Brain Research Bulletin 206 (2024) 110846

Contents lists available at ScienceDirect

Brain Research Bulletin

journal homepage: www.elsevier.com/locate/brainresbull

Research report

Discriminative analysis of schizophrenia patients using an integrated model combining 3D CNN with 2D CNN: A multimodal MR image and connectomics analysis

Haiman Guo [a,1], Shuyi Jian [a,1], Yubin Zhou [a], Xiaoyi Chen [a], Jinbiao Chen [a], Jing Zhou [b,d,e], Yuanyuan Huang [c,d], Guolin Ma [h], Xiaobo Li [i], Yuping Ning [c,d], Fengchun Wu [c,d,*], Kai Wu [a,f,g,j,**]

[a] School of Biomedical Sciences and Engineering, South China University of Technology, Guangzhou International Campus, Guangzhou 511442, China
[b] School of Material Sciences and Engineering, South China University of Technology, Guangzhou 510610, China
[c] The Affiliated Brain Hospital of Guangzhou Medical University, Guangzhou 510370, China
[d] Guangdong Engineering Technology Research Center for Translational Medicine of Mental Disorders, Guangzhou 510370, China
[e] Guangdong Engineering Technology Research Center for Diagnosis and Rehabilitation of Dementia, Guangzhou 510500, China
[f] National Engineering Research Center for Tissue Restoration and Reconstruction, South China University of Technology, Guangzhou 510006, China
[g] Guangdong Province Key Laboratory of Biomedical Engineering, South China University of Technology, Guangzhou 510006, China
[h] Department of Radiology, China-Japan Friendship Hospital, Beijing 100029, China
[i] Department of Biomedical Engineering, New Jersey Institute of Technology, Newark, NJ, USA
[j] Department of Nuclear Medicine and Radiology, Institute of Development, Aging and Cancer, Tohoku University, Sendai 980-8575, Japan

ARTICLE INFO

Keywords:
Schizophrenia
Brain connectivity
Integrated model
Convolutional neural network
Classification

ABSTRACT

Objective: Few studies have applied deep learning to the discriminative analysis of schizophrenia (SZ) patients using the fusional features of multimodal MRI data. Here, we proposed an integrated model combining a 3D convolutional neural network (CNN) with a 2D CNN to classify SZ patients.
Method: Structural MRI (sMRI) and resting-state functional MRI (rs-fMRI) data were acquired for 140 SZ patients and 205 normal controls. We computed structural connectivity (SC) from the sMRI data as well as functional connectivity (FC), amplitude of low-frequency fluctuation (ALFF), and regional homogeneity (ReHo) from the rs-fMRI data. The 3D images of T1, ReHo, and ALFF were used as the inputs for the 3D CNN model, while the SC and FC matrices were used as the inputs for the 2D CNN model. Moreover, we added squeeze and excitation blocks (SE-blocks) to each layer of the integrated model and used a support vector machine (SVM) to replace the softmax classifier.
Results: The integrated model proposed in this study, using the fusional features of the T1 images, and the matrices of FC, showed the best performance. The use of the SE-blocks and SVM classifiers significantly improved the performance of the integrated model, in which the accuracy, sensitivity, specificity, area under the curve, and F1-score were 89.86%, 86.21%, 92.50%, 89.35%, and 87.72%, respectively.
Conclusions: Our findings indicated that an integrated model combining 3D CNN with 2D CNN is a promising method to improve the classification performance of SZ patients and has potential for the clinical diagnosis of psychiatric diseases.

1. Introduction

Schizophrenia (SZ) is a major psychiatric disease that results in chronic impairments in cognition, emotion, and behavior (Rossler et al., 2005; Tost and Meyer-Lindenberg, 2012). Recently, multimodal magnetic resonance imaging (MRI) has been widely used to study abnormalities in the brain structure and function of SZ patients. A large number of structural MRI (sMRI) studies of SZ patients have indicated a

* Corresponding author at: The Affiliated Brain Hospital of Guangzhou Medical University, Guangzhou 510370, China.
** Corresponding author at: School of Biomedical Sciences and Engineering, South China University of Technology, Guangzhou International Campus, Guangzhou 511442, China.
E-mail addresses: 13580380071@163.com (F. Wu), kaiwu@scut.edu.cn (K. Wu).
[1] Co-first authors.

https://doi.org/10.1016/j.brainresbull.2023.110846
Received 30 March 2023; Received in revised form 20 November 2023; Accepted 12 December 2023
Available online 15 December 2023

reduction in gray matter volume in the hippocampal and bilateral thalamus (Adriano et al., 2012; Fornito et al., 2009), as well as lower fractional anisotropy (FA) in the whole brain and individually in the frontal, parietal, occipital, and temporal lobes (White et al., 2009). Previous resting-state functional MRI (rs-fMRI) studies have demonstrated that SZ patients showed an increased amplitude of low-frequency fluctuation (ALFF) in the right putamen, decreased ALFF in the right postcentral gyrus (Cui et al., 2016), increased regional homogeneity (ReHo) in the left medial frontal gyrus (MFG) and left inferior parietal lobule, and decreased ReHo in the left MFG and right precentral gyrus (Zhao et al., 2019). Moreover, structural connectivity (SC) and functional connectivity (FC) can be measured to generate human brain connectomics (Farras-Permanyer et al., 2015), which is widely used to quantitatively analyze global and regional network topology in the human brain (Jiang et al., 2020; van den Heuvel and Fornito, 2014). Previous studies have demonstrated that SZ can be conceptualized as a dysconnection syndrome (Kong et al., 2021; Nelson et al., 2017; Cui et al., 2019), in which there is decreased SC between the frontal regions and other cortices, as well as decreased global efficiency (Su et al., 2015) and disrupted rich-club organization of the functional brain networks (Collin et al., 2017; Karlsgodt et al., 2009; Camchong et al., 2009).

In recent decades, traditional machine learning (ML) algorithms have been applied to the classification of SZ patients, such as linear discriminant analysis (LDA) (Caan et al., 2006), support vector machine (SVM) (Wang et al., 2018), and random forest (RF) (Venkataraman et al., 2012). Previous discriminative studies demonstrated that an SVM classifier based on sMRI data achieved an accuracy of 83.5% (Chin et al., 2018), and an SVM classifier in conjunction with intrinsic discriminant analysis based on rs-fMRI data achieved an accuracy of 80.9% (Yu et al., 2013). High-dimensional neuroimaging data can lead to biased performance estimates of ML and have become a challenging problem (Qureshi et al., 2019), whereas deep learning can automatically learn the critical features from the neuroimaging data and it achieved better performance in recent discriminative studies of psychiatric diseases (Campese et al., 2020; Wang et al., 2019; Lei et al., 2020). Convolutional neural networks (CNNs), as a popular deep learning method, have the advantages of automatic feature learning, noting subtle significant details during convolution operations, and eliminating the subjectivity of selecting relevant spatial features. Campese et al. applied a 3D CNN model to the classification of SZ patients with sMRI data and obtained an average accuracy of 86.3% (Campese et al., 2020). Wang et al. obtained an accuracy of 82.7% with a 3D CNN model using rs-fMRI data for the classification of SZ patients (Wang et al., 2019). Moreover, Lei et al. applied a 2D CNN model to the classification of SZ patients with rs-fMRI data and obtained an average accuracy of 81.0% (Lei et al., 2020).

Previous SZ studies of 3D CNN or 2D CNN have mainly used the features of unimodal data, such as sMRI and rs-fMRI (Kim et al., 2016; Islam and Zhang, 2017; Xiao et al., 2018). Recent studies have indicated that multimodal data have the potential for the improvement of classification performance (Hu et al., 2022; Shi et al., 2018; Cabral et al., 2016). Hu et al. indicated that 3D CNN models using multimodal fusion features extracted from sMRI and diffusion tensor imaging have advantages over unimodal data (Hu et al., 2022). Shi et al. proposed a stacked deep polynomial network to distinguish patients with Alzheimer's disease from normal controls (NCs) and demonstrated that the classification performance using multimodal data (sMRI and positron emission tomography) is significantly better than that of the model using unimodal data (Shi et al., 2018).

In this study, we proposed an integrated model combining 3D CNN with 2D CNN using the features of 3D images and 2D connectivity matrices from sMRI and rs-fMRI data to classify SZ patients from NCs. Specifically, the T1, ReHo, and ALFF images were used as the inputs for the 3D CNN model, while the SC and FC matrices were used as the inputs for the 2D CNN model. In addition, the attention module and the SVM classifier were added to the integrated model.

2. Methods and materials

2.1. Subjects

A total of 140 SZ patients and 205 NCs were included in this study. The diagnosis of SZ was made by clinical psychiatrists based on trained experience and structured clinical interviews that met the criteria of the Diagnostic and Statistical Manual of Mental Disorders: Fourth Edition, Text Revision (DSM-IV-TR). All subjects were recruited from the local community and the Affiliated Brain Hospital of Guangzhou Medical University. Each subject or their legal guardian was fully informed of the experimental details and signed an informed consent form prior to enrollment. The study was conducted according to the Declaration of Helsinki and was approved by the Ethics Committees of the Affiliated Brain Hospital of Guangzhou Medical University. The demographic and clinical characteristics of all participants are presented in Table 1. Note that the values are expressed as the mean ± standard deviation.

2.2. Data acquisition and image processing

MRI data of all subjects were collected using a Philips 3 T MR device system (Philips, Achieva, Netherlands) in the Affiliated Brain Hospital of Guangzhou Medical University. The participants were instructed to keep their eyes closed, relax but not fall asleep and to move as little as possible. For each subject, the sMRI data were collected using a sagittal three-dimensional gradient-echo TI-weighted sequence (repetition time (TR) = 8.2 ms, echo time (TE) = 3.7 ms, field of view (FOV) = 256 × 256 mm^2, flip angle = 7°, spatial resolution = 1 × 1×1 mm^3, matrix = 256 × 256 × 188). The rs-MRI data were generated using an echo-planar imaging (EPI) sequence (TR = 2000 ms, TE = 30 ms, FOV = 220 ×220 mm^2, flip angle = 90°, spatial resolution = 3.44 × 3.44 ×4 mm^3, matrix = 64 × 64 × 36).

The image preprocessing steps of multimodal MRI data were the same as those in our previous studies (Kong et al., 2021; Wu et al., 2018; Zang et al., 2021; Huang et al., 2022; Chen et al., 2023; Wang et al., 2022). We performed a comprehensive preprocessing of all MRI data using SPM8 (http://www.fil.ion.ucl.ac.uk/spm; Institute of Neurology, University College London) in MATLAB and the Data Processing & Analysis for Brain Imaging (DPABI) toolbox (Yan et al., 2016). For T1-weighted images, we employed the new segmentation algorithm from SPM8, which classifies the tissue into three distinct maps: gray matter, white matter, and cerebrospinal fluid. Following this, we normalized the gray matter segmentations to Montreal Neurological Institute (MNI) space using a high-dimensional 'DARTEL' approach. The Gray Matter Volume (GMV) map was subsequently obtained after reslicing with an isotropic voxel size of 2 mm^3 using nearest neighbor interpolation. We intentionally avoided spatial smoothing to prevent inducing artifactual signal overlap among spatially adjacent regions.

For fMRI preprocessing, we began by discarding the first 10 volumes of the time series to eliminate unstable initial signals. We ensured that

Table 1
Demographic information of the participants.

	SZ group	NC group	t value (χ^2)*	P value
Sample Size	140	205	-	-
Age (years)	34.22 ± 8.23	32.51 ± 8.39	-1.870	0.062
Education (years)	10.69 ± 3.32	12.84 ± 2.83	6.260	< 0.001
Gender (M/F)	95/45	110/95	6.378	0.012
PANSS Total	85.55 ± 18.55	-	-	-
PANSS Positive	23.14 ± 5.25	-	-	-
PANSS Negative	22.53 ± 7.48	-	-	-
PANSS General	39.88 ± 9.59	-	-	-

* Note: The comparison of age and years of education was analyzed by a two-sample t-test. The gender comparison was performed using a χ^2 test. PANSS: Positive and Negative Syndrome Scale.

all participants did not exhibit excessive head motion (2 mm or 2° during realignment) and mean motion (2 mm during framewise assessment). After slice-timing adjustment, the remaining images were realigned to the first volume to correct for head motion. We then regressed out the nuisance signals (Rigid-body 6 motion parameters, the white matter signal, and the cerebrospinal fluid signal) from the data. The warped data were then spatially normalized and resampled to a spatial isotropic resolution of 3 mm. The resampled data were spatially smoothed using a Gaussian kernel with a full-width at half-maximum value of 4 mm and bandpass filtered to isolate the low-frequency signal in each voxel (0.01–0.08 Hz). ALFF maps as well as the ReHo maps was computed from the rs-fMRI data using SPM8 and DPABI after the preprocessing process. This comprehensive and meticulous preprocessing approach ensured the quality and reliability of our input data for the subsequent analysis.

2.3. Structural and functional connectivity analysis

The construction of structural and functional brain networks using multimodal MRI data was the same as the pipeline used in our previous studies (Kong et al., 2021; Chen et al., 2023). The individual SC matrix was constructed based on the Kullback—Leibler divergence-based similarity (KLS) method using GM images, and the individual FC matrix was computed based on the Pearson's correlation coefficient (PCC) method using rs-fMRI data.

2.4. Discriminative analysis

In search of discriminative biomarkers and use them to improve the model performance of automatically extracting features, we construct an integrated model combining 3D CNN with 2D CNN. The architecture of our integrated model, including the 3D CNN and 2D CNN branches, was determined based on the characteristics of the input data and the task at hand. The initial inputs of the two branches are the 2D connectivity matrix and 3D image. All the experiments were conducted on a 4 x NVIDIA Geforce GTX 3080 Graphics Processing Unit (GPU) setup. The deep learning model was implemented using the TensorFlow API version 2.7.0, compatible with the GPU, and scripted in Python. An overview of our integrated model architecture is shown in Fig. 1.

For the 3D CNN branch of our model, we initially experimented with architectures having deeper convolutional layers. However, due to the high dimensions of three-dimensional neural images, these architectures often led to overfitting and graphics memory issues, with a very low classification accuracy. After reducing the number of convolutional layers, we found that an architecture with three 3D convolutional layers provided the best classification performance. 3D convolutions can be performed to compute features from multiple dimensions, which can capture the brain region information more effectively. Formally, the value of the unit at position (x, y, z) in the jth feature map in the ith layer is given by (Ji et al., 2012):

$$v_{ij}^{xyz} = \tanh\left(b_{ij} + \sum_{m}\sum_{p=0}^{P_i-1}\sum_{q=0}^{Q_i-1}\sum_{r=0}^{R_i-1} w_{ijm}^{pqr} v_{(i-1)m}^{(x+p)(y+q)(z+r)}\right) \quad (1)$$

where R_i is the size of the 3D kernel along the temporal dimension and w_{ijm}^{pqr} is the (p, q, r)th value of the kernel connected to the mth feature map in the previous layer.

For the 2D CNN branch of our model, given the smaller dimensions and relatively less information in the 2D connectivity matrix, we found that the results were quite similar across different architectures. Therefore, to optimize resource consumption and training efficiency, we chose an architecture with two convolutional layers for the 2D CNN branch. Two-dimensional convolutions are applied to the 2D connectivity matrix to extract features from the local neighborhood. Formally, the value of the unit at position (x, y) in the jth feature map in the ith layer, denoted as v_{ij}^{xy}, is given by (Zhang et al., 2020; LeCun et al., 1998):

$$v_{ij}^{xy} = \tanh\left(b_{ij} + \sum_{m}\sum_{p=0}^{P_i-1}\sum_{q=0}^{Q_i-1} w_{ijm}^{pq} v_{(i-1)m}^{(x+p)(y+q)}\right) \quad (2)$$

where $\tanh(\bullet)$ is the hyperbolic tangent function, b_{ij} is the bias for the feature map, m indices over the set of feature maps in the $(i-1)$th layer connected to the current feature map, w_{ijk}^{pq} is the value at position (p, q) of the kernel connected to the kth feature map, and P_i and Q_i are the height and width of the kernel, respectively. The parameters b_{ij} and the kernel weight w_{ijk}^{pq} are trained with the supervised approach in the experiment.

To extract the most valuable features after each convolutional layer, we added a channel attention layer, also called SE-Block, to screen further the features extracted from the convolutional layer. Next, for the 3D CNN branch, we added global average pooling after the last convolutional layer. After global average pooling, we connected two fully connected layers. For the 2D CNN branch, we used a flattening layer and added two fully connected layers after the flattening layer. To unify the feature layer dimension, we set the number of nodes in each branch's last fully connected layer to 64. We then used the concatenated layer to combine the features extracted from the two fully connected layers. The

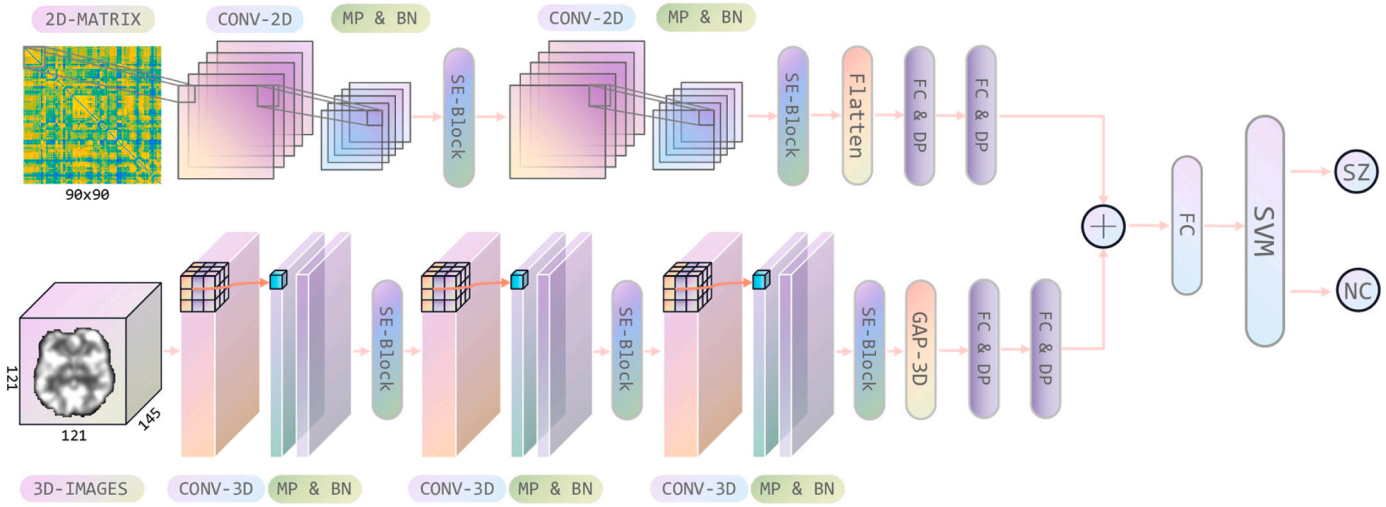

Fig. 1. Model Architecture Overview. MP: Max Pooling; BN: Batch Normalization; FC: Fully Connected; DP: Dropout; GAP-3D: Global Average Pooling 3D.

hyperparameters of the model, such as the number of nodes in the fully connected layer and the size of the 3D kernel, were set based on empirical studies and fine-tuned during the training process to achieve the best performance. The detailed structure of our integrated model is shown in Fig. 2.

2.5. Optimization module

To improve the efficiency of feature extraction and reduce the redundant noise generated in the feature fusion step, we added SE-blocks to our model based on SENet (Hu et al., 2018), and the structure is as follows.

The first step is the squeeze operation, where the dimension of the original feature map is $H \times W \times L \times C$, where H is the height, W is the width, L is the length, C is the number of channels, and i, j, k corresponds to each value in the input three-dimensional feature matrix. The formula is as follows:

$$\mathbb{z}_c = F_{sq}(u_c) = \frac{1}{H \times W \times L} \sum_{i=1}^{H} \sum_{j=1}^{W} \sum_{k=1}^{L} u_c(i,j,k) \tag{3}$$

The second step is the excitation operation, which is realized by two fully connected layers; the first fully connected layer compresses the C channels into C/r channels to reduce the computation, and the second fully connected layer is restored to the C channels, where σ represents the sigmoid function, δ represents the ReLU function, r refers to the ratio of compression, and $W_1 \in R^{\frac{C}{r} \times C}, W_2 \in R^{C \times \frac{C}{r}}$. The formula is as follows:

$$\mathbb{s} = F_{ex}(\mathbb{z}, W) = \sigma(\mathrm{g}(\mathbb{z}, W)) = \sigma(W_2 \delta(W_1 \mathbb{z})) \tag{4}$$

Finally, there is a Reweight operation. We compute the output of the softmax operation. These outputs will serve as weights to represent the importance of each feature channel after feature selection. Then, we reweight the feature channels by multiplying them element-wise with the corresponding weight. This operation enhances the more relevant features and suppresses the less relevant ones. After obtaining s, the final output of the SE block can be obtained by the following equation:

$$\begin{aligned} \widetilde{x}_c &= F_{scale}(u_c, s_c) = s_c u_c \\ \widetilde{X} &= [\widetilde{x}_1, \widetilde{x}_2, \ldots, \widetilde{x}_C] \end{aligned} \tag{5}$$

where $\widetilde{X} = [\widetilde{x}_1, \widetilde{x}_2, \ldots, \widetilde{x}_c]$ and $F_{scale}(u_c, s_c)$ refer to channelwise multiplication between the feature map $u_c \in R^{H \times W \times L}$ and the scalar s_c.

Moreover, based on literature suggesting that SVM classifiers more effectively categorize higher-dimensional vectors (Tang, 2013; Agarap, 2018; Alalshekmubarak and Smith, 2013), we aimed to boost our model's generalization capabilities and improve binary classification

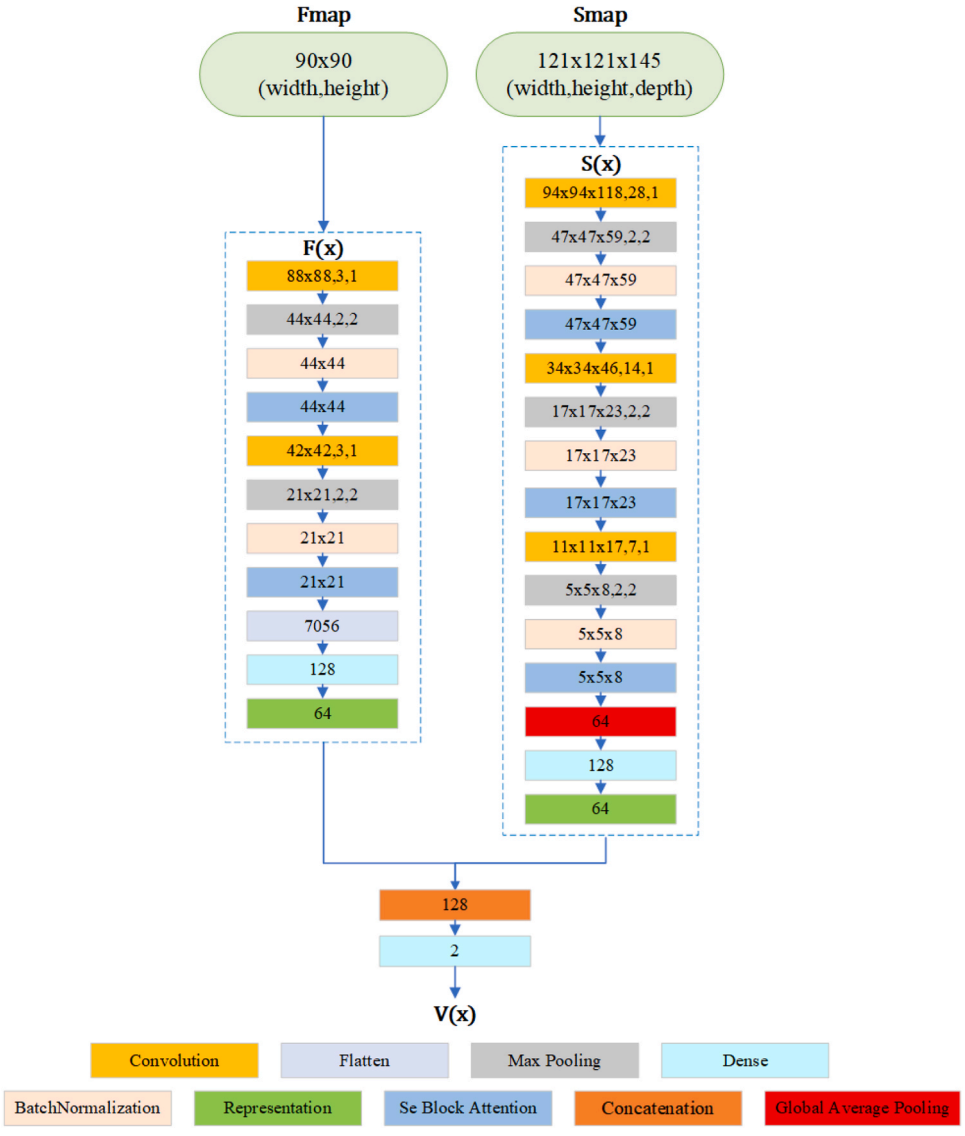

Fig. 2. The multimodal architecture framework. The dimensions inside each box represent the output dimensions, kernel size, and stride length.

performance. To achieve this, we fed the features extracted from the concatenated layer into the SVM for further training.

3. Results

3.1. Experimental Setup

The binary cross entropy loss function was adopted in this study for the binary classification task. In addition, to prevent the learning rate from being too large, causing the loss to fluctuate up and down when converging to the global optimum, the Adam optimizer and the approach of exponential decay of the learning rate were employed for the training model, which means that the learning rate was reduced to $e^{-\beta}$ times that of the original after iterating T steps. The initial learning rate α_0 was set to 0.001, the decay steps to 100,000, and the decay rate to 0.96. For the attention modules, considering about that as the network goes deeper, the features become more abstract and high-level, and to avoid the information bottleneck and the vanishing gradient, we set more channels in deeper layers to represent and preserve the features. The detailed configuration is shown in Table 2. In the fully connected layer, a dropout rate of 0.3 was applied. The batch size was 4, and all models were trained for 200 epochs. In terms of the subsequent SVM classifier, the penalty parameter, kernel function and maximum number of iterations were set to 3.0, radial basis function and 20, respectively.

3.2. Comparison of single- and multibranch architectures

For the three fMRI features and sMRI data mentioned in the previous section, we split our integrated model into two parts, a 3D CNN model and a 2D CNN model, to classify the SZ patients with features from different modalities. First, for the 3D ReHo and ALFF data of size $61 \times 73 \times 61$ obtained by rs-fMRI data, we fed them into a 3D CNN model. Second, we input the 3D T1 and 3D GM data of size $121 \times 145 \times 121$ obtained from the sMRI data into another 3D CNN model. Finally, we sent the 2D FC matrix obtained from the rs-fMRI data and the 2D SC matrix obtained from the sMRI data into a 2D CNN model. Fig. 3 shows the testing results for the different features from the different modalities.

Next, we combined the 3D CNN model with the 2D CNN to form an integrated model and tested its performance. The results were presented in Fig. 4:

After that, we added SE-Blocks to each convolutional layer to test the performance of the above features and combinations that perform well. The results were presented in Fig. 5:

Finally, we tested the above best-performing model with an SVM classifier and evaluated its different performance metrics. The overall results are shown in Table 3.

Additionally, the training loss and accuracy curves of the proposed model are shown in Fig. 6. Although the accuracy of the training set was close to 100%, the accuracy of the test set was only hovering at approximately 85%, indicating that our model had a certain overfitting problem. This may be because the sample size of our dataset was relatively small, only 345, and only 276 samples were used for training. In addition, the dimension of each sample was very high, especially the dimension of 3D sMRI data, which was up to 121×145×121, so the model was prone to overfitting.

Table 2
Parameters of the SE Block.

SE_Blocks	Ratios	Channels
2d_1	4	6
2d_2	8	16
3d_1	4	16
3d_2	8	32
3d_3	16	64

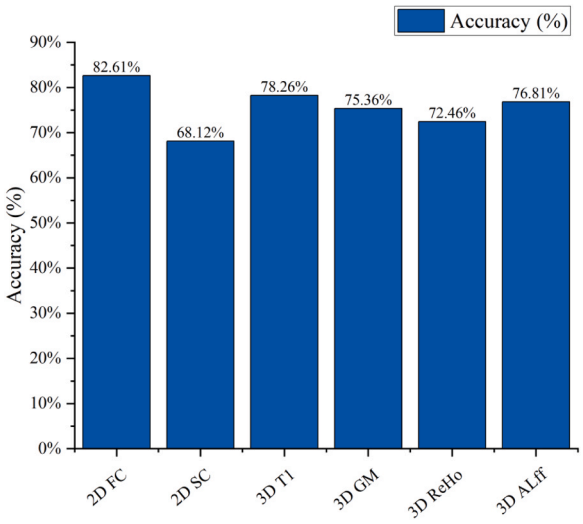

Fig. 3. Comparison of the performance of different unimodal features.

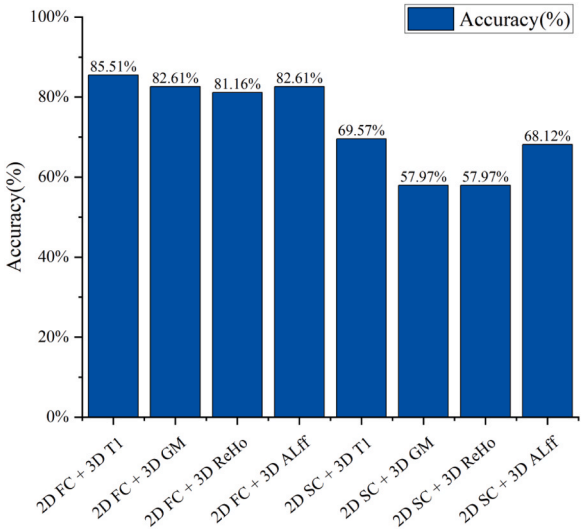

Fig. 4. Accuracy of different combinations of features.

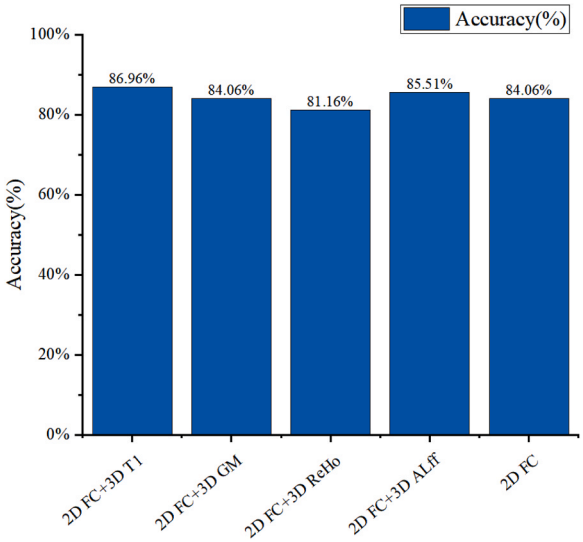

Fig. 5. Accuracy after adding SE-Blocks.

Table 3
Result of our multibranch model.

Metrics	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC (%)	F1-score (%)
	89.86	86.21	92.50	89.35	87.72

*Note: Accuracy: the proportion of correctly predicted samples to all samples; Sensitivity: the probability of a positive test given that the patient has the disease; Specificity: the probability of a negative test given that the patient is well; AUC: Area Under Curve, this curve always means ROC curve; F1-score: the harmonic means of specificity and recall.

3.3. Comparisons with other methods

To demonstrate the superiority of our integrated model, we compared it with some traditional 2D CNN-based state of the art (SOTA) models, including Inception_V3, Inception_ResNet_V2, MobileNet_V2, ResNet152_V2, ResNet50 and Xception. Since the above 2D CNN models only accepted 2D images as inputs and their inputs must be RGB images with three color channels, we extended the channels of the 2D FC matrix by a pseudo-RGB method and then input it into the 2D CNN model for training. Fig. 7 shows the performance of the different evaluation metrics, including the accuracy, sensitivity, specificity, AUC and f1-score of our proposed SZ classifier using multimodal data and other SOTA models. Compared with the best performing 2D CNN model, InceptionResNetV2, our model had 8.7% higher classification accuracy, 14.2% higher classification precision, and 6.9% higher recall, highlighting the superiority of our multimodal model. Additionally, Fig. 8 shows the ROC curves of the various models.

3.4. Confusion matrix

The confusion matrix based on our integrated model is also presented in Fig. 9. Our proposed multimodal classifier made 69 subjective predictions, and among these subjects, the classifier predicted that there were 28 SZ patients and 41 NC subjects. However, in the original dataset, there were 29 SZ patients and 40 NC subjects. From the confusion matrix, we can see that the overall misclassification rate is approximately 10.14%, and the true positive and true negative rates are 96.55% and 5.0%. The confusion matrix indicates that our model demonstrates strong acceptability.

4. Discussion

Inspired by the recent success of multibranch models, we propose an integrated model combining a 3D CNN with a 2D CNN using the fusional features of multimodal MRI data to classify SZ patients (Masoudi and Danishvar, 2022). Our main findings included the following: (1) the integrated model combined 3D CNN with 2D CNN using the fusional features of the T1 images, and the matrix of FC showed the best performance; (2) the use of SE-Block and SVM classifier could significantly improve the performance of the integrated model.

Recent studies have indicated that the performance of a multibranch model is better than that of a single-branch model. Aslani et al. designed a multibranch 2D CNN model, and their results showed that the performance of a multibranch model using multimodal data was significantly better than that of a single-branch model with all modalities stacked (Aslani et al., 2019). Zou et al. used two 3D CNN branches and showed similar results (Zou et al., 2017). Ge et al. proposed a hyperspectral image classification method based on 2D combined 3D CNN and multibranch feature fusion, and their results showed that the proposed model extracts more discriminative features through the combination of 2D CNN and 3D CNN, which significantly improves the final classification performance compared to 2D CNN or 3D CNN alone (Ge et al., 2020). In our multibranch model, the 3D CNN branch can extract spatial topology information from the 3D brain map (Hosseini-Asl et al., 2016), while the 2D CNN branch can extract the brain network connection information from the 2D connectivity matrix (Wang et al., 2018). The combination of 3D CNN and 2D CNN can form a complementary relationship, which makes our model perform better than the separate 3D CNN or 2D CNN models.

In previous studies, rs-fMRI and sMRI data were usually analyzed separately in the CNN-based adjunctive diagnosis of schizophrenia, and the fusional information of multimodal MRI data was not fully explored

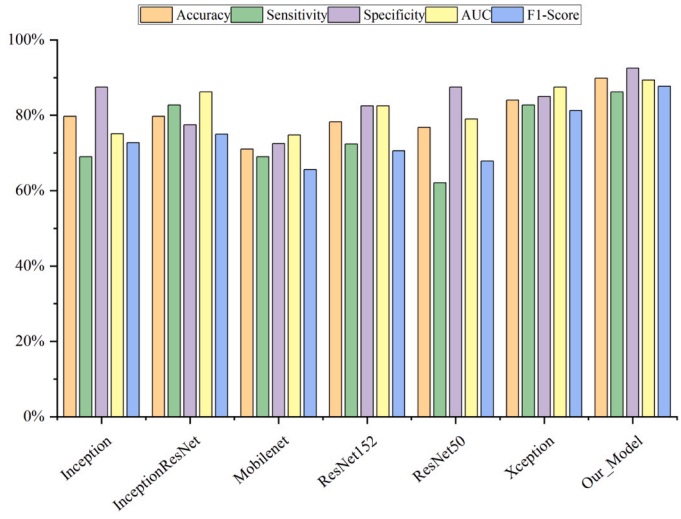

Fig. 7. Comparison of different models.

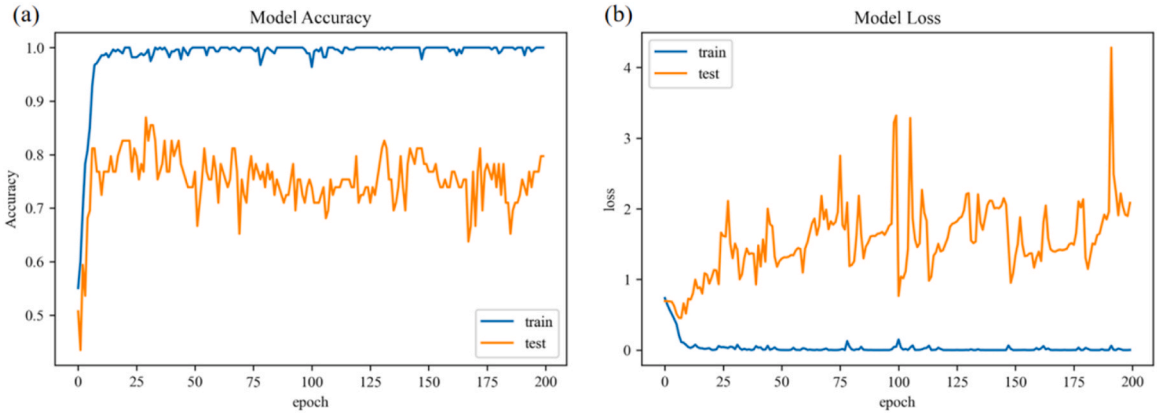

Fig. 6. (a) Accuracy curve when training and testing the model. (b) Loss curve when training and testing the model.

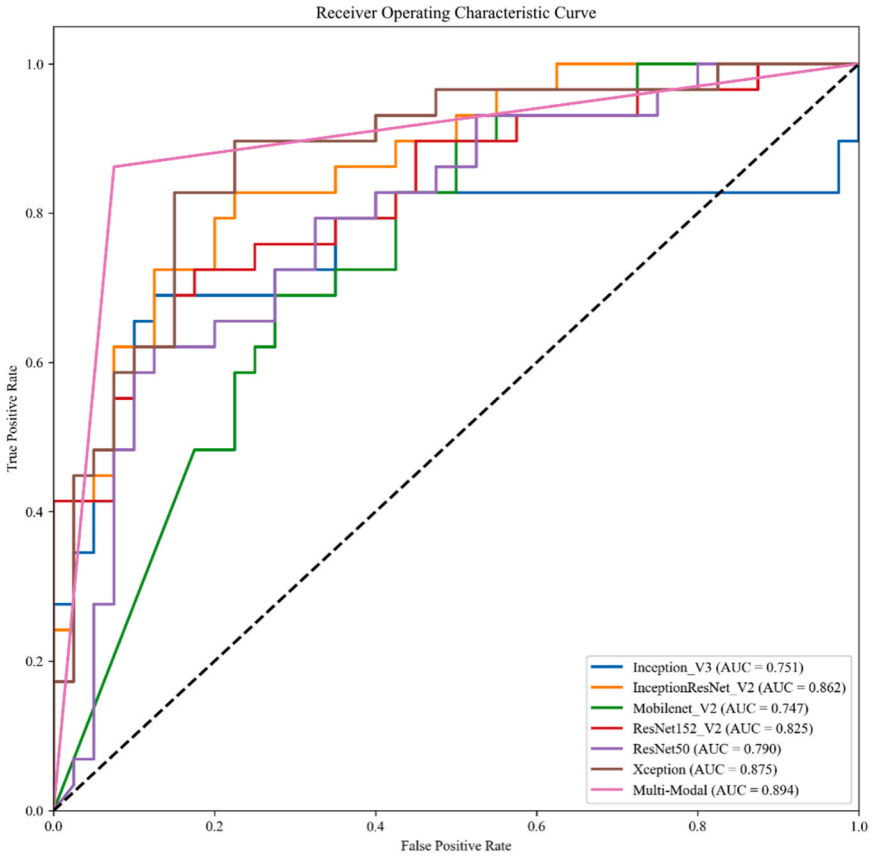

Fig. 8. ROC curves of the various models.

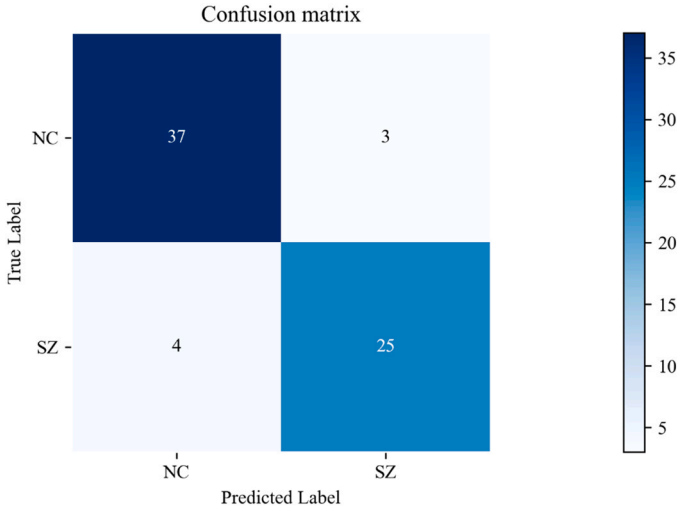

Fig. 9. Confusion matrix.

(Wang et al., 2019; Lei et al., 2020; Kim et al., 2016; Islam and Zhang, 2017; Xiao et al., 2018; Campese et al., 2019). However, numerous classification studies of neuropsychiatric diseases have shown that the use of fusional features of multimodal data shows better performance than that of unimodal data (Zou et al., 2017; Venugopalan et al., 2021; Huang et al., 2019; Cheng and Liu, 2017). Venugopalan et al. designed a multimodal deep learning model for the early detection of Alzheimer's disease, and their results showed that the integration of multiple modalities significantly improved the prediction accuracy (Venugopalan et al., 2021). Cheng et al. proposed cascaded CNNs to discriminate Alzheimer's disease and indicated that using multimodal data can further improve the classification accuracy (Cheng and Liu, 2017). Huang et al. proposed a 3D CNN model using the modalities of MRI and PET and indicated that a combination of two modality imaging data generates better results (Huang et al., 2019). Kanyal et al. proposed a multimodal deep learning model using a combination of imaging genetics features as input and indicated that this feature combination significantly improved SZ classification accuracy (Kanyal et al., 2023). Masoudi et al. proposed a deep-multi-modal fusion based on the fusion of different neuroimaging features and achieved a high classification accuracy (Masoudi and Danishvar, 2022). Our results also showed that a model using multimodal data showed better performance than a model using unimodal data, which suggested that discriminative analysis with fusional multimodal features was a promising direction for exploring neuroimaging biomarkers of SZ.

In recent years, attention modules have been widely used in deep learning for various tasks, such as natural language processing and image classification (Bahdanau et al., 2014; Xu et al., 2015; Vaswani et al., 2017). The attention module emphasized the contribution of different brain regions to the SZ classification and considered the interactions between brain regions, which was consistent with the neural mechanisms of cognitive processes in the brain (Zhang et al., 2020). However, the use of attention modules in SZ classification is relatively rare. Our results showed that the attention module in the integrated model could learn the specificity and interaction patterns between different brain regions, which could significantly improve the classification results.

In the field of deep learning, previous studies have mainly used sigmoid or softmax functions as the classifier of the model (Bahdanau et al., 2014; Xu et al., 2015). However, many studies have investigated an alternative classification method to the Softmax classifier, the SVM (Tang, 2013; Agarap, 2018; Alalshekmubarak and Smith, 2013). Their

results showed that using SVM as a classifier has a better performance than using the traditional Softmax classifier. Therefore, we adopted a SVM classifier in our integrated model, and the results showed that it significantly improved the accuracy of the model by approximately 3%, indicating that using SVM as a classifier would be a better choice.

Despite the above advantages, there were several methodological issues to be addressed in the integrated model proposed in this study. The base model, lacking attention modules and an SVM classifier, is prone to overfitting. This issue could stem from the high dimensionality of individual data points and limited sample size, hindering the model's ability to extract high-level features. Thus, it is worth exploring how to downscale the MRI data and decrease the model's parameters that affect its classification performance.

5. Conclusion

In this study, we proposed an integrated model combining a 3D CNN with a 2D CNN for the classification of SZ patients, in which 3D T1 images and 2D FC matrices derived from rs-fMRI data were used as the inputs for 3D CNN and 2D CNN, respectively. The integrated model achieved an accuracy of 89.86%, showing better performance than other SOTA methods. Our results also indicated that the use of attention modules and SVM classifiers could significantly improve the classification performance and highlighted potential advantages of the integrated model in classifying SZ patients using multimodal MRI data. Our findings provide new insights into the clinical diagnosis of psychiatric diseases from the perspective of deep learning.

Data availability statement

The datasets presented in this article are not readily available because this is a private dataset. Requests to access the datasets should be directed to KW.

CRediT authorship contribution statement

Guo Haiman: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Jian Shuyi:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Zhou Yubin:** Data curation, Formal analysis, Validation, Writing – review & editing. **Chen Xiaoyi:** Data curation, Formal analysis, Validation, Writing – review & editing. **Chen Jinbiao:** Data curation, Formal analysis, Validation, Writing – review & editing. **Zhou Jing:** Conceptualization, Methodology, Resources, Writing – review & editing. **Huang Yuanyuan:** Data curation, Resources, Writing – original draft, Writing – review & editing. **Ma Guolin:** Data curation, Formal analysis, Validation, Writing – review & editing. **Li Xiaobo:** Formal analysis, Writing – review & editing. **Ning Yuping:** Data curation, Resources, Writing – original draft, Writing – review & editing. **Wu Fengchun:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Methodology, Project administration, Resources, Supervision, Writing – review & editing. **WU Kai:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Methodology, Project administration, Resources, Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors declare that there are no conflicts of interest that could inappropriately influence the results of this paper.

Data availability

The authors do not have permission to share data.

Acknowledgments

This work was supported by the Key Research and Development Program of Guangdong (2020B0101130020), the National Natural Science Foundation of China (72174082, 82271953, 81971585, 82301688), Guangdong Basic and Applied Basic Research Foundation Outstanding Youth Project (2021B1515020064), the Science and Technology Program of Guangzhou (201903010032, 202103000032, 202206060005, 202206080005, 202206010077, 202206010034, 202201010093, 2023A03J0856, 2023A03J0839), and Key Laboratory Program of Guangdong Provincial Education Department (2020KSYS001).

References

Adriano, F., Caltagirone, C., Spalletta, G., 2012. Hippocampal volume reduction in first-episode and chronic schizophrenia: a review and meta-analysis. Neuroscientist 18, 180–200. https://doi.org/10.1177/1073858410395147.

Agarap, A.F.M., 2018. A neural network architecture combining gated recurrent unit (GRU) and support vector machine (SVM) for intrusion detection in network traffic data. Proc. 2018 10th Int. Conf. Mach. Learn. Comput. 26–30.

Alalshekmubarak, A., Smith, L.S., 2013. A novel approach combining recurrent neural network and support vector machines for time series classification. 2013 9th International Conference on Innovations in Information Technology (IIT). IEEE,, pp. 42–47.

Aslani, S., Dayan, M., Storelli, L., Filippi, M., Murino, V., Rocca, M.A., Sona, D., 2019. Multi-branch convolutional neural network for multiple sclerosis lesion segmentation. NeuroImage 196, 1–15. https://doi.org/10.1016/j.neuroimage.2019.03.068.

Bahdanau, D., Cho, K., Bengio, Y., 2014. Neural machine translation by jointly learning to align and translate. arXiv Prepr. arXiv 1409, 0473. https://doi.org/10.48550/arXiv.1409.0473.

Caan, M.W., Vermeer, K.A., van Vliet, L.J., Majoie, C.B., Peters, B.D., den Heeten, G.J., Vos, F.M., 2006. Shaving diffusion tensor images in discriminant analysis: a study into schizophrenia. Med Image Anal. 10, 841–849. https://doi.org/10.1016/j.media.2006.07.006.

Cabral, C., Kambeitz-Ilankovic, L., Kambeitz, J., Calhoun, V.D., Dwyer, D.B., von Saldern, S., Urquijo, M.F., Falkai, P., Koutsouleris, N., 2016. Classifying schizophrenia using multimodal multivariate pattern recognition analysis: evaluating the impact of individual clinical profiles on the neurodiagnostic performance. Schizophr. Bull. 42, S110–S117. https://doi.org/10.1093/schbul/sbw053.

Camchong, J., Lim, K.O., Sponheim, S.R., MacDonald III, A.W., 2009. Frontal white matter integrity as an endophenotype for schizophrenia: diffusion tensor imaging in monozygotic twins and patients' nonpsychotic relatives. Front. Hum. Neurosci. 35 https://doi.org/10.3389/neuro.09.035.2009.

Campese, S., Lauriola, I., Scarpazza, C., Sartori, G., Aiolli, F., 2019. Psychiatric disorders classification with 3D convolutional neural networks. INNS Big Data and Deep Learning Conference. Springer,, pp. 48–57.

Campese, S., Lauriola, I., Scarpazza, C., Sartori, G., Aiolli, F., 2020. Psychiatric Disorders Classification with 3D Convolutional Neural Networks. Springer International Publishing,, Cham, pp. 48–57.

Chen, X., Zhou, J., Ke, P., Huang, J., Xiong, D., Huang, Y., Ma, G., Ning, Y., Wu, F., Wu, K., 2023. Classification of schizophrenia patients using a graph convolutional network: A combined functional MRI and connectomics analysis. Biomed. Signal Process. Control 80, 104293. https://doi.org/10.1016/j.bspc.2022.104293.

Cheng, D., Liu, M., 2017. CNNs based multi-modality classification for AD diagnosis. 2017 10th International Congress on Image and Signal Processing, Biomedical Engineering and Informatics (CISP-BMEI). IEEE,, pp. 1–5.

Chin, R., You, A.X., Meng, F., Zhou, J., Sim, K., 2018. Recognition of Schizophrenia with Regularized Support Vector Machine and Sequential Region of Interest Selection using Structural Magnetic Resonance Imaging. Sci. Rep. 8, 13858 https://doi.org/10.1038/s41598-018-32290-9.

Collin, G., Scholtens, L.H., Kahn, R.S., Hillegers, M.H.J., van den Heuvel, M.P., 2017. Affected Anatomical Rich Club and Structural-Functional Coupling in Young Offspring of Schizophrenia and Bipolar Disorder Patients. Biol. Psychiatry 82, 746–755. https://doi.org/10.1016/j.biopsych.2017.06.013.

Cui, L.B., Wei, Y., Xi, Y.B., Griffa, A., De Lange, S.C., Kahn, R.S., Yin, H., Van den Heuvel, M.P., 2019. Connectome-Based Patterns of First-Episode Medication-Naive Patients With Schizophrenia. Schizophr. Bull. 45, 1291–1299. https://doi.org/10.1093/schbul/sbz014.

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., Wang, H.-N., Yin, H., 2016. Putamen-related regional and network functional deficits in first-episode schizophrenia with auditory verbal hallucinations. Schizophr. Res. 173, 13–22. https://doi.org/10.1016/j.schres.2016.02.039.

Farras-Permanyer, L., Guardia-Olmos, J., Pero-Cebollero, M., 2015. Mild cognitive impairment and fMRI studies of brain functional connectivity: the state of the art. Front. Psychol. 6, 1095. https://doi.org/10.3389/fpsyg.2015.01095.

Fornito, A., Yucel, M., Pantelis, C., 2009. Reconciling neuroimaging and neuropathological findings in schizophrenia and bipolar disorder. Curr. Opin. Psychiatry 22, 312–319. https://doi.org/10.1097/YCO.0b013e32832a1353.

Ge, Z., Cao, G., Li, X., Fu, P., 2020. Hyperspectral Image Classification Method Based on 2D–3D CNN and Multibranch Feature Fusion. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 13, 5776–5788. https://doi.org/10.1109/JSTARS.2020.3024841.
van den Heuvel, M.P., Fornito, A., 2014. Brain networks in schizophrenia. Neuropsychol. Rev. 24, 32–48. https://doi.org/10.1007/s11065-014-9248-7.
Hosseini-Asl, E., Keynton, R., El-Baz, A., 2016. Alzheimer's disease diagnostics by adaptation of 3D convolutional network. IEEE Int. Conf. Image Process. (ICIP) 2016, 126–130.
Hu, J., Shen, L., Sun, G., 2018. Squeeze-and-excitation networks. Proc. IEEE Conf. Comput. Vis. Pattern Recognit. 7132–7141.
Hu, M.J., Qian, X., Liu, S.W., Koh, A.J., Sim, K., Jiang, X.D., Guan, C.T., Zhou, J.H., 2022. Structural and diffusion MRI based schizophrenia classification using 2D pretrained and 3D naive Convolutional Neural Networks. Schizophr. Res. 243, 330–341. https://doi.org/10.1016/j.schres.2021.06.011.
Huang, J., Ke, P., Chen, X., Li, S., Zhou, J., Xiong, D., Huang, Y., Li, H., Ning, Y., Duan, X., Li, X., Zhang, W., Wu, F., Wu, K., 2022. Multimodal Magnetic Resonance Imaging Reveals Aberrant Brain Age Trajectory During Youth in Schizophrenia Patients. Front Aging Neurosci. 14, 823502 https://doi.org/10.3389/fnagi.2022.823502.
Huang, Y., Xu, J., Zhou, Y., Tong, T., Zhuang, X., Initiative, As.D.N., 2019. Diagnosis of Alzheimer's disease via multi-modality 3D convolutional neural network. Front Neurosci. 13, 509. https://doi.org/10.3389/fnins.2019.00509.
Islam, J., Zhang, Y., 2017. A novel deep learning based multi-class classification method for Alzheimer's disease detection using brain MRI data. International Conference on Brain Informatics. Springer,, pp. 213–222.
Ji, S., Xu, W., Yang, M., Yu, K., 2012. 3D convolutional neural networks for human action recognition. IEEE Trans. Pattern Anal. Mach. Intell. 35, 221–231. https://doi.org/10.1109/TPAMI.2012.59.
Jiang, H., Cao, P., Xu, M., Yang, J., Zaiane, O., 2020. Hi-GCN: A hierarchical graph convolution network for graph embedding learning of brain network and brain disorders prediction. Comput. Biol. Med. 127, 104096 https://doi.org/10.1016/j.compbiomed.2020.104096.
Kanyal, A., Kandula, S., Calhoun, V., Ye, D.H., 2023. Multi-Modal Deep Learning on Imaging Genetics for Schizophrenia Classification (June). In: 2023 IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops (ICASSPW), 2023. IEEE,, pp. 1–5 (June).
Karlsgodt, K.H., Niendam, T.A., Bearden, C.E., Cannon, T.D., 2009. White matter integrity and prediction of social and role functioning in subjects at ultra-high risk for psychosis. Biol. Psychiatry 66, 562–569. https://doi.org/10.1016/j.biopsych.2009.03.013.
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. https://doi.org/10.1016/j.neuroimage.2015.05.018.
Kong, L.Y., Huang, Y.Y., Lei, B.Y., Ke, P.F., Li, H.H., Zhou, J., Xiong, D.S., Li, G.X., Chen, J., Li, X.B., Xiang, Z.M., Ning, Y.P., Wu, F.C., Wu, K., 2021. Divergent Alterations of Structural-Functional Connectivity Couplings in First-episode and Chronic Schizophrenia Patients. Neuroscience 460, 1–12. https://doi.org/10.1016/j.neuroscience.2021.02.008.
LeCun, Y., Bottou, L., Bengio, Y., Haffner, P., 1998. Gradient-based learning applied to document recognition. Proc. IEEE 86, 2278–2324. https://doi.org/10.1109/5.726791.
Lei, D., Pinaya, W.H.L., van Amelsvoort, T., Marcelis, M., Donohoe, G., Mothersill, D.O., Corvin, A., Gill, M., Vieira, S., Huang, X., Lui, S., Scarpazza, C., Young, J., Arango, C., Bullmore, E., Qiyong, G., McGuire, P., Mechelli, A., 2020. Detecting schizophrenia at the level of the individual: relative diagnostic value of whole-brain images, connectome-wide functional connectivity and graph-based metrics. Psychol. Med 50, 1852–1861. https://doi.org/10.1017/S0033291719001934.
Masoudi, B., Danishvar, S., 2022. Deep multi-modal schizophrenia disorder diagnosis via a GRU-CNN architecture. Neural Netw. World 32 (3), 147.
Nelson, B.G., Bassett, D.S., Camchong, J., Bullmore, E.T., Lim, K.O., 2017. Comparison of large-scale human brain functional and anatomical networks in schizophrenia. NeuroImage. Clin. 15, 439–448. https://doi.org/10.1016/j.nicl.2017.05.007.
Qureshi, M.N.I., Oh, J., Lee, B., 2019. 3D-CNN based discrimination of schizophrenia using resting-state fMRI. Artif. Intell. Med 98, 10–17. https://doi.org/10.1016/j.artmed.2019.06.003.
Rossler, W., Salize, H.J., van Os, J., Riecher-Rossler, A., 2005. Size of burden of schizophrenia and psychotic disorders. Eur. Neuropsychopharmacol. 15, 399–409. https://doi.org/10.1016/j.euroneuro.2005.04.009.
Shi, J., Zheng, X., Li, Y., Zhang, Q., Ying, S.H., 2018. Multimodal Neuroimaging Feature Learning With Multimodal Stacked Deep Polynomial Networks for Diagnosis of Alzheimer's Disease. Ieee J. Biomed. Health 22, 173–183. https://doi.org/10.1109/Jbhi.2017.2655720.
Su, T.W., Hsu, T.W., Lin, Y.C., Lin, C.P., 2015. Schizophrenia symptoms and brain network efficiency: A resting-state fMRI study. Psychiatry Res. 234, 208–218. https://doi.org/10.1016/j.pscychresns.2015.09.013.
Tang, Y., 2013. Deep learning using linear support vector machines. arXiv Prepr. arXiv 1306, 0239. https://doi.org/10.48550/arXiv.1306.0239.
Tost, H., Meyer-Lindenberg, A., 2012. Puzzling over schizophrenia: schizophrenia, social environment and the brain. Nat. Med 18, 211–213. https://doi.org/10.1038/nm.2671.
Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Ł. Kaiser, I., 2017. Polosukhin, Attention is all you need. Adv. Neural Inf. Process. Syst. 30.
Venkataraman, A., Whitford, T.J., Westin, C.F., Golland, P., Kubicki, M., 2012. Whole brain resting state functional connectivity abnormalities in schizophrenia. Schizophr. Res 139, 7–12. https://doi.org/10.1016/j.schres.2012.04.021.
Venugopalan, J., Tong, L., Hassanzadeh, H.R., Wang, M.D., 2021. Multimodal deep learning models for early detection of Alzheimer's disease stage. Sci. Rep. 11, 3254 https://doi.org/10.1038/s41598-020-74399-w.
Wang, J., Ke, P.F., Zang, J.Y., Wu, F.C., Wu, K., 2022. Discriminative Analysis of Schizophrenia Patients Using Topological Properties of Structural and Functional Brain Networks: A Multimodal Magnetic Resonance Imaging Study. Front Neurosci. -Switz. (15 ().
Wang, S., Zhang, Y., Lv, L., Wu, R., Fan, X., Zhao, J., Guo, W., 2018. Abnormal regional homogeneity as a potential imaging biomarker for adolescent-onset schizophrenia: a resting-state fMRI study and support vector machine analysis. Schizophr. Res. 192, 179–184. https://doi.org/10.1016/j.schres.2017.05.038.
Wang, Y., Yang, Y., Guo, X., Ye, C., Gao, N., Fang, Y., Ma, H.T., 2018. A novel multimodal MRI analysis for Alzheimer's disease based on convolutional neural network. In: 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2018. IEEE,, pp. 754–757.
Wang, Z.J., Sun, Y.R., Shen, Q.Z., Cao, L., 2019. Dilated 3D Convolutional Neural Networks for Brain MRI Data Classification. Ieee Access 7, 134388–134398. https://doi.org/10.1109/Access.2019.2941912.
White, T., Magnotta, V.A., Bockholt, H.J., Williams, S., Wallace, S., Ehrlich, S., Mueller, B.A., Ho, B.-C., Jung, R.E., Clark, V.P., Lauriello, J., Bustillo, J.R., Schulz, S. C., Gollub, R.L., Andreasen, N.C., Calhoun, V.D., Lim, K.O., 2009. Global White Matter Abnormalities in Schizophrenia: A Multisite Diffusion Tensor Imaging Study. Schizophr. Bull. 37, 222–232. https://doi.org/10.1093/schbul/sbp088.
Wu, F., Zhang, Y., Yang, Y., Lu, X., Fang, Z., Huang, J., Kong, L., Chen, J., Ning, Y., Li, X., Wu, K., 2018. Structural and functional brain abnormalities in drug-naive, first-episode, and chronic patients with schizophrenia: a multimodal MRI study. Neuropsychiatr. Dis. Treat. 14, 2889–2904. https://doi.org/10.2147/NDT.S174356.
Xiao, Z., Wang, C., Jia, N., Wu, J., 2018. SAE-based classification of school-aged children with autism spectrum disorders using functional magnetic resonance imaging. Multimed. Tools Appl. 77, 22809–22820. https://doi.org/10.1007/s11042-018-5625-1.
Xu, K., Ba, J., Kiros, R., Cho, K., Courville, A., Salakhudinov, R., Zemel, R., Bengio, Y., 2015. Show, attend and tell: Neural image caption generation with visual attention. International conference on machine learning. PMLR,, pp. 2048–2057.
Yan, C.G., Wang, X.D., Zuo, X.N., Zang, Y.F., 2016. DPABI: Data Processing & Analysis for (Resting-State) Brain Imaging. Neuroinformatics 14, 339–351. https://doi.org/10.1007/s12021-016-9299-4.
Yu, Y., Shen, H., Zeng, L.L., Ma, Q., Hu, D., 2013. Convergent and divergent functional connectivity patterns in schizophrenia and depression. Plos One 8, e68250. https://doi.org/10.1371/journal.pone.0068250.
Zang, J., Huang, Y., Kong, L., Lei, B., Ke, P., Li, H., Zhou, J., Xiong, D., Li, G., Chen, J., Li, X., Xiang, Z., Ning, Y., Wu, F., Wu, K., 2021. Effects of Brain Atlases and Machine Learning Methods on the Discrimination of Schizophrenia Patients: A Multimodal MRI Study. Front Neurosci. 15, 697168 https://doi.org/10.3389/fnins.2021.697168.
Zhang, J., Wei, F., Feng, F., Wang, C., 2020. Spatial–Spectral Feature Refinement for Hyperspectral Image Classification Based on Attention-Dense 3D–2D-CNN. Sensors 20, 5191. https://doi.org/10.3390/s20185191.
Zhang, T., Li, C.B., Li, P.Y., Peng, Y.H., Kang, X.D., Jiang, C.Y., Li, F.L., Zhu, X.Y., Yao, D. Z., Biswal, B., Xu, P., 2020. Separated Channel Attention Convolutional Neural Network (SC-CNN-Attention) to Identify ADHD in Multi-Site Rs-fMRI Dataset. Entropy-Switz. 22 https://doi.org/10.3390/e22080893.
Zhao, X., Yao, J., Lv, Y., Zhang, X., Han, C., Chen, L., Ren, F., Jin, Z., Li, Y., Sui, Y., 2019. Abnormalities of regional homogeneity and its correlation with clinical symptoms in Naïve patients with first-episode schizophrenia. Brain Imaging Behav. 13, 503–513. https://doi.org/10.1007/s11682-018-9882-4.
Zou, L., Zheng, J.N., Mia, C.Y., Mckeown, M.J., Wang, Z.J., 2017. 3D CNN based automatic diagnosis of attention deficit hyperactivity disorder using functional and structural MRI. Ieee Access 5, 23626–23636. https://doi.org/10.1109/ACCESS.2017.2762703.