



# Mapping relationships among schizophrenia, bipolar and schizoaffective disorders: A deep classification and clustering framework using fMRI time series

Weizheng Yan<sup>a,b,1</sup>, Min Zhao<sup>b,c,1</sup>, Zening Fu<sup>a</sup>, Godfrey D. Pearlson<sup>d</sup>, Jing Sui<sup>a,b,c,\*</sup>, Vince D. Calhoun<sup>a,\*\*</sup>

<sup>a</sup> Tri-Institutional Center for Translational Research in Neuroimaging and Data Science (TReNDS), Georgia State University, Georgia Institute of Technology, Emory University, Atlanta 30303, GA, USA

<sup>b</sup> Brainnetome Center and National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, China

<sup>c</sup> School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049, China

<sup>d</sup> Department of Psychiatry and Neurobiology, Yale University School of Medicine, New Haven, CT, USA

## ARTICLE INFO

### Article history:

Received 21 September 2020

Received in revised form 11 February 2021

Accepted 12 February 2021

Available online 3 March 2021

### Keywords:

Deep learning

fMRI

Schizophrenia

Bipolar disorder

Schizoaffective disorder

## ABSTRACT

**Background:** Psychiatric disorders are categorized using self-report and observational information rather than biological data. There is also considerable symptomatic overlap between different types of psychiatric disorders, which makes diagnostic categorization and multi-class classification challenging.

**Methods:** In this work, we propose a unified framework for supervised classification and unsupervised clustering of psychotic disorders using brain imaging data. A new multi-scale recurrent neural network (MsRNN) model was developed and applied to fMRI time courses (TCs) for multi-class classification. The high-level representations of the original TCs were then submitted to a tSNE clustering model for visualizing the group differences between disorders. A leave-one-feature-out approach was used for disorder-related biomarker identification.

**Results:** When studying fMRI from schizophrenia, psychotic bipolar disorder, schizoaffective disorder, and healthy individuals, the accuracy of a 4-class classification reached 46%, significantly above chance. The hippocampus, supplementary motor area and paracentral lobule were discovered as the most contributing regional TCs in the multi-class classification. Beyond this, visualization of the tSNE clustering suggested that the disease severity can be captured and schizoaffective disorder (SAD) may be separated into two subtypes. SAD cluster1 has significantly higher Positive And Negative Syndrome Scale (PANSS) scores than SAD cluster2 in PANSS negative2 (emotional withdrawal), general2 (anxiety), general3 (guilt feelings), general4 (tension).

**Conclusions:** The proposed deep classification and clustering framework is not only able to identify psychiatric disorders with high accuracy, but also interpret the correlation between brain networks and specific psychiatric disorders, and reveal the relationship between them. This work provides a promising way to investigate a spectrum of similar disorders using neuroimaging-based measures.

© 2021 Elsevier B.V. All rights reserved.

## 1. Introduction

At present, psychiatric disorders are diagnosed based on symptoms and course of illness, according to the classifications in the Diagnostic and Statistical Manual of Mental Disorders (APA, 2013; Tandon et al., 2009). Finding the biological or physiological biomarkers, rather than purely relying on behavioral symptoms and signs, might provide a

more stable and precise diagnosis. In addition, biomarkers could inform the type, timing and course of interventions, and they could allow disorders to be subtyped based on physiological criteria, creating a more personalized approach to psychiatric treatments. However, identifying biomarkers using current methods is challenging (Singh and Rose, 2009). Long conceptualized as distinct diagnostic categories, major psychotic disorders, consisting of schizophrenia (SZ), bipolar disorder with psychosis (BDP), and schizoaffective disorder (SAD), share substantial biological features in common as implicated by converging lines of evidence from genetic, molecular, histological, and neuroimaging studies (Clementz et al., 2020; Tamminga et al., 2013; Tamminga et al., 2014). Consequently, to improve diagnostic and treatment precision, it is essential to revise the classification of these disorders based on molecular/biomarker data rather than merely base on symptoms on clinical phenomenology.

\* Correspondence to: J. Sui, Institute of Automation, Chinese Academy of Sciences, Beijing, China.

\*\* Correspondence to: V.D. Calhoun, Tri-institutional Center for Translational Research in Neuroimaging and Data Science (TReNDS) Center, Georgia State University, Georgia Institute of Technology, Emory University, Atlanta 30303, GA, USA.

E-mail addresses: [jing.sui@nlpr.ia.ac.cn](mailto:jing.sui@nlpr.ia.ac.cn) (J. Sui), [vcalhoun@gsu.edu](mailto:vcalhoun@gsu.edu) (V.D. Calhoun).

<sup>1</sup> Weizheng Yan and Min Zhao contribute equally as co-first authors.

The Bipolar-Schizophrenia Consortium on Intermediate Phenotypes (B-SNIP study), a multi-site study of SZ, BDP and SAD that measured over 50 biomarkers across multiple domains in a standardized manner, provides a unique window into psychotic illnesses (Tamminga et al., 2013). Among other issues, the B-SNIP project investigated the following questions: a) which are the most discriminative biomarkers associated with a specific DSM disorder; b) which diagnostic groups (e.g., SZ, BDP, SAD, and HCs) are most separated from one another. Many novel data-driven methods have been proposed and utilized in B-SNIP resting-state functional MRI (rs-fMRI), to seek innovative diagnostic imaging biomarkers (Clementz et al., 2020; Du et al., 2015). We formulated the following two hypotheses. Hypothesis 1: The use of well-trained data-driven methods on the B-SNIP fMRI data will be more sensitive and specific in differentiating multiple DSM diagnostic groups. Hypothesis 2: Individual differences in the subcortical area such as the hippocampus, amygdala, insula and cortical area including the anterior cingulate cortex and prefrontal gyrus may play an essential role in differentiating SZ, BDP, SAD and healthy controls (Downar et al., 2016; Yu et al., 2016). Several previous studies on B-SNIP data also examined dynamic functional network connectivity. However, challenges exist when applying traditional classification and clustering methods to multi-class data (Du et al., 2020).

Deep learning models have recently made significant advances in classification, outperforming regular machine learning classification models in a range of sample sizes and multiple problem domains, including psychiatry (Durstewitz et al., 2019; LeCun et al., 2015; Abrol et al., 2021; Yan et al., 2018). The essence of deep learning methods is their ability to extract features automatically through multiple layers. Their power of automatic feature learning and extraction depends on both big data and high computational power. Previous work using 389 structural MRI data sets from SZ and HC showed that class separation performance improved with the addition of more layers, and that deeper networks better separated patients and control groups on both training and validation data (Plis et al., 2014). In the same study, despite not providing the algorithm with any information about severity during the model fitting stage, the disease severity of Huntington's disease was well captured. The findings indicate such flexible models can go well beyond simple prediction, providing new information and visualization of complex relationships in the data. Besides, the recently proposed multi-scale convolutional recurrent neural network (MsRNN) model is an excellent solution for solving the 2-class fMRI classification problem (Yan et al., 2019). The model achieved high accuracy in both multi-site pooling and cross-site classification tasks using time courses as input. Besides, the leave-one-feature-out approach can identify the most discriminative schizophrenia-related brain networks, which were consistent with the previous clinical study. These results highlight the potential of deep learning and support the need for models that can capture the complexity of mental illness.

The clustering task is challenging but crucial in investigating the relationship between different psychiatric disorders (Drysdale et al., 2017; Xie et al., 2016; Zeng et al., 2013; Yao et al., 2020). When the signal-noise-ratio is low, the models are likely to be misled by noise or by low-level features (e.g., sites, age, gender), which are not of primary interest for the study of psychiatric disorders. To prevent the network from focusing on such low-level information, adding constraints to the optimization function is essential. The constraints should: a) encourage distances between data points in the new feature space to be similar to distances in original space; b) maximize the information-theoretic dependency between data and their predicted discrete representations. For fMRI data, because of the low signal-noise-ratio, using whole-brain connectivity features for clustering is challenging and cannot always obtain satisfying performance. Deep learning has an advantage in extracting the features from low-level to high-level. Therefore, we propose to use the supervised network for the guidance of the feature extraction, then use the extracted hidden

layer features for further clustering. In this way, noisy confounds (such as age, gender, site effects) are inhibited for better clustering.

In this work, we propose a deep classification and clustering framework. An advanced MsRNN was first improved for 4-class classification. The leave-one-feature-out method is used for discovering the most discriminative brain networks. The extracted feature representations by MsRNN are further used for clustering to visualize the relationship between different psychiatric disorders.

## 2. Materials and methods

Fig. 1 presents an overview of the proposed deep classification and clustering framework. After preprocessing the resting-state fMRI data using standard procedure, 693 subjects from B-SNIP study were selected for further analysis (Du et al., 2015; Tamminga et al., 2013). Time courses for each subject were extracted using a spatially constrained ICA approach called group information guided ICA (GIG-ICA) (Du and Fan, 2013). Hence each subject was then represented by the TC features (No. time points \* No. ICs). The improved MsRNN was directly applied to the TCs to identify four different types of psychiatric disorders. The framework consists of three functions: Fig. 1a is the classification function for accurate diagnosis of psychiatric disorders; Fig. 1c is the interpretation for biomarker discovery; Fig. 1b is the clustering for comparing group distances among psychiatric disorder groups.

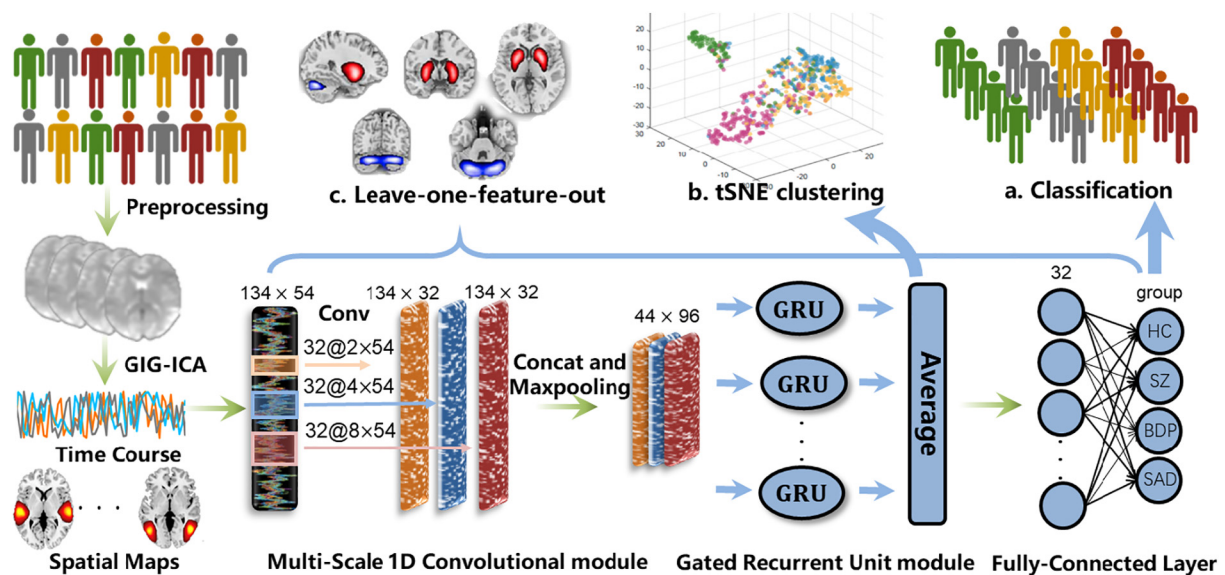
### 2.1. Image acquisition

In this study, we analyzed resting-state fMRI data of 693 subjects, including 229 HCs, 176 SZ patients, 140 BDP patients, and 129 SAD patients from the multi-site B-SNIP study (Tamminga et al., 2013). Table 1 lists the demographic and clinical information of all 693 subjects. The scanning period for the rest state fMRI data collection was ~5 min for all subjects. The detailed scanning information for each site is shown in Supplementary Table S1. All subjects provided informed consent and were clinically stable. They were taking stable medications for at least 30 days at the time of the study. During the scanning, participants were asked to rest with their eyes open and to stay awake. Patients were classified diagnostically using DSM-IV-TR criteria ascertained using the SCID (Spitzer et al., 2002).

### 2.2. Data preprocessing and IC extraction

The rsfMRI data were preprocessed with the Data Processing Assistant for Resting-State fMRI (DPARSF) toolbox (Yan and Zang, 2010) based on the statistical parametric mapping software (SPM <https://www.fil.ion.ucl.ac.uk/spm/>). The first six volumes of each scan time series were discarded to ensure the magnetization equilibrium. Then the remaining images were slice-time corrected and realigned to the first volume for head-motion correction. For each subject, the translation of head motion was less than 3 mm, and the rotation of head motion did not exceed 3° in all axis through the whole scanning process. Besides, the mean framewise displacement (FD) for HC, SZ, BDP, SAD groups was compared. The mean FD has no significant group differences between the four groups. Subsequently, the images were spatially normalized to the Montreal Neurological Institute (MNI) EPI template (Friston et al., 1995), resliced to 3 mm × 3 mm × 3 mm voxels, and smoothed with a Gaussian kernel with a full-width at half-maximum (FWHM) of 8 mm.

Next, the group ICA of fMRI Toolbox (GIFT, <https://trendscenter.org/software/gift/>) was used to analyze the rsfMRI data (Calhoun and Adali, 2012; Calhoun et al., 2001). Subject-specific data reduction by principal component analysis (PCA) retained 150 principal components (PCs) by preserving the variance higher than 99% using a standard economy-size decomposition (Allen et al., 2011; Erhardt et al., 2011; Liu et al., 2010). The Infomax algorithm (Bell and Sejnowski, 1995) was then repeated



**Fig. 1.** The overview of the deep classification and clustering framework. The data were preprocessed using standard procedure and then GIG-ICA for decomposition. (a) classification for accurate diagnosis of psychiatric disorder; (b) tSNE clustering for discovering the similarity and differences between different psychiatric disorder groups; (c) the leave-one-feature-out method for discovering the most discriminative features.

20 times using ICASSO (<http://www.cis.hut.fi/projects/ica/icasso>) and the centroid run was selected to improve the reproducibility of the decomposition, resulting in 100 group independent components (ICs). Individual subject spatial maps and time courses were back-reconstructed using GIG-ICA run on each subject and initialized with the group maps (Du and Fan, 2013). This approach has been shown to be robust to artifacts and outperforms a single-subject ICA-based de-noising approach (Du et al., 2016). 54 ICs were characterized as intrinsic connectivity networks (ICNs) after removing the ICs corresponding to physiological, movement-related or imaging artifacts, and their spatial maps (SMs) are listed in the Supplementary file Fig. S1. Nuisance covariates were all regressed out (Fox et al., 2005; Satterthwaite et al., 2013; Yan et al., 2013). Next, the time courses were stacked to form a matrix with dimensions of [No. Subjects]  $\times$  [No. Time courses]  $\times$  [No. Independent components] which was then used to calculate the functional network connectivity (FNC) matrix or to train the MsRNN model directly. Two covariates (age and gender) which may have potential confounding effects were also regressed out.

### 2.3. MsRNN for supervised classification

As shown in Fig. 1, the rsfMRI data were analyzed using group and back-reconstructed using GIG-ICA to acquire the respective independent component and time courses. The MsRNN model received time courses directly as the input. The multi-scale spatial features were extracted from the TCs by using three different scales of convolutional

filters. A gated recurrent unit (GRU) module was then applied to extract the temporal information, and the averaged layer was further used for integrating the temporal information throughout (Yan et al., 2019).

#### 2.3.1. Multi-scale convolutional layer

Multi-scale convolution layers are very helpful in extracting and integrating features from different scales (Yan et al., 2019). Thereby different scales (from seconds to minutes) of brain activity can be captured by using multi-scale convolutional layers. To synchronously capture the spatial correlation between brain regions, 1D convolutional filters whose length equals the number of ICs was applied. The width of the 1D convolutional filters was drawn from a logarithmic instead of linear scale. Therefore, multiple scales of convolutional filters, including  $54 \times 2$ ,  $54 \times 4$ , and  $54 \times 8$ , were used in the experiment. 54 is the number of ICs, [2, 4, 8] are different time spans. The architecture allows the network to extract multi-scale spatial information from the fMRI time courses. The outputs of the filters were then concatenated among the depth axis. Max-pooling operation was then performed along the time dimensions. The outputs of the max-pooling layer were then submitted to the following GRU layers.

#### 2.3.2. Densely connected GRU layer

Gated recurrent unit was designed to solve the gradient exploding/vanishing problem when training time sequence model (Chung et al., 2014). The dense connection which connects different layers in a feed-forward way was also been proved as a solution for maintaining gradients (Huang et al., 2017). In this work, a two-layer densely connected GRU was used for integrating high-level temporal information from fMRI. The size of the GRU hidden state was set at 32.

#### 2.3.3. Averaged layer

Even with the best experimental fMRI design, it is not possible to equalize the random thoughts of subjects during resting-state fMRI scanning because they depend on too many unobservable subject-specific factors. Similarly, the time series across subjects are not synchronized (Morioka et al., 2020). Therefore, we combined the fMRI steps by averaging all of the GRU outputs. In this way, all activities of the brain during scanning were leveraged to obtain better classification performance.

**Table 1**  
Demographics and clinical details of all 693 subjects.

Categories	HC (n = 229)	SZ (n = 176)	BDP (n = 159)	SAD (n = 129)
Age (year)	38.3 (12.5)	35.2 (11.9)	36.1 (12.5)	36.5 (12.2)
Gender (M/F)	98/131	123/53	52/107	57/72
PANSS (Positive)	None	15.9 (6.5)	12.4 (4.8)	18.1 (5.6)
PANSS (Negative)	None	15.7 (6.9)	11.6 (4.0)	15.4 (5.1)
PANSS (General)	None	30.4 (11.1)	28.0 (9.1)	34.3 (10.4)
PANSS (Total)	None	62.0 (22.0)	52.0 (15.6)	67.5 (19.1)



In summary, the *MsRNN* classification model consists of multiple-scale Conv1D layers, stacked GRU layers that are densely connected in a feed-forward manner, an averaged layer that integrates the context of the whole sequence, and fully-connected layers. More detailed information about the *MsRNN* model can be found in Supplementary Fig. S2.

#### 2.4. Estimating the discriminative power of independent components (leave-one-feature-out)

The ultimate goal of fMRI classification studies is to identify a collection of statistical features that can serve as reliable imaging biomarkers for disease diagnosis and are reproducible across multiple datasets. Despite extraordinary classification performance in some cases, the lack of interpretability often restricts the application of deep learning methods (Kim et al., 2015; Kohoutová et al., 2020; Yan et al., 2019). Some previously proposed models, such as LRP (Yan et al., 2017), weight-activation product (Polyn et al., 2005), and leave-one-feature-out (Yan et al., 2019), provide good strategies for model interpretation. As for leave-one-feature-out, the basic idea is that the features whose elimination leads to the most significant damage to classification performance should be regarded as the top contributing features. *MsRNN* learns dynamic information and temporal dependency from the time courses. If one IC's time course is replaced with its average value, there would be no useful temporal information that can be learned by *MsRNN*. More specifically, each subject is represented with a  $T \times D$  matrix, where  $T$  is the length of time courses, and  $D$  is the number of independent components (ICs). A specific element in the matrix can be denoted by  $e_{ti}$ . To quantify the classification contribution of the  $i_{th}$  IC, we replace the time courses of  $i_{th}$  IC with its averaged value ( $\sum_{t=1}^T e_{ti}/T$ ) while retaining the other ICs' time courses. Since the *MsRNN* model learns temporal dependency of BOLD signals, this is equivalent to eliminating the contribution of the  $i_{th}$  component. A detailed description of the procedure can be found in Supplementary methods.

All the testing samples are processed in the same way and subsequently fed to the trained *MsRNN* model. The classification performance of the models trained with reduced features may decrease compared to those using all features. The variation of the classification performances (e.g., accuracy, sensitivity, specificity) when removing  $i_{th}$  dimension are recorded and sorted. The features which maximize the decrease of the classification performance are further selected as the most discriminative features. Specifically, the 693 samples were randomly split into five folds. 554 samples (four folds) were used for optimizing the parameters of *MsRNN*, and 139 samples (one-fold) were used for further finding the contribution of each IC during each cross-validation. The procedures are as follow: 1) After optimizing the trained model with 554 samples, the parameters of the trained model were saved; 2) The time courses of 139 subjects without removing any component were fed to the model to obtain a baseline classification performance; 3) The 139 subjects which have removed the contribution of one specific IC were fed to the model to obtain the classification performance repeatedly. The change of accuracy/precision/recall when removing specific features was recorded and sorted; 4) Repeat step 3 until each IC has been removed once.

#### 2.5. Unsupervised clustering from the selected feature representations

Data clustering based on the original fMRI data is exceptionally challenging because the brain fMRI is high dimensional and low in signal-noise-ratio. Consequently, the number of dimensions must be reduced to avoid the "curse of dimensionality". Another challenge is that the confounds may mislead the clustering orientation. Clustering results are dependent on the dimensional representation selected for analysis. However, there is no established standard for selecting appropriate fMRI dimensional representations. The tSNE approach embeds high-dimensional data into a low-dimensional space while preserving the

pairwise distances of the data points (Maaten and Hinton, 2008). To overcome the effects of confounds (e.g., age, gender, sites) which may mislead the unsupervised clustering method, we first extracted the output of the last hidden layer of the *MsRNN*, then the extracted feature representations were submitted to the tSNE model for unsupervised learning.

#### 2.6. Model training

The time courses of ICs described above were used as the inputs for training the *MsRNN* model. The models are trained by minimizing the cross-entropy loss using Adam optimizer. The training batch size is set as 32. The learning rate started from 0.001 and decayed after each epoch with a decay rate of  $10^{-2}$ . To improve the generalization performance of the model and overcome overfitting, dropout ( $dropout = 0.5$ ) and  $L_{1,2}$ -norm regularization ( $l_1 = 5e-4$ ,  $l_2 = 5e-4$ ) were also applied for regulating the model parameters. The training process was stopped when the validation loss stopped decreasing for 50 epochs or when the maximum epochs (1000 epochs) had been executed. The intermediate models which achieved the highest accuracy on the validation dataset were reserved for testing (Yan et al., 2019). Besides, the proposed models are implemented on the platform of Keras (<https://keras.io/>) and ScikitLearn (<https://scikit-learn.org/>). All the above models were implemented on a desktop computer (Intel(R) Xeon (R) CPU E5-1650 v4 @ 3.60GHz, 6 CPU cores) with a single GPU (12GB NVIDIA GTX TITAN 12GB).

### 3. Results

#### 3.1. Four-class HC/SZ/BDP/SAD classification

We compared the modified *MsRNN* model with three classical machine learning classifiers (SVM (Guyon et al., 2002), Adaboost (Zhu et al., 2009), and Random Forest (Breiman, 2001)). Note that the three conventional classification methods usually work on the FNC matrix computed using the correlation of TCs of selected components instead of the TCs themselves. As a result, for performance comparison, FNCs were used as the input of conventional methods while TCs were used as the input of *MsRNN* method. In the deep learning *MsRNN* classification frameworks, we used four folds as the training set (10% samples of the training set were further selected randomly as validation dataset), and one-fold as the testing dataset. For conventional classification models (SVM, Adaboost, and Random Forest), four folds were used for training and one-fold for testing (Yan et al., 2019). Table 2 is the confusion matrix of 4-class classification achieved by four methods in multi-site pooling classification. As shown in Fig. 2, the *MsRNN* achieved an accuracy of 46% in 4-class classification task using DSM labels, which is significantly better than the results obtained by using SVM, Adaboost and Random Forest. In our experiment, the training time for *MsRNN* was around 1.5 min, while the testing time for a new subject was around 0.1 s. More details about the model complexities are listed in Supplementary Files Table S2-S6.

#### 3.2. Two-class classification using *MsRNN* (5-fold cross-validation, average)

To further investigate common and specific impairments of mental disorders, we compared the *MsRNN* with three traditional popular classifiers (SVM, Adaboost, Random Forest) in two-class classification tasks. When training the *MsRNN* model, four folds were as the training set (10% samples of the training set were further selected randomly as validation dataset), and one-fold was as the testing dataset. As for conventional classification models, four folds were used for training and one-fold for testing. Table 3 lists the classification results: the accuracy of SZ vs. HC was 78.5%, the accuracy of BDP vs. HC was 71.6%, the accuracy of SAD vs. HC was 70.4%.

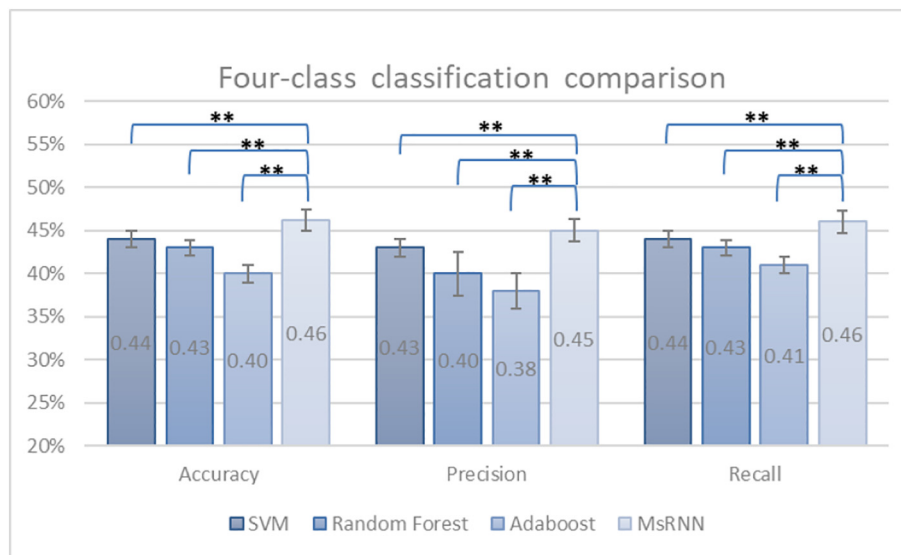
**Table 2**  
Confusion matrix of four-class classification.

	SVM				Random Forest			
	HC	SZ	BDP	SAD	HC	SZ	BDP	SAD
HC	153(67%)	28(12%)	18(8%)	30(13%)	196(86%)	26(11%)	5(2%)	2(1%)
SZ	35(20%)	73(41%)	33(19%)	35(20%)	74(42%)	83(47%)	16(9%)	3(2%)
BDP	41(26%)	47(30%)	41(26%)	30(18%)	86(54%)	56(35%)	15(9%)	2(2%)
SAD	31(24%)	30(23%)	29(23%)	39(30%)	75(58%)	43(33%)	7(5%)	4(4%)

	AdaBoost				MsRNN			
	HC	SZ	BDP	SAD	HC	SZ	BDP	SAD
HC	148(65%)	27(12%)	33(14%)	21(9%)	144(63%)	35(15%)	38(17%)	12(5%)
SZ	46(26%)	66(38%)	39(22%)	25(14%)	34(19%)	80(45%)	43(25%)	19(11%)
BDP	51(32%)	45(28%)	49(31%)	14(9%)	33(21%)	40(25%)	74(47%)	12(7%)
SAD	47(36%)	33(26%)	30(23%)	19(15%)	36(28%)	41(32%)	30(23%)	22(17%)

Note: a(b%): a is the amount of accurately classified data, b% is the ratio of the accurately classified data. For instance, using SVM classifier, 67% HC were predicted correctly, 12% HC were misclassified as SZ. True positives are on the diagonal, false negatives are on the upper diagonal. The classification results of MsRNN are well above chance.



**Fig. 2.** Four-class classification comparison. The MsRNN achieved significantly higher classification performance comparing the other machine learning methods. The bars represent the averaged value of accuracy, precision and recall from all 5-fold cross-validations. \*/\*\* denote respectively that the proposed MsRNN is significantly better than the conventional model with P value = 0.05/0.01.

In general, compared with the traditional machine learning methods, MsRNN achieved higher accuracy by balancing sensitivity and specificity. Notably, in SAD vs. HC classification task, the SVM performed better than the MsRNN in accuracy and specificity. However, it was not able to balance the sensitivity and specificity effectively.

### 3.3. Estimating the most discriminating independent components

Here, we used the leave-one-feature-out approach, which leaves one IC's time course out, and used the remaining 53 IC's time course to train the model. After that, we compared the alteration of classification by looping all 54 ICs. The trained 4-class classification and 2-class

classification models were all analyzed and interpreted one by one using the leave-one-feature-out method. The top five components that contributed most to the respective classification tasks are shown in Fig. 3 and Table 4. The results demonstrated that the discriminating regions that contribute to the 4-class classification were mainly located in hippocampus, supplementary motor area, paracentral lobule, precentral, and insula. The middle frontal gyrus was specific for SZ/HC classification. The cerebellum was specific for BDP/HC classification. The right middle temporal gyrus, inferior frontal gyrus and paracentral lobule were specific for SAD/HC classification.

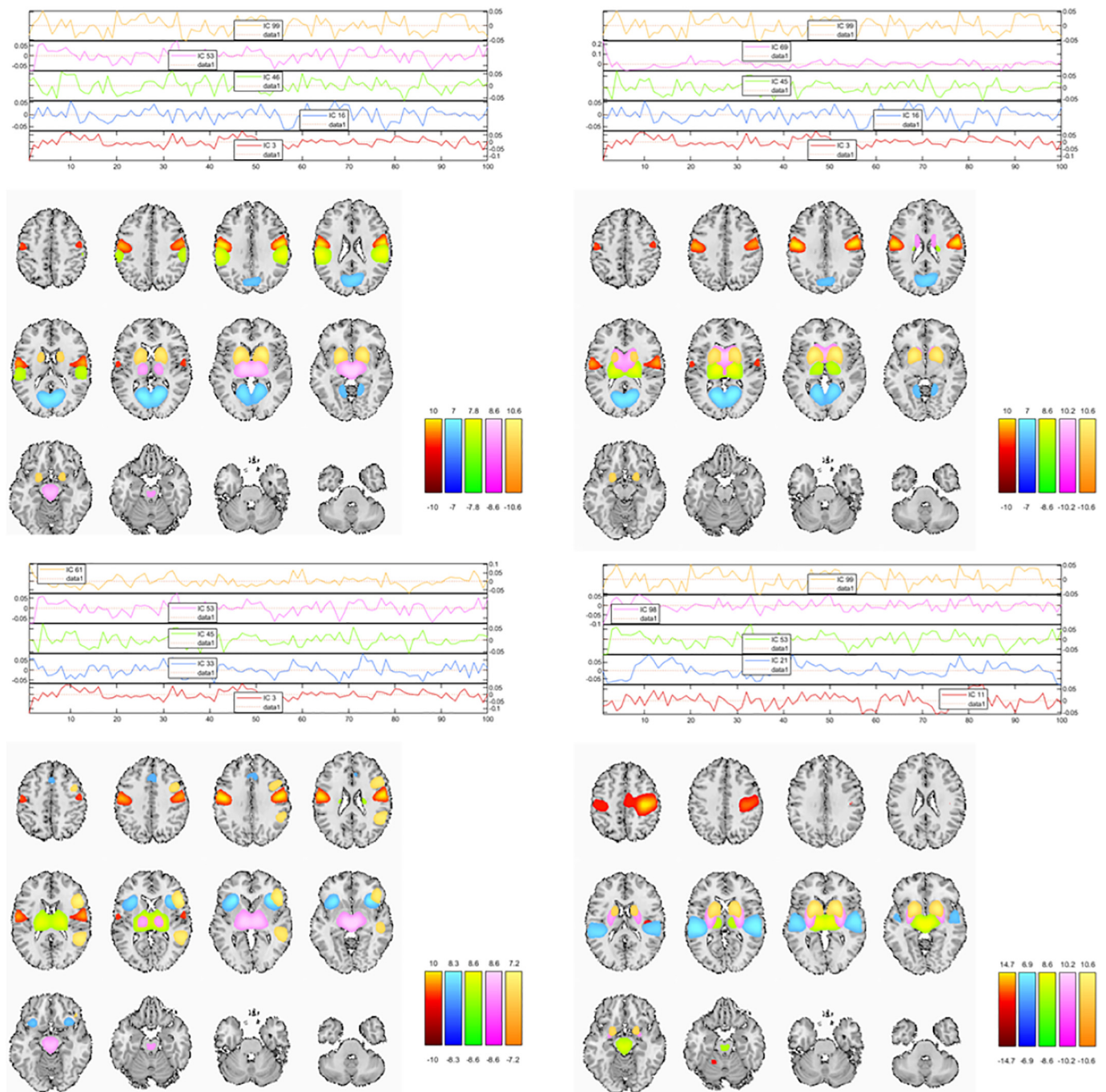
### 3.4. Clustering using the feature representation extracted using MsRNN

Fig. 4(a) shows the results for clustering psychiatric disorders based on DSM label using tSNE. Interestingly, the SAD group separated into two distinct clusters. One is the top left cluster (SAD Cluster 1), the other (SAD Cluster 2) is mixed with SZ and BDP groups. We then examined clinical psychosis rating scores in these two SAD clusters. Fig. 4(b, c, d) shows the PANSS positive, PANSS negative and PANSS total score values respectively. The 30-item PANSS, firstly proposed by Kay et al., is conceived as an operationalized, drug-sensitive instrument that provides a balanced representation of positive and negative symptoms and gauges their relationship to one another and global psychopathology

**Table 3**  
Two-class classification comparison.

Methods	SZ vs. HC			BDP vs. HC			SAD vs. HC		
	ACC	SEN	SPE	ACC	SEN	SPE	ACC	SEN	SPE
SVM	74.8	65.3	82.1	70.4	47.8	<b>86.0</b>	<b>71.2</b>	38.8	<b>89.5</b>
Random Forest	74.8	58.5	<b>87.3</b>	69.3	35.2	93.0	67.9	22.5	93.4
AdaBoost	73.0	68.2	76.9	62.4	49.1	71.6	65.9	45.7	77.3
MsRNN	<b>78.5</b>	<b>74.4</b>	81.6	<b>71.6</b>	<b>65.4</b>	76.0	70.4	<b>58.9</b>	76.9

The best classification performance of each column is in bold.



**Fig. 3.** (a) Top 5 contributing components (spatial map and time courses) for 4-class classification. (b) Top 5 contributing components (spatial map and time courses) for SZ/HC classification; (c) Top 5 contributing components (spatial map and time courses) for BDP/HC classification; (d) Top 5 contributing components (spatial map and time courses) for SAD/HC classification. The discriminating regions that contribute to the 4-class classification are mainly located in hippocampus, supplementary motor area, paracentral lobule, precentral, and insula. The middle frontal gyrus is specific to SZ vs. HC classification. The cerebellum is specific for BDP vs. HC classification. The right middle temporal gyrus, inferior frontal gyrus, and paracentral lobule are specific for SAD vs. HC classification.

(Kay et al., 1987). The PANSS scale is an evaluation scale of 30 disparate items from 1 to 7 for psychopathological symptoms observed in patients presenting psychotic syndromes, especially schizophrenia states. Three scores obtained with this evaluation tool are generally calculated for evaluating three dimensions of the syndrome: positive, negative, and general psychopathology, as part of a categorical or dimensional perspective. As shown in Fig. 5, SAD cluster1 has significantly higher scores than SAD cluster2 in PANSS negative2 (emotional withdrawal), general2 (anxiety), general3 (guilt feelings), general4 (tension). Besides, to minimize the effects of confounds, we mapped the age, gender and site information to the tSNE map. As shown in Fig. 6, these

confounds were not systematically associated with any groups, showing a random pattern.

#### 4. Discussion

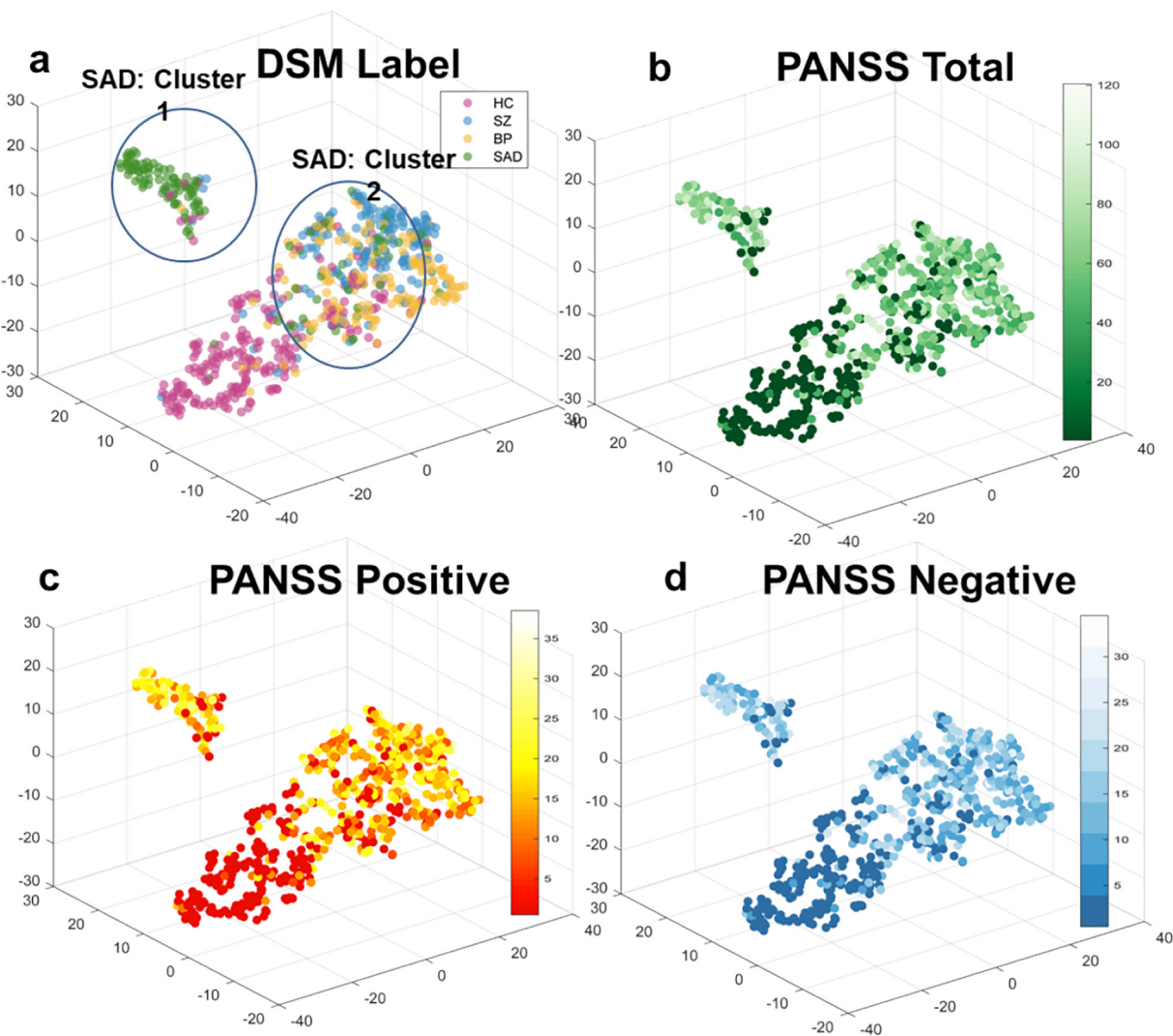
Lacking clinical biomarkers, the symptomatic overlap between different psychiatric disorders makes diagnosis challenging. Using computational strategies to discover the most informative biologic fingerprint is a promising strategy to uncover mechanisms in psychosis. Investigations of symptom-related psychiatric disorders using the temporal information from fMRI are rising (Dvornek et al., 2017; Yan et al., 2019),



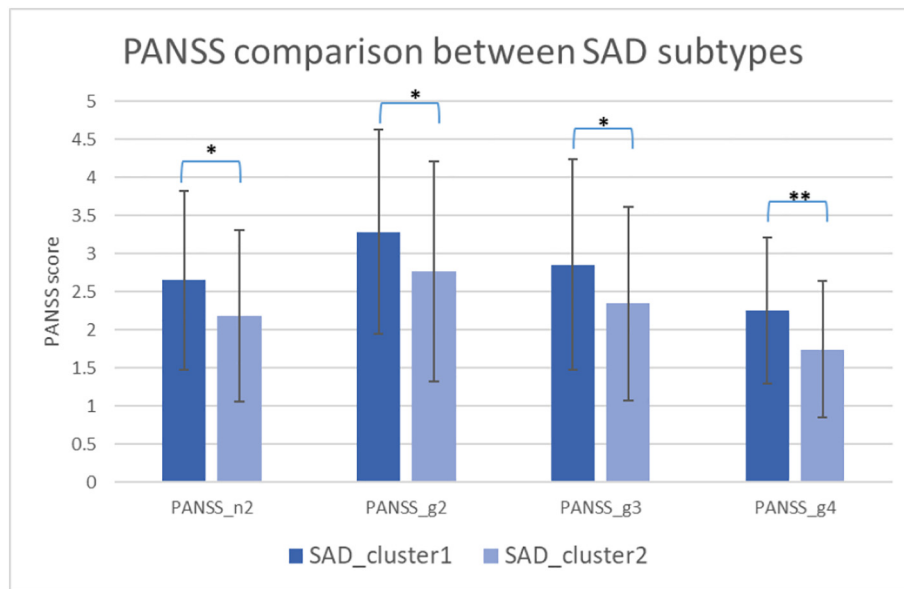
**Table 4**  
Top 5 components contributing to the MsRNN classification tasks.

4-class classification		SZ/HC classification	
Brian regions	Network	Brian regions	Network
Insula	CON/HPN	Paracentral lobule	SMN
Hippocampus	HPN	Supplementary motor area	CON
Superior parietal lobule	SMN	Precentral	SMN
Supplementary motor area	CON	Middle frontal gyrus	CON
Paracentral lobule	SMN	Insula	CON/HPN
BDP/HC classification		SAD/HC classification	
Brian regions	Network	Brian regions	Network
Hippocampus	HPN	Hippocampus	HPN
Paracentral lobule	SMN	Supplementary motor area	SMN
Cerebellum	CER	R MTG+IFG	CON
Precentral	SMN	Paracentral lobule	SMN
Precuneus	DMN	Superior temporal gyrus	AUD

Note: R MTG + IFG: right middle temporal gyrus+inferior frontal gyrus. The blue cells are the common components shared by various disorders. The Arabic numbers in the bracket are the indexes of the components.



**Fig. 4.** (a) The clustering results of psychiatric disorders based on DSM label; (b) PANSS total; (c) PANSS positive; (d) PANSS negative. Notes: The healthy controls have no PANSS scores, values are set at 0. The figure shows that the gradients are from low to high, and a different but similar gradient appears in the separate SAD clusters.



**Fig. 5.** PANSS comparison between SAD subtypes. \*/\*\* denote respectively that there exists a significant difference between two groups with  $P = 0.05/0.01$ . PANSS\_n2 represents emotional withdrawal, PANSS\_g2 represents anxiety, PANSS\_g3 represents guilt feelings, PANSS\_g4 represents tension.

and may lead to further insights into the underlying brain biology of psychiatric disorders. In this work, we propose a framework that combines the classification and clustering approaches for classification, interpreting disease-related networks, and discovering the relationships among various psychiatric disorders. A large dataset that consists of 693 subjects was used for analysis. By using the most advanced *MsRNN* model, an accuracy of 46% was achieved in 4-class classification, significantly above chance. Using a leave-one-feature-out approach, the most discriminative brain networks for specific psychiatric disorder diagnoses were identified. Also, the tSNE clustering approach was used for visualizing the relationships among multiple psychiatric disorders. To the best of our knowledge, this is the first framework that incorporates classification, interpretation, and clustering tasks at the same time using TC features.

Comparing with the traditional methods such as SVM, Adaboost and Random Forest, *MsRNN* achieved higher classification accuracy while balancing precision and recall metrics. Based on the confusion matrix we do not see a situation in which one category is severely misclassified. The confusion matrix also indicates the difficulty level of the classification tasks. HC is the most straightforward category to be identified, followed by BDP and SZ. SAD is the most challenging category to

identify. SAD is more likely to be misidentified as SZ. Our finding is consistent with the choice to assign SAD into the same diagnosis class as SZ in the current DSM-5 (Heckers et al., 2013; Malaspina et al., 2013). We also used 2-class classification tasks to identify disorder-specific brain networks. Results show the accuracies for *MsRNN* in SZ vs HC is 78.5%, BDP vs. HC is 71.6%, SAD vs. HC is 70.4%. Based on this, the SZ is most easily distinguished from HC, while SAD and BDP were more challenging to separate from HC.

There has been increasing interest in identifying disorder-related brain regions following biologic classification. The leave-one-feature-out results demonstrate that the discriminating regions that contribute to the 4-class classification are mainly located in hippocampus, supplementary motor area, paracentral lobule, precentral, and insula. The middle frontal gyrus is specific to SZ vs. HC classification. The cerebellum is specific for BDP vs. HC classification. The right middle temporal gyrus, inferior frontal gyrus, and paracentral lobule are specific for SAD vs. HC classification. Some previous studies support the role of the insula (Mikolas et al., 2016; Wylie and Tregellas, 2010), hippocampus (Heckers, 2001), cerebellum (Andreasen and Pierson, 2008), supplementary motor area (Northoff et al., 2020), and middle frontal gyrus (Kikinis et al., 2010) for SZ identification. The cerebellum plays an



**Fig. 6.** tSNE map of the confounds effects. (a) site effects; (b) gender effects; (c) age effects. The confounds are not systematically associated with any groups, showing a random pattern.



important role in psychiatric disorders, such as bipolar disorder. A growing body of evidence showed structural (DelBello et al., 1999; Lippmann et al., 1982) and functional (Liu et al., 2012; Wang et al., 2015) abnormalities of cerebellum in bipolar disorder. For example, Shinn et al. found reduced cerebrocerebellar functional connectivity in somatomotor, ventral attention, salience, and frontoparietal control networks in patients with BDP (Shinn et al., 2017), and Yates et al. found a greater rate of cerebellar atrophy in patients with BDP (Yates et al., 1987). As for motor areas, Northoff et al. demonstrated that psychomotor mechanisms and their underlying biochemical modulation are operative in both healthy subjects as well as in MDD, BDP, and SZ subjects, the only difference consists in the fact that these mechanisms are abnormally balanced and thus manifest in extreme values in psychiatric disorders. Psychomotor mechanisms and their biochemical modulation can be considered paradigmatic examples of a dimensional approach as suggested in RDoC and the recently introduced spatiotemporal psychopathology (Northoff et al., 2020).

As for group relationships, our results separated SAD into two clusters, which may hint at different subtypes. One is clearly demarcated (SAD Cluster1), and the other (SAD Cluster2) is interspersed among other psychiatric disorders. Whether SAD should be considered a separate category and what is the relationship among these disorders is still controversial (Keshavan et al., 2011; Tamminga et al., 2013). In our case, SAD cluster1 has significantly higher scores than SAD cluster2 in PANSS negative2 (emotional withdrawal), general2 (anxiety), general3 (guilt feelings), general4 (tension).

Several aspects of the current work may need further refinement in the future. As for the clustering approaches, clustering is still a challenging problem because the clustering directions/clues are always affected by many effects such as gender, age, and sites. Therefore, effective guidance is necessary for clustering. In the future, a more advanced pure clustering method should be developed. Deep clustering may be a feasible solution (Chang et al., 2017; Xie et al., 2016). Another strategy might use an ensemble of approaches (e.g., multi-model and multi-modalities) to achieve even better classification performance. Further studies of these subclusters as related to prognostic and retreatment-response profiles, or biotype profiles (Clementz et al., 2015), would provide more evidence about SAD subtypes. In addition, the use of dynamic functional connectivity which can both capture the spatial and temporal information from the fMRI may provide additional insights into the data. In the future study, we plan to use dynamic functional connectivity as the input to a recurrent neural network model.

In summary, to the best of our knowledge, this is the first attempt to integrate recurrent neural network, clustering, and interpretation for 4 groups based on resting-state fMRI time series. The framework incorporates the strength of the *multi-scale RNN* model which can efficiently capture the spatial-temporal features based on time courses directly, and then use a leave-one-feature-out approach for interpretation. The proposed deep classification and clustering framework can reveal the relationships among multiple psychotic psychiatric disorders and provides a promising approach that can be used to investigate a spectrum of similar disorders using neuroimaging-based measures.

### Role of the funding source

No funders played a role in the study.

### CRediT authorship contribution statement

The current research was designed by WY, JS and VDC. The data was analyzed by WY, MZ and ZF. The draft was written by WY, JS, VDC, GDP. All authors contributed to the discussion of the results and have approved the final manuscript to be published.

### Declaration of competing interest

The authors report no biomedical financial interests or potential conflicts of interest.

### Acknowledgments

This work was supported by Natural Science Foundation of China (82022035, 61773380), National Institute of Health (R01MH117107, R01MH118695 and R01EB024047), the NIMH support of the Bipolar-Schizophrenia Network for Intermediate Phenotypes (Grant R01MH077851, MH078113, MH077945, MH096942, and MH096957), and Beijing Municipal Science and Technology Commission (Z181100001518005). All authors report no biomedical financial interest or potential conflicts of interest.

### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.schres.2021.02.007>.

### References

- Abrol, A., Fu, Z., Salman, M., Silva, R., Du, Y., Plis, S., Calhoun, V., 2021. Deep learning encodes robust discriminative neuroimaging representations to outperform standard machine learning. *Nat. Comm.* 12.
- Allen, E.A., Erhardt, E.B., Damaraju, E., Gruner, W., Segall, J.M., Silva, R.F., Havlicek, M., Rachakonda, S., Fries, J., Kalyanam, R., Michael, A.M., Caprihan, A., Turner, J.A., Eichele, T., Adelsheim, S., Bryan, A.D., Bustillo, J., Clark, V.P., Feldstein Ewing, S.W., Filbey, F., Ford, C.C., Hutchison, K., Jung, R.E., Kiehl, K.A., Kodituwakku, P., Komesu, Y.M., Mayer, A.R., Pearson, G.D., Phillips, J.P., Sadek, J.R., Stevens, M., Teuscher, U., Thoma, R.J., Calhoun, V.D., 2011. A baseline for the multivariate comparison of resting state networks. *Front. Syst. Neurosci.* 5.
- Andreasen, N.C., Pierson, R., 2008. The role of the cerebellum in schizophrenia. *Biol. Psychiatry* 64, 81–88.
- APA, 2013. *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*. American Psychiatric Pub.
- Bell, A.J., Sejnowski, T.J., 1995. An information-maximization approach to blind separation and blind deconvolution. *Neural Comput.* 7, 1129–1159.
- Breiman, L., 2001. Random forests. *Mach. Learn.* 45, 5–32.
- Calhoun, V.D., Adali, T., 2012. Multisubject independent component analysis of fMRI: a decade of intrinsic networks, default mode, and neurodiagnostic discovery. *IEEE Rev. Biomed. Eng.* 5, 60–73.
- Calhoun, V.D., Adali, T., Pearson, G.D., Pekar, J.J., 2001. A method for making group inferences from functional MRI data using independent component analysis. *Hum. Brain Mapp.* 14, 140–151.
- Chang, J., Wang, L., Meng, G., Xiang, S., Pan, C., 2017. Deep adaptive image clustering. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5879–5887.
- Chung, J., Gulcehre, C., Cho, K., Bengio, Y., 2014. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. *arXiv preprint. arXiv:1412.3555*.
- Clementz, B.A., Sweeney, J.A., Hamm, J.P., Ivleva, E.I., Ethridge, L.E., Pearson, G.D., Keshavan, M.S., Tamminga, C.A., 2015. Identification of distinct psychosis biotypes using brain-based biomarkers. *Am. J. Psychiatry* 173, 373–384.
- Clementz, B.A., Trotti, R.L., Pearson, G.D., Keshavan, M.S., Gershon, E.S., Keedy, S.K., Ivleva, E.I., McDowell, J.E., Tamminga, C.A., 2020. Testing psychosis phenotypes from bipolar-schizophrenia network for intermediate phenotypes for clinical application: biotype characteristics and targets. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* 5, 808–818.
- DelBello, M.P., Strakowski, S.M., Zimmerman, M.E., Hawkins, J.M., Sax, K.W., 1999. MRI analysis of the cerebellum in bipolar disorder: a pilot study. *Neuropsychopharmacology* 21, 63–68.
- Downar, J., Blumberger, D.M., Daskalakis, Z.J., 2016. The neural crossroads of psychiatric illness: an emerging target for brain stimulation. *Trends Cogn. Sci.* 20, 107–120.
- Drysdale, A.T., Grosenick, L., Downar, J., Dunlop, K., Mansouri, F., Meng, Y., Fetcho, R.N., Zebley, B., Oathes, D.J., Etkin, A., Schatzberg, A.F., Sudheimer, K., Keller, J., Mayberg, H.S., Gunning, F.M., Alexopoulos, G.S., Fox, M.D., Pascual-Leone, A., Voss, H.U., Casey, B.J., Dubin, M.J., Liston, C., 2017. Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nat. Med.* 23, 28–38.
- Du, Y., Fan, Y., 2013. Group information guided ICA for fMRI data analysis. *Neuroimage* 69, 157–197.
- Du, Y., Pearson, G.D., Liu, J., Sui, J., Yu, Q., He, H., Castro, E., Calhoun, V.D., 2015. A group ICA based framework for evaluating resting fMRI markers when disease categories are unclear: application to schizophrenia, bipolar, and schizoaffective disorders. *Neuroimage* 122, 272–280.
- Du, Y., Allen, E.A., He, H., Sui, J., Wu, L., Calhoun, V.D., 2016. Artifact removal in the context of group ICA: a comparison of single-subject and group approaches. *Hum. Brain Mapp.* 37, 1005–1025.
- Du, Y., Hao, H., Wang, S., Pearson, G.D., Calhoun, V.D., 2020. Identifying commonality and specificity across psychosis sub-groups via classification based on features from dynamic connectivity analysis. *Neuroimage Clin.* 27, 102284.
- Durstewitz, D., Koppe, G., Meyer-Lindenberg, A., 2019. Deep neural networks in psychiatry. *Mol. Psychiatry* 24, 1583–1598.
- Dvornek, N.C., Ventola, P., Pelphrey, K.A., Duncan, J.S., 2017. Identifying autism from resting-state fMRI using long short-term memory networks. *Mach. Learn. Med. Imaging 362–370 MLMI (Workshop)* 10541.

- Erhardt, E.B., Rachakonda, S., Bedrick, E.J., Allen, E.A., Adali, T., Calhoun, V.D., 2011. Comparison of multi-subject ICA methods for analysis of fMRI data. *Hum. Brain Mapp.* 32, 2075–2095.
- Fox, M.D., Snyder, A.Z., Vincent, J.L., Corbetta, M., Van Essen, D.C., Raichle, M.E., 2005. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc. Natl. Acad. Sci. U. S. A.* 102, 9673.
- Friston, K.J., Ashburner, J., Frith, C.D., Poline, J.B., Heather, J.D., Frackowiak, R.S.J., 1995. Spatial registration and normalization of images. *Hum. Brain Mapp.* 3, 165–189.
- Guyon, I., Weston, J., Barnhill, S., Vapnik, V., 2002. Gene selection for cancer classification using support vector machines. *Mach. Learn.* 46, 389–422.
- Heckers, S., 2001. Neuroimaging studies of the hippocampus in schizophrenia. *Hippocampus* 11, 520–528.
- Heckers, S., Barch, D.M., Bustillo, J., Gaebel, W., Gur, R., Malaspina, D., Owen, M.J., Schultz, S., Tandon, R., Tsuang, M., Van Os, J., Carpenter, W., 2013. Structure of the psychotic disorders classification in DSM-5. *Schizophr. Res.* 150, 11–14.
- Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q., 2017. Densely connected convolutional networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4700–4708.
- Kay, S.R., Fiszbein, A., Opler, L.A., 1987. The Positive and Negative Syndrome Scale (PANSS) for schizophrenia. *Schizophr. Bull.* 13, 261–276.
- Keshavan, M.S., Morris, D.W., Sweeney, J.A., Pearson, G., Thaker, G., Seidman, L.J., Eack, S.M., Tamminga, C., 2011. A dimensional approach to the psychosis spectrum between bipolar disorder and schizophrenia: the Schizo-Bipolar Scale. *Schizophr. Res.* 133, 250–254.
- Kikinis, Z., Fallon, J.H., Niznikiewicz, M., Nestor, P., Davidson, C., Bobrow, L., Pelavin, P.E., Fischl, B., Yendiki, A., McCarley, R.W., Kikinis, R., Kubicki, M., Shenton, M.E., 2010. Gray matter volume reduction in rostral middle frontal gyrus in patients with chronic schizophrenia. *Schizophr. Res.* 123, 153–159.
- Kim, J., Calhoun, V.D., Shim, E., Lee, J.H., 2015. Deep neural network with weight sparsity control and pre-training extracts hierarchical features and enhances classification performance: evidence from whole-brain resting-state functional connectivity patterns of schizophrenia. *Neuroimage* 124, 127–146.
- Kohoutová, L., Heo, J., Cha, S., Lee, S., Moon, T., Wager, T.D., Woo, C.-W., 2020. Toward a unified framework for interpreting machine-learning models in neuroimaging. *Nat. Protoc.* 15, 1399–1435.
- LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. *Nature* 521, 436–444.
- Lippmann, S., Manshadi, M., Baldwin, H., Drasin, G., Rice, J., Alrajeh, S., 1982. Cerebellar vermis dimensions on computerized tomographic scans of schizophrenic and bipolar patients. *Am. J. Psychiatry* 139 (5), 667–668.
- Liu, J., Hutchison, K., Perrone-Bizzozero, N., Morgan, M., Sui, J., Calhoun, V., 2010. Identification of genetic and epigenetic marks involved in population structure. *PLoS one* 5 (10), e13209. <https://doi.org/10.1371/journal.pone.0013209>.
- Liu, C.-H., Li, F., Li, S.-F., Wang, Y.-J., Tie, C.-L., Wu, H.-Y., Zhou, Z., Zhang, D., Dong, J., Yang, Z., 2012. Abnormal baseline brain activity in bipolar depression: a resting state functional magnetic resonance imaging study. *Psychiatry Res. Neuroimaging* 203, 175–179.
- Maaten, V., Hinton, G., 2008. Visualizing data using t-SNE. *J. Mach. Learn. Res.* 9.
- Malaspina, D., Owen, M.J., Heckers, S., Tandon, R., Bustillo, J., Schultz, S., Barch, D.M., Gaebel, W., Gur, R.E., Tsuang, M., Van Os, J., Carpenter, W., 2013. Schizoaffective disorder in the DSM-5. *Schizophr. Res.* 150, 21–25.
- Mikolas, P., Melicher, T., Skoch, A., Matejka, M., Slovakova, A., Bakstein, E., Hajek, T., Spaniel, F., 2016. Connectivity of the anterior insula differentiates participants with first-episode schizophrenia spectrum disorders from controls: a machine-learning study. *Psychol. Med.* 46, 2695–2704.
- Morioka, H., Calhoun, V., Hyvärinen, A., 2020. Nonlinear ICA of fMRI reveals primitive temporal structures linked to rest, task, and behavioral traits. *Neuroimage* 218, 116989.
- Northoff, G., Hirjak, D., Wolf, R.C., Magioncalda, P., Martino, M., 2020. All roads lead to the motor cortex: psychomotor mechanisms and their biochemical modulation in psychiatric disorders. *Mol. Psychiatry* 26, 92–102.
- Plis, S.M., Hjelm, D.R., Salakhutdinov, R., Allen, E.A., Bockholt, H.J., Long, J.D., Johnson, H.J., Paulsen, J.S., Turner, J.A., Calhoun, V.D., 2014. Deep learning for neuroimaging: a validation study. *Front. Neurosci.* 8, 229.
- Polyn, S.M., Natu, V.S., Cohen, J.D., Norman, K.A., 2005. Category-specific cortical activity precedes retrieval during memory search. *Science* 310, 1963.
- Satterthwaite, T.D., Elliott, M.A., Gerraty, R.T., Ruparel, K., Loughead, J., Calkins, M.E., Eickhoff, S.B., Hakonarson, H., Gur, R.C., Gur, R.E., Wolf, D.H., 2013. An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. *Neuroimage* 64, 240–256.
- Shinn, A.K., Roh, Y.S., Ravichandran, C.T., Baker, J.T., Öngür, D., Cohen, B.M., 2017. Aberrant cerebellar connectivity in bipolar disorder with psychosis. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* 2, 438–448.
- Singh, I., Rose, N., 2009. Biomarkers in psychiatry. *Nature* 460, 202–207.
- Spitzer, R.L., Gibbon, M.E., Skodol, A.E., Williams, J.B., First, M.B., 2002. DSM-IV-TR Casebook: A Learning Companion to the Diagnostic and Statistical Manual of Mental Disorders, Text Rev. American Psychiatric Publishing, Inc.
- Tamminga, C.A., Ivleva, E.I., Keshavan, M.S., Pearson, G.D., Clementz, B.A., Witte, B., Morris, D.W., Bishop, J., Thaker, G.K., Sweeney, J.A., 2013. Clinical phenotypes of psychosis in the bipolar-schizophrenia network on intermediate phenotypes (B-SNIP). *Am. J. Psychiatry* 170, 1263–1274.
- Tamminga, C.A., Pearson, G., Keshavan, M., Sweeney, J., Clementz, B., Thaker, G., 2014. Bipolar and schizophrenia network for intermediate phenotypes: outcomes across the psychosis continuum. *Schizophr. Bull.* 40, 131–137.
- Tandon, R., Nasrallah, H.A., Keshavan, M.S., 2009. Schizophrenia, “just the facts” 4. Clinical features and conceptualization. *Schizophr. Res.* 110, 1–23.
- Wang, Z., Meda, S.A., Keshavan, M.S., Tamminga, C.A., Sweeney, J.A., Clementz, B.A., Schretlen, D.J., Calhoun, V.D., Lui, S., Pearson, G.D., 2015. Large-scale fusion of gray matter and resting-state functional MRI reveals common and distinct biological markers across the psychosis spectrum in the B-SNIP cohort. *Front. Psychiatry* 6, 174.
- Wylie, K.P., Tregellas, J.R., 2010. The role of the insula in schizophrenia. *Schizophr. Res.* 123, 93–104.
- Xie, J., Girshick, R., Farhadi, A., 2016. Unsupervised deep embedding for clustering analysis. *International Conference on Machine Learning*, pp. 478–487.
- Yan, C., Cheung, B., Kelly, C., Colcombe, S., Craddock, R.C., Di Martino, A., Li, Q., Zuo, X.-N., Castellanos, F.X., Milham, M.P., 2013. A comprehensive assessment of regional variation in the impact of head micromovements on functional connectomics. *Neuroimage* 76, 183–201.
- Yan, W., Plis, S., Calhoun, V.D., Liu, S., Jiang, R., Jiang, T.Z., Sui, J., 2017. Discriminating schizophrenia from normal controls using resting state functional network connectivity: a deep neural network and layer-wise relevance propagation method. 2017 IEEE 27th International Workshop on Machine Learning for Signal Processing (MLSP), pp. 1–6.
- Yan, C., Zang, Y., 2010. DPARSF: A MATLAB Toolbox for “Pipeline” Data Analysis of Resting-state fMRI. *Front. Syst. Neurosci.* 4.
- Yan, W., Zhang, H., Sui, J., Shen, D., 2018. Deep chroconnectome learning via full bidirectional long short-term memory networks for MCI diagnosis. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, pp. 249–257.
- Yan, W., Calhoun, V., Song, M., Cui, Y., Yan, H., Liu, S., Fan, L., Zuo, N., Yang, Z., Xu, K., Yan, J., Lv, L., Chen, J., Chen, Y., Guo, H., Li, P., Lu, L., Wan, P., Wang, H., Wang, H., Yang, Y., Zhang, H., Zhang, D., Jiang, T., Sui, J., 2019. Discriminating schizophrenia using recurrent neural network applied on time courses of multi-site fMRI data. *EBioMedicine* 47, 543–552.
- Yao, D., Sui, J., Wang, M., Yang, E., Jiaerken, Y., Luo, N., Yap, P., Liu, M., Shen, D., 2020. A Mutual Multi-Scale Triplet Graph Convolutional Network for Classification of Brain Disorders Using Functional or Structural Connectivity. *IEEE Transactions on Medical Imaging* <https://doi.org/10.1109/TMI.2021.3051604> In press.
- Yates, W.R., Jacoby, C.G., Andreasen, N.C., 1987. Cerebellar atrophy in schizophrenia and affective disorder. *Am. J. Psychiatry* 144 (4), 465–467.
- Yu, Q., Wu, L., Bridwell, D.A., Erhardt, E.B., Du, Y., He, H., Chen, J., Liu, P., Sui, J., Pearson, G., Calhoun, V.D., 2016. Building an EEG-fMRI multi-modal brain graph: a concurrent EEG-fMRI study. *Front. Hum. Neurosci.* 10, 476.
- Zeng, L.-L., Shen, H., Liu, L., Hu, D., 2013. Unsupervised classification of major depression using functional connectivity MRI. *Hum. Brain Mapp.* 35, 1630–1641.
- Zhu, J., Zou, H., Rosset, S., Hastie, T., 2009. Multi-class AdaBoost. *Stat Interface.* vol. 2, pp. 349–360.