

Functional network connectivity (FNC)-based generative adversarial network (GAN) and its applications in classification of mental disorders



Jianlong Zhao^{a,b}, Jinjie Huang^{a,*,**}, Dongmei Zhi^{b,c}, Weizheng Yan^{b,c}, Xiaohong Ma^{d,e}, Xiao Yang^{d,e}, Xianbin Li^f, Qing Ke^g, Tianzi Jiang^{b,c,i}, Vince D. Calhoun^h, Jing Sui^{b,c,i,*}

^a Department of Automation, Harbin University of Science and Technology, Harbin, 150080, China

^b Brainnetome Center and National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, 100190, China

^c University of Chinese Academy of Sciences, Beijing, 100190, China

^d Psychiatric Laboratory and Mental Health Center, the State Key Laboratory of Biotherapy, West China Hospital of Sichuan, 610041, China

^e Huaxi Brain Research Center, West China Hospital of Sichuan University, 610041, China

^f Beijing Key Lab of Mental Disorders, Beijing Anding Hospital, Capital Medical University, Beijing, China

^g Department of Neurology, the First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China

^h Tri-Institutional Center for Translational Research in Neuroimaging and Data Science (TReNDS) Center, Georgia State University, Georgia Institute of Technology, and Emory University, Atlanta, GA, 30303, USA

ⁱ CAS Center for Excellence in Brain Science, Institute of Automation, Chinese Academy of Sciences, Beijing, 100190, China

ARTICLE INFO

Keywords:

Resting-state fMRI
Generative adversarial networks (GAN)
Deep learning
Classification
Major depressive disorders
Schizophrenia

ABSTRACT

As a popular deep learning method, generative adversarial networks (GAN) have achieved outstanding performance in multiple classifications and segmentation tasks. However, the application of GANs to fMRI data is relatively rare. In this work, we proposed a functional network connectivity (FNC) based GAN for classifying psychotic disorders from healthy controls (HCs), in which FNC matrices were calculated by correlation of time courses derived from non-artefactual fMRI independent components (ICs). The proposed GAN model consisted of one discriminator (real FNCs) and one generator (fake FNCs), each has four fully-connected layers. The generator was trained to match the discriminator in the intermediate layers while simultaneously a new objective loss was determined for the generator to improve the whole classification performance. In a case for classifying 269 major depressive disorder (MDD) patients from 286 HCs, an average accuracy of 70.1% was achieved in 10-fold cross-validation, with at least 6% higher compared to the other 6 popular classification approaches (54.5–64.2%). In another application to discriminating 558 schizophrenia patients from 542 HCs from 7 sites, the proposed GAN model achieved 80.7% accuracy in leave-one-site-out prediction, outperforming support vector machine (SVM) and deep neural net (DNN) by 3%–6%. More importantly, we are able to identify the most contributing FNC nodes and edges with the strategy of leave-one-FNC-out recursively. To the best of our knowledge, this is the first attempt to apply the GAN model on the FNC-based classification of mental disorders. Such a framework promises wide utility and great potential in neuroimaging biomarker identification.

1. Introduction

Mental disorders cause high socioeconomic burdens and many exhibit comorbidity between each other (Kessler et al., 2012). Machine learning methods based on neuroimaging data have been widely applied in the classification of mental disorder in order to facilitate the diagnosis objectively e.g., support vector machine (SVM), linear discriminant analysis (LDA), and nearest neighbors (NN) (Gao et al., 2018;

Jie et al., 2018; Sato et al., 2015; Zhang et al., 2015). In most cases, the raw fMRI data needs to be extracted and processed in order to acquire less redundant and more informative features. Recently, deep learning has achieved remarkable performance in many research field (Arbabshirani et al., 2017; Calhoun and Sui, 2016; LeCun et al., 2015; Yan et al., 2019a; Yang et al., 2019) due to its ability to automatically learns the patterns from the big data without feature selection. For instance, (Kim et al., 2016) adopted a stacked autoencoder to initialize

* Corresponding author at: Institute of Automation, Chinese Academy of Sciences, Beijing, China.

** Corresponding author at: Department of Automation, Harbin University of Science and Technology, No. 52, Xuefu Road, Nangang District, Harbin, 150080, China.

E-mail addresses: [jjhuang@hrbust.edu.cn](mailto:jjhhuang@hrbust.edu.cn) (J. Huang), kittysj@gmail.com (J. Sui).

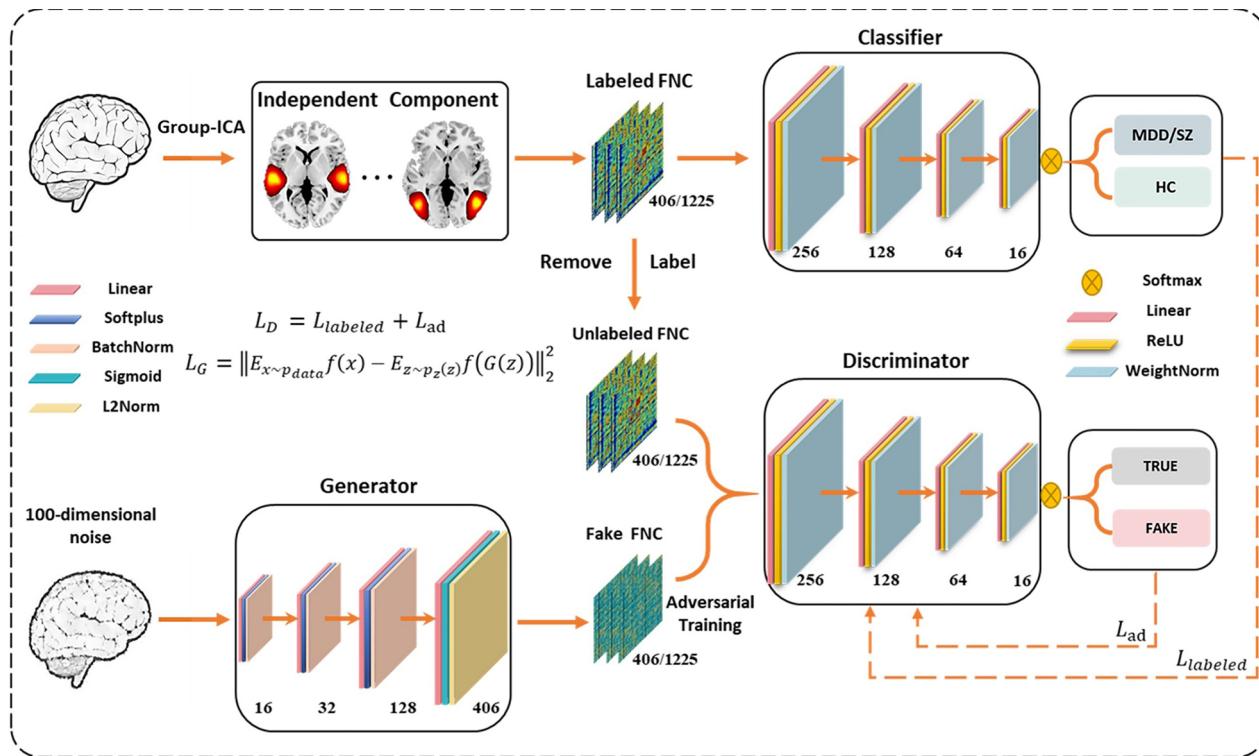


Fig. 1. Overview of the GAN model for MDD vs HC and SZ vs HC classification.

Figure 1 The GAN model was composed of a discriminator and a generator with four fully-connected layers. FNC estimated for each subject based on group-ICA was used as the input of the GAN module.

its own weight based on pre-training to increase schizophrenia classification accuracy. (Zeng et al., 2018) investigated the discriminant autoencoder network for multi-site classification of schizophrenia with fMRI. (Yan et al., 2019a) proposed a multi-scale RNN model to enable schizophrenia classification by using time courses of fMRI independent components (ICs) directly. Basically, deep learning methods have demonstrated powerful discriminating ability for classifying mental disorders using fMRI data.

In particular, generative adversarial networks (GAN) have drawn increasing attention due to their capability to perform data generation and have been widely used in the fields of image synthesis, reconstruction, segmentation, and classification (Yi et al., 2019). It has been recently proven successful on several standard classification benchmark tasks (Feng et al., 2019; Salimans et al., 2016; Springenberg, 2015; Vandenbende et al., 2019) with the fact that the generated samples promote the GAN's classification ability through adversarial learning, especially in the case of small samples size. However, the high-dimensional data in medical imaging is often with only a limited sample size, which makes it challenging for the classifiers to learn a good decision boundary. In order to improve GAN's classification performance, feature matching has been implemented as state-of-the-art approaches in many GAN models. Specifically, here feature matching is setting a new objective loss for the generator which is trained to match the feature value in the intermediate layers of the discriminator. It can generate fake samples within the high-density region in feature space, which can split the bounds of different classes because of its continuity to further enhance classification performance (Dai et al., 2017; Salimans et al., 2016).

Till now, the application of GAN on fMRI data is relatively rare. Inspired by this, in this work, we proposed a functional network connectivity (FNC)-based GAN with feature matching for classifying mental disorders from healthy controls (HCs), in which FNC matrices were calculated by correlation of time courses derived from non-artefactual fMRI independent components (ICs). The reason we used FNC is that it

is able to reflect functional interactions among structurally segregated brain regions, known as intrinsic functional connectivity networks (ICN) (Calhoun et al., 2014) and reduce the feature dimension greatly. Furthermore, the FNC estimated by independent component analysis (group-ICA) also demonstrated to be more reliable and sensitive in biomarker detection for psychosis (Du et al., 2015; Sui et al., 2018) and can be applied to new individuals (Du et al., 2018).

To the best of our knowledge, this is the first attempt to apply GAN on FNC features to discriminate psychiatric patients from controls. We tried on two kinds of disorders: 1) MDD vs HC (555 subjects, 269 MDD, 286 HCs); 2) SZ vs HC (1100 subjects, 558 SZ, 542 HCs). The proposed GAN model combines fake FNC generation and classification together into a unified optimization framework. To increase result interpretability, leave-one-FNC-out looping was adopted to identify the most contributing FNCs.

2. Materials and methods

2.1. Generative adversarial network with feature matching

The proposed GAN model consisted of one discriminator (real FNCs) and one generator (fake FNCs), each has four fully-connected layers. Furthermore, feature matching, the generator was trained to match the feature value in intermediate layers of the discriminator, was implemented by specifying a new objective loss for the generator to improve classification performance. FNC estimated for each subject based on group-ICA was used as the input of the GAN module. To verify that adversarial training (AD) and feature matching (FM) in GAN do directly improve the learning capacity of the GAN method, the ablation experiments were performed in the paper. Equations are listed in the Supplementary files for a precise definition of the GAN model with feature matching.

2.1.1. Generative adversarial network

GAN may learn a better decision boundary than the traditional methods because the generated samples promote the discriminator's classification ability through adversarial learning, especially in the case of a small sample size. As shown in Fig. 1, a GAN architecture was applied for classification. The GAN model was composed of a discriminator and a generator with four fully-connected layers, optimizing its network parameters through continuous competitions. The discriminator's output layer has $K+1$ classes, where $K=2$ for the real class from data x and $K+1$ class for the generated image. It is worthy to note that the discriminator can realize multiclass discriminative ability by a soft-max classifier. Compared to the unsupervised GAN model, the proposed GAN incorporated both labeled and generated data into the loss function.

2.1.2. Feature matching

Feature matching was able to address the instability of GAN by specifying a new objective loss for the generator, which can prevent it from overtraining on the discriminator (Salimans et al., 2016). Instead of directly maximizing the output of the discriminator, the new objective enabled the generator to generate data that matched the statistics of the real data, where we only used the discriminator to specify the statistics that were worth matching. The generator was trained to match the feature value in intermediate layers of the discriminator. The intermediate layer we set in this study is the third hidden layer. Through feature matching by the generator, discriminative features can be found to distinguish real data from generated data by training the discriminator. It can generate fake samples within the high-density region in feature space, which can split the bounds of different classes because of its continuity to further enhance classification performance (Dai et al., 2017).

Activations in an intermediate layer of the discriminator were denoted as $f(x)$, our new objective for the generator can be defined as follows:

$$\|E_{x \sim P_{\text{data}}} f(x) - E_{z \sim p_z(z)} f(G(z))\|_2^2 \quad (1)$$

Where x represents the input data, P_{data} and $p_z(z)$ represent the distribution of the input data and the distribution of the input noise, G is the generator, and D is the discriminator.

2.2. GAN model implementation

The GAN model was trained and evaluated by using Theano and Scikit-learn (<https://scikit-learn.org/>). The GAN model consisted of one discriminator and one generator that had four layers respectively. The output layer of the discriminator has $K+1$ classes, where $K=2$ for the input pattern from HC and patient group, and the $[K+1]_{\text{th}}$ class is for generating images. The high-dimensional features of each subject may lead to overfitting on the training set. In order to reduce the overfitting susceptibility, L2 normalization ($L2 = 0.1$) and batch normalization were added to the generator to improve modal generalization. In addition, weight normalization (weight normalization = 0.1) was used to the output of each layer of the discriminator to prevent overfitting.

The Adam optimizer was adopted as minimizing the loss of the GAN model and a standard error back-propagation algorithm was used by training the GAN model with multiple layers. The batch size was set as 120 in the training process. In addition, to overcome overfitting, the weights were controlled with weight norm regularization. Different layers were attempted to the constructed architecture of the GAN model and results revealed that using four layers could obtain the optimal classification performance. The learning rate was set as 0.0003. In this experiment, the training time for GAN was about one hour, while the testing time for a new subject is less than 1 s. All trainings and experiments were completed on a standard workstation (Intel(R) Xeon(R) CPU E5-1650 v4 @ 3.60 GHz, 6 CPU cores, 12GB NVIDIA GTX TITAN).

2.3. GAN classification framework

In this study, k-fold multi-site pooling classification and leave-one-site-out classification were conducted. In order to verify the validity of the proposed GAN method, five conventional methods SVM, NN, Gaussian Process, Naive Bayes, and a deep learning method DNN were used as a comparison in this study.

10-fold cross-validation was applied to verify the generalization ability of the classifiers in the multi-site pooling classification. The training dataset and the testing dataset were embedded in nested 10-fold cross-validation cycles. As for all the models, we used nine folds as the training set, and one fold for the testing dataset. A specific imaging site was applied as the testing set and the sample of other sites as the training set in the leave-one-site-out transfer classification. Accuracy (ACC), sensitivity (SEN), specificity (SPE), F-score (F1) and area under curve (AUC) verify the classification performance. The formula is defined as follows:

$$\begin{aligned} \text{ACC} &= \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \text{SPE} = \frac{\text{TN}}{\text{TN} + \text{FP}}, \text{SEN} = \frac{\text{TP}}{\text{TP} + \text{FN}} \\ \text{PPV} &= \frac{\text{TP}}{\text{TP} + \text{FP}}, \text{F1} = 2 \frac{\text{SEN} * \text{PPV}}{\text{SEN} + \text{PPV}} \end{aligned}$$

Where true positive, true negative, false positive, false negative and positive predictive value was denoted as TP, TN, FP, FN, PPV respectively. The means and standard deviations of ACC, SEN, SPE, F1 and AUC were obtained by 10-time 10-fold cross-validation and a two-sample t-test was adopted as the comparison of the performance of different classification models.

2.4. Estimate the contributing FNC (Leave-one-FNC-out)

The basic idea is that potential biomarkers are such that the removal of these features leads to the most significant degradation of accuracy. The specific implementation is as follows:

In order to calculate the most contributing FNC, we substitute the d_{th} FNC with default value 0 while keeping other FNC as they were. This is equivalent to erasing the contribution of d_{th} FNC. The classification accuracy with reduced FNC may decrease compared to that using all FNCs. Compared to that using all features, the classification accuracy of the trained model which is fed with reduced features may decrease. Based on the leave-one-feature-out strategy, we sorted the classification accuracy when getting rid of each FNC. Then the FNCs which cause the top 3% decrease of classification accuracy were recorded and considered as the most contributing features. Similarly, for the most contributing FNC node, we summarized the top 30% contributing FNCs and discovered the top 10% nodes that linking most of the FNCs (higher degrees). Moreover, the selected features are visualized with BrainnetViewer (<https://www.nitrc.org/projects/bnv/>) and BRANT(<http://brant.brainnetome.org/>).

The brain regions in MDD were grouped into seven functional subsystems for visualization based on their anatomical and functional properties: auditory (AUD), cognitive control (CC), sensorimotor network (SMN), default-mode network (DMN), frontal network (FN), visual (VIS), and cerebellum (CB).

The brain regions in SZ were grouped into eight functional subsystems for visualization based on their anatomical and functional properties: sub-cortical (SC), auditory (AUD), Sensorimotor network (SMN), visual (VIS), default-mode network (DMN), attentional network (ATN), frontal network (FN), and cerebellum (CB). long term, we should look at both analyzed via neuromark so we can directly compare different mental disorders such as MDD and SZ.

3. Participants

In this study, 555 Chinese Han subjects including 269 MDD patients

Table 1
Demographic information of the MDD/HC database.

	MDD	HC	P-value
Number of scans	269	286	NA
Age (mean \pm std, yrs)	32.8 \pm 10.6	31.7 \pm 10.4	0.22
Gender (M/F)	105/164	105/181	0.18
SYMPTOM SEVERITY			
HDRS	18.3 \pm 8.3	NA	NA
BDI	20.4 \pm 6.9	NA	NA

Table 1 Note: MDD: major depressive disorder; HC: healthy control; yrs: years; BDI: beck depression rating scale; std: standard deviation; HAMD: Hamilton depression scale; F: female; M: male; NA: not applicable.

and 286 HCs derived from 4 sites were used, including the Henan Mental Hospital of Xinxiang (Site 1), the West China Hospital of Sichuan (Site 2), the Anding Hospital of Beijing (Site 3), and the First Affiliated Hospital of Zhejiang (Site 4). No significant group difference between HC and MDD was obtained in age or gender (age: $p = 0.22$; gender: $p = 0.18$). The DSM-IV based on SCID-P interviews was used to diagnose patients. HCs were interviewed using SCID-I/NP, and the first-degree relatives with any mental illness were excluded. Each site acquired approval from their respective research ethics boards, and informed consent was received from each subject prior to scanning by each site's Institutional Review Boards. Patients were excluded if they had a history of head injury which cause significant non-psychiatric medical illness, or unconsciousness for longer than a few seconds. These individuals were also excluded with obsessive-compulsive disorder, posttraumatic stress disorder or an active substance use disorder. **Table 1** provides a statistical information, please see more information in Table S1.

For SZ data, 1100 Chinese subjects (542 HCs and 558 SZ patients) from seven sites were used, including the Peking University sixth Hospital (Site 1), the Huilongguan Hospital (Site 2), the Henan Mental Hospital (Site 3), the Henan Mental Hospital (Site 4), the Xijing Hospital (Site 5), the Renmin Hospital of Wuhan University (Site 6), and the Zhumadian Psychiatric Hospital (Site 7). Each site received approval from the relevant Ethics Committees and all study participants obtained written informed consents. All the SZ patients were evaluated based on the SCID and diagnosed by experienced psychiatrists by using DSM-IV-TR. All the HCs which were free of Axis I or II disorders were recruited from the same local geographical areas as the patients cohort through local advertisement. The exclusion criteria were suitable for all subjects included substance abuse or dependence, pregnancy, prior electroconvulsive therapy, and current or past neurological illness. **Table 2** provides a statistical information, please see more information in Table S2.

3.1. Data acquisition

The resting-state fMRI data for MDD patients and demographically matched HCs were collected on a 3 T Philips scanner at Site 2 and on a 3 T Siemens scanner (Verio, Germany) at site 1, site 3 and site 4. The field of view (FOV) was 240 \times 240 mm (64 \times 64 matrix) at site 2, 220 \times 220 mm (64 \times 64 matrix) at site 1 and site 4, 200 \times 200 mm (64 \times 64 matrix) at site 3 and all sites have 240 volumes of echo planar images

Table 2
Demographic information of the SZ/HC database.

	HC	SZ
Number of scans	542	558
Age (mean \pm std, yrs)	28.0 \pm 7.2	27.6 \pm 7.1
Gender (M/F)	276/266	292/266

Table 2 Note: SZ: schizophrenia; HC: healthy control; yrs: years; std: standard deviation; F: female; M: male.

obtained by the same parameters: repetition time (TR) = 2000 ms; echo time (TE) = 30 ms; flip angle (FA) = 90°. Site 1 was 33 sequential ascending axial slices of 4 mm thickness, Site 3 was 33 sequential ascending axial slices of 3.5 mm thickness and site 2, and site 4 were 38 sequential ascending axial slices of 4 mm thickness. During scanning, head movement and scanner noise were minimized by foam padding and earplugs and subjects followed the instructions to keep eyes closed and stay awake when lying still.

The resting-state fMRI data for SZ patients and age- and gender-matched HCs were collected on a 3.0 T Siemens Trio Tim Scanner in Site 1, site 2 and site 5, a 3.0 T Siemens Verio Scanner at site 3, and a 3.0 T Signa HDx GE Scanner at site 4, site 6, and site 7. During scanning, head movement and scanner noise were minimized by foam padding and earplugs and subjects followed the instructions to keep eyes closed and stay awake while lying still.

3.2. Data preprocessing and FNC measure

The SPM8 (<http://www.fil.ion.ucl.ac.uk/spm/>) software is used by the preprocessing of the resting-state fMRI data and the preprocessing flow is as follows: The first 10 volumes were discarded to exclude T1 equilibration effects, slice timing correction, motion correction, normalization in the standard Montreal Neurological Institute (MNI) space (3 mm isotropic voxels), denoising and spatially smoothing with an 8 mm full-width half max Gaussian kernel. Each voxel time course was z-scored to normalize variance across space. In addition, to limit the impact of head motion, a maximum translation of > 2 mm or rotation of > 2 mm or framewise displacements (FD) > 1 mm was excluded in the subjects. Results indicate that mean FD for all subjects were < 0.5 mm and there was no significant group difference between HC vs SZ patients and a significant group difference between HC vs MDD patients on mean FD (HC: 0.14 ± 0.070 , MDD: 0.12 ± 0.066 , two-sample t-test: $p = 0.001$, HC: 0.137 ± 0.071 , SZ: 0.142 ± 0.085 , two-sample t-test: $p = 0.98$). For MDD and HC, since the mean FD showed a significant group difference, we further regressed out the mean FD from each functional network connectivity to eliminate the potential group difference (Yuan et al., 2016). Moreover, we regressed age, gender in order to eliminate the effects of these variables.

For MDD patients and HCs, the fMRI data were decomposed into subject-specific spatial ICs and its corresponding time courses using a spatially constrained ICA back-reconstruction approach called group ICA implemented in the GIFT software (<http://trendscenter.org/software/gift>) which is robust to artifacts (Du et al., 2016), resulting in 29 selected ICNs from 100 group independent components. More details of spatial maps are listed in the Supplementary Figure. S3. ICA was applied first to decompose the 4D fMRI into useful ICNs and some noise ICs that were thrown off, which is also a denoising step. Due to the time courses of the selected ICNs may still contain remaining noise sources, the additional post-processing including detrending, controlling covariates by regression, and bandpass filtering was applied (Allen et al., 2012). The time courses of selected ICNs were post-processed by detrending linear, quadratic and cubic trends, regressing out 6 realignment parameters and their temporal derivatives, despiking, and bandpass filtering between [0.01–0.15] Hz using a 5th order Butterworth filter. FNC was computed as the pairwise correlation between any two ICN time courses for each subject, which was further used as input feature of the GAN model.

For SZ patients and HCs, the fMRI data were decomposed into subject-specific spatial ICs and its time courses by performing group-ICA within the GIFT software (Du et al., 2016), resulting in 50 selected ICNs from 100 group independent components. More details of spatial maps are listed in the Supplementary Figure. S4. The time courses (TCs) of selected ICNs are post-processed by detrending, regressing out head motion, despiking and lowpass filtering (< 0.15 Hz) and the FNC matrices are calculated as the Pearson's correlation in each pair of ICs, which was further used as input feature of the GAN model.

Table 3

Performance of different methods in MDD/HC classification with 10-fold cross-validation.

Method	ACC(%)	SEN(%)	SPE(%)	f1(%)	AUC(%)
Nearest Neighbors	54.5(0.9)	55.9(0.9)	53.1(0.9)	56.1(0.9)	54.5(0.9)
AdaBoost	54.6(1.7)	55.8(1.5)	53.3(1.8)	56.6(1.8)	54.5(1.7)
Naive Bayes	59.2(0.9)	62.1(1.0)	56.8(0.8)	57.3(1.0)	59.4(0.9)
Gaussian Process	60.4(0.6)	61.1(0.8)	59.6(0.5)	62.4(0.4)	60.3(0.6)
Linear SVM	62.8(0.7)	61.9(0.6)	64.2(0.9)	66.7(0.7)	62.5(0.7)
Deep Neural Net	64.2(0.9)	64.4(1.0)	64.1(1.2)	66.3(1.2)	64.1(0.9)
GAN	70.1(0.6)	73.5(4.7)	66.5(4.7)	71.7(1.5)	70.3(0.9)

Table 3 The performance of different methods in MDD/HC classification with 10-fold cross-validation. SVM, support vector machine; GAN, Generative Adversarial Networks; ACC, Accuracy; SEN, sensitivity; SPE, specificity; F1, F-score; AUC, area under curve; MDD: major depressive disorder; HC: healthy control.

4. Results

4.1. Ten-fold and leave-one-site-out classification in MDD/HC

To demonstrate the performance of the proposed method, we compared the proposed method with other state-of-art methods, including five conventional methods SVM, NN, Gaussian Process, Naive Bayes, and AdaBoost and a deep learning method DNN. The training dataset and the testing dataset were embedded in nested 10-fold cross-validation cycles. As for all the models, we used nine folds as the training set, and one fold for the testing dataset. In order to test the generalization of the model, we train the different models with the leave-one-site-out method. The means and standard deviations of ACC, SEN, SPE, F1 and AUC were obtained by 10-time 10-fold cross-validation and a two-sample t-test was adopted as the comparison of the performance of different classification models.

The classification results of the GAN model compared with the other six methods were summarized in Table 3, Table 4, and Table S3. The accuracy of $70.1 \pm 0.6\%$ and $64.3 \pm 2.9\%$ was obtained by using the GAN model in the 10-fold cross-validation and leave-one-site-out method, which is significantly higher than those obtained by using those traditional classification methods and DNN approach. For the binary classification problem, the results suggested that the proposed GAN method significantly outperformed five conventional methods SVM, NN, Gaussian Process, Naive Bayes, and AdaBoost and a deep learning method DNN in terms of ACC, SEN, SPE, F1, and AUC measures ($P < 0.05$, two-sample t-test), and achieved good generalization results in the leave-one-site-out method (Table 3, Table 4, Table S3, Fig. 2A and B) on account of our GAN model makes full use of adversarial learning and feature matching that alleviate the small size problem of FNC images and mitigate the instability of GAN. The t-distributed stochastic neighbor embedding (t-SNE) was used to visualize the GAN classification performance. The t-SNE visualization result as

Table 4

Performance of the leave-one-site-out method in MDD/HC classification.

Method	ACC(%)	SEN(%)	SPE(%)	f1(%)	AUC(%)
Nearest Neighbors	55.5(2.0)	57.3(2.5)	53.9(4.6)	55.5(4.4)	55.6(1.2)
AdaBoost	47.9(1.4)	49.3(5.2)	47.1(1.5)	42.1(1.6)	48.3(2.0)
Naive Bayes	55.1(6.8)	59.5(9.5)	52.8(6.9)	48.2(12.7)	55.6(6.3)
Gaussian Process	56.9(5.0)	59.2(4.6)	55.0(6.8)	55.8(8.9)	57.1(2.6)
Linear SVM	55.9(6.2)	59.0(2.7)	53.7(9.0)	52.4(15.8)	56.1(4.8)
Deep Neural Net	55.0(3.4)	60.8(5.0)	52.4(3.1)	44.7(2.9)	55.6(1.8)
GAN	64.3(2.9)	61.5(3.1)	70.8(6.5)	70.5(1.5)	63.8(3.4)

Table 4 The performance of the leave-one-site-out method in MDD/HC classification. SVM, support vector machine; GAN, Generative Adversarial Networks; ACC, Accuracy; SEN, sensitivity; SPE, specificity; F1, F-score; AUC, area under curve; MDD: major depressive disorder; HC: healthy control.

shown in Fig. S1, indicating that the proposed GAN model shows better separating power and can provide intuition interpretation for the increase of classification performance. In addition, the result that generated samples by the GAN method in comparison with the true FNC as shown in Fig. S2. The analysis of the comparison is listed in the Supplementary files, revealing that the distribution of FNC generated by GAN is similar to the distribution of real FNC.

4.2. The classification of ten-fold and leave-one-site-out in SZ/HC

To validate the effectiveness of the GAN model, we evaluated the proposed GAN method on the task of SZ classification (542 HCs and 558 SZ). The proposed method was compared with six state-of-art methods using 10-fold cross-validation and leave-one-site-out methods. Results were summarized in Table 5, Table 6, and Table S4, the accuracy of $82.1 \pm 0.7\%$ and $80.7 \pm 3.8\%$ was obtained by using the GAN model in the 10-fold cross-validation and leave-one-site-out method, demonstrating that the classification performance of GAN was significantly better than other methods (Table 5, Table 6, Table S4, Fig. 2C and Fig. 2D). The proposed GAN model can be generalized to classify the new site. In addition, the t-SNE visualization result is shown in Fig. S1.

4.3. Most contributing FNC for MDD-HC and SZ-HC – classification

The research intention of the fMRI classification is that a specific set of features may be used to diagnose the mental disorders as stable biomarkers and to ensure different datasets have a consistent result. In addition, the interpretability of deep learning has always been one of the factors limiting the application of deep learning in the medical field instead of classification performance. Recently, some research has evaluated the interpretability of deep learning (Kim et al., 2016; Yan et al., 2019b; Zeng et al., 2018). The basic idea is that potential biomarkers are such that the removal of these features leads to the most significant degradation of accuracy. Here the most contributing FNC for classification was analyzed in the multi-site pooling strategy with the leave-one-FNC-out method. It was found that the most important FNCs for MDD are presented in Fig S5.a and Table 7. We also evaluated the proposed method on the classification of SZ (542 HCs and 558 SZ) using the leave-one-FNC-out method. It was found that the most important FNCs for MDD are presented in Fig S5.b and Table 8. Moreover, to present the functional "network" connectivity feature, the BRANT (Xu et al., 2018) was applied to plot these brain maps which contain the most contributing FNCs related to a specific network and these brain maps were presented in Fig. 3 and 4.

5. Ablation study

To verify that adversarial training (AD) and feature matching (FM) in GAN do directly improve the learning capacity of the GAN method, we performed the ablation experiments by removing either adversarial training or feature matching, similar as did in (Huang et al., 2016). Specifically, we compared the performance of GAN models with adversarial training vs no adversarial training (No-AD) (Fig. 5a, b) and feature matching vs no feature matching (No-FM) (Fig. 5c,d). We obtained the means and standard deviations of ACC, SEN, SPE, F1 and AUC by 10-time 10-fold cross-validation and a two-sample t-test was adopted as the comparison of the classification performance. It is clear that the GAN model with adversarial training outperformed the GAN with No-AD in almost all evaluation criteria, with a 6% increase in MDD. Moreover, the GAN model with FM also achieved better classification accuracy compared to the GAN with No-FM (70% vs 69% in MDD, 82 vs. 81% in SZ).

6. Discussion

In this work, we developed a novel GAN model using resting-state

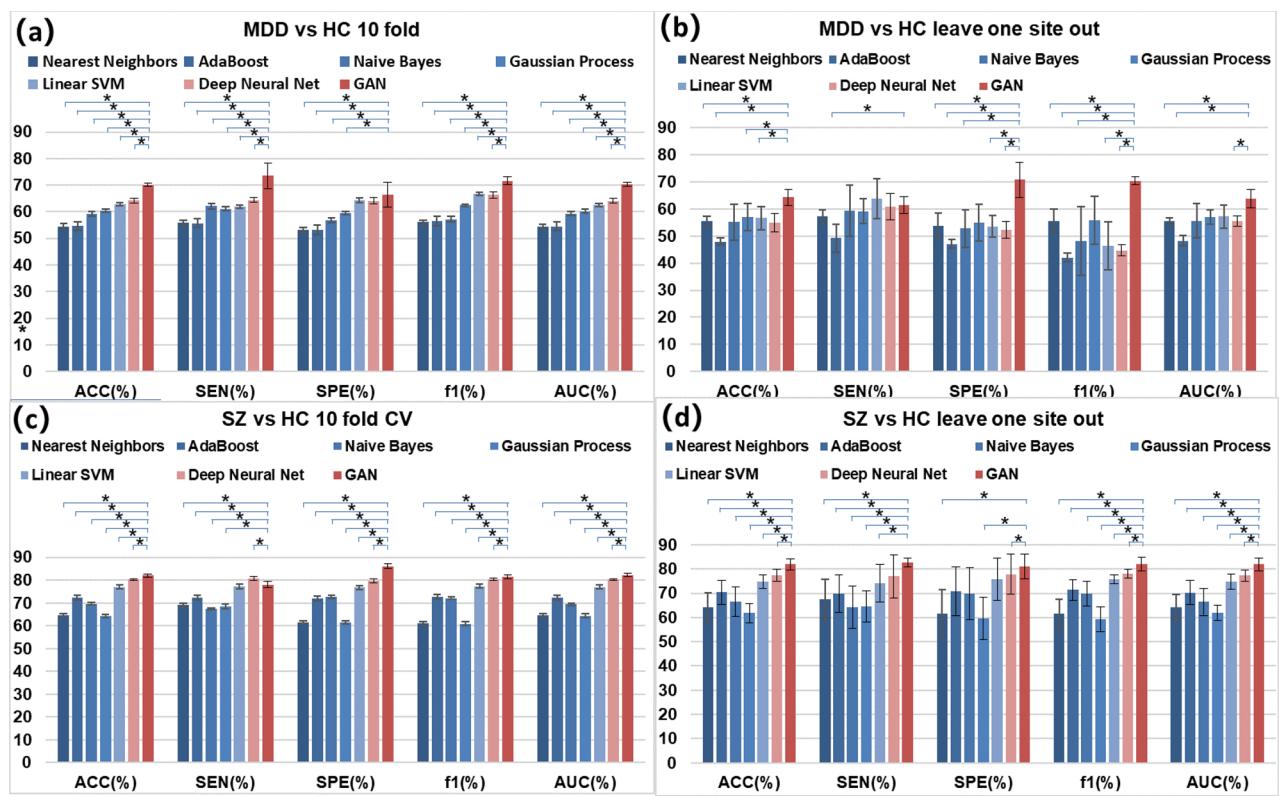


Fig. 2. The results of multi-site pooling classification and leave-one-site-out transfer classification for MDD vs HC and SZ vs HC. Figure 2 (a) The MDD classification results of 10-fold multi-site pooling classification. (b) The MDD classification results of leave-one-site-out transfer classification. (c) The SZ classification results of 10-fold multi-site pooling classification. (d) The SZ classification results of leave-one-site-out transfer classification. In all cases, * means $p < 0.05$ in two-sample t-tests.

Table 5
Performance of different methods in SZ/HC classification with 10-fold cross-validation.

Method	ACC(%)	SEN(%)	SPE(%)	f1(%)	AUC(%)
Nearest Neighbors	64.6(0.6)	69.1(0.8)	61.6(0.6)	61.0(0.8)	64.8(0.6)
AdaBoost	72.3(1.1)	72.4(1.1)	72.1(1.1)	72.8(1.0)	72.2(1.1)
Naive Bayes	69.6(0.5)	67.5(0.4)	72.7(0.6)	72.1(0.5)	69.5(0.5)
Gaussian Process	64.3(0.8)	68.5(1.0)	61.4(0.7)	60.9(0.9)	64.4(0.8)
Linear SVM	77.1(0.9)	77.4(1.0)	76.8(0.9)	77.4(0.9)	77.1(0.9)
Deep Neural Net	80.3(0.5)	80.8(0.9)	79.8(0.9)	80.5(0.5)	80.3(0.5)
GAN	82.1(0.7)	78.1(1.5)	86.2(1.1)	81.6(0.8)	82.3(0.7)

Table 5 The performance of different methods in SZ/HC classification with 10-folds cross-validation. SVM, support vector machine; No FM, No feature matching; GAN, Generative Adversarial Networks; ACC, Accuracy; SEN, sensitivity; SPE, specificity; F1, F-score; AUC, area under curve; SZ: schizophrenia; HC: healthy control.

Table 6
Performance of the leave-one-site-out methods in SZ/HC classification.

Method	ACC(%)	SEN(%)	SPE(%)	f1(%)	AUC(%)
Nearest Neighbors	64.3(5.9)	67.7(8.2)	61.8(9.8)	61.7(6.0)	64.4(5.2)
AdaBoost	70.4(5.0)	70.0(7.8)	70.8(10.0)	71.4(4.2)	70.3(5.0)
Naive Bayes	66.6(6.0)	64.4(8.8)	69.9(10.9)	69.8(5.1)	66.4(5.7)
Gaussian Process	61.8(4.0)	64.6(6.5)	59.7(8.7)	59.2(5.2)	61.9(3.1)
Linear SVM	74.9(2.9)	74.1(7.9)	75.8(8.6)	75.8(1.9)	74.9(3.1)
Deep Neural Net	77.4(2.5)	77.0(9.0)	77.8(8.3)	78.0(1.8)	77.3(2.4)
GAN	80.7(3.8)	81.8(3.7)	79.7(9.7)	80.8(3.9)	80.7(3.4)

Table 6 The performance of the leave-one-site-out method in SZ/HC classification. SVM, support vector machine; GAN, Generative Adversarial Networks; ACC, Accuracy; SEN, sensitivity; SPE, specificity; F1, F-score; AUC, area under curve; SZ: schizophrenia; HC: healthy control.

fMRI to discriminate mental disorders from HCs in large, multi-site datasets. Our model alleviates the small size problem of FNC images by making full use of generated FNC samples. In addition, the proposed GAN model defines adversarial objections between the generators and discriminator, which uses adversarial learning and feature matching to further improve the classification performance of the discriminator. Results showed that the GAN model achieved an average 70.1% accuracy with 10-fold cross-validation in MDD vs HC, at least 6% higher than five conventional methods and deep neural net (DNN) (54.5–64.2%). To validate the effectiveness of the GAN model, we further applied it to a large-scale multi-site schizophrenia (SZ) dataset including 558 patients and 542 HCs from seven sites, achieving 80.7% accuracy in leave-one-site prediction, outperforming SVM and DNN by 3–6%, demonstrating the efficacy of the GAN approach. Subsequently, to increase result interpretability, leave-one-FNC-out looping was adopted, and results suggested the most contributing FNCs were primarily located in the FN, DMN, CB, SMN, and CC for MDD and in the DMN, SC, VIS, and SMN for SZ. Thus, such a framework promises wide utility and may have great potential in neuroimaging biomarker identification. In this study the good classification performance may be attributed to the following aspects:

This is the first attempt to apply GAN on FNC-based classification, results from ablation experiments suggest that the generated samples in adversarial training did promote the classification ability for FNC-based GAN, especially in the case of a small sample size. The potential reason might be that GAN acts as a regularizer for the decision surface of the classifier to improve the classification performance (Vandenbende et al., 2019). Comparing with the conventional DNN network without the adversarial item in the cost function, we found that the accuracy of the GAN is significantly better than the conventional DNN network with a 6% increase for MDD and HC.

Table 7

The most contributing FNC with the leave-one-FNC-out method for MDD vs HC.

Brain Region	FNC node	Brain Region	FNC node
AUD	inferior temporal gyrus (ITG)	FN	middle frontal gyrus (MFG)
AUD	middle temporal gyrus (MTG)	FN	superior frontal gyrus (SFG)
DMN	anterior cingulate (AC)	CC	inferior parietal lobule (IPL)
DMN	anterior cingulate (AC)	CC	superior temporal gyrus (STG)
SMN	postcentral gyrus (PoCG)	CC	inferior parietal lobule (IPL)
DMN	precuneus	CB	fusiform gyrus (FFG)
SMN	medial frontal gyrus (MeFG)	CB	declive

Table 7 The most contributing FNC and brain region with the leave-one-FNC-out method for MDD vs HC. For corresponding connections, the brain region and FNC node are presented. AUD, auditory; CC, cognitive control; SMN, sensorimotor network; DMN, default-mode network; FN, frontal network; and CB, cerebellum.

Table 8

The most contributing FNC with the leave-one-FNC-out method for SZ vs HC.

Brain Region	FNC node	Brain Region	FNC node
SMN	postcentral gyrus (PoCG)	DMN	precuneus
VIS	middle occipital gyrus (MOG)	DMN	precuneus
VIS	middle occipital gyrus (MOG)	DMN	angular gyrus (AG)
SC	putamen	DMN	precuneus
SMN	precentral gyrus (PreCG)	DMN	angular gyrus (AG)
AUD	superior temporal gyrus (STG)	VIS	fusiform gyrus (FFG)

Table 8 The most contributing FNC and brain region with the leave-one-FNC-out method for SZ vs HC. For corresponding connections, the brain region and FNC node are presented. SMN, sensorimotor network; DMN, default-mode network; VIS, visual; SC, sub-cortical; and AUD, auditory.

Second, feature matching (FM) in GAN could directly improve the classification performance of the GAN model. We compared the performance of GAN models with feature matching vs no feature matching (No-FM). As shown in Fig. 5, the GAN model showed increased performance from almost all evaluation criteria. In this study, the leave-one-site-out transfer classification also gained promising accuracies indicating that the classification models are independent with imaging sites. Maybe due to no prior knowledge from the testing site, the leave-one-site-out transfer classification accuracies were lower than the accuracies of multi-site pooling classification to some extent. However, the present results suggest that the GAN method proposed may provide a potential realistic solution for individual diagnostic classification with MDD vs HC and SZ vs HC across independent imaging sites.

As to the most contributing FNC, results showed that the most contributing features were primarily located in FN, DMN, CB, SMN, and CC in the classification of MDD patients involving in cognitive function

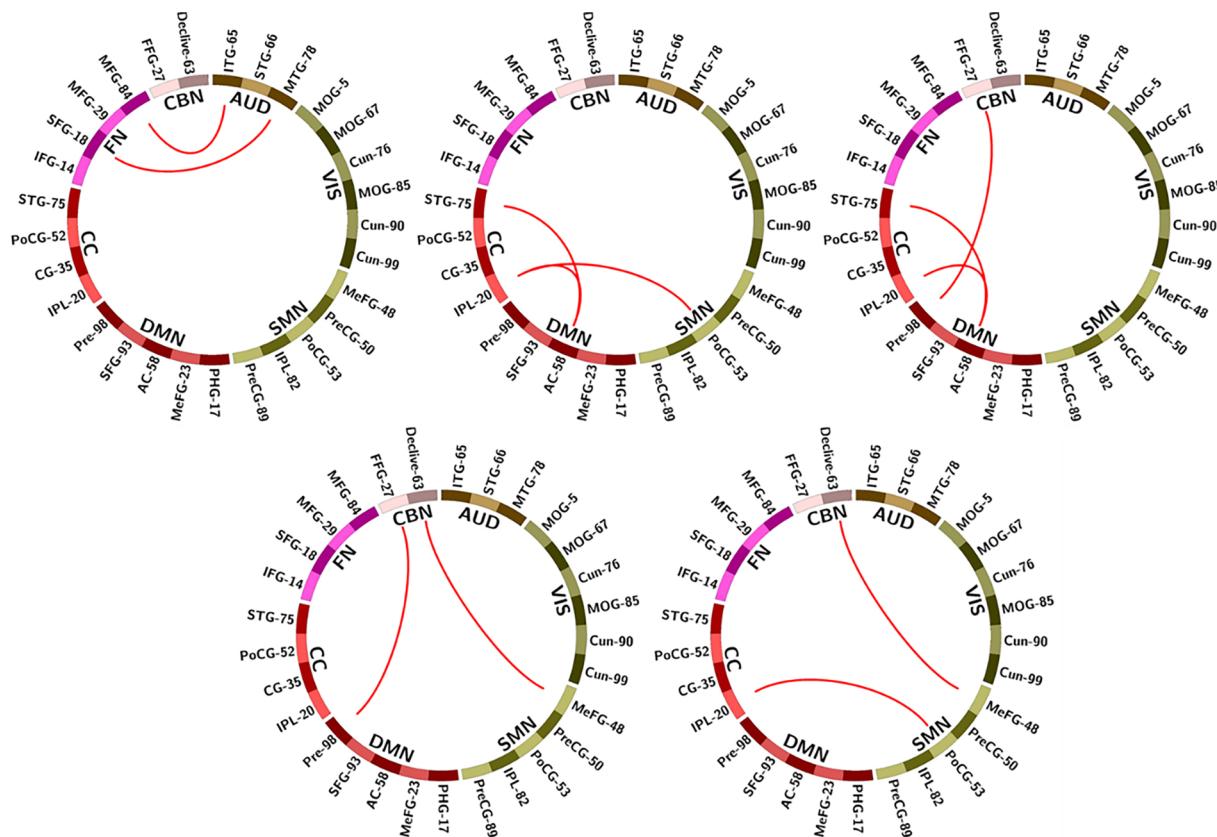


Fig. 3. Overview of the most contributing FNC related to a specific network for MDD vs HC. AUD: auditory; CC: cognitive control; SMN: sensorimotor network; DMN: default-mode network; FN: frontal network; VIS: visual; CB: cerebellum; ITG: inferior temporal gyrus; MFG: middle frontal gyrus; MTG: middle temporal gyrus; SFG: superior frontal gyrus; AC: anterior cingulate; IPL: inferior parietal lobule; STG: superior temporal gyrus; PoCG: postcentral gyrus; Precun: precuneus; MeFG: medial frontal gyrus; FFG: fusiform gyrus.

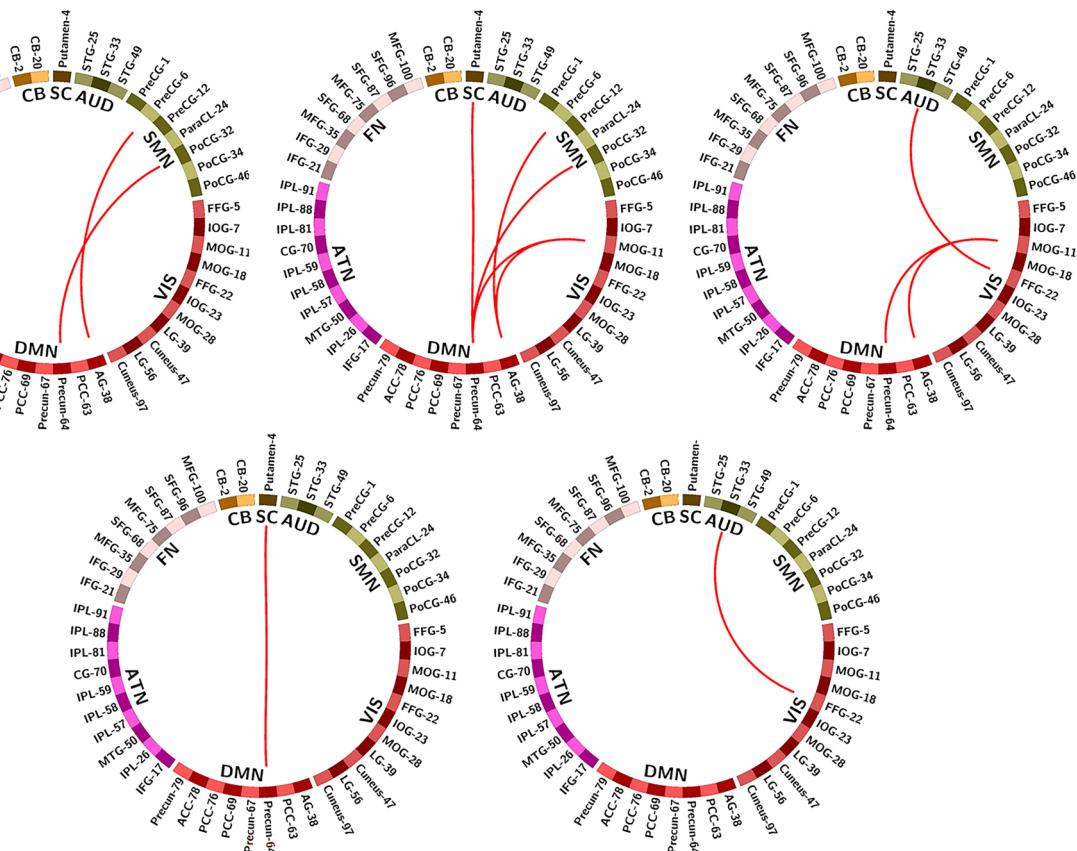


Fig. 4. Overview of the most contributing FNC related to a specific network for SZ vs HC. SC: sub-cortical; AUD: auditory; SMN: Sensorimotor network; VIS: visual; DMN: default-mode network; ATN: attentional network; FN: frontal network; CB: cerebellum; PoCG: postcentral gyrus; MOG: middle occipital gyrus; AG: angular gyrus; PreCG: precentral gyrus; STG: superior temporal gyrus; FFG: fusiform gyrus.

and emotion regulation (Kaiser et al., 2015; Mulders et al., 2015). The research (Mulders et al., 2015) is consistent with the result of the current study that patients of the MDD observed abnormal FNCs between DMN and CC, CB, which are related to frontal, parietal, and postcentral gyrus. The DMN and cerebellum are related to self-processing and emotional (Drevets et al., 2008; Sheline et al., 2010), and the FN network has an important influence on cognitive ability and mood regulation (Drevets et al., 2008; Steele and Lawrie, 2004). The imbalance within externally-directed attention and control systems was caused by the aberrant connectivity of the MDD between anterior cingulate (AC) and inferior parietal lobule (IPL) (Kaiser et al., 2015) and the perception and emotional facial expressions processing was related to the aberrant FNCs in AUD and FN located in inferior temporal gyrus (ITG), superior frontal gyrus (SFG) and middle frontal gyrus (MFG) (Haxby et al., 2000). Abnormal reactivity caused by FNCs in precuneus and fusiform gyrus (FFG) might accept as a depression early biomarker as described in (Hahn et al., 2011). The SFG is related to executive memory emotions (Fuster, 2001; Jaworska et al., 2015) and the motor, premotor and prefrontal networks are related to executive memory. Executive function abnormalities in MDD patients were implied by the aberrant FNCs with SFG in FN and ITG (Fuster, 2001; Wang et al., 2008).

Results suggested that the most contributing FNCs were primarily located in DMN, SC, VIS, and SMN in the classification of SZ patients involving in deficits in a multitude of cognitive domains (Sheffield and Barch, 2016). Abnormality of the DMN has been linked to SZ in past studies (Sui et al., 2013). The aberrant FNC is associated with angular gyrus (AG) and precuneus in the DMN, which may cause episodic memory deficits and auditory hallucinations in SZ patients (Mondino et al., 2015; Vercammen et al., 2010). Recent study manifest that the deficiency of working memory in SZ is also linked with deficient DMN

and emotion regulation (Kaiser et al., 2015; Mulders et al., 2015). The research (Mulders et al., 2015) is consistent with the result of the current study that patients of the MDD observed abnormal FNCs between DMN and CC, CB, which are related to frontal, parietal, and postcentral gyrus. The DMN and cerebellum are related to self-processing and emotional (Drevets et al., 2008; Sheline et al., 2010), and the FN network has an important influence on cognitive ability and mood regulation (Drevets et al., 2008; Steele and Lawrie, 2004). The imbalance within externally-directed attention and control systems was caused by the aberrant connectivity of the MDD between anterior cingulate (AC) and inferior parietal lobule (IPL) (Kaiser et al., 2015) and the perception and emotional facial expressions processing was related to the aberrant FNCs in AUD and FN located in inferior temporal gyrus (ITG), superior frontal gyrus (SFG) and middle frontal gyrus (MFG) (Haxby et al., 2000). Abnormal reactivity caused by FNCs in precuneus and fusiform gyrus (FFG) might accept as a depression early biomarker as described in (Hahn et al., 2011). The SFG is related to executive memory emotions (Fuster, 2001; Jaworska et al., 2015) and the motor, premotor and prefrontal networks are related to executive memory. Executive function abnormalities in MDD patients were implied by the aberrant FNCs with SFG in FN and ITG (Fuster, 2001; Wang et al., 2008).

Several limitations should be mentioned for this study. One is we used only resting fMRI, however, classifying complex disorders may become more reliable and precise by applying multi-modal neuroimaging fusion (Cetin et al., 2016; Sui et al., 2020, 2011). Second, the generalizability of the classification framework to scanners outside our harmonized study can not be evaluated. Third, it is difficult to assess the medication effect due to the limited sample size and the information inadequacy of antipsychotic or emotion stabilizing medications. Finally, data preprocessing may be vital for machine learning, and each preprocessing step of brain imaging may have a potential effect on the final accuracy in deep learning. Thus, it is important to evaluate the influence of data preprocessing for deep learning in future, and the dynamics of functional connectivity may be another option to improve classification (Meng et al., 2017; Rashid et al., 2016; Zhi et al., 2018)

In summary, to the best of our knowledge, this is the first attempt to apply GAN on FNC-based classification, which integrates both adversarial training and feature matching. Application on discriminations of MDD-HC and SZ-HC verified the effectiveness of the proposed method. Compared with five traditional classification methods and a DNN

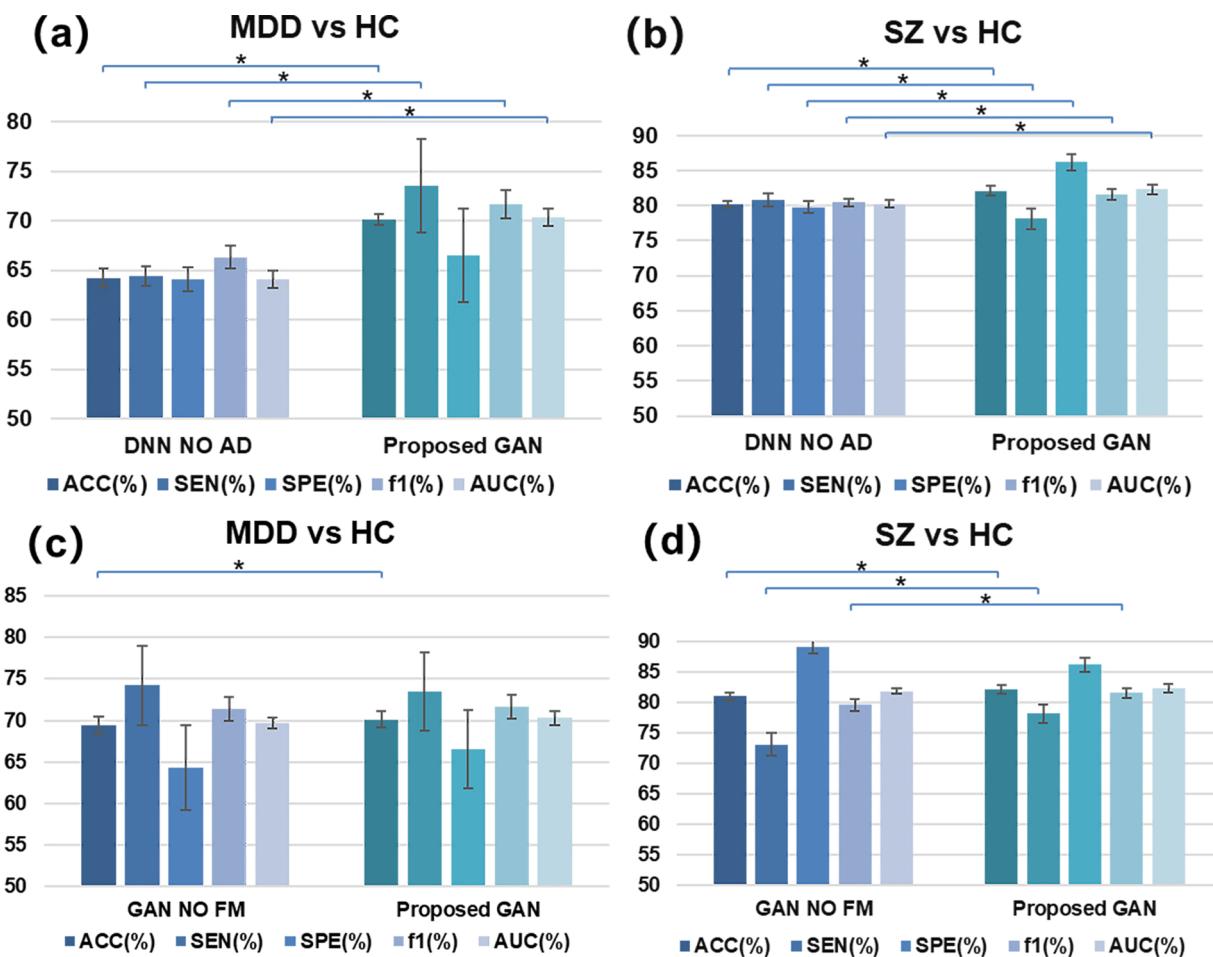


Fig. 5. Effect of adversarial training and feature matching.

Figure 5 (a) The comparison of AD vs NO-AD in MDD classification. (b) The comparison of AD vs NO-AD in SZ classification. (c) The comparison of FM vs NO-FM in MDD classification. (d) The comparison of FM vs NO-FM in SZ classification. FNC, Functional network connectivity; AD: adversarial training; FM: feature matching. In all cases, * means $p < 0.05$ in two-sample t-tests.

approach, the GAN model achieved at least 6% higher in MDD classification and 2% higher in SZ classification, suggesting its utility as a potentially powerful tool to aid in discriminative mental disorder detection.

Credit author statement

J.Sui, and J.Zhao designed the study. J.Zhao performed the analysis. J.Sui, J.Zhao, D.Zhi, J.Huang, and V. Calhoun wrote the paper. W.Yan, D.Zhi contributed to the data preprocessing. The acquisition of data was performed by T.Jiang, X.Ma, X.Yang, X.Li, and Q.Ke. All authors contribute to the discussion and the interpretation of the results.

Acknowledgments

This work was supported by the Natural Science Foundation of China (No.61773380), the Strategic Priority Research Program of the Chinese Academy of Sciences (No. XDB32040100), Beijing Municipal Science and Technology Commission (Z181100001518005), the National Institute of Health (R01MH117107, R01EB005846, and P20GM103472) and the National Science Foundation (1539067). The authors report no financial interests or potential conflicts of interest.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the

online version, at doi:<https://doi.org/10.1016/j.jneumeth.2020.108756>.

References

- Allen, E.A., Damaraju, E., Plis, S.M., Erhardt, E.B., Eichele, T., Calhoun, V.D., 2012. Tracking whole-brain connectivity dynamics in the resting state. *Cereb. Cortex* 24, 663–676.
- Arbabshirani, M.R., Plis, S., Sui, J., Calhoun, V.D., 2017. Single subject prediction of brain disorders in neuroimaging: promises and pitfalls. *Neuroimage* 145, 137–165.
- Braff, D.L., 1993. Information processing and attention dysfunctions in schizophrenia. *Schizophr. Bull.* 19, 233–259.
- Calhoun, V.D., Miller, R., Pearson, G., The Chronnectome, Adali T., 2014. Time-varying connectivity networks as the next frontier in fMRI data discovery. *Neuron* 84, 262–274.
- Calhoun, V.D., Sui, J., 2016. Multimodal fusion of brain imaging data: a key to finding the missing link(s) in complex mental illness. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* 1, 230–244.
- Camchong, J., MacDonald III, A.W., Bell, C., Mueller, B.A., Lim, K.O., 2009. Altered functional and anatomical connectivity in schizophrenia. *Schizophr. Bull.* 37, 640–650.
- Cetin, M.S., Houck, J.M., Rashid, B., Agacoglu, O., Stephen, J.M., Sui, J., Canive, J., Mayer, A., Aine, C., Bustillo, J.R., 2016. Multimodal classification of schizophrenia patients with MEG and fMRI data using static and dynamic connectivity measures. *Front. Neurosci.* 10, 466.
- Cui, L.-B., Liu, K., Li, C., Wang, L.-X., Guo, F., Tian, P., Wu, Y.-J., Guo, L., Liu, W.-M., Xi, Y.-B.B., Wang, L.-X.-N., Yin, H., 2016. Putamen-related regional and network functional deficits in first-episode schizophrenia with auditory verbal hallucinations. *Schizophr. Res.* 173, 13–22.
- Dai, Z., Yang, Z., Yang, F., Cohen, W.W., Salakhutdinov, R., 2017. Good semi-supervised learning that requires a bad GAN. In: Proceedings of the 31st International Conference on Neural Information Processing Systems. Curran Associates Inc.: Long Beach, California, USA. pp. 6513–6523.
- Drevets, W.C., Price, J.L., Furey, M.L., 2008. Brain structural and functional abnormalities

- in mood disorders: implications for neurocircuitry models of depression. *Brain Struct. Funct.* 213, 93–118.
- 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., Fryer, S.L., Lin, D., Sui, J., Yu, Q., Chen, J., Stuart, B., Loewy, R.L., Calhoun, V.D., Mathalon, D.H., 2018. Identifying functional network changing patterns in individuals at clinical high-risk for psychosis and patients with early illness schizophrenia: a group ICA study. *Neuroimage Clin.* 17, 335–346.
- 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.
- Feng, J., Yu, H., Wang, L., Cao, X., Zhang, X., Jiao, L., 2019. Classification of hyperspectral images based on multiclass spatial-spectral generative adversarial networks. *IEEE Trans. Geosci. Remote. Sens.*
- Fuster, J.M., 2001. The prefrontal cortex—an update: time is of the essence. *Neuron* 30, 319–333.
- Gao, S., Calhoun, V.D., Sui, J., 2018. Machine learning in major depression: from classification to treatment outcome prediction. *CNS Neurosci. Ther.* 24, 1037–1052.
- Hahn, T., Marquand, A.F., Ehli, A.-C., Dresler, T., Kittel-Schneider, S., Jarczok, T.A., Lesch, K.-P., Jakob, P.M., Mourao-Miranda, J., Brammer, M.J., 2011. Integrating neurobiological markers of depression. *Arch. Gen. Psychiatry* 68, 361–368.
- Haxby, J.V., Hoffman, E.A., Gobbini, M.I., 2000. The distributed human neural system for face perception. *Trends Cogn. Sci. (Regul. Ed.)* 4, 223–233.
- Hooker, C.I., Bruce, L., Fisher, M., Verosky, S.C., Miyakawa, A., Vinogradov, S., 2012. Neural activity during emotion recognition after combined cognitive plus social cognitive training in schizophrenia. *Schizophr. Res.* 139, 53–59.
- Huang, C., Li, Y., Change Loy, C., Tang, X., 2016. Learning deep representation for imbalanced classification. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* 5375–5384.
- Jaworska, N., Yang, X.-R., Knott, V., MacQueen, G., 2015. A review of fMRI studies during visual emotive processing in major depressive disorder. *World J. Biol. Psychiatry* 16, 448–471.
- Jie, N.-F., Osuch, E.A., Zhu, M.-H., Wammes, M., Ma, X.-Y., Jiang, T.-Z., Sui, J., Calhoun, V.D., Discriminating Bipolar Disorder from MaX-Y, 2018. Depression using Whole-Brain Functional Connectivity: a Feature Selection Analysis with SVM-FoBa Algorithm. *J. Signal Process. Syst.* 90, 259–271.
- Kaiser, R.H., Andrews-Hanna, J.R., Wager, T.D., Pizzagalli, D.A., 2015. Large-scale network dysfunction in major depressive disorder: a meta-analysis of resting-state functional connectivity. *JAMA Psychiatry* 72, 603–611.
- Kaufmann, T., Skåtun, K.C., Alnæs, D., Doan, N.T., Duff, E.P., Tønnesen, S., Roussos, E., Ueland, T., Aminoff, S.R., Lagerberg, T.V., Agartz, I., Melle, I.S., Smith, S.M., Andreassen, O.A., Westlye, L.T., 2015. Disintegration of Sensorimotor Brain Networks in Schizophrenia. *Schizophrenia Bulletin* 41, 1326–1335.
- Kessler, R.C., Petukhova, M., Sampson, N.A., Zaslavsky, A.M., Wittchen, H.U., 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.
- Kim, J., Calhoun, V.D., Shim, E., Lee, J.-H., 2016. 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.
- LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. *nature* 521, 436.
- Meng, X., Jiang, R., Lin, D., Bustillo, J., Jones, T., Chen, J., Yu, Q., Du, Y., Zhang, Y., Jiang, T., 2017. Predicting individualized clinical measures by a generalized prediction framework and multimodal fusion of MRI data. *Neuroimage* 145, 218–229.
- Mondino, M., Jardri, R., Suaud-Chagny, M.-F., Saoud, M., Poulet, E., Brunelin, J., 2015. Effects of fronto-temporal transcranial direct current stimulation on auditory verbal hallucinations and resting-state functional connectivity of the left temporo-parietal junction in patients with schizophrenia. *Schizophr. Bull.* 42, 318–326.
- Mulders, P.C., van Eijndhoven, P.F., Schene, A.H., Beckmann, C.F., Tendolkar, I., 2015. Resting-state functional connectivity in major depressive disorder: a review. *Neurosci. Biobehav. Rev.* 56, 330–344.
- Pu, W., Luo, Q., Palaniyappan, L., Xue, Z., Yao, S., Feng, J., Liu, Z., 2016. Failed cooperative, but not competitive, interaction between large-scale brain networks impairs working memory in schizophrenia. *Psychol. Med.* 46, 1211–1224.
- Rashid, B., Arbabshirani, M.R., Damaraju, E., Cetin, M.S., Miller, R., Pearson, G.D., Calhoun, V.D., 2016. Classification of schizophrenia and bipolar patients using static and dynamic resting-state fMRI brain connectivity. *NeuroImage* 134, 645–657.
- Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., Chen, X., 2016. Improved techniques for training GANs. *arXiv e-prints*.
- Sato, J.R., Moll, J., Green, S., Deakin, J.F., Thomaz, C.E., Zahn, R., 2015. Machine learning algorithm accurately detects fMRI signature of vulnerability to major depression. *Psychiatry Res. Neuroimaging* 233, 289–291.
- Sheffield, J.M., Barch, D.M., 2016. Cognition and resting-state functional connectivity in schizophrenia. *Neurosci. Biobehav. Rev.* 61, 108–120.
- Sheline, Y.I., Price, J.L., Yan, Z., Mintun, M.A., 2010. Resting-state functional MRI in depression unmasks increased connectivity between networks via the dorsal nexus. *Proc. Natl. Acad. Sci.* 107, 11020–11025.
- Springenberg, J.T., 2015. Unsupervised and semi-supervised learning with categorical generative adversarial networks. *arXiv e-prints*.
- Steele, J.D., Lawrie, S.M., 2004. Segregation of cognitive and emotional function in the prefrontal cortex: a stereotactic meta-analysis. *NeuroImage* 21, 868–875.
- Sui, J., He, H., Yu, Q., chen, J., Rogers, J., Pearson, G., Mayer, A., Bustillo, J., Canive, J., Calhoun, V., 2013. Combination of resting state fMRI, DTI, and sMRI data to discriminate schizophrenia by N-way MCCA + jICA. *Front. Hum. Neurosci.* 7.
- Sui, J., Jiang, R., Bustillo, J., Calhoun, V., 2020. Neuroimaging-based individualized prediction of cognition and behavior for mental disorders and health: methods and promises. *Biol. Psychiatry*.
- Sui, J., Pearson, G., Caprihan, A., Adali, T., Kiehl, K.A., Liu, J., Yamamoto, J., Calhoun, V.D., 2011. Discriminating schizophrenia and bipolar disorder by fusing fMRI and DTI in a multimodal CCA + joint ICA model. *NeuroImage* 57, 839–855.
- Sui, J., Qi, S., van Erp, T.G., Bustillo, J., Jiang, R., Lin, D., Turner, J.A., Damaraju, E., Mayer, A.R., Cui, Y., 2018. Multimodal neuromarkers in schizophrenia via cognition-guided MRI fusion. *Nat. Commun.* 9, 3028.
- Unschuld, P.G., Buchholz, A.S., Varvaris, M., van Zijl PCM, Ross C.A., Pekar, J.J., Hock, C., Sweeney, J.A., Tamminga, C.A., Keshavan, M.S., Pearson, G.D., Thaker, G.K., Schretlen, D.J., 2013. Prefrontal brain network connectivity indicates degree of both schizophrenia risk and cognitive dysfunction. *Schizophr. Bull.* 40, 653–664.
- Vandenbende, S., De Brabandere, B., Neven, D., Van Gool, L., 2019. A three-player GAN: generating hard samples to improve classification networks. *arXiv e-prints*.
- Vercammen, A., Knegtering, H., den Boer, J.A., Liemburg, E.J., Aleman, A., 2010. Auditory hallucinations in schizophrenia are associated with reduced functional connectivity of the temporo-parietal area. *Biol. Psychiatry* 67, 912–918.
- Wang, L., LaBar, K.S., Smoski, M., Rosenthal, M.Z., Dolcos, F., Lynch, T.R., Krishnan, R.R., McCarthy, G., 2008. Prefrontal mechanisms for executive control over emotional distraction are altered in major depression. *Psychiatry Res. Neuroimaging* 163, 143–155.
- Xu, K., Liu, Y., Zhan, Y., Ren, J., Jiang, T., 2018. BRANT: a versatile and extensible resting-state fMRI toolkit. *Front. Neuroinform.* 12, 52.
- Yan, W., Calhoun, V., Song, M., Cui, Y., Yan, H., Liu, S., Fan, L., Zuo, N., Yang, Z., Xu, K., 2019a. Discriminating schizophrenia using recurrent neural network applied on time courses of multi-site FMRI data. *EBioMedicine* 47, 543–552.
- 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., 2019b. Discriminating schizophrenia using recurrent neural network applied on time courses of multi-site FMRI data. *EBioMedicine*.
- Yang, Z., Wu, J., Xu, L., Deng, Z., Tang, Y., Gao, J., Hu, Y., Zhang, Y., Qin, S., Li, C., Wang, J., 2019. Individualized psychiatric imaging based on inter-subject neural synchronization in movie watching. *NeuroImage*, 116227.
- Yi, X., Walia, E., Babyn, P., 2019. Generative adversarial network in medical imaging: a review. *Med. Image Anal.* 58, 101552.
- Yuan, B.-K., Zang, Y.-F., Liu, D.-Q., 2016. Influences of head motion regression on high-frequency oscillation amplitudes of resting-state fMRI signals. *Front. Hum. Neurosci.* 10, 243.
- Zeng, L.-L., Wang, H., Hu, P., Yang, B., Pu, W., Shen, H., Chen, X., Liu, Z., Yin, H., Tan, Q., Wang, K., Hu, D., 2018. Multi-site diagnostic classification of schizophrenia using discriminant deep learning with functional connectivity MRI. *Evidence-Based Medicine* 30, 74–85.
- Zhang, J., Zhou, L., Wang, L., Li, W., 2015. Functional brain network classification with compact representation of SICE matrices. *IEEE Trans. Biomed. Eng.* 62, 1623–1634.
- Zhi, D., Calhoun, V.D., Lv, L., Ma, X., Ke, Q., Fu, Z., Du, Y., Yang, Y., Yang, X., Pan, M., Qi, S., Jiang, R., Yu, Q., 2018. Sui J. Aberrant dynamic functional network connectivity and graph properties in major depressive disorder. *Front. Psychiatry* 9.