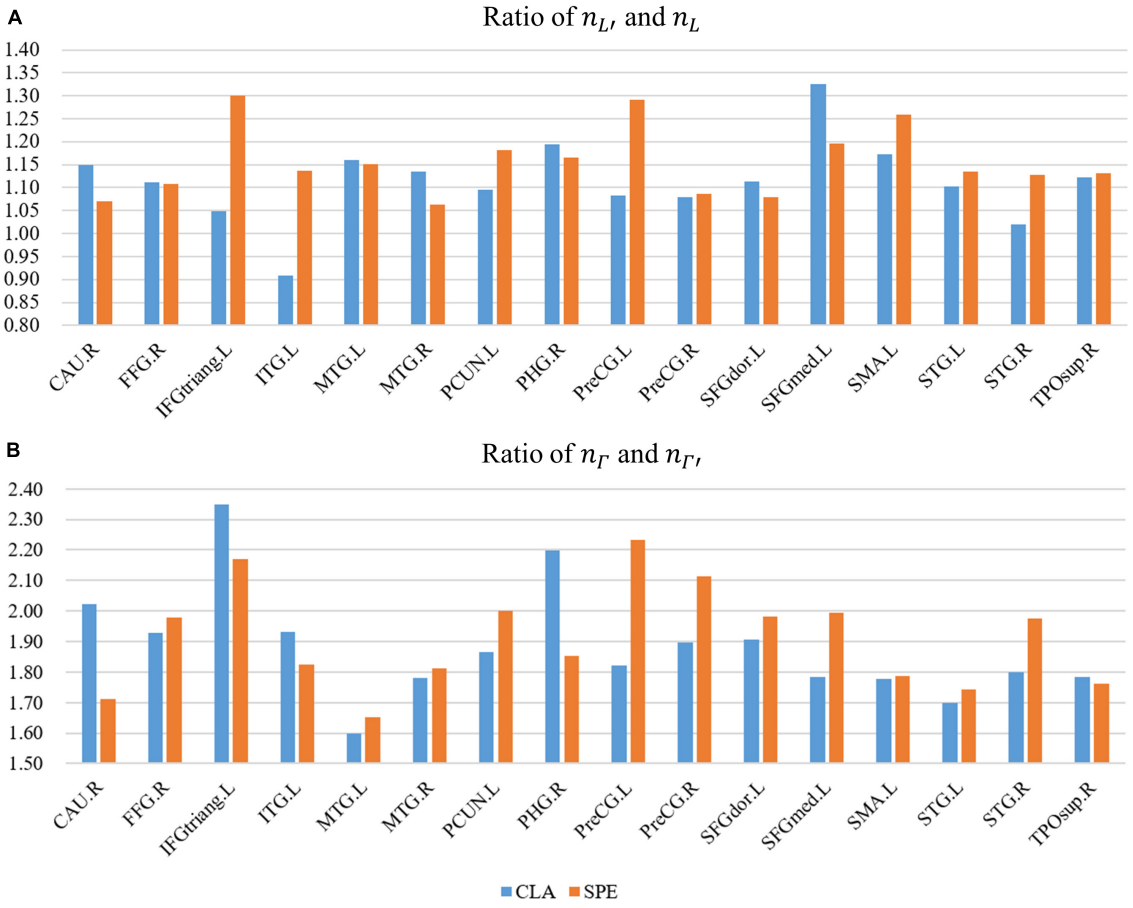


FIGURE 12
The ratio of higher frequency kernels to lower frequency kernels. **(A)** The ratio of the ratio of $n_{L'}$ (higher) to n_L (lower). **(B)** The ratio of n_{Γ} (higher) to $n_{\Gamma'}$ (lower).

tasks. As shown in **Figure 10**, the proportion of VOS is more than 50% in almost all brain regions. Especially, the proportions of VOS exceed 55% on MTG. L, MTG. R, STG. L, STG. R, and CAU. R in both CLA and SPE audio tasks. The brain regions with a higher proportion of VOS, which mostly spread over the temporal lobe, such as MTG.L, MTG.R, STG. L, and STG. R, are consistent with common perception and research (Whitehead and Armony, 2018). Besides, the proportions of VOS on MTG.L, STG.L, FFG.R, TPOsup. R, IFGtriang. L, and SFGdor. L in CLA task are 4% greater than that in SPE task, indicating that CLA signals have more unique and significant characteristics than SPE signals in these regions. In addition, the proportions of VOS brain regions outside the temporal lobe, such as FFG. R and CAU. R, are more than 55% in CLA task, which may help reveal the non-auditory function of these brain regions.

To achieve a more intuitive understanding and interpreting of the significant voxels, we selected and visualized the VOS in each brain area in **Figure 11**, where the blue represents each region, and the red represents the collection of VOS in each region. We can see that most of the red areas spread over the temporal lobe, such as MTG. L, MTG. R, STG. L, and STG. R. We can also see that red area of MTG. L is about 5% larger than that of MTG. R for CLA signals, indicating the activation characteristics of MTG. L region (on left brain) were more consistent and dominant than MTG. R region (on right brain), as shown in **Figures 10**, **11**. This further supports the phenomenon that the auditory function of the left brain is greater than that of the right brain. Besides, FFG. R also has distinct activation characteristics. Although existing studies could not conclusively prove the exact auditory-related function of FFG. R, its performance obtained by our framework may help reveal it.

Relationship between high-level semantic features and low-level acoustic features

In order to analyze the high-level semantic features in N-fMRI, we calculated the frequency domain features of voxel signals. As introduced in section "Exploration of relationship between high-level semantic features and low-level acoustic features *via* one-dimensional convolutional neural network

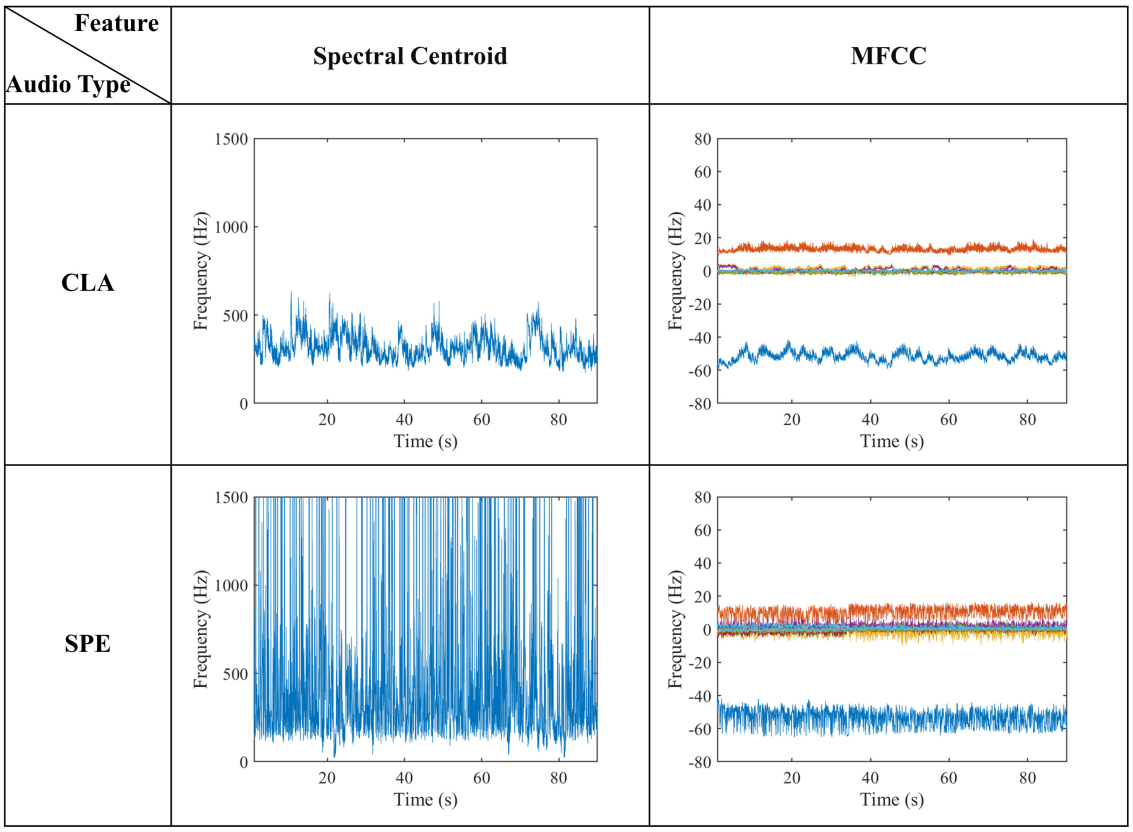


FIGURE 13
Low-level acoustic features of naturalistic audio.

model," the convolution kernels containing key features are divided into six types according to the frequency domain distribution. Since kernels of V and V' shape contain both low- and high-frequency domains, the analysis of V and V' shape would be more complex than that of other four types with a single frequency state (higher or lower). Therefore, we only focus on the four shapes except for V and V' shape in this section.

Firstly, the number of convolution kernels with higher and lower frequency domains was counted, respectively, represented as n_{Γ} (higher), $n_{L'}$ (higher), n_{L} (lower), and $n_{\Gamma'}$ (lower). Then, we calculated the ratio of n_{Γ} (higher) to $n_{\Gamma'}$ (lower), the ratio of $n_{L'}$ (higher) to n_{L} (lower), which represents the ratio of higher frequency kernels to lower frequency kernels, as shown in Figure 12. From Figure 12A, we find that the ratios in all regions except ITG. L are greater than 1, indicating the number of high-frequency convolution kernels is more than that of low-frequency convolution kernels in almost all ROS. From Figure 12B, we can see that ratios in all regions are greater than 1.6, even than 2 in the region of IFGtriang. L in both CLA and SPE. These results indicate that the high-frequency features of voxel signals are richer than the low-frequency features.

Furthermore, the ratio of $n_{L'}$ (higher) to n_{L} (lower) in SPE is 4.1% higher than that in CLA in Figure 12A. The ratio of n_{Γ} (higher) to $n_{\Gamma'}$ (lower) in SPE is 2.8% higher than that in CLA in Figure 12B, indicating that signals of ROS in SPE have more high-frequency characteristics than that in CLA in total. Besides, in regions of PCUN. L, PreCG. L, PreCG. R, STG. L, and STG. R, the ratios in Figures 12A,B of SPE are both higher than that of CLA, where the values of difference are between 0.04 and 0.4. This result further discloses that voxel signals of SPE have consistently richer high-frequency characteristics than that of CLA in these regions.

Figure 13 shows two typical acoustic features extracted from naturalistic audios. As can be seen from Figure 13, the average Spectral Centroid of the SPE and CLA audios is about 900–1100 Hz and 400–500 Hz, respectively. For the MFCCs features, frequency of CLA audios is concentrated in several frequency bands with a range of no more than 10 Hz. Although the frequency of SPE audios is also concentrated in several frequency bands, the band oscillates a lot with a maximum range over 20 Hz. These two typical acoustic features indicate that the frequency domain of the SPE is not only higher than that of the CLA but also more volatile. By comparing low-level acoustic features with high-level semantic features, we find that

the difference between SPE and CLA audio is consistent. We hypothesize that the audio with the larger Spectral Centroid leads to more intense brain activity. In terms of MFCCs features, the change in a frequency band may lead to more high-frequency characteristics in voxel signals.

Conclusion

In this work, we propose a novel hybrid learning framework that comprehensively studies the brain spatial (*via* Predictive Model) and temporal (*via* CNN model) characteristics during N-fMRI. By integrating spatial and temporal characteristics, ROS are obtained *via* the Predictive Model, and VOS are further interpreted *via* 1D-CNN model. Experiment results show that the proposed framework can achieve promising classification performance of audio categories and identify meaningful characteristics of the high-level semantic features. Especially for the classic and speech audios, the accuracy of classification is up to 92%. Furthermore, the relationship between high-level semantic features and low-level acoustic features is proved to be consistent in the frequency domain. In conclusion, the proposed framework provides novel insights on characterizing spatiotemporal patterns from the N-fMRI and effectively studying the high-level semantic features under naturalistic stimuli, which will further benefit the understanding of the brain working mechanism and the advance of naturalistic stimuli clinical application.

Data availability statement

The original contributions presented in the study are included in the article/**Supplementary material**, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by the Institutional Review Board (IRB), University of Georgia (UGA). The patients/participants provided their written informed consent to participate in this study.

Author contributions

SY, SZhang, and XJ: conception and design. SY: analysis and interpretation. TL: data collection. SY, SZhang, and XJ writing the manuscript. SY, SZhang, ES, RW, SZhao and XJ: critical revision of the manuscript. SZhang: overall responsibility. All authors contributed to the article and approved the submitted version.

Funding

SZhang was supported by the National Natural Science Foundation of China (62006194), the Fundamental Research Funds for the Central Universities (Grant No. 3102019QD005), and High-level researcher start-up projects (Grant No. 06100-22GH0202178). XJ was supported by the National Natural Science Foundation of China (61976045), Sichuan Science and Technology Program (2021YJ0247), and the Special Fund for Basic Scientific Research of Central Colleges (ZYGX2021J036).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnhum.2022.944543/full#supplementary-material

References

Anwar, S. M., Majid, M., Qayyum, A., Awais, M., Alnowami, M., and Khan, M. K. (2018). Medical image analysis using convolutional neural networks: a review. *J. Med. Syst.* 42:226. doi: 10.1007/s10916-018-1088-1

Çelik, E., Dar, S. U. H., Yılmaz, Ö, Keleş, Ü, and Çukur, T. (2019). Spatially informed voxelwise modeling for naturalistic fMRI experiments. *NeuroImage* 186, 741–757. doi: 10.1016/j.neuroimage.2018.11.044

Chen, S., and Hu, X. (2018). Individual identification using the functional brain fingerprint detected by the recurrent neural network. *Brain Connect.* 8, 197–204. doi: 10.1089/brain.2017.0561

Churchill, N. W., Raamana, P., Spring, R., and Strother, S. C. (2017). Optimizing fMRI preprocessing pipelines for block-design tasks as a function of age. *NeuroImage* 154, 240–254. doi: 10.1016/j.neuroimage.2017.02.028

Craddock, R. C., Holtzheimer, P. E. III, Hu, X. P., and Mayberg, H. S. (2009). Disease state prediction from resting state functional connectivity. *Magn. Reson. Med.* 62, 1619–1628. doi: 10.1002/mrm.22159

Erlikhman, G., and Caplovitz, G. P. (2017). Decoding information about dynamically occluded objects in visual cortex. *NeuroImage* 146, 778–788. doi: 10.1016/j.neuroimage.2016.09.024

Evans, A. C., Janke, A. L., Collins, D. L., and Baillet, S. (2012). Brain templates and atlases. *NeuroImage* 62, 911–922. doi: 10.1016/j.neuroimage.2012.01.024

Géron, A. (2019). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. Sebastopol, CA: O'Reilly Media, Inc.

Grama, L., and Rusu, C. (2017). "Choosing an accurate number of mel frequency cepstral coefficients for audio classification purpose," in *Proceedings of the International Symposium on Image and Signal Processing and Analysis, ISPA*, (Ljubljana: IEEE), 225–230. doi: 10.1109/ISPA.2017.8073600

Gupta, S., Jaafar, J., wan Ahmad, W. F., and Bansal, A. (2013). Feature extraction using Mfcc. *SIPIJ* 4, 101–108. doi: 10.5121/sipij.2013.4408

Hu, X., Guo, L., Han, J., and Liu, T. (2015). Decoding semantics categorization during natural viewing of video streams. *IEEE Trans. Auton. Ment. Dev.* 7, 201–210. doi: 10.1109/TAMD.2015.2415413

Huang, H., Hu, X., Zhao, Y., Makkie, M., Dong, Q., Zhao, S., et al. (2018). Modeling task fMRI data via deep convolutional autoencoder. *IEEE Tran. Med. Imaging* 37, 1551–1561. doi: 10.1109/TMI.2017.2715285

Jenkinson, M., and Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Med. Image Anal.* 5, 143–156. doi: 10.1016/S1361-8415(01)00036-6

Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., and Smith, S. M. (2012). FSL. *NeuroImage* 62, 782–790. doi: 10.1016/j.neuroimage.2011.09.015

Jiang, M., Yang, S., Yan, J., Zhang, S., Liu, H., Zhao, L., et al. (2020). "Exploring functional difference between gyri and sulci via region-specific 1D convolutional neural networks," in *Machine Learning in Medical Imaging*, (Cham: Springer International Publishing), 250–259. doi: 10.1007/978-3-030-59861-7_26

Jiang, X., Zhang, T., Hu, X., Lu, L., Han, J., Guo, L., et al. (2012). "Music/speech classification using high-level features derived from fmri brain imaging," in *Proceedings of the 20th ACM international conference on Multimedia MM '12*, (New York, NY: Association for Computing Machinery), 825–828. doi: 10.1145/2393347.2396322

Kohoutová, L., Heo, J., Cha, S., Lee, S., Moon, T., Wager, T. D., et al. (2020). Toward a unified framework for interpreting machine-learning models in neuroimaging. *Nat. Protoc.* 15, 1399–1435. doi: 10.1038/s41596-019-0289-5

Lad, A., and Patel, R. (2021). "Decoding with purpose: improving image reconstruction from fMRI with multitask learning," in *Proceedings of the 2021 IEEE 4th International Conference on Computing, Power and Communication Technologies, GUCON 2021*, (Kuala Lumpur: IEEE), doi: 10.1109/GUCON50781.2021.9573575

Lahnakoski, J., Glerean, E., Salmi, J., Jääskeläinen, I., Sams, M., Hari, R., et al. (2012). Naturalistic fMRI mapping reveals superior temporal sulcus as the hub for the distributed brain network for social perception. *Front. Hum. Neurosci.* 6:233. doi: 10.3389/fnhum.2012.00233

Lin, M., Chen, Q., and Yan, S. (2013). Network in network. *arXiv[Preprint]* doi: 10.48550/arXiv.1312.4400

Liu, H., Zhang, S., Jiang, X., Zhang, T., Huang, H., Ge, F., et al. (2019). The cerebral cortex is bisectionally segregated into two fundamentally different functional units of gyri and sulci. *Cereb. Cortex* 29, 4238–4252. doi: 10.1093/cercor/bhy305

Martinez-Garcia, F., Puelles, L., Ten Donkelaar, H., and González, A. (2012). Adaptive function and brain evolution. *Front. Neuroanat.* 6:17. doi: 10.3389/fnana.2012.00017

Obuchowski, N. A., and Bullen, J. A. (2018). Receiver operating characteristic (ROC) curves: review of methods with applications in diagnostic medicine. *Phys. Med. Biol.* 63:07TR01. doi: 10.1088/1361-6560/aab4b1

Ozcelik, F., Choksi, B., Mozafari, M., Reddy, L., and VanRullen, R. (2022). Reconstruction of perceived images from fMRI patterns and semantic brain exploration using instance-conditioned GANs. *arXiv* [Preprint]. arXiv: 2202.12692

Prasetio, B. H., Widasari, E. R., and Tamura, H. (2021). "Automatic multiscale-based peak detection on short time energy and spectral centroid feature extraction for conversational speech segmentation," in *Proceedings of the ACM International Conference Proceeding Series*, (New York, NY: ACM), 44–49. doi: 10.1145/3479645.3479675

Raposo, F. A., Martins de Matos, D., and Ribeiro, R. (2022). Learning low-dimensional semantics for music and language via multi-subject fMRI. *Neuroinform* doi: 10.1007/s12021-021-09560-5 [Online ahead of print]

Ray, S. (2019). "A quick review of machine learning algorithms," in *Proceedings of the 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*, Faridabad, 35–39. doi: 10.1109/COMITCon.2019.8862451

Ren, Y., Tao, Z., Zhang, W., and Liu, T. (2021). "Modeling hierarchical spatial and temporal patterns of naturalistic fmri volume via volumetric deep belief network with neural architecture search," in *Proceedings of the International Symposium on Biomedical Imaging*, (Nice: IEEE), 130–134. doi: 10.1109/ISBI48211.2021.9433811

Rolls, E. T., Huang, C.-C., Lin, C.-P., Feng, J., and Joliot, M. (2020). Automated anatomical labelling atlas 3. *NeuroImage* 206:116189. doi: 10.1016/j.neuroimage.2019.116189

Rosenberg, M. D., Casey, B. J., and Holmes, A. J. (2018). Prediction complements explanation in understanding the developing brain. *Nat. Commun.* 9:589. doi: 10.1038/s41467-018-02887-9

Saarimäki, H. (2021). Naturalistic stimuli in affective neuroimaging: a review. *Front. Hum. Neurosci.* 15:675068. doi: 10.3389/fnhum.2021.675068

Scheinost, D., Noble, S., Horien, C., Greene, A. S., Lake, E. M. R., Salehi, M., et al. (2019). Ten simple rules for predictive modeling of individual differences in neuroimaging. *NeuroImage* 193, 35–45. doi: 10.1016/j.neuroimage.2019.02.057

Shain, C., Blank, I. A., van Schijndel, M., Schuler, W., and Fedorenko, E. (2020). fMRI reveals language-specific predictive coding during naturalistic sentence comprehension. *Neuropsychologia* 138:107307. doi: 10.1016/j.neuropsychologia.2019.107307

Simony, E., and Chang, C. (2020). Analysis of stimulus-induced brain dynamics during naturalistic paradigms. *NeuroImage* 216:116461. doi: 10.1016/J.NEUROIMAGE.2019.116461

Sonkusare, S., Breakspear, M., and Guo, C. (2019). Naturalistic stimuli in neuroscience: critically acclaimed. *Trends Cogn. Sci.* 23, 699–714. doi: 10.1016/J.TICS.2019.05.004

Vanderwal, T., Eilbott, J., and Castellanos, F. X. (2019). Movies in the magnet: naturalistic paradigms in developmental functional neuroimaging. *Dev. Cogn. Neurosci.* 36:100600. doi: 10.1016/J.DCN.2018.10.004

Wang, J., Ren, Y., Hu, X., Nguyen, V. T., Guo, L., Han, J., et al. (2017). Test–retest reliability of functional connectivity networks during naturalistic fMRI paradigms. *Hum. Brain Mapp.* 38, 2226–2241. doi: 10.1002/hbm.23517

Wang, L., Hu, X., Liu, H., Zhao, S., Guo, L., Han, J., et al. (2020). "Functional brain networks underlying auditory saliency during naturalistic listening experience," in *Proceedings of the IEEE Transactions on Cognitive and Developmental Systems*, (New Jersey, NJ: IEEE), doi: 10.1109/TCDS.2020.3025947

Wen, D., Wei, Z., Zhou, Y., Li, G., Zhang, X., and Han, W. (2018). Deep learning methods to process fMRI data and their application in the diagnosis of cognitive impairment: a brief overview and our opinion. *Front. Neuroinform.* 12:23. doi: 10.3389/fninf.2018.00023

Whitehead, J. C., and Armony, J. L. (2018). Singing in the brain: neural representation of music and voice as revealed by fMRI. *Hum. Brain Mapp.* 39, 4913–4924. doi: 10.1002/hbm.24333

Woolrich, M. W., Ripley, B. D., Brady, M., and Smith, S. M. (2001). Temporal autocorrelation in univariate linear modeling of FMRI data. *NeuroImage* 14, 1370–1386. doi: 10.1006/nimg.2001.0931

Xia, M., Wang, J., and He, Y. (2013). BrainNet viewer: a network visualization tool for human brain connectomics. *PLoS One* 8:e68910. doi: 10.1371/journal.pone.0068910

Yan, W., Calhoun, V., Song, M., Cui, Y., Yan, H., Liu, S., et al. (2019). Discriminating schizophrenia using recurrent neural network applied on time courses of multi-site FMRI data. *EBioMedicine* 47, 543–552. doi: 10.1016/j.ebiom.2019.08.023

Yuen, N. H., Osachoff, N., and Chen, J. J. (2019). Intrinsic frequencies of the resting-state fMRI signal: the frequency dependence of functional connectivity and the effect of mode mixing. *Front. Neurosci.* 13:900. doi: 10.3389/fnins.2019.00900

Zhang, S., Liu, H., Huang, H., Zhao, Y., Jiang, X., Bowers, B., et al. (2019). Deep learning models unveiled functional difference between cortical gyri and sulci. *IEEE Trans. Biomed. Eng.* 66, 1297–1308. doi: 10.1109/TBME.2018.2872726

Zhang, T., Gao, J. S., Çukur, T., and Gallant, J. L. (2021). Voxel-based state space modeling recovers task-related cognitive states in naturalistic fMRI experiments. *Front. Neurosci.* 14:565976. doi: 10.3389/fnins.2020.565976

Zhang, Y., Hu, X., He, C., Wang, X., Ren, Y., Liu, H., et al. (2019). "A two-stage DBN-based method to exploring functional brain networks in naturalistic paradigm FMRI," in *Proceedings of the 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, (Venice: IEEE), 1594–1597. doi: 10.1109/ISBI.2019.8759376

Zhao, S., Jiang, X., Han, J., Hu, X., Zhu, D., Lv, J., et al. (2014). "Decoding auditory saliency from FMRI brain imaging," in *Proceedings of the 22nd ACM International conference On Multimedia MM '14*, (New York, NY: Association for Computing Machinery), 873–876. doi: 10.1145/2647868.2655039

frontiers | Frontiers in Neuroimaging

TYPE Original Research
PUBLISHED 04 January 2023
DOI 10.3389/fnimg.2022.1007668

OPEN ACCESS

EDITED BY
Yao Wu,
Children's National Hospital,
United States

REVIEWED BY
Mario Ganau,
Oxford University Hospitals NHS Trust,
United Kingdom
Dongren Yao,
Massachusetts Eye and Ear Infirmary
and Harvard Medical School,
United States

*CORRESPONDENCE
Ayan Banerjee
abanerj3@asu.edu

SPECIALTY SECTION
This article was submitted to
Brain Imaging Methods,
a section of the journal
Frontiers in Neuroimaging

RECEIVED 30 July 2022
ACCEPTED 24 October 2022
PUBLISHED 04 January 2023

CITATION
Banerjee A, Kamboj P, Wyckoff SN, Sussman BL, Gupta SKS and Boerwinkle VL (2023) Automated seizure onset zone locator from resting-state functional MRI in drug-resistant epilepsy.
Front. Neuroimaging 1:1007668.
doi: 10.3389/fnimg.2022.1007668

Automated seizure onset zone locator from resting-state functional MRI in drug-resistant epilepsy

Ayan Banerjee [1]*, Payal Kamboj[1], Sarah N. Wyckoff[2], Bethany L. Sussman[2], Sandeep K. S. Gupta[1] and Varina L. Boerwinkle[3]

[1]School of Computing and Augmented Intelligence, Arizona State University, Tempe, AZ, United States, [2]Division of Neuroscience, Barrow Neurological Institute at Phoenix Children's Hospital, Phoenix, AZ, United States, [3]Division of Child Neurology, University of North Carolina Department of Neurology, Chapel Hill, NC, United States

Objective: Accurate localization of a seizure onset zone (SOZ) from independent components (IC) of resting-state functional magnetic resonance imaging (rs-fMRI) improves surgical outcomes in children with drug-resistant epilepsy (DRE). Automated IC sorting has limited success in identifying SOZ localizing ICs in adult normal rs-fMRI or uncategorized epilepsy. Children face unique challenges due to the developing brain and its associated surgical risks. This study proposes a novel SOZ localization algorithm (EPIK) for children with DRE.

Methods: EPIK is developed in a phased approach, where fMRI noise-related biomarkers are used through high-fidelity image processing techniques to eliminate noise ICs. Then, the SOZ markers are used through a maximum likelihood-based classifier to determine SOZ localizing ICs. The performance of EPIK was evaluated on a unique pediatric DRE dataset ($n = 52$). A total of 24 children underwent surgical resection or ablation of an rs-fMRI identified SOZ, concurrently evaluated with an EEG and anatomical MRI. Two state-of-art techniques were used for comparison: (a) least squares support-vector machine and (b) convolutional neural networks. The performance was benchmarked against expert IC sorting and Engel outcomes for surgical SOZ resection or ablation. The analysis was stratified across age and sex.

Results: EPIK outperformed state-of-art techniques for SOZ localizing IC identification with a mean accuracy of 84.7% (4% higher), a precision of 74.1% (22% higher), a specificity of 81.9% (3.2% higher), and a sensitivity of 88.6% (16.5% higher). EPIK showed consistent performance across age and sex with the best performance in those $<$ 5 years of age. It helped achieve a ~5-fold reduction in the number of ICs to be potentially analyzed during pre-surgical screening.

Significance: Automated SOZ localization from rs-fMRI, validated against surgical outcomes, indicates the potential for clinical feasibility. It eliminates the need for expert sorting, outperforms prior automated methods, and is consistent across age and sex.

KEYWORDS

resting state fMRI, seizure onset zone, resting state network, drug resistant epilepsy, expert knowledge driven classification

Introduction

Epilepsy is devastating, affecting 50 million people worldwide (WHO). One in 150 children have epilepsy (Aaberg et al., 2017; Epilepsy Foundation, 2018), with 30% having drug-resistant epilepsy (DRE; Wieser et al., 2001; Kwan and Sander, 2004; Kwan and Brodie, 2010), which causes significant morbidity and mortality (Sillanpää and Shinnar, 2010; Laxer et al., 2014; Engel, 2016). A consensus proposal by the *ad-hoc* Task Force of the International League Against Epilepsy (ILAE) proposed the following definition for DRE: "a failure of adequate trials of two tolerated, appropriately chosen, and used antiepileptic drug schedules (whether as monotherapies or in combination) to achieve sustained seizure freedom (considered as freedom from all seizures, including auras) for at least 12 months" (Kwan et al., 2010).

Early diagnosis and treatment of DRE can potentially deflect complications such as evolution into status epilepticus (Prisco et al., 2020) and Sudden Unexplained Death in Epilepsy (SUDEP), wherein the individual dies due to cardio-respiratory failure from presumed nocturnal seizure activity (Sillanpää and Shinnar, 2010). Moreover, in children, timely diagnosis, intensive management, and treatment are pivotal in minimizing neurological damage (Prisco et al., 2020). Further, the earliest onset of severe epilepsy in the neonatal population can lead to nearly constant life-threatening seizures requiring an urgent need for surgical evaluation early in life (Russ et al., 2021).

Surgery for DRE

The most effective treatment for DRE is surgery (Luders et al., 2006; Luckett et al., 2022). Early surgery is key: "minimally invasive surgical treatment can be a life-changing option for DRE patients; hence management of the SOZ requiring disconnecting techniques (Young et al., 2020), or deep sited lesions requiring excision should be considered earlier rather than later (Chibbaro et al., 2017)." Notably, recent findings showed that ultra-early (before 3 months old) surgical intervention in children evaluated to have DRE after trials of an average of four anti-seizure drugs, although seldom performed, has excellent epilepsy outcomes and leads to a decrease in usage of anti-seizure drugs, without any increased risk of surgery-related permanent morbidity (Roth et al., 2021).

Brain imaging for pre-surgical screening

Surgical intervention in DRE requires accurate localization of the seizure onset zone (SOZ) for success. We make a distinction between the epileptic network (EN) and the SOZ. The EN denotes regions where seizure propagates and may be more extensive than the SOZ. As such, it may be difficult as well as unnecessary to surgically eliminate the EN since it can incorporate sensitive areas of the brain. Several brain imaging techniques have been explored to identify the ictal seizure onset zone, propagation zone (i.e., EN), and interictal activity (Table 1). This can be done with nuclear medicine-based imaging techniques such as positron emission tomography (PET) or ictal single-photon emission computerized tomography (SPECT; Desai et al., 2013). Recent studies suggest some SOZ identification capabilities for PET and SPECT in both adults and children; however, their accuracy heavily depends on the timing of the scan. Delay in drug infusion can result in the detection of the EN instead of the SOZ. Invasive modalities such as intracranial EEG (iEEG) are considered the gold standard for SOZ identification and have shown excellent accuracy for both adults and pediatric DRE. Stereo-electroencephalography (SEEG) is minimally invasive, uses a three-dimensional configuration of depth electrodes to localize epileptiform activity, and has shown some SOZ identification capability recently (Satzer et al., 2022).

However, traditional analysis of PET, SPECT, or SEEG is relatively, temporally, and spatially restricted, whereas functional interpolation of brain activity might allow for a non-invasive three-dimensional representation of epileptiform activity and avoid pitfalls inherent of other modalities (Table 1). Recently, magnetoencephalography (MEG) and functional magnetic resonance imaging (fMRI)-based non-invasive techniques have been analyzed for DRE in both adults and children and show decent SOZ identification capability. A combination of MEG and fMRI imaging has also been proposed

TABLE 1 Summary of brain imaging techniques and their application in SOZ identification for adults and children.

Imaging technique	Invasive/non-invasive	Number of imaging sessions	Spatial resolution	Temporal resolution	Brain area	SOZ ID in adults with DRE	SOZ ID in children with DRE
PET	Invasive	1	6 mm	5 mins	Surface	Identified in 66.7% of subjects (Mayoral et al., 2019)	Only detects EN in 6.6% of patients and not SOZ (Bansal et al., 2016)
Ictal SPECT	Invasive	2, needs difference between ictal and inter-ictal SPECT (Desai et al., 2013)	6 mm	15 mins	Surface	Detects EN and not SOZ (Kaiboriboon et al., 2002; von Oertzen, 2018)	Detects EN and not SOZ (Van Paesschen et al., 2007)
MEG	Non-invasive	1	10–20 mm	<10 ms	Cortex	75% accurate (Foley et al., 2021) in eight patients, 60% with ML in a large cohort (Nissen et al., 2018)	Poor concordance with SOZ. Identified onset zone is > 20 mm distance from ground truth SOZ (Ntolkeras et al., 2022)
iEEG (gold standard)	Invasive	Long-term monitoring (14 days)	6–8 cm	<5 ms	Cortex	92.3% accurate (Nagahama et al., 2018)	>90% accurate (Nagahama et al., 2018)
fMRI	Non-invasive	1	3 mm	2–5s	Central brain	Sensitivity 83% and specificity 67% (Chen et al., 2017)	89% accurate SOZ identification (Boerwinkle et al., 2019)

PET, positron emission tomography; SPECT, single-photon emission computerized tomography; MEG, magnetoencephalography; iEEG, intracranial electroencephalography; fMRI, functional magnetic resonance imaging.

for accurate SOZ identification (Berger et al., 2021). However, a major drawback of such brain imaging-based SOZ identification techniques is the heavy reliance on manual sorting of images and their components, which not only increases cost but also reduces accessibility and repeatability.

Unfortunately, <1% of patients with DRE are evaluated for surgery and only 25% of those undergo surgery (Engel, 2016), partly due to the high cost of diagnostic and surgical treatment (>$200,000/patient) and the risk of debilitating impairment (Murray et al., 1996; Begley et al., 2000). Of the 1% evaluated, surgical failure rates are 30–70% despite the use of non-invasive SOZ-localization biomarkers such as anatomical MRI, scalp EEG, simultaneous EEG-fMRI, and magnetoencephalography, which are then often confirmed by invasive iEEG (McIntosh et al., 2004; Luders et al., 2006; Sillanpää and Shinnar, 2010; Bulacio et al., 2012; Laxer et al., 2014; Engel, 2016; Epilepsy Foundation, 2018). Hence, for surgery to be safe and efficient for wide acceptance (England et al., 2012), accessible, minimally invasive, and accurate SOZ localization is essential.

One of the newer methods showing promise, to this end, is resting-state functional MRI (rs-fMRI). Rs-fMRI has been shown to have an accurate SOZ-localization capacity through various analysis approaches (Bandt et al., 2014; van Houdt et al., 2015; Malmgren and Edelvik, 2017; Boerwinkle et al., 2018), but only independent component analysis (ICA; Gonzalez-Martinez et al., 2007) has provided Level 1 evidence and has led to improvement in surgical outcomes (Malmgren and Edelvik, 2017; Chakraborty et al., 2020) and candidacy (Boerwinkle et al., 2019) in DRE. However, expert interpretation of independent components (IC) into sources of noise, normal resting state networks (RSN), and SOZs (Hunyadi et al., 2014; Boerwinkle et al., 2017, 2020) limits reproducibility and availability. An automated whole-brain data-driven SOZ-localizing IC identification technique that is rigorously validated against surgical destruction outcomes, reproducible, equally effective across age and sex, and applicable to all epilepsy subtypes may greatly improve epilepsy care feasibility, morbidity, and mortality.

fMRI-based screening

Functional MRI (fMRI) is a popular imaging technique originally used to identify brain activity in terms of blood oxygenation level change in different parts of the brain for a given mental task (Figure 1). However, for SOZ detection, it is required to identify blood oxygenation changes due to the onset of seizure. Hence, an important step is to remove other sources of brain activity such as mental tasks, fMRI noise, and head motion. Rs-fMRI requires the subject to be in a resting state, which is achieved in a majority of children through sedation. Even if any mental task is eliminated, there is still the presence of resting-state brain activity in subjects, which manifests as RSN

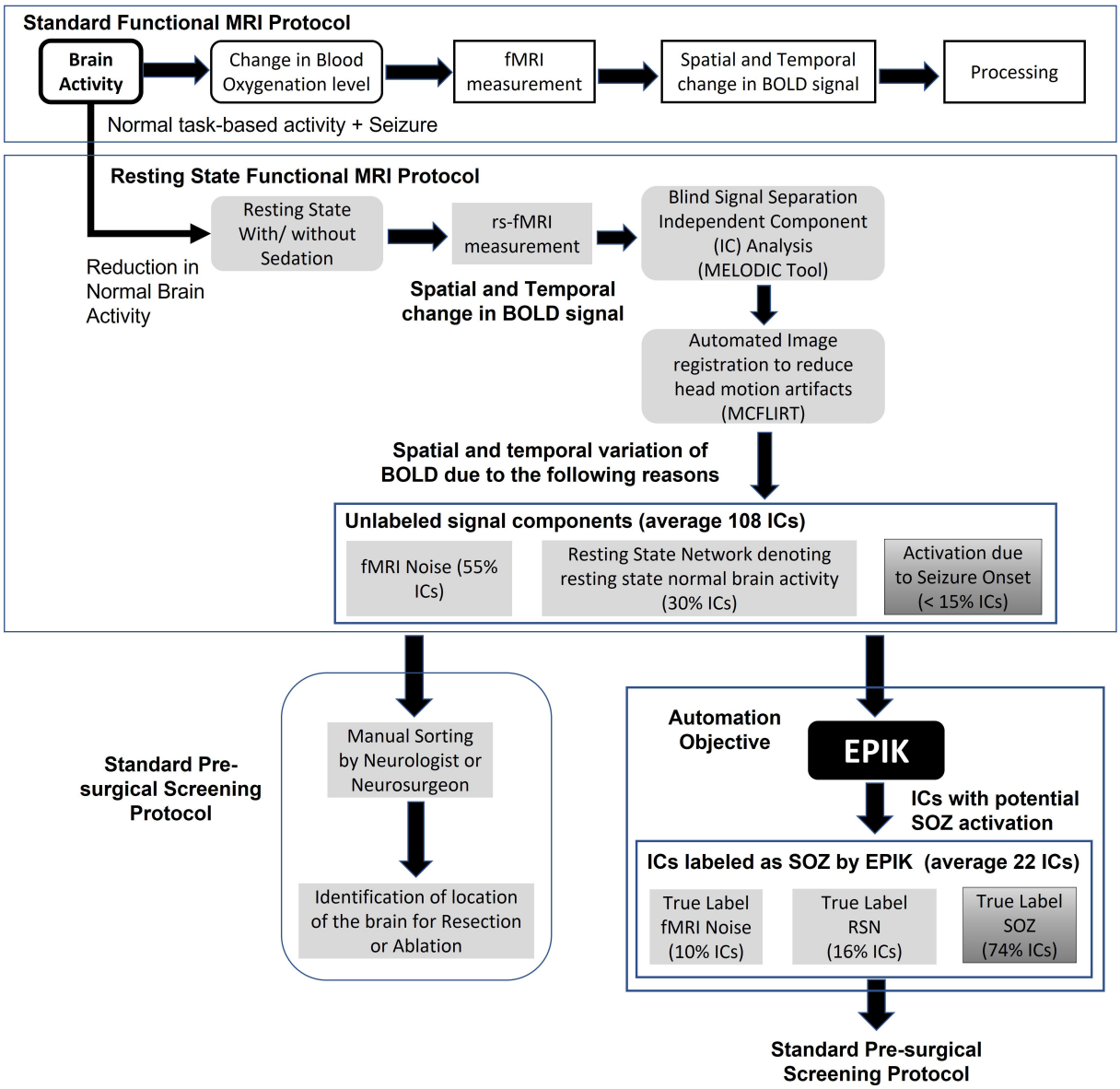


FIGURE 1
The standard task-based fMRI protocol (top panel), resting state fMRI and data processing pipeline (middle panel), standard rs-fMRI-based pre-surgical evaluation of SOZ location (lower left panel), automation objective of EPIK (lower right panel), and advantages of EPIK in terms of reduction in manual processing requirements.

brain activity. Head motion is a significant source of noise. Even if head motion is limited to <1 mm, it still can pose a significant amount of noise in the rs-fMRI measurement. Automated image registration is used to reduce head motion artifacts in rs-fMRI (Figure 1 middle panel). The resulting rs-fMRI captures brain activity due to several sources including (a) noise (fMRI machine noise and head motion), (b) RSN (resting-state activity of the brain), and (c) SOZ (change in blood oxygenation due to seizure onset). To decouple the effects of noise, RSN, and SOZ in rs-fMRI signals, ICA is used to recover mutually independent fMRI signal components (ICs) that potentially only capture brain activity from one of the three sources.

Rs-fMRI ICA results in ∼100 ICs. Each IC is a spatial-temporal distribution of regions of synchronous activity. In ICA of those with DRE, there are three IC categories: (1) RSNs which are well-described and validated in the literature; (2) SOZ which is, currently, highly dependent on expert sorting; and (3) noise, which is also well-understood, resulting from cardiovascular, cerebral-spinal-fluid-pulsation, or scanner artifacts [see Boerwinkle et al. (2017) for details and examples].

In standard rs-fMRI-based pre-surgical screening for children with DRE, the entire set of ICs is analyzed by a neurosurgeon or neurologist to determine which ICs capture blood oxygenation changes due to seizure onset. Such ICs are referred to as SOZ localizing IC. The neurosurgeon then determines the location of seizure onset in the brain using the SOZ localizing IC and a recommendation for a surgical procedure such as resection or ablation or neurostimulation is made.

Given that ICA results in $>$ 100 ICs and only $<$ 10% are SOZ localizing ICs, manual sorting of rs-fMRI ICs to search for SOZ localizing IC is a significant time commitment by the neurosurgeon, resulting in increased cost, reduced availability, and a higher chance of false positives (Figure 1). This study focuses on automating the task of IC sorting and reducing the number of ICs to be analyzed by the neurosurgeon for pre-surgical evaluation for children with DRE.

Automation of fMRI-based screening

Artificial Intelligence (AI) has been employed on rs-fMRI to automatically identify several brain disorders including Attention Deficit Hyperactivity Disorder (ADHD), Alzheimer's disease, White Matter Hyperintensity (WMH; Bharath et al., 2019), and major depressive disorder (Nguyen et al., 2021). Recent studies considered two automation objectives in epilepsy with rs-fMRI (Table 2): (a) classification of subjects with or without epilepsy by identifying epilepsy networks using rs-fMRI blood oxygen level-dependent (BOLD) signal z-score latency maps (Lopes et al., 2012; Bharath et al., 2019; Nguyen et al., 2021), and (b) localization of the seizure onset zone using rs-fMRI ICs (Hunyadi et al., 2014, 2015; Shah et al., 2019). Epilepsy networks indicate the areas of the brain that are affected by the propagation of a seizure. As such, they may not indicate the origin of the seizure, which is encapsulated by the SOZ. Our research focus in this study tackles the second automation objective of SOZ localization.

Automated classification of rs-fMRI ICs as SOZ or RSN has been explored using supervised shallow machine learning (ML; Nozais et al., 2021) and using deep learning (DL) in healthy adults to identify the typical RSNs and is yet to be tested in epilepsy (Zhang et al., 2019; Table 2). Supervised ML indicates that the DRE population has to be divided into two parts: (a) a training set, which is used to configure the ML, and (b) a testing set, which is used to test the performance of the ML. Some supervised ML can also choose to utilize a validation set as mentioned in a previous study (Nguyen et al., 2021). The performance of the ML technique on the validation set is used to update the training process and improve the performance in the validation set. Hence, the performance on the validation set is excluded from the analysis in Table 2 and only the test set performance is reported. Recent automated (Luckett et al., 2022) methods to classify adult rs-fMRI into RSN, SOZ, and noise ICs are of three types: (1) voxel-based network measures quantifying the number of connections to each voxel in an IC, called voxel degree connectivity (VDC), as indicators for SOZ (Hunyadi et al., 2014; Lee et al., 2014). Such approaches have a small sample size ($n \approx 20$) and show a maximum reported sensitivity of 77% and a specificity of 57% (Table 2); (2) ML-based classification, with a sensitivity of 40% and a specificity of 77% (Hunyadi et al., 2014); and (3) DL approaches for only identifying RSN and noise, but not SOZs, for normal and non-DRE patients with epilepsy [accuracy 92% (Nozais et al., 2021) in Table 2].

To date, automated approaches have not been successful in the classification of RSN, noise, and SOZ, in rs-fMRI for pediatric patients with DRE due to the following challenges: (1) Lack of normalized pathological rs-fMRI RSN data for children (Zhang et al., 2019); (2) databases with balanced instances of RSN, SOZ, and noise, large enough for DL techniques to effectively recognize the three IC categories that are not available; (3) the potentially inadequate performance of SOZ identification in children with DRE can indicate a high risk of developmental disorders post-surgery. Given that each patient only has 5% ICs as SOZ, a 40% sensitivity (Hunyadi et al., 2014) indicates that only two out of five SOZ ICs are correctly identified but 14 of them are wrongly identified as SOZ; and (4) fMRI-based pre-surgical mapping is more complicated for children with DRE due to developmental changes during cognitive maturation (Jiang et al., 2018; Bouyssi-Kobar et al., 2019), the impairment experienced due to DRE, and the normal representation of memory function during development (Michels et al., 2012; Darki and Klingberg, 2015; Cui et al., 2018; Kasradze et al., 2021), which may differ from adults (Faghiri et al., 2017; Lee et al., 2019; DeGeorge et al., 2021; Moncrief et al., 2021). Hence, the efficacy of fMRI classification techniques on adults needs to be reexamined for children with DRE.

Most current studies (Table 2) focused on adult epilepsy with an unknown effect of the degree of hypothesized network disruption effect on localization. Currently available automated IC sorting techniques either only identify SOZ or RSN localizing ICs. Hunyadi et al. (2015), the first major work to attempt SOZ localizing ICs identification, used supervised ML but could only achieve a specificity of 77% and a sensitivity of 40% on a subset of the adult patient population. A more recent technique by Nozais et al. (2021) used DL to identify only RSN in healthy adults and reports an accuracy of 92%. The major drawback of DL techniques is the requirement for labeled data on all three IC categories. Table 2 shows that such labeled data is rarely available, even if we combine datasets from different authors, IC data labeled as RSN and SOZ are only available from 212 children with DRE. For DL to successfully recognize SOZ, it will need at least a balanced distribution of RSN and SOZ. The DL works in this domain utilize RSN data from 2,000 healthy subjects for appropriate training (Table 2; Nozais et al., 2021;

TABLE 2 Comparison of related research highlighting innovative aspects of the proposed research.

Problem	References	Epilepsy type	Epileptic zone indicators	Task objective	Validation modality	Machine learning method	Supervised (sup)/ unsupervised (U)	Performance	*N*	A adult C child	*N* of epilepsy Subjects (*N* with DRE)	*N* of test epilepsy subjects	*N* of subjects with surgical outcomes
Detection of epilepsy	Nguyen et al., 2021	DRE	Epilepsy Network (EN)	Epilepsy classification using fMRI *z* score latency	Seizures	CNN	Sup	Accuracy in Epilepsy identification = 74% Sensitivity = 85% Specificity = 71%	322	C	63 (63)	13	0
	Lopes et al., 2012	Focal	Not Specified (NS)	Epilepsy classification using BOLD time series	Seizures	Time series analysis	U	Accuracy = 87.5%	15	A	15 (0)	15 subjects with 40 events	0
	Bharath et al., 2019	Focal Temporal Lobe	Hand classification EN	Epilepsy classification using ICA	Seizures	SVM	Sup	Accuracy = 97.5% Sensitivity = 100% Specificity = 94.4%	132	A	42 (0)	0 (No test data, cross validation accuracy)	0
Pre-surgical screening to determine seizure onset zone (SOZ)	Boerwinkle et al., 2017	DRE	RSN and SOZ	Manual SOZ localization	iEEG and post-op seizure	No automation	Sup	89% accuracy	40	C	40 (40)	33	40
	Shah et al., 2018	DRE	NS	Finding correlation between fMRI *z* score latency and seizure freedom	Post-op seizure	Statistical correlation measures	U	25 out of 26 subjects have temporal lobe signal latency	26	C	26 (26)	26	26 (21 seizure free)
	Shah et al., 2019	DRE	NS	fMRI *z* score latency-based seizure foci lateralization	Manual lateralization	Statistical correlation measures	U	Mean accuracy 70% Mean sensitivity 85% Mean Specificity 65%	38	C	38 (38)	38	38 (14 seizure free)
	Hunyadi et al., 2015	DRE	SOZ	Automated SOZ identification (ID)	EEG-fMRI	LS-SVM	Sup	40% sensitivity 77% specificity	18	A	18 (18)	10	Not Specified (NS)
	Zhang et al., 2015	DRE	SOZ	SOZ localization with the manual determination of brain boundary	Concordance with surgery resection	Statistical methods using thresholds	U	77.7% sensitivity* 57% specificity**	9	A	9 (9)	9	NS
	Lee et al., 2014	Intractable partial	SOZ	VDC-based SOZ localization	iEEG.	Statistical methods on time series	U	72.4% sensitivity***	29	A	29 (29)	21	29 (2 seizure free)
	Nozais et al., 2021	Healthy	RSN	DL-based RSN ID	Manual IC Sorting	MLP	Sup	92% accuracy	2000	A	0 (0)	0	0
	Luckett et al., 2022	Focal Temporal Lobe	RSN and SOZ	DL-based SOZ hemisphere ID	Hemisphere lateralization of SOZ, and RSNs	CNN	Sup	90.6% accuracy*****	2164	A	32 (0)	32	15 (11 seizure free)
	EPIK (current study)	DRE	RSN and SOZ	Fully automated unsupervised method (EPIK)	Manual IC sorting Comparison of Hunyadi et al. and Nozais et al. on the same dataset; and post-op seizures	Rule guided noise elimination, maximum likelihood based classification	U	RSN success: Specificity: 73.7% Sensitivity: 72% SOZ success: Specificity: 81.9% Sensitivity: 88.6%	52	C	52 (52)	52	24 (18 seizure free)

DL, deep learning; DRE, drug-resistant or intractable epilepsy; EEG-fMRI, simultaneous EEG and functional MRI; IC, independent component; ID, identification; iEEG, intracranial EEG; ML, machine learning; LS-SVM, least squares support vector machine; VDC, voxel degree connectivity; CNN, convolutional neural network.

*Zhang et al. (2015) only mentions concordance with surgical resection. Concordance is assumed to be true positive, and failure is assumed to be false negative. Hence percentage concordance is assumed to be sensitivity.

**Zhang et al. (2015) mentions success in rejecting non-epilepsy related IC. Rejection of non-epilepsy IC is assumed to be true negative and failure to reject is assumed to be false positive. Hence success rate is specificity.

***Lee et al. (2014) defines accordance with IC EEG. Accordance is assumed to be true positive, and failure is assumed to be false negative. Hence, percentage accordance is assumed to be sensitivity.

*****Sensitivity and specificity not mentioned. The high accuracy could also be due to the presence of a large number of true negatives.

T epileptic networks were those found to be altered compared to healthy controls, however, were not identified as being causative of epilepsy or a seizure onset zone.

10.3389/fnimg.2022.1007668

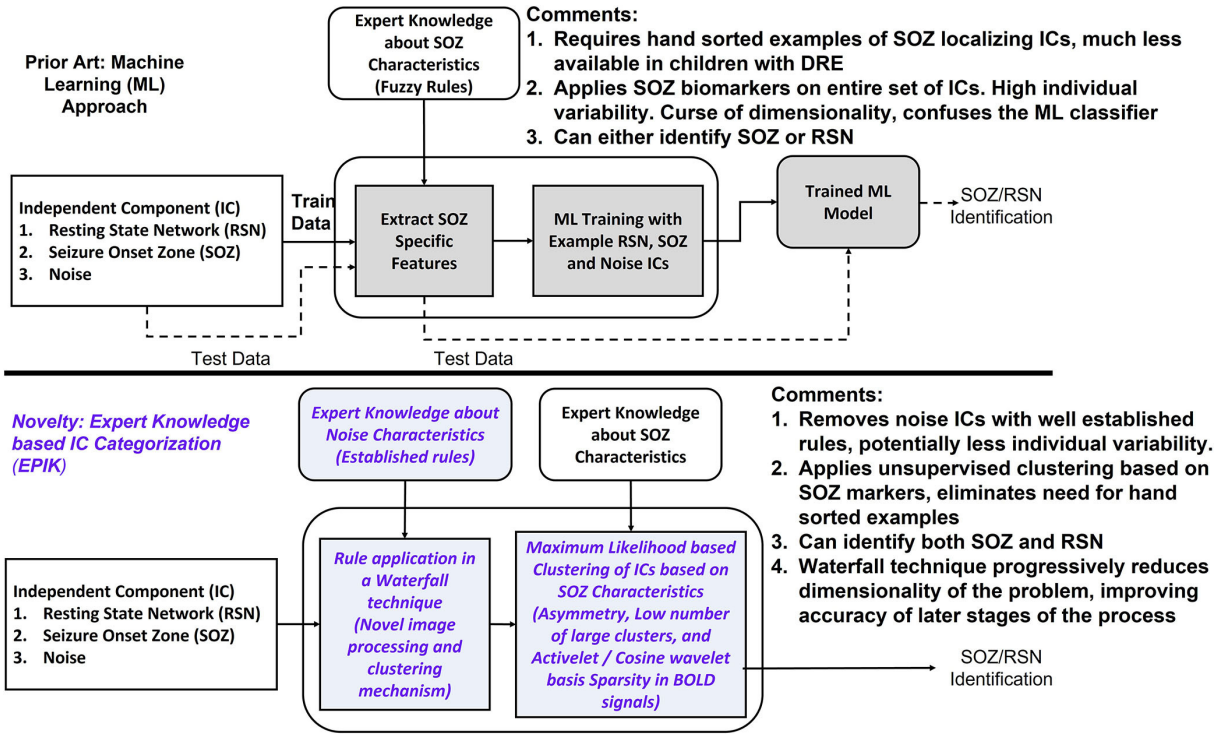


FIGURE 2
Innovations in EPIK. Compared to prior ML techniques, EPIK first purges ICs with noise markers by employing rules compiled from experts. The ICs that pass the initial purge, are then classified into RSN and SOZ based on the maximum likelihood-based clustering mechanism. The italicized text marks the innovations in this work.

Luckett et al., 2022). Hence, to achieve balanced data, we would need SOZ from at least 2,000 subjects, a sample size that is currently not available.

There has been one prior unsupervised approach by Zhang et al. (2015); however, it was applied to DRE adults and achieved a sensitivity of 78% and a specificity of 57%. We cannot replicate that study for this paper, because specific information about parameter settings was not discussed in Zhang et al. (2015).

The difference between EPIK and supervised ML

In this study, we present a novel, unsupervised technique to identify SOZ localizing ICs that require no prior dataset for training and classify ICs by encoding expert knowledge. The unsupervised nature of our algorithm implies that the entire dataset is used as a test set and no training dataset is required. Our algorithm is tested on the largest number of children with DRE among the recent studies on automated SOZ identification mechanisms with rs-fMRI listed in Table 2. Figure 2 illustrates differences from Figure 1. ML techniques (Figure 2) utilize examples of SOZ and RSN ICs to learn a model in the training phase, which is subsequently used for the identification of SOZ on previously unseen rs-fMRI signals. Such techniques have not been successful, possibly for the following reasons: (1) SOZ biomarkers are not precise and exhibit significant individual variances (Hunyadi et al., 2015; Boerwinkle et al., 2017) and (2) patients have low numbers of SOZ localizing ICs as compared to noise and RSN, leading to an imbalance in data and potential overfitting of ML models.

In the current study, EPIK (ExPert Knowledge-based IC categorization; Figure 2) is used with an alternative approach. Instead of directly learning SOZ-related features from training data, EPIK first used expert rules in a waterfall technique to purge noise ICs. Noise markers used by EPIK such as clusters outside brain boundaries or overlapping white matter or arteries are well-established, evidenced by consistency across several publications (Kelly et al., 2010; Griffanti et al., 2014, 2017). It then used SOZ-specific spatial and temporal markers in a maximum likelihood-based clustering to further classify the ICs into RSN and SOZ. Clustering was unsupervised and did not implement training with prior data to tune its parameters.

To illustrate differences compared to prior, we replicated the shallow learning strategy of Hunyadi et al. (2014) and implemented a Convolution Neural Network (CNN) based DL technique (Krizhevsky et al., 2012; Cui et al., 2018; Nozais et al., 2021) for the identification of SOZ localizing ICs from rs-fMRI, thereby providing a preliminary comparative study of all three approaches on the same dataset of children with DRE

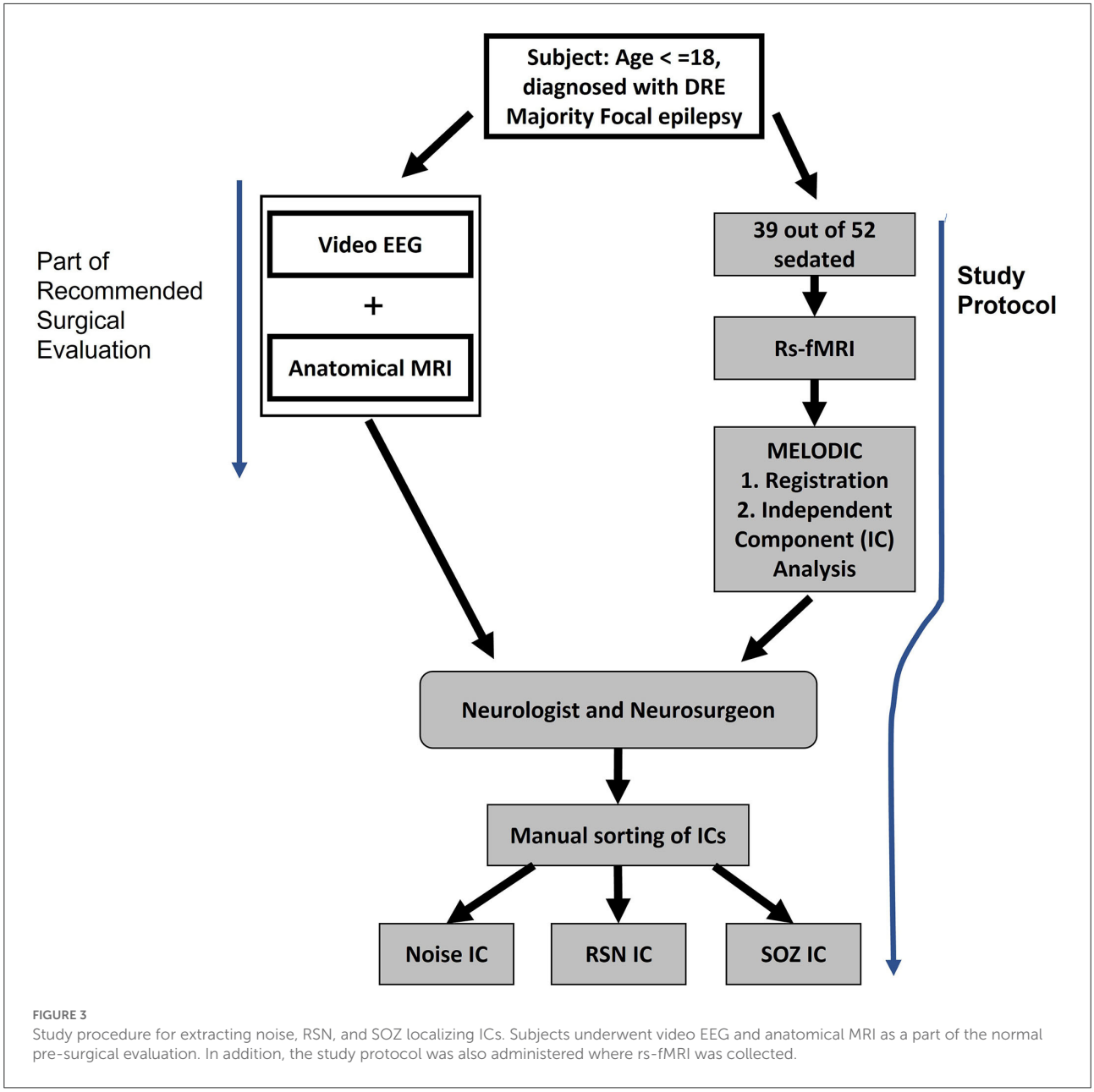


FIGURE 3
Study procedure for extracting noise, RSN, and SOZ localizing ICs. Subjects underwent video EEG and anatomical MRI as a part of the normal pre-surgical evaluation. In addition, the study protocol was also administered where rs-fMRI was collected.

using the standard metrics of accuracy, precision, specificity, and sensitivity. We hypothesize that EPIK will perform at least equally well as prior methods and consistently across age and sex, due to being informed by developmental- and sex-informed expert sorting in the pediatric DRE population.

Materials and methods

Inclusion criteria

Patients who were determined to have DRE by a treating epileptologist and received surgery evaluation were included. Most of the patients had focal epilepsy; however, rapid generalization of epileptiform activity from an epileptogenic focus may appear to be generalized epilepsy when evaluated using surface EEG. Hence, generalized epilepsy was not an exclusion criterion.

Data collection method

The rs-fMRI data from 52 children with DRE aged 3 months−18 years old, who were under the care of a treating epileptologist at Phoenix Children's Hospital (PCH), were

selected in descending alphabetical order from the PCH clinical database (Age and sex distribution provided in Table 3). The diagnosis of DRE was according to the treating epileptologist's documented medical record notes. The children received rs-fMRI, video EEG, and anatomical MRI as part of standard clinical MRI SOZ localization for epilepsy surgery evaluation (Figure 3). For rs-fMRI, patients who were determined to require conscious sedation, received a propofol infusion as a part of standard care determined by the institution's policies. Of the 52 children, 39 required conscious sedation. The dataset included patients who had <1 mm head motion in any direction during scanning. For children who received sedation, propofol administered at levels to produce conscious sedation (80–110 micrograms/kilogram/minute), avoiding higher dosages typical of general anesthesia, was utilized. Propofol administered at levels producing conscious sedation reduces the BOLD signal strength by ∼10%, still allowing for complete network detection (Vanderby et al., 2010; Schrouff et al., 2011). General anesthesia causes gross loss of ability to detect the large-scale cortical networks and, was, therefore avoided.

As part of the standard of care, the children also received inpatient video EEG and anatomical MRI. This data also aided the manual identification of SOZ localizing ICs in rs-fMRI (Figure 3).

The MRI images were acquired using a 3T MRI unit from Ingenuity Philips Medical systems. It has a 32-channel head coil. The resting state fMRI parameters were set at TR 2,000 ms, TE 30 ms, matrix size 80 × 80, flip angle 80, number of slices 46, slice thickness 3.4 mm with no gap, in-plane resolution 3 × 3 mm, interleaved acquisition, and number of total volumes 600, in two 10-min runs, with a total time of 20 mins.

rs-fMRI pre-processing

Oxford Centre FMRIB (Functional MRI of the Brain) Software Library tool MELODIC (Beckmann and Smith, 2004) was used to analyze the rs-fMRI and extract ICs as detailed in a previous study (Boerwinkle et al., 2019). Pre-processing included deletion of the first five volumes to remove T1 saturation effects, passing through a high-pass filter at 100 s, slice time correction, spatial smoothing of 1-mm full-width at half maximum, and motion corrected by MCFLIRT (Jenkinson et al., 2002), with non-brain structures removed.

Linear registration was performed between the individual functional scans and patients' high-resolution anatomical scans (Jenkinson and Smith, 2001), which was further optimized using boundary-based registration (Greve and Fischl, 2009). Individual rs-fMRI data sets then underwent independent component analysis (ICA) as previously reported (Boerwinkle et al., 2017).

Expert rs-fMRI evaluation methodology

The SOZ was evaluated by the expert epilepsy surgery conference team and deemed to be consistent with the other acquired data (video EEG and anatomical MRI) with high enough evidence to surgically target the SOZ. Further, the confirmation that the SOZ was deemed true by the treatment team was evidenced by the Engel I and II scores 1 year post-operatively.

The ICA results were viewed by two blinded reviewers (one neurologist and one neurosurgeon) who sorted the ICs into three categories—noise, resting-state network, and rs-fMRI SOZ—by the criteria below. In case of disagreement between the first two reviewers, the opinion of a third reviewer (a neurologist) was used to make the final determination. In this study, there was no disagreement between the blinded reviewers for the selected subjects.

Rs-fMRI was categorized into noise, resting state network (RSN), and SOZ using the following criteria.

Noise category

Consistent noise markers in rs-fMRI are reported in the literature (Hunyadi et al., 2015; Boerwinkle et al., 2019). The noise markers reported in different manuscripts are summarized in Table 4.

RSN category

These are activations in the MRI images that are spatially located in established anatomical regions. Such regions are highlighted in literature (Boerwinkle et al., 2019) and include "primary sensory motor networks located in the bilateral face area, the bilateral leg area, and the unilateral right- and left-hand regions; language networks primarily located within the left and right inferior frontal gyrus, posterior–superior temporal gyrus, posterior–superior temporal sulcus, posterior–middle temporal gyrus, and the supramarginal gyrus; parietal networks primarily located within the bilateral homologous parietal gyri; frontal networks primarily located within the bilateral premotor, and homologous bilateral frontal gyri; temporal networks primarily located within the bilateral homologous anterior and posterior temporal regions; visual networks located within the bilateral homologous primary and secondary visual association cortices; the default mode network located primarily within the bilateral posterior cingulate gyrus, precuneus, inferior parietal lobules, and medial prefrontal cortex; and the deep gray networks located with the bilateral putamen and bilateral mesial thalami."

TABLE 3 Patient distribution and information about the data set.

Number of subjects	52
Age $\leq$ 5 years	20
Age $>$ 5 and $\leq$ 13	18
Age $>$ 13 and $\leq$ 18	14
Men/Women	23/29
Prior surgery	2
Surgery post resting-state fMRI	24 (ablation 15, resection 7, disconnection 2)
Seizure free post-surgery, and rs-fMRI SOZ is the same location as the region destroyed determined by expert review of pre-operative rs-fMRI SOZ and post-operative imaging	16 (ablation 10, resection 6, disconnection 1)

SOZ category

SOZ characteristics consist of two types of features: (a) spatial features and (b) temporal features.

Spatial features

The activation must be located within the gray matter while not overlapping with the RSN spatial patterns. It must have a bullseye pattern, where two or more overlapping abnormal neuronal IC can be identified, may have an alternating activation and deactivation pattern that does not overlap noise zones, (noise IC characteristics 2, 3, and 4 in Table 4), may extend to ventricles through white matter, and may have irregular borders.

Temporal features

The SOZ BOLD signal power spectra must contain dominant frequencies $>$0.073 Hz, the rs-fMRI SOZ must have power spectra at higher frequencies than RSN, and the BOLD time series may have irregular patterns.

The rs-fMRI IC were sorted by an expert and reported to the clinical epilepsy surgery evaluation team. The data includes ICs extracted using the MELODIC module in FSL (Beckmann and Smith, 2004). Table 3 provides the age and sex distribution and surgical outcome statistics.

Ethics statement

Institutional IRB for retrospective analysis for this project was approved by the PCH Institutional Review Board (20–358), who determined that, since the retrospective rs-fMRI for these subjects was collected as part of a standard-preoperative MRI, no additional consent procedures were required.

Data/code availability statement

The data were deidentified according to the National Institutes of Health (NIH) Privacy Rule permits and made available for research application. Further, in accordance with the open science policy, we will provide interested researchers access to EPIK to enable them to reproduce our results.

EPIK method

EPIK (Figure 4) considers noise markers for ICs in an rs-fMRI, as documented in several studies including Griffanti et al. (2014, 2017). The method applies rules in a waterfall technique to classify an IC as noise (Figure 4). If an IC is not noisy, then it classifies the IC as either an RSN or a SOZ. In detail, there are six expert-derived rules for IC noise markers, combined from Boerwinkle and Hunyadi's works (Table 4). Automated application of such rules necessitates the development of the following key components:

a) Voxel cluster detection algorithm: A density-based scanning approach is undertaken to derive voxel clusters (upper panel Figure 4). The algorithm takes two configurable inputs: neighborhood, which includes a distance metric and a value ϵ, and the minimum number of nearby voxels $\nu_{\min}$. If a voxel has more than $\nu_{\min}$ voxels in the ϵ neighborhood, then it is marked as a core point of a cluster. If a voxel is not a core point but is in ϵ neighborhood of a core point, then it is identified as a border point. All other points are ignored from clusters. Core points, that are in ϵ neighborhood of each other, are combined into one cluster, and border points are assigned to the cluster of the nearest core point. The output of this step is the set of clusters in each IC slice.
b) Brain boundary/periphery detection: Contours in the brain are derived using a Sobel filter-based edge detection technique (Figure 4; Chakraborty et al., 2020). The lowest intensity contour is most likely the outer contour of the brain. However, the cerebrospinal fluid and blood vessels also present as low-intensity contours. The method searches for the contour that encompasses all other contours, which gives us the brain periphery.
c) White matter detection: The white matter manifests as the brightest contour in the brain. The blood vessels and cerebrospinal fluid in the white matter contour are discarded.
d) Blood vessel detection: The major basal-region blood vessels present themselves as low-intensity contours encompassed in the brain periphery contour.
e) Noise IC classification: Utilizing the a, b, c, and d steps, an IC can be classified as noise (Figure 5). From each slice of an IC, the clusters and the contours are extracted. An overlapping cluster can cause the contour detection algorithm to fail

in extracting the peripheral, the white matter, and the blood vessel contours. In the initial pass through the ICs, EPIK obtains a version of each slice devoid of clusters, which is subsequently used to identify contours. The algorithm then reruns through each slice of an IC and performs cluster detection. It then evaluates the overlap of the largest cluster with the brain boundary (first row in Figure 5) and the intersection of the largest cluster with the white matter and blood vessels (third row in Figure 5). The output of the first stage classifier (upper panel in Figure 4) is a statistic for each slice on the cluster size, the percentage (%) overlap with the brain boundary, the blood vessels, and the white matter for each cluster in a slice.

Each IC has multiple slices (around 55 for PCH dataset). The second stage classifier sorts the slices in decreasing order of cluster size (lower panel of Figure 2). It selects the top 10 slices and checks the percentage overlap to determine noise slices. If the majority of the top 10 slices are noise, the IC is classified as a noise IC. If the IC passes through the majority evaluation, it is passed to the second-level classifier, which determines if it is a normal RSN or an SOZ (Figure 4). The SOZ classification is based on expert guidance on the SOZ markers in ICs, as documented in Hunyadi et al. (2013; 2015; Table 4).

f) BOLD signal feature extraction: The BOLD signal was first divided into windows of length 256 samples. Four levels of activelet transformation coefficients using "*à trous*" algorithm with exponential-spline wavelets were extracted from each window. Sparsity in the activelet coefficients was evaluated using the Gini Index metric (Lerman and Yitzhaki, 1984). A Gini Index of >0.75 is sparse. If an IC is classified as white matter noise, then it can be classified as an SOZ if the Gini Index in the BOLD signal is >0.75. In addition, sparsity in matching pursuit using a sine dictionary limited to frequencies between 0.01 and 0.1 Hz was also evaluated using the Gini index. If an IC was classified as white noise, then it can be classified as SOZ if the BOLD signal Gini Index in the sine dictionary matching pursuit is >1.72.

TABLE 4 Noise markers in fMRI IC.

Noise independent component (IC) characteristics
1. A large number of small voxel clusters
2. Cluster peaks in the white matter
3. High overlap with the white matter, the cerebrospinal fluid, or the blood vessels
4. Crescent shape aligning with the brain boundary
5. Sudden changes in the oscillation pattern in the BOLD signal
6. Located within area of signal loss

DL strategy for SOZ localization

Nozais et al. (2021) recently proposed a DL-based technique where a multi-layer perceptron (MLP) is trained on 12,690 RSNs from 282 participants. As such, it does not incorporate

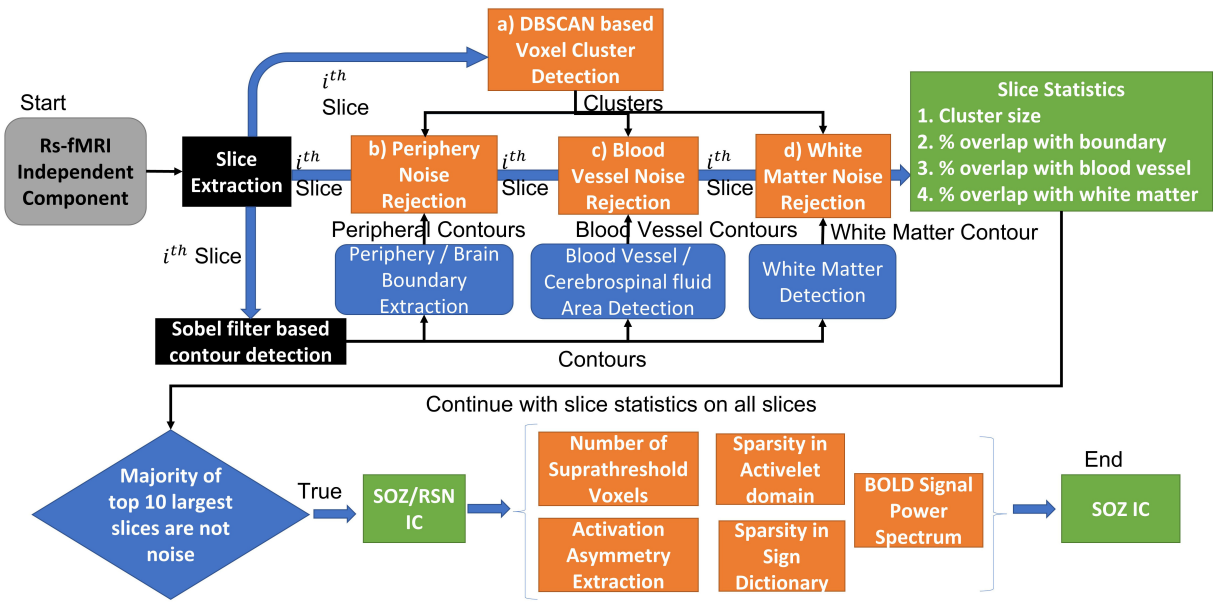


FIGURE 4
EPIK: Unsupervised approach for SOZ classification. The EPIK method is applied on individual IC to classify it as SOZ or RSN or noise.

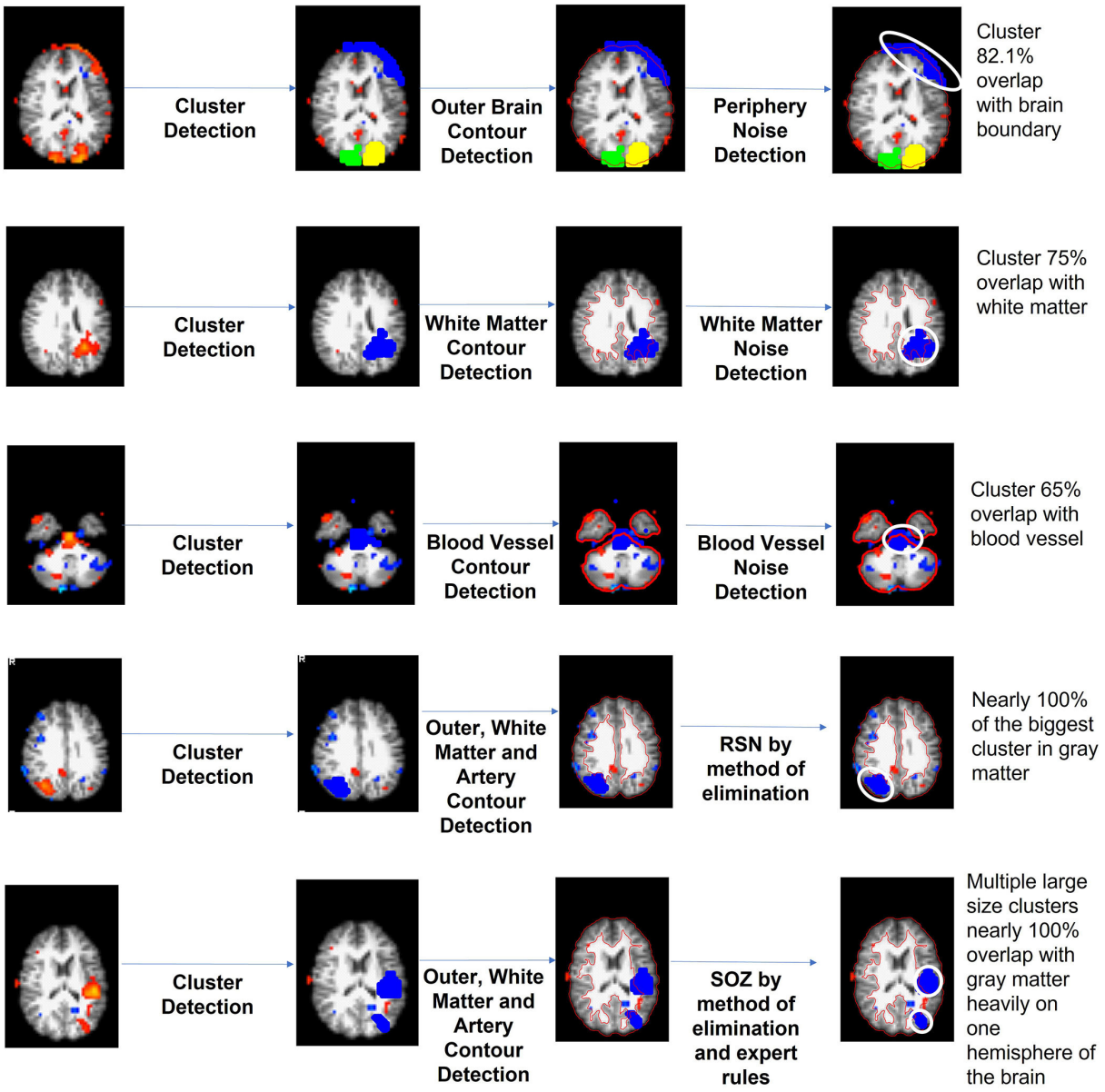


FIGURE 5
Demonstration of unsupervised IC sorting mechanism.

any expert knowledge but instead attempts to build its own hypothesis from examples. The technique has not been used to classify SOZ and can currently only identify RSNs. We implemented CNN-based DL for SOZ localization.

For the CNN technique, hyperparameter tuning is one of the most important steps. A KerasTuner was implemented to get the optimal values of the hyperparameters. We used a hyperband algorithm with the objective of least validation loss to select the best model of CNN by optimizing the following hyperparameters:

- Number of layers: [3; 4; 5]
- Number of units/filters per layer: min_value = 32, max_value = 512, default = 128.
- Learning rate: [10^{-2}; 10^{-3}; 10^{-4}]
- Dropout rate: [0; 0.2; 0.33; 0.4; 0.5; 0.66].

We used 4,212 ICs for training and 1,404 ICs for validation in the hyperparameter tuning process. The input shape of the IC image was downsampled from 1,006 × 709 × 3 to 270 × 400 × 3 during preprocessing. Binary cross-entropy was used as a loss function, and Adam was used as an optimizer. To avoid the

overfitting problem, "dropout" and "early stopping" strategies were implemented. "ReLU" (rectified linear unit), being more computationally efficient, was used as an activation function for the input and hidden layers, and the "Sigmoid" activation function was used for the output layer. For CNN, weights were initialized using the "He uniform" initializer.

Shallow learning strategy

The technique proposed by Hunyadi et al. (2013, 2014, 2015) was replicated. The rs-fMRI image and BOLD signal features were extracted from the IC images. From the entire pool of ICs, 60% of the data were randomly sampled to be used as training data. The remaining 40% were used for testing. The features extracted from the rs-fMRI image and BOLD signal were used to train a Least Squares Support Vector Machine (LS-SVM), as described by Hunyadi et al. (2013, 2014, 2015).

The following features were extracted from each IC following the study of Griffanti et al. (2014): (a) number of clusters greater than a pixel size of 135; (b) asymmetry of an IC using the difference in the z-scored value of the voxels in the left hemisphere and their contralateral voxels in the right hemisphere; (c) sparsity in activelet basis using the Gini index metric; (d) sparsity in sine basis using the Gini index metric. The SVM was then trained for the two-class classification task [(either RSN vs. noise) or (SOZ vs. non-SOZ)]. We utilized two kernels: radial basis function (RBF) and linear kernel. The performance for the linear and RBF kernels was similar and, hence, followed Occam's Razor theory; in this manuscript, we only report the performance for the linear kernel.

Metrics and statistical analysis method

We evaluated the performance of each approach for two objectives: (a) noise IC removal and (b) SOZ localizing IC identification. For the first objective, we defined true positives (TP) as ICs that are classified as RSN or SOZ by both expert and the automated approach, true negatives (TN) as ICs that are classified as noise by both the expert and the automated approach, false positives (FP) as ICs classified as noise by the expert but RSN or SOZ by the automated approach, and false negatives (FN) as ICs classified as RSN or SOZ by the expert but noise by the automated approach. For the second objective, we define TP as ICs classified by both the expert and the automated approach as SOZ or RSN, TN as ICs classified by both the expert and the automated approach as not SOZ, FP as ICs classified as non-SOZ by the expert but SOZ by the automated approach, and FN as ICs classified as SOZ by the expert but non-SOZ by the automated approach. From these, we derived accuracy, precision, sensitivity, and specificity.

We evaluated the statistical significance of a difference in performance metrics between the two approaches by utilizing a one-sided paired t-test. The alternate hypothesis was that there is a positive non-zero difference between EPIK and any other approach (LS-SVM or CNN). The alternate hypothesis was rejected if the p-value for the paired t-test was <0.05.

We also evaluated the effect of age and gender on each approach using a mixed-effects model with each parameter as the observation variable and age or gender as the predictor variable. A random effect on the patient ID was also introduced. For each algorithm, a separate mixed-effects model was generated for each metric and for each predictor variable, i.e., age/gender.

Results

Overall identification results

We compared the performance of EPIK with two competing ML-based approaches: shallow learning (LS-SVM) and deep learning (CNN). For the ML-based approach, training data was used from every subject. This is also known as the user-dependent (Bhakta et al., 2020) supervised classification approach and gave us the best performance metrics. For EPIK, no such training set is needed. The results in Table 5 show that EPIK outperforms both LS-SVM and CNN approaches for SOZ localizing IC identification tasks. The CNN approach is more accurate in noise removal but performs poorly in the SOZ identification task.

EPIK has high sensitivity in the SOZ identification task with a low number of FNs. This implies that EPIK rarely misses any SOZ localizing IC. The LS-SVM approach is poor in noise removal, but its performance improves for the SOZ identification task. The confidence interval is specified as [a,b] for metrics with a p-value < 0.5.

Performance variation with age and gender

Table 6 shows the variation of the performance metrics for EPIK, LS-SVM, and CNN with respect to age and gender. The accuracy, precision, and sensitivity of EPIK for noise removal do not have a statistically stable dependence on age or gender. The specificity of EPIK for noise removal decreases with age, resulting in more FPs, where noise is categorized as RSN or SOZ. For the SOZ identification task, there is a statistically significant trend for sensitivity to increase and specificity to decrease with age. This implies that, as age progresses, EPIK tends to classify more RSN or noise as SOZs; however, fewer SOZs are ignored as noise. Consequently, EPIK is observed to have an

TABLE 5 Overall RSN or SOZ identification results for the three approaches.

Approach	RSN or SOZ vs. noise				SOZ vs. non-SOZ (RSN or noise)				Key observations
	Accuracy	Precision	Sensitivity	Specificity	Accuracy	Precision	Sensitivity	Specificity	
EPIK (this paper)	71.7%	73.1%	72%	73.7%	84.7%	74.1%	88.6%	81.9%	Best performance for SOZ identification
LS-SVM (Hunyadi et al., 2014)	61.8%	52.2%	43%	73.6%	80.7%	52.2%	72.1%	78.7%	High false positives and false negatives Significant variance across patients
One sided t-test for +ve difference between EPIK and LS-SVM	p-value = ∼0 [5, 15.2]	p-value = ∼0 [18.7, 27.4]	p-value = ∼0 [27.4, 45.9]	Rejected p-value = 0.9	p-value = ∼0 [2, 6.5]	p-value = ∼0 [20.7, 29.1]	p-value = ∼0 [14.1 25.3]	Rejected p-value = 0.06	
CNN (Nozais et al., 2021)	82.45%	82.7%	82.1%	81.5%	73.5%	28.5%	97.7%	42.85%	Best RSN identification performance. Poor SOZ performance due to lack of hand sorted SOZ IC examples.
One sided t-test for +ve difference between EPIK and CNN	Negative change P-value ∼ 0 [−5.1, −13.2]	Rejected P-value = 0.6	Negative change P-value ∼0 [−7, −12.1]	Negative change P-value = 0.02 [−4.1, −9]	P-value ∼0 [8.3, 15.7]	P-value ∼ 0 [51.2, 60]	Negative change P-value ∼ 0 [−4.2, −11.3]	P-value = 0.001 [31.6, 45.2]	

10.3389/fnimg.2022.1007668

TABLE 6 Age- and sex-segregated metrics for the unsupervised IC classification algorithm, the LS-SVM approach by Hunyadi et al., and the CNN deep learning approach.

Metric	Algorithm	0 <Age ≤5 ($N = 20$)	5 < Age ≤13 ($N = 18$)	13 < Age ≤18 ($N = 14$)	*P*-value fixed effects on age	Men ($N = 23$)	Women ($N = 29$)	*P*-value fixed effects on sex	Key observations
Noise vs. network/SOZ performance metrics									
Accuracy	EPIK	**69.4% (±9%)**	**74% (±8.2%)**	**71.7% (±5.8%)**	**0.32**	**73.8% (±5.3%)**	**70% (±9.4%)**	**0.04**†	CNN gives the best RSN identification accuracy for all age categories. followed closely by EPIK
	LS-SVM	55.8% (±11.5%)	63.7% (±7.7%)	65.3% (±8.4%)	0.004†	63.6% (±9.5%)	59.1% (±10.6%)	0.06	LS-SVM is poorest in identifying RSN since it only considers SOZ markers in ICs.
	CNN	73.2% (±4.5%)	76.1% (±0.6%)	80.2% (±5.8%)	∼0	72.8% (±8.2%)	77.4% (±4.7%)	0.09	Success of CNN can be attributed to availability of a significant number of normal RSN ICs ($n = 2,427$)
Precision	EPIK	**74.9% (±16.2%)**	**73.6% (±13.7%)**	**66.5% (±10%)**	**0.048**	**73.5% (±11.5%)**	**71.1% (±16%)**	**0.27**	
	LS-SVM	55.6% (±32.4%)	52.8% (±15.9%)	46.5% (±18%)	0.3	54.8% (±22.2%)	50.1% (±25.4%)	0.24	
	CNN	68.2% (±11.7%)	75.2% (±1.5%)	75.4% (±7.3%)	∼0	69.2% (±13%)	75.91% (±15.1%)	0.3	
Sensitivity	EPIK	**63% (±18%)**	**76.6% (±9.3%)**	**76.8% (±9.7%)**	**0.001**†	**75.2% (±10.7%)**	**68.43% (±16.9%)**	**0.047**	
	LS-SVM	27.5% (±25.9%)	50.8% (±26.9%)	55.1% (±22%)	0.001†	52.6% (±28.9%)	35.4% (±24.9%)	0.012†	
	CNN	86.09%	81.5%	85.96%		78%	79.5%		
Specificity	EPIK	**79% (±13.8%)**	**72.7% (±17.1%)**	**68.2% (±8.4%)**	**0.01**†	**73.4% (±13.5%)**	**74.2% (±15.2%)**	**0.41**	
	LS-SVM	80.5% (±17.4%)	68.8% (±23%)	70% (±10%)	0.035†	71.3% (±17%)	75.5% (±19.9%)	0.2	
	CNN	60.5%	70.8%	75%		67.9%	75.31%		
SOZ identification metrics									
Accuracy	EPIK	**87.5% (±7.6%)**	**83.5% (±9.6%)**	**82.2% (±6.1%)**	**0.025**†	**84.6% (±6.7%)**	**84.7% (±9.3%)**	**0.48**	EPIK has the best performance for SOZ localizing IC identification
	LS-SVM	85.3% (±6.6%)	77.2% (±9.4%)	78.6% (5.7%)	0.008†	79.5% (±8.8%)	81.6% (±7.7%)	0.17	EPIK has consistent performance across age.
	CNN	75.5% (±27.7%)	75.3% (±26.6%)	76.5% (±21%)	0.8	71% (±28.2%)	73% (±30.2%)	0.44	EPIC has the best performance for children of age <5 years. This is a key benefit because it is known that earlier surgery for epilepsy yields better surgical and developmental outcomes.
Precision	EPIK	**76.7% (±16.3%)**	**75.2% (±14.4%)**	**69.2% (±9.9%)**	**0.07**	**76.3% (±10.7%)**	**72.5% (±16.5%)**	**0.17**	

(Continued)

10.3389/fnimg.2022.1007668

TABLE 6 (Continued)

Metric	Algorithm	0 <Age ≤5 (N = 20)	5 < Age ≤13 (N = 18)	13 < Age ≤18 (N = 14)	*P*-value fixed effects on age	Men (N = 23)	Women (N = 29)	*P*-value fixed effects on sex	Key observations
	LS-SVM	62.5% (±17.2%)	56.9% (±15%)	51.4% (±15%)	0.06	54.7% (±15.7%)	55% (±16.3%)	0.2	
	CNN	53.8% (±50.2%)	50% (±51.4%)	45% (±50%)	0.035	45.9% (±49.9%)	54.4% (±50%)	0.54	
Sensitivity	EPIK	**86.8% (±8.8%)**	**89.4% (±6.8%)**	**90.4% (±7.6%)**	**0.085**†	88.1% (±7.1%)	89.1% (±8.5%)	0.34	
	LS-SVM	58.8% (±33.9%)	74.6% (±19.5%)	87.8% (±23%)	0.001†	78.83% (±25.6%)	66.7% (±30.4%)	0.065	
	CNN	11.1% (±3%)	0 (±0)	0 (±0)	0.001	20% (±5%)	3.44% (±3%)	0.002	
Specificity	EPIK	**86.6% (±12.8%)**	**79.9 % (±15.6%)**	**77.8% (±8%)**	**0.02**†	81.7% (±12.1%)	82% (±14.2%)	0.47	
	LS-SVM	86.4% (±14.1%)	73.2% (±21.8%)	74.9% (±9.3%)	0.015†	75.3% (±17%)	81.5% (±17%)	0.09	
	CNN	74.9% (±27.6%)	75.3% (±26.6%)	76.1% (±22%)	0.4	75.37% (±28.2%)	76.59% (±30%)	0.8	

†Indicates that the result has a p-value of < 0.05 and is statistically significant. Bold value refers to our technique EPIK's results.

accuracy >85% at ages below 5, which is higher than those previously reported.

The LS-SVM approach had consistently better performance for the SOZ identification task than noise removal. It also had the same pattern of increasing sensitivity and decreasing specificity with age. The LS-SVM approach had a higher variance in performance across subjects. This indicates that the hand-crafted features chosen by Hunyadi et al. may be less applicable to specific scenarios of DRE in children.

The CNN approach outperformed EPIK and LS-SVM for all age groups for noise removal. However, it had a lower performance for SOZ identification. In the training data, there were only 318 SOZ localizing ICs as opposed to 2427 RSN IC. This may have led to an underfitting of the CNN technique for SOZ identification. For the CNN technique in noise removal, both sensitivity and specificity increased with age. This potentially indicates that the CNN technique is finding novel hidden features from the ICs that are characteristic of RSN but not SOZ.

Overall, EPIK provided a consistent performance across the three age categories considered in this study compared to prior reported methods. Whereas, the ML techniques of Hunyadi et al. (2015) and CNN have significantly higher variance, possibly indicating inconsistent performance.

Performance on subjects undergoing surgery

Out of the 52 subjects considered in this study, 24 underwent surgery. The surgical outcomes were varied with 16 subjects becoming seizure-free (Engel I) after surgically destroying an expert-identified SOZ using rs-fMRI and seven having reduced post-operative seizure frequency (Engel II; Table 7). We focused on EPIK and LS-SVM for the SOZ identification task on the 24 subjects that underwent surgery because CNN had significantly poorer performance than the other two.

Table 7 shows that, for subjects whose post-operative outcomes are either seizure-free or have significantly reduced frequency, the agreement between EPIK and expert-hand classification is significantly high (88.9% sensitivity and 79% specificity). Although the LS-SVM approach has nearly similar accuracy as EPIK, the sensitivity is far lower in LS-SVM, with significant individual variance. To better understand the difference between EPIK and the LS-SVM approach, Figure 6 shows the receiver operating characteristics (ROC) curve for both EPIK and LS-SVM. EPIK exhibits higher sensitivity and specificity than LS-SVM, which appears to possibly sacrifice one for the other.

For patients undergoing ablation surgery, the specificity for EPIK was 82.9%, while the sensitivity was 88%. This

TABLE 7 Performance of EPIK and LS-SVM approaches for subjects undergoing surgery.

Age, years (months)	Sex	Pre-surgery frequency (per month)	Post-surgery frequency (per month)	Procedure	Method	Accuracy	Precision	Sensitivity	Specificity
18 (0)	W	1	0	A*	EPIK	82.8%	71.7%	82.5%	82.9%
					LS-SVM	78.7%	58.8%	76.9%	79.4%
14 (8)	M	1	0	A	EPIK	85.2%	75.5%	90.2%	82.1%
					LS-SVM	68.2%	48%	92.3%	58.1%
14 (7)	W	3	1 (66% reduced)	A	EPIK	86.5%	78.2%	95.6%	79.7%
					LS-SVM	71.4%	52%	100%	58.6%
14 (10)	M	1	0	A	EPIK	86.3%	77.1%	86.1%	86.4%
					LS-SVM	82%	35.7%	100%	80%
16 (3)	W	210	0	A	EPIK	86.5%	67.9%	97.4%	82.4%
					LS-SVM	91.3%	37.5%	100%	63.5%
10 (1)	W	240	0	A	EPIK	74.2%	60.5%	95.8%	60.5%
					LS-SVM	76%	60%	100%	62.5%
8 (2)	M	4	2 (50% reduced)	A	EPIK	83.3%	77%	94.1%	73.1%
					LS-SVM	85.4%	71.4%	83.3%	86.2%
15 (6)	W	12	0	A	EPIK	86%	62.9%	78.6%	88%
					LS-SVM	87.3%	0	0	97.9%
10 (5)	W	60	8 (87% reduced)	A	EPIK	85.5%	69.6%	91.4%	82.9%
					LS-SVM	83%	40%	66.7%	85.4%
3 (2)	W	60	1 (98% reduced)	A	EPIK	87.4%	78%	86.5%	87.8%
					LS-SVM	88.9%	81.8%	75%	94%
17 (9)	M	2	0	A	EPIK	74.3%	50%	88.9%	69.2%
					LS-SVM	75%	25%	66.7%	76%
11 (8)	M	1	0	R*	EPIK	80.6%	62.5%	65.8%	85.8%
					LS-SVM	82.8%	16.7%	16.7%	90.4%
4 (9)	W	2	0	R	EPIK	84.1%	72.7%	84.2%	84%
					LS-SVM	80.4%	14.3%	25%	85.7%
18 (1)	W	5	0	R	EPIK	72%	55.6%	87%	64.4%
					LS-SVM	71.4%	50%	75%	70%
10 (5)	W	8	0	R	EPIK	57.5%	44.1%	93.8%	38.7%
					LS-SVM	63.1%	58.8%	100%	22%
13 (8)	M	120	1 (99% reduced)	R	EPIK	82.9%	72.9%	93.5%	75.4%
					LS-SVM	77.8%	43.7%	87.5%	75.7%
2 (7)	W	2	0	D*	EPIK	84.5%	53.2%	100%	81.3%
					LS-SVM	81.8%	38.5%	100%	79.5%
2 (7)	M	720	0	A	EPIK	89.1%	81.3%	88.6%	89.3%
					LS-SVM	90.4%	50%	60%	93.6%
0 (3)	W	90	30 (66% reduced)	A	EPIK	98.8%	100%	92.9%	100%
					LS-SVM	97.1%	100%	50%	100%
2 (10)	W	300	0	R	EPIK	84.8%	74.5%	97.4%	75.5%
					LS-SVM	89.2%	71.4%	100%	85.2%
2 (11)	M	4	0	D	EPIK	69.1%	61%	96.2%	44.8%
					LS-SVM	81.8%	78.9%	100%	42.8%
2 (1)	W	3,000	0	A	EPIK	88.4%	45%	60%	91.6%
					LS-SVM	88.1%	25%	20%	94.4%
3 (6)	M	30	0	A	EPIK	88.6%	78.6%	91.7%	87%
					LS-SVM	78.6%	68.4%	100%	60%
1 (4)	F	180	0	R	EPIK	94.1%	94.3%	84.6%	98%
					LS-SVM	88.9%	50%	16.7%	98%

(Continued)

TABLE 7 (Continued)

Approach	Mean accuracy (SD)	Mean precision (SD)	Mean sensitivity (SD)	Mean specificity (SD)
Agreement with expert hand classification for seizure-free/reduced post-operative outcome				
EPIK	83.3% (8.43%)	69.45% (13.6%)	88.9% (9.6%)	79% (14.4%)
LS-SVM	80.45% (9.8%)	47.6% (23.9%)	72.5% (32.6%)	75.5% (19.5%)
Agreement with expert hand classification for ablation procedures				
EPIK	85.5% (5.8%)	71.6% (13.4%)	88% (9.3%)	82.9% (9.6%)
LS-SVM	82.8% (8.1%)	50.2% (24.9%)	72.7% (30.5%)	79.3% (15.3%)
Agreement with expert hand classification for resection procedures				
EPIK	79.4% (11.6%)	68.1% (16%)	86.6% (10.5%)	74.5% (18.9%)
LS-SVM	79.1% (9.4%)	43.6% (21.1%)	60.1% (39.1%)	75.3% (25.2%)

The table also shows the difference in agreement between the automated approach and the expert hand classification for subjects with seizure-free/reduced frequency post-operative outcomes and without any change in seizure frequency.
*A, Ablation; R, Resection; D, Disconnection.

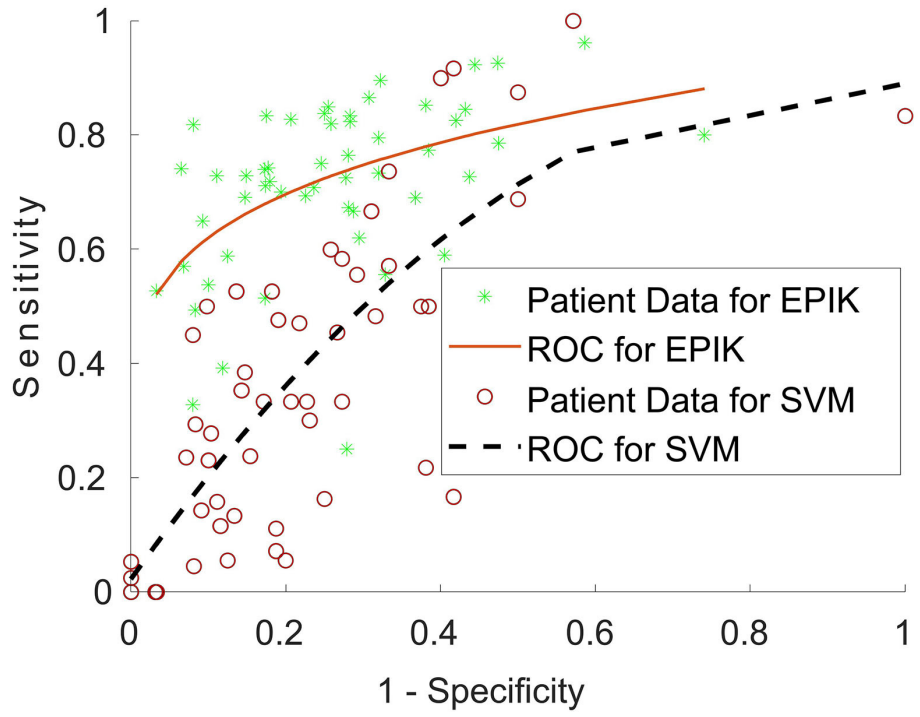


FIGURE 6
Receiver operating characteristics (ROC) curve for EPIK and the LS-SVM approach for patients undergoing surgery. A curve close to the top left-hand corner of the graph is favorable and shows a balance between sensitivity and specificity.

is preliminarily an encouraging result, given that ablation is minimally invasive and thus largely accepted as less risky than resection. The specificity and sensitivity in EPIK for patients undergoing resection reduce to 79.5 and 86.6%, respectively. Of the 15 subjects who underwent the ablation procedure, 10 were seizure-free (Engel 1 outcome), which is slightly better than recently reported statistics [66% in this study vs. 60.4% reported in Kanner et al. (2022)]. Supplementary Table 1 gives the SOZ location and fMRI evidence of SOZ for all subjects in the study.

Reduction in IC sorting effort for the neurosurgeon/neurologist

The ICs marked as SOZ by the EPIK method can be supplied to the neurosurgeon or neurologist for localization of SOZs in the brain. The number of SOZ classifications in EPIK per subject is 22 (±4). Out of 22, 16 are true positive SOZ ICs, two are noise ICs, and four are RSN ICs. These ICs are then evaluated by the neurosurgeon or neurologist for determining SOZ in the brain.

This implies that there is ∼5 times reduction in the number of ICs to be analyzed by the neurosurgeon or neurologist. This can significantly aid in presurgical screening by reducing the cognitive burden of the neurosurgeon or neurologist and improving the accuracy of the SOZ identification.

Discussion

A strength of EPIK, which may increase its utility, is that it does not require any prior training data and hence it uses a plug-n-play IC sorting method. EPIK combines spatial and temporal markers specific for RSN and SOZ, which results in possibly equivalent or better performance than prior methods. The waterfall technique removes the number of noise ICs using well-established expert rules; hence, it may reduce false positives and increase true positives of SOZ localizing ICs.

For subjects with good postoperative outcomes, there was excellent agreement between expert hand sorting and EPIK-based SOZ localizing IC identification. Also, EPIK appeared to perform well in those <5 years of age, in whom surgery yields improved developmental outcomes (Pindrik et al., 2019; Perry and Shandley, 2021).

The LS-SVM approach did not perform as well for the noise identification task but did show a drastic improvement in performance for the SOZ identification task. This was expected because the hand-selected features proposed by Hunyadi et al. (2013, 2015) are specifically geared toward the SOZ identification task. However, LS-SVM exhibits significant variance in performance across subjects, resulting in inconsistent accuracy in this study. EPIK had a higher and more consistent balance in the identification of all three categories of IC compared to LS-SVM herein.

The CNN approach had a lower performance for SOZ identification. However, there was a significant improvement in the performance of the noise identification task. This can be explained by the difference in data availability for the two tasks. This gives confidence that CNN can perform better if given an adequate number of training-SOZ-localizing ICs; this could be an avenue for future research.

The general assumption in supervised machine learning is that elements from each class come from a unique distribution specific to the class. The ML technique then attempts to learn the differences in the distribution of each class and evaluate the best fit distribution for the test data. The fundamental limitation of the LS-SVM approach is that SVM is inherently a two-class classifier. Although there are multi-class versions of SVM, the multi-class classification is performed in stages, where each stage is a two-class classifier. For rs-fMRI sorting, this would mean that the RSN and noise class will have to be combined into one composite class, while the SOZ ICs are labeled as the class of interest. In rs-fMRI, noise ICs are composed of several different categories of noise such as peripheral noise, white matter noise, and artery noise. Each such noise characteristic has different feature distributions; however, they are considered to be the same class by the supervised ML technique. Moreover, the noise class is combined with the RSN class to make a non-SOZ composite class. Hence, the non-SOZ composite class for rs-fMRI ICs has a composite distribution. As such, it is very difficult for the supervised ML classifier to learn the unique distribution of the non-SOZ IC class. A way around this is to learn each kind of noise and RSN separately. However, that requires data for each kind of noise from each patient. This cannot be guaranteed in a practical real-life setting.

The performance of the CNN-based DL strategy suffered because of the differences in the size of the three classes. RSN and noise classes had a nearly balanced data size, and the CNN strategy had good performance in distinguishing between them. However, since there are very limited SOZ IC examples, the CNN strategy could not reliably identify them.

The unsupervised technique utilizes expert knowledge and image processing algorithms to detect each kind of noise without the need for training a machine. Hence, it learns the noise characteristics without utilizing noise data from each patient. This capability of the unsupervised technique to employ specific algorithms for each type of noise and RSN is one of the major reasons for its success in separating noise, RSN, and SOZ ICs.

Limitations and future directions

This study ($n = 52$) evaluated a small group of data, and prior automated methods perform well on small samples but have reduced performance on larger datasets; hence, EPIK needs large set validation, which is a future direction. Larger datasets for focused performance evaluation within each age bracket, including young and separately older adults, and the very young vs. middle childhood are needed. Subtypes of epilepsy—acquired, congenital/genetic—and surgical approaches' success metrics should be statistically evaluated with acceptable power. Last, repeat studies in the same individuals over time would increase knowledge of the validity and reproducibility of the tool.

The majority of the subjects in this study received propofol infusion for sedation as part of a standard of clinical care for epilepsy surgery evaluation. Head motion maximum was <1 mm of frame wise displacement in any direction. Although propofol use has minimal effect on the overall rs-fMRI BOLD signal, it puts small but additional risks on the child (Pizoli et al., 2011). Several research studies proposed alternate methods of reducing head motion by engaging the child with videos and post-processing by measuring and accounting for head movements through computational methods (Dosenbach et al., 2017; Greene et al., 2018; D'Andrea et al., 2022). An area of future study is to evaluate the effect of sedation on the EPIK SOZ identification accuracy and integration of live motion monitoring and reduction-based approaches toward the elimination of head movement artifacts.

Conclusion

1. EPIK identified seizure onset zone (SOZ) localizing resting-state fMRI-independent components in children with drug-resistant epilepsy with an accuracy of 84.7% in this preliminary study.
2. EPIK can reduce the number of potential ICs to be analyzed by the neurosurgeon by ∼5-fold, hence significantly reducing the time commitment for pre-surgical evaluation.
3. EPIK is unsupervised and does not need any prior example of SOZ and works by codifying expert knowledge about fMRI noise and SOZ markers.
4. EPIK had consistent performance across age and gender and has been validated with surgical outcomes.
5. EPIK appeared to perform best for those under 5 years of age and thus may enable successful surgeries early in their life, potentially improving long-term postoperative outcomes.
6. EPIK preliminarily performed as well or better than shallow and deep learning systems for the identification of SOZ localizing ICs in a resting-state fMRI.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by Phoenix Children's Hospital. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

AB was responsible for designing (with inputs from SKSG), implementing, writing, and statistical analysis of the EPIK methodology. PK was responsible for implementing, writing, and analyzing the DL technique for the automated classification of noise, RSN, and SOZ ICs. SW was responsible for rs-fMRI data's administrative, technical, and material support. BS was responsible for subject identification and data collection. SG was responsible for the revision of this manuscript and validating the authenticity of this study. VB conceived the project, evaluated the stepwise methodological design and intermediate results, responsible for providing the motivation and perspective of this work with respect to epilepsy care, analysis, and interpretation of rs-fMRI data, and providing explanation and guidelines of noise, RSN, and SOZ biomarkers. All authors have read and approved the final manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnimg.2022.1007668/full#supplementary-material

References

Aaberg, K. M., Gunnes, N., Bakken, I. J., Søraas, C. L., Berntsen, A., Magnus, P., et al. (2017). Incidence and prevalence of childhood epilepsy: a nationwide cohort study. *Pediatrics* 139, e20163908. doi: 10.1542/peds.2016-3908

Bandt, S. K., Bundy, D. T., Hawasli, A. H., Ayoub, K. W., Sharma, M., Hacker, C. D., et al. (2014). The role of resting state networks in focal neocortical seizures. *PLoS ONE* 9, e107401. doi: 10.1371/journal.pone.0107401

Bansal, L., Miller, I., Hyslop, A., Bhatia, S., Duchowny, M., and Jayakar, P. (2016). PET hypermetabolism in medically resistant childhood epilepsy: incidence, associations, and surgical outcome. *Epilepsia* 57, 436–444. doi: 10.1111/epi.13311

Beckmann, C. F., and Smith, S. M. (2004). Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE Trans. Medical Imag.* 23, 137–152. doi: 10.1109/TMI.2003.822821

Begley, C. E., Famulari, M., Annegers, J. F., Lairson, D. R., Reynolds, T. F., Coan, S., et al. (2000). The cost of epilepsy in the United States: an estimate from population-based clinical and survey data. *Epilepsia* 41, 342–351. doi: 10.1111/j.1528-1157.2000.tb00166.x

Berger, A., Cohen, N., Fahoum, F., Medvedovsky, M., Meller, A., Ekstein, D., et al. (2021). Preoperative localization of seizure onset zones by magnetic source imaging, EEG-correlated functional MRI, and their combination. *J. Neurosurg.* 134, 1037–1043. doi: 10.3171/2020.3.JNS192794

Bhakta, K., Camargo, J., Donovan, L., Herrin, K., and Young, A. (2020). Machine learning model comparisons of user independent & dependent intent recognition systems for powered prostheses. *IEEE Robot. Automat. Lett.* 5, 3007480. doi: 10.1109/LRA.2020.3007480

Bharath, R. D., Panda, R., Raj, J., Bhardwaj, S., Sinha, S., Chaitanya, G., et al. (2019). Machine learning identifies "rsfMRI epilepsy networks" in temporal lobe epilepsy. *Eur. Radiol.* 29, 3496–3505. doi: 10.1007/s00330-019-5997-2

Boerwinkle, V. L., Cediel, E. G., Mirea, L., Williams, K., Kerrigan, J. F., Lam, S., et al. (2019). Network targeted approach and postoperative resting state functional MRI are associated with seizure outcome. *Ann. Neurol.* 86, 344–356. doi: 10.1002/ana.25547

Boerwinkle, V. L., Foldes, S. T., Torrisi, S. J., Temkit, H., Gaillard, W. D., Kerrigan, J. F., et al. (2018). Subcentimeter epilepsy surgery targets by resting state functional magnetic resonance imaging can improve outcomes in hypothalamic hamartoma. *Epilepsia* 59, 2284–2295. doi: 10.1111/epi.14583

Boerwinkle, V. L., Mirea, L., Gaillard, W. D., Sussman, B. L., Larocque, D., Bonnell, A., et al. (2020). Resting-state functional MRI connectivity impact on epilepsy surgery plan and surgical candidacy: prospective clinical work. *J. Neurosurg. Pediatr.* 2020, 1–8. doi: 10.3171/2020.1.PEDS19695

Boerwinkle, V. L., Mohanty, D., Foldes, S. T., Guffey, D., Minard, C. G., Vedantam, A., et al. (2017). Correlating resting-state functional magnetic resonance imaging connectivity by independent component analysis-based epileptogenic zones with intracranial electroencephalogram localized seizure onset zones and surgical outcomes in prospective pediatric in. *Brain Connectivity* 7, 424–442. doi: 10.1089/brain.2016.0479

Bouyssi-Kobar, M., De Asis-Cruz, J., Murnick, J., Chang, T., and Limperopoulos, C. (2019). Altered functional brain network integration, segregation, and modularity in infants born very preterm at term-equivalent age. *J. Pediatr.* 213, 13–21. doi: 10.1016/j.jpeds.2019.06.030

Bulacio, J. C., Jehi, L., Wong, C., Gonzalez-Martinez, J., Kotagal, P., Nair, D., et al. (2012). Long-term seizure outcome after resective surgery in patients evaluated with intracranial electrodes. *Epilepsia* 53, 1722–1730. doi: 10.1111/j.1528-1167.2012.03633.x

Chakraborty, A. R., Almeida, N. C., Prather, K. Y., O'Neal, C. M., Wells, A. A., Chen, S., et al. (2020). Resting-state functional magnetic resonance imaging with independent component analysis for presurgical seizure onset zone localization: a systematic review and meta-analysis. *Epilepsia* 61, 1958–1968. doi: 10.1111/epi.16637

Chen, Z., An, Y., Zhao, B., Yang, W., Yu, Q., Cai, L., et al. (2017). The value of resting-state functional magnetic resonance imaging for detecting epileptogenic zones in patients with focal epilepsy. *PLoS ONE* 12, e172094. doi: 10.1371/journal.pone.0172094

Chibbaro, S., Cebula, H., Scholly, J., Todeschi, J., Ollivier, I., Timofeev, A., et al. (2017). Pure endoscopic management of epileptogenic hypothalamic hamartomas. *Neurosurg. Rev.* 40, 647–653. doi: 10.1007/s10143-017-0822-3

Cui, Y., Song, Y., Sun, C., Howard, A., and Belongie, S. (2018). "Large scale fine-grained categorization and domain-specific transfer learning." in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Salt Lake City), 4109–4118. doi: 10.1109/CVPR.2018.00432

D'Andrea, C. B., Kenley, J. K., Montez, D. F., Mirro, A. E., Miller, R. E., Earl, E. A., et al. (2022). Real-time motion monitoring improves functional MRI data quality in infants. *Dev. Cogn. Neurosci.* 55, e101116. doi: 10.1016/j.dcn.2022.101116

Darki, F., and Klingberg, T. (2015). The role of fronto-parietal and fronto-striatal networks in the development of working memory: a longitudinal study. *Cereb. Cortex* 25, 1587–1595. doi: 10.1093/cercor/bht352

DeGeorge, E. G., Fullen, C., Gess, J., Kleiner, J., and Larson-Prior, L. (2021). Effects of age of onset and medication on cognitive performance and quality of life in patients with epilepsy. *Epilepsy Behav.* 121, 108008. doi: 10.1016/j.yebeh.2021.108008

Desai, A., Bekelis, K., Thadani, V. M., Roberts, D. W., Jobst, B. C., Duhaime, A. C., et al. (2013). Interictal PET and ictal subtraction SPECT: sensitivity in the detection of seizure foci in patients with medically intractable epilepsy. *Epilepsia* 54, 341–350. doi: 10.1111/j.1528-1167.2012.03686.x

Dosenbach, N. U. F., Koller, J. M., and Earl, E. A. (2017). Real-time motion analytics during brain MRI improve data quality and reduce costs. *Neuroimage* 161, 80–93. doi: 10.1016/j.neuroimage.2017.08.025

Engel, J. (2016). What can we do for people with drug-resistant epilepsy? The 2016 Wartenberg Lecture. *Neurology* 87, 2483–2489. doi: 10.1212/WNL.0000000000003407

England, M. J., Liverman, C. T., Schultz, A. M., and Strawbridge, L. M. (2012). Summary: a reprint from epilepsy across the spectrum: promoting health and understanding. *Epilepsy Curr.* 12, 245–253. doi: 10.17226/13379

Epilepsy Foundation (2018). *Drug Resistant Epilepsy*. Available online at: https://www.epilepsy.com/learn/drug-resistant-epilepsy (accessed February 12, 2022).

Faghiri, A., Stephen, J. M., Wang, Y. P., Wilson, T. W., and Calhoun, V. D. (2017). Changing brain connectivity dynamics: from early childhood to adulthood. *Hum. Brain Map.* 39, 1108–1117. doi: 10.1002/hbm.23896

Foley, E., Quitadamoa, L. R., Richard Walsh, A., Bill, P., Hillebrand, A., Reijneveld, J. V., et al. (2021). MEG detection of high frequency oscillations and intracranial-EEG validation in pediatric epilepsy surgery. *Clin. Neurophysiol.* 132, 2136–2145. doi: 10.1016/j.clinph.2021.06.005

Gonzalez-Martinez, J. A., Srikijvilaikul, T., Nair, D., and Bingaman, W. E. (2007). Long-term seizure outcome in reoperation after failure of epilepsy surgery. *Neurosurgery* 60, 873–880. doi: 10.1227/01.NEU.0000255438.13871.FA

Greene, D. J., Koller, J. M., Hampton, J. M., Wesevich, V., Van, A. N., Nguyen, A. L., et al. (2018). Behavioral interventions for reducing head motion during MRI scans in children. *Neuroimage* 171, 234–245. doi: 10.1016/j.neuroimage.2018.01.023

Greve, D. N., and Fischl, B. (2009). Accurate and robust brain image alignment using boundary-based registration. *Neuroimage* 48, 63–72. doi: 10.1016/j.neuroimage.2009.06.060

Griffanti, L., Douaud, G., Bijsterbosch, J., Evangelisti, S., Alfaro-Almagro, F., Glasser, M. F., et al. (2017). Hand classification of fMRI ICA noise components. *Neuroimage* 154, 188–205. doi: 10.1016/j.neuroimage.2016.12.036

Griffanti, L., Salimi-Khorshidi, G., Beckmann, C. F., Auerbach, E. J., Douaud, G., Sexton, C. E., et al. (2014). ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging. *Neuroimage* 95, 232–247. doi: 10.1016/j.neuroimage.2014.03.034

Hunyadi, B., Tousseyn, S., Dupont, P., Van Huffel, S., De Vos, M., and Van Paesschen, M. (2015). A prospective fMRI-based technique for localising the epileptogenic zone in presurgical evaluation of epilepsy. *Neuroimage* 113, 329–339. doi: 10.1016/j.neuroimage.2015.03.011

Hunyadi, B., Tousseyn, S., Dupont, P., Van Huffel, S., Van Paesschen, W., and De Vos, M. (2014). Automatic selection of epileptic independent fMRI components. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* 2014, 3853–3856. doi: 10.1109/EMBC.2014.6944464

Hunyadi, B., Tousseyn, S., Mijović, B., Dupont, P., Van Huffel, S., Van Paesschen, W., et al. (2013). ICA extracts epileptic sources from fMRI in EEG-negative patients: a retrospective validation study. *PLoS ONE* 8, e78796. doi: 10.1371/journal.pone.0078796

Jenkinson, M., Bannister, P., Brady, M., and Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* 17, 825–841. doi: 10.1006/nimg.2002.1132

Jenkinson, M., and Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Medical Image Anal.* 5, 143–156. doi: 10.1016/S1361-8415(01)00036-6

Jiang, P., Vuontela, V., Tokoriev, M., Lin, H., Aronen, E. T., Ma, Y., et al. (2018). Functional connectivity of intrinsic cognitive networks during resting state and task performance in preadolescent children. *PLoS ONE* 13, e205690. doi: 10.1371/journal.pone.0205690

Kaiboriboon, K., Lowe, V. J., Chantarujikapong, S. I., and Hogan, R. E. (2002). The usefulness of subtraction ictal SPECT coregistered to MRI in single- and dual-headed SPECT cameras in partial epilepsy. *Epilepsia* 43, 408–414. doi: 10.1046/j.1528-1157.2002.21201.x

Kanner, A. M., Irving, L. T., Cajigas, I., and Saporta, A. (2022). Long-term seizure and psychiatric outcomes following laser ablation of mesial temporal structures. *Epilepsia* 63, 812–823. doi: 10.1111/epi.17183

Kasradze, S., Lomidze, G., Helen Cross, J., Kvernadze, D., Alkhidze, M., Gagoshidze, T., et al. (2021). A six-year longitudinal study of neurocognitive problems in children with epilepsy. *Brain Dev.* 43, 833–842. doi: 10.1016/j.braindev.2021.03.007

Kelly, R. E., Alexopoulos, G. S., Wang, Z., Gunning, F. M., Murthy, C. F., Morimoto, S. S., et al. (2010). Visual inspection of independent components: defining a procedure for artifact removal from fMRI data. *J. Neurosci. Methods.* 189, 233–45. doi: 10.1016/j.jneumeth.2010.03.028

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Adv. Neural Informat. Process. Syst.* 25, 84–90. doi: 10.1145/3065386

Kwan, P., Arzimanoglou, A., Berg, A. T., Brodie, M. J., Allen Hauser, W., Mathern, G., et al. (2010). Definition of drug resistant epilepsy: consensus proposal by the ad hoc Task Force of the ILAE Commission on Therapeutic Strategies. *Epilepsia* 51, 1069–1077. doi: 10.1111/j.1528-1167.2009.02397.x

Kwan, P., and Brodie, M. J. (2010). Definition of refractory epilepsy: defining the indefinable? *Lancet Neurol.* 9, 27–29. doi: 10.1016/S1474-4422(09)70304-7

Kwan, P., and Sander, J. W. (2004). The natural history of epilepsy: an epidemiological view. *J. Neurol. Neurosurg. Psychiatr.* 75, 1376–1381. doi: 10.1136/jnnp.2004.045690

Laxer, K. D., Trinka, E., Hirsch, L. J., Cendes, F., Langfitt, J., Delanty, N., et al. (2014). The consequences of refractory epilepsy and its treatment. *Epilepsy Behav.* 37, 59–70. doi: 10.1016/j.yebeh.2014.05.031

Lee, H. W., Arora, J., Papademetris, X., Tokoglu, F., Negishi, M., Scheinost, D. (2014). Altered functional connectivity in seizure onset zones revealed by fMRI intrinsic connectivity. *Neurology* 83, 2269–77. doi: 10.1212/WNL.0000000000001068

Lee, K. H., Lee, Y. J., and Westerveld, M. (2019). Epilepsy surgery in children versus adults. *J. Kor. Neurosurg. Soc.* 62, 26. doi: 10.3340/jkns.2019.0026

Lerman, R. I., and Yitzhaki, S. (1984). A note on the calculation and interpretation of the Gini index. *Econ. Lett.* 15, 363–368. doi: 10.1016/0165-1765(84)90126-5

Lopes, R., Lina, J. M., Fahoum, F., and Gotman, J. (2012). Detection of epileptic activity in fMRI without recording the EEG. *Neuroimage* 60, 83. doi: 10.1016/j.neuroimage.2011.12.083

Luckett, P. H., Maccotta, L., Lee, J. J., Park, K. Y., Dosenbach, N. U. F., Ances, B. M., et al. (2022). Deep learning resting state functional magnetic resonance imaging lateralization of temporal lobe epilepsy. *Epilepsia* 63, 1542–1555. doi: 10.1111/epi.17233

Luders, H. O., Najm, I., Nair, D., Widdess-Walsh, P., and Bingmann, W. (2006). The epileptogenic zone: general principles. *Epileptic Disord.* 8, S1–9.

Malmgren, K., and Edelvik, A. (2017). Long-term outcomes of surgical treatment for epilepsy in adults with regard to seizures, antiepileptic drug treatment and employment. *Seizure* 44, 217–224. doi: 10.1016/j.seizure.2016.10.015

Mayoral, M., Niñerola-Baizán, A., Marti-Fuster, B., Donaire, A., Perissinotti, A., Rumià J., et al. (2019). Pileptogenic zone localization with FDG PET using a new dynamic parametric analysis. *Front. Neurol.* 10, 380. doi: 10.3389/fneur.2019.00380

McIntosh, A. M., Kalnins, R. M., Anne Mitchell, L., Fabinyi, G. C. A., Briellmann, R. S., and Berkovic, S. F. (2004). Temporal lobectomy: long-term seizure outcome, late recurrence and risks for seizure recurrence. *Brain* 127, 2018–2030. doi: 10.1093/brain/awh221

Michels, L., Lüchinger, R., Koenig, T., Martin, E., and Brandeis, D. (2012). Developmental changes of BOLD signal correlations with global human EEG power and synchronization during working memory. *PLoS ONE* 7, e39447. doi: 10.1371/journal.pone.0039447

Moncrief, G. G., Aita, S. L., Tyson, B. T., Abecassis, M., Roth, R. M., Caller, T. A., et al. (2021). Self-rated executive dysfunction in adults with epilepsy and effects of a cognitive-behavioral intervention (HOBSCOTCH). *Epilepsy Behav.* 121, 108042. doi: 10.1016/j.yebeh.2021.108042

Murray, M. I., Halpern, M. T., and Leppik, I. E. (1996). Cost of refractory epilepsy in adults in the USA. *Epilepsy Res.* 23, 139–148. doi: 10.1016/0920-1211(95)00090-9

Nagahama, Y., Schmitt, A. J., Nakagawa, D., Vesole, A. S., Kamm, J., Kovach, C. K., et al. (2018). Intracranial EEG for seizure focus localization: evolving techniques, outcomes, complications, and utility of combining surface and depth electrodes. *J. Neurosurg.* 130, 1–13. doi: 10.3171/2018.1.JNS171808

Nguyen, R. D., Smyth, M. D., Zhu, L., Pao, L. P., Swisher, S. K., Kennady, E. H., et al. (2021). A comparison of machine learning classifiers for pediatric epilepsy using resting-state functional MRI latency data. *Biomed. Rep.* 15, 77. doi: 10.3892/br.2021.1453

Nissen, I. A., Stam, C. J., van Straaten, E. C. W., and Wottschel, V. (2018). Localization of the epileptogenic zone using interictal MEG and machine learning in a large cohort of drug-resistant epilepsy patients. *Front. Neurol.* 9, e647. doi: 10.3389/fneur.2018.00647

Nozais, V., Boutinaud, P., Verrecchia, V., Gueye, M. F., Hervé, P. Y., Tzourio, C., et al. (2021). Deep learning-based classification of resting-state fMRI independent-component analysis. *Neroinformatics* 19, 619–637. doi: 10.1007/s12021-021-09514-x

Ntolkeras, G., Tamilia, E., Al Hilani, M., Bolton, J., Grant, P. E., Prabhu, S. P., et al. (2022). Presurgical accuracy of dipole clustering in MRI-negative pediatric patients with epilepsy: validation against intracranial EEG and resection. *Clin Neurophysiol.* 141, 126–148. doi: 10.1016/j.clinph.2021.01.036

Perry, M. S., and Shandley, S. (2021). Surgical evaluation in children < 3 years of age with drug resistant epilepsy: patient characteristics, diagnostic utilization and potential for treatment delays. *Int. League Against Epilepsy.* 2021, 17124. doi: 10.1111/epi.17124

Pindrik, J., Hoang, N., Smith, L., Halverson, M., Wojnaroski, M., McNally, K., et al. (2019). Preoperative evaluation and surgical management of infants and toddlers with drug-resistant epilepsy. *J. Neurosurg.* 45, 18220. doi: 10.3171/2018.7.FOCUS18220

Pizoli, C. E., Shah, M. N., and Snyder, A. Z. (2011). Resting-state activity in development and maintenance of normal brain function. *Proc. Natl. Acad. Sci. Biol. Sci.* 108, 11638–11643. doi: 10.1073/pnas.1109144108

Prisco, L., Ganau, M., Aurangzeb, S., Moswela, O., Hallett, C., Raby, S., et al. (2020). A pragmatic approach to intravenous anaesthetics and electroencephalographic endpoints for the treatment of refractory and super-refractory status epilepticus in critical care. *Seizure* 75, 153–164. doi: 10.1016/j.seizure.2019.09.011

Roth, J., Constantini, S., Ekstein, M., Weiner, S. L., Chandra, P. S., Cossu, M., et al. (2021). Epilepsy surgery in infants up to 3 months of age: safety, feasibility, and outcomes: a multicenter, multinational study. *Epilepsia* 62, 1897–1906. doi: 10.1111/epi.16959

Russ, J. B., Simmons, R., and Glass, H. C. (2021). Neonatal encephalopathy: beyond hypoxic-ischemic encephalopathy. *Neoreviews* 22, e148–162. doi: 10.1542/neo.22-3-e148

Satzer, D., Esengul, Y. T., Warnke, P. C., Issa, N. P., and Nordli, D. R. (2022). SEEG in 3D: interictal source localization from intracerebral recordings. *Front. Neurol.* 2022. 782880. doi: 10.3389/fneur.2022.782880

Schrouff, J., Perlbarg, V., Boly, M., Marrelec, G., Boveroux, P., Vanhaudenhuyse, A., et al. (2011). Brain functional integration decreases during propofol-induced loss of consciousness. *Neuroimage* 57, 198–205. doi: 10.1016/j.neuroimage.2011.04.020

Shah, M. N., Mitra, A., Goyal, M. S., Snyder, A. Z., Zhang, J., Shimony, J. S., et al. (2018). Resting state signal latency predicts laterality in pediatric medically refractory temporal lobe epilepsy. *Childs Nervs. Syst.* 34, 901–910. doi: 10.1007/s00381-018-3770-5

Shah, M. N., Nguyen, R. D., Pao, L. P., Zhu, L., CreveCoeur, T. S., Mitra, A., et al. (2019). Role of resting state MRI temporal latency in refractory pediatric extratemporal epilepsy lateralization. *JMRI.* 49, 1347–1355. doi: 10.1002/jmri.26320

Sillanpää, M., and Shinnar, S. (2010). Long-term mortality in childhood-onset epilepsy. *N. Engl. J. Med.* 363, 2522–2529. doi: 10.1056/NEJMoa0911610

van Houdt, P. J., Ossenblok, P. P. W., Colon, A. J., Hermans, K. H. M., Verdaasdonk, R. M., Boon, P. A. J. M., et al. (2015). Are epilepsy-related fMRI components dependent on the presence of interictal epileptic discharges in scalp EEG? *Brain Topogr.* 28, 606–618. doi: 10.1007/s10548-014-0407-1

Van Paesschen, W., Dupont, P., Sunaert, S., Goffin, K., and Van Laere, K. (2007). The use of SPECT and PET in routine clinical practice in epilepsy. *Curr. Opin. Neurol.* 20, 194–202. doi: 10.1097/WCO.0b013e328042baf6

Vanderby, S. A., Babyn, P. S., Carter, M. W., Jewell, S. M., and McKeever, P. D. (2010). Effect of anesthesia and sedation on pediatric MR imaging patient flow. *Radiology* 256, 229–237. doi: 10.1148/radiol.10091124

von Oertzen, T. J. (2018). PET and ictal SPECT can be helpful for localizing epileptic foci. *Curr. Opin. Neurol.* 31, 184–191. doi: 10.1097/WCO.0000000000000527

Wieser, H. G., Blume, W. T., Fish, D., Goldensohn, E., Hufnagel, A., King, D., et al. (2001). ILAE commission report. Proposal for a new classification of outcome with respect to epileptic seizures following epilepsy surgery. *Epilepsia* 42, 282–286. doi: 10.1046/j.1528-1157.2001.4220282.x

Young, C. C., Williams, J. R., Feroze, A. H., McGrath, M., Ravanpay, A. C., Ellenbogen, R. G., et al. (2020). Pediatric functional hemispherectomy: operative techniques and complication avoidance. *Neurosurg. Focus* 48, E9. doi: 10.3171/2020.1.FOCUS19889

Zhang, C. H., Lu, Y., Brinkmann, B., Welker, K., Worrell, G., and He, B. (2015). Lateralization and localization of epilepsy related hemodynamic foci using presurgical fMRI. *Clin. Neurophysiol.* 126, 27–38. doi: 10.1016/j.clinph.2014.04.011

Zhang, H., Shen, D., and Lina, W. (2019). Resting-state functional MRI studies on infant brains: a decade of gap-filling efforts. *Neuroimage* 185, 664–684. doi: 10.1016/j.neuroimage.2018.07.004

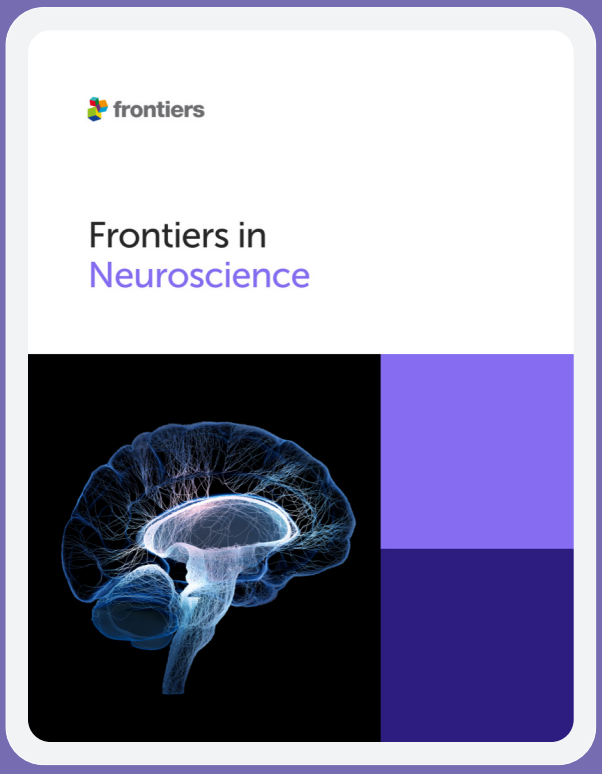
Frontiers in Neuroscience

Provides a holistic understanding of brain function from genes to behavior

Part of the most cited neuroscience journal series which explores the brain - from the new eras of causation and anatomical neurosciences to neuroeconomics and neuroenergetics.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact