Applied Neuropsychology: Adult

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/hapn21

Exploring brain effective connectivity of early MCI with GRU_GC model on resting-state fMRI

Lei Wang, Weiming Zeng, Le Zhao & Yuhu Shi

To cite this article: Lei Wang, Weiming Zeng, Le Zhao & Yuhu Shi (21 Mar 2024): Exploring brain effective connectivity of early MCI with GRU_GC model on resting-state fMRI, Applied Neuropsychology: Adult, DOI: 10.1080/23279095.2024.2330100

To link to this article: https://doi.org/10.1080/23279095.2024.2330100

Published online: 21 Mar 2024.

APPLIED NEUROPSYCHOLOGY: ADULT
https://doi.org/10.1080/23279095.2024.2330100

Exploring brain effective connectivity of early MCI with GRU_GC model on resting-state fMRI

Lei Wang, Weiming Zeng, Le Zhao and Yuhu Shi

Lab of Digital Image and Intelligent Computation, Shanghai Maritime University, Shanghai, China

ABSTRACT

Background: Investigating the functional interactions between different brain regions and revealing the transmission of information by computing brain connectivity have great potential and significance in the diagnosis of early Mild Cognitive Impairment (EMCI).

Methods: The Granger causality with Gate Recurrent Unit (GRU_GC) model is a suitable method that allows the detection of a nonlinear causal relationship and solves the limitation of fixed time lag, which cannot be detected by the classical Granger method. The model can transmit time series signals with any transmission delay length, and the time series can be screened and learned through the gate model.

Results: The classification experiment of 89 EMCI and 73 neurologically healthy controls (HC) shows that the accuracy reached 87.88%. Compared with multivariate variables GC (MVGC) and Long Short-Term Memory-based GC (LSTM_GC), the GRU_GC significantly improved the estimation of brain connectivity communication. Constructing a difference network to explore the brain effective connectivity between EMCI and HC.

Conclusions: The GRU_GC can discover the abnormal brain regions, including the parahippocampal gyrus, the posterior cingulate gyrus. The method can be used in clinical applications as an effective brain connectivity analysis tool and provides auxiliary means for the medical diagnosis of EMCI.

KEYWORDS

Difference network; early mild cognitive impairment; effective connectivity; gate recurrent unit; random forest

Introduction

Functional magnetic resonance imaging (fMRI) mainly refers to blood oxygen levels-dependent fMRI and has the advantages of noninvasive, repeatable, and high spatial resolution. It is not only a tool for generating brain activation maps but also a means to study neural network dynamics by tracking the response characteristics on different temporal and spatial scales (Logothetis et al., 2001). Another merit of fMRI is that it can track signal changes in real-time and obtain time series of brain activities (Friston et al., 1995). The fMRI has been widely used to study pathogenesis, locate brain functional regions, and detect mental diseases such as mild cognitive impairment (MCI) and Alzheimer's disease, with MCI further divided into early MCI (EMCI) and late MCI (LMCI) stages (Deng et al., 2018; Shi et al., 2015, 2020; Zhao et al., 2020).

Brain Connectivity (BC) based on fMRI provides broader research directions for further revealing the internal working mechanism of the brain and the occurrence of diseases. BC can be divided into Functional Connectivity (FC) and Effective Connectivity (EC) (Friston, 1995). FC is the spatial connectivity of brain functional regions, describing whether there is connectivity or interactive information between brain regions, commonly used methods such as independent component analysis (ICA) (Shi et al., 2018). EC refers to the directional connectivity of brain neurons or brain regions to form brain networks, and its edges represent the directional weight of neurons or brain regions to another, reflecting the correlation and directionality between brain regions (Schlösser et al., 2008). Compared with FC, EC can reflect the flow of information between neurons or brain regions, and the causality influence of one brain region (neuron system) on another. It is a trend to reveal the information transmission and interaction between brain regions and study the working mechanism of the brain. The methods used by EC to study the brain include the Structural Equation Model (SEM), Dynamic Causality Model (DCM) (Xin & Biswal, 2014), Granger Causality Analysis (GCA) (Riley et al., 2018). GCA was defined by Wiener in 1956 and Granger proposed in 1969 (Granger, 1969; Porta & Faes, 2016). K. J. Friston introduced it into the field of neuroscience to measure the brain electrical activity between two brain regions and afterward achieved outstanding application effects. GCA requires no prior knowledge and only relies on the characteristics of time series. It is a statistical method to study the flow of information between time series (Friston et al., 2013). Therefore, it is widely used in different data related to time series, such as electroencephalogram (EEG) and fMRI (Dimitriadis et al., 2012).

The original GCA is a linear method based on a time-domain vector autoregressive model. Although this

CONTACT W. Zeng zengwm86@163.com Lab of Digital Image and Intelligent Computation, Shanghai Maritime University, Shanghai, 201306, China.

method improves the understanding of brain activity and cognitive function, its drawbacks are obvious. Linear Granger Causality Model (LGCM) cannot simply judge virtual causality because the nervous system of the brain has potential variable problems (Seth et al., 2013). At the same time, it turns out that the time series of scanned fMRI data is often highly non-linear (Logothetis et al., 2001). In addition, the premise of LGCM requires the time series to be random and widely stable which requires that the mean and variance are constant. Otherwise, it is easy to get incorrect regression values (Bressler & Seth, 2011). Non-stationary data can be processed using sliding time windows to obtain local stability. Scholars have put forward different optimization solutions to the limitations of the above LGCM, such as the Partial Granger Causality (PGC) to eliminate the influence of latent variables in the nervous system (Guo et al., 2008). Then, in Multivariate Variables Granger Causality (MVGC), the bivariate is extended to multivariate to evaluate the causality between time series to measure the directional power transfer in the frequency domain, which is called the directional transfer function (DTF) (Deshpande et al., 2010). Linearity, however, is a potential drawback of DTF. In recent years, with the rise of deep learning, Granger causality and neural networks have been combined. Granger causality based on the neural network directly characterizes the relationship between brain signals, replacing vector autoregression (Farokhzadi et al., 2018; Khadem & Hossein-Zadeh, 2014; Montalto et al., 2015). Therefore, it is suitable for non-stationary fMRI time series, and the performance has been improved (Guo et al., 2020). Although it is capable in non-linear situations, it usually requires a long stationary signal and is susceptible to noise (Pereda et al., 2005). And more importantly, the transmission of signals in brain regions is not certain to be completed in a fixed time, so the limitation of the methods is fixed by the past time lag. They used the past lags of time series to predict current information, while the model needs to be set artificially. Meanwhile, it makes little full use of the time sequence information of the data.

The introduction of complex learning networks such as Recurrent Neural Networks (RNN) further increases the exploration ability of brain connectivity in fMRI data. These deep learning algorithms can automatically study from the input of fMRI data to discover high-level information hidden in it. The RNN network is more focused on the correlation of data in the time periods and also considers the data of past time points, which just meets the needs of learning fMRI data in time series. Long Short-Term Memory (LSTM) is a special RNN that can learn long-term dependent information (Hochreiter & Schmidhuber, 1997). LSTM removes or adds information to the cell state through a well-designed structure called "gate". The LSTM has now achieved advanced results in various sequence processing tasks (Graves, 2012), such as speech recognition, image captioning and brain connectivity analysis (Greff, 2016; Wang et al., 2018). The shortcoming of LSTM is computationally higher (Group, Nlc, 2017) and largely depends on the hardware requirements. As a variant of the LSTM network, GRU not only has similar performance to the LSTM network but also is easier to calculate than LSTM (Cho et al., 2014). These advanced recurrent units in GRU are better than the most of traditional recurrent units such as tanh units (Chung et al., 2014). According to the above, the GRU_GC model used in this paper is focused on overcoming the main disadvantages of the non-modified Granger method based on linear autoregressive models and proposing an alternative for existing nonlinear methods (Rosoł et al., 2022). The GRU_GC and LSTM_GC were applied to estimate multivariate brain connectivity and experimental results show that GRU has better estimation performance in brain effective connectivity than LSTM. The model takes time series of arbitrary lag time as the input and the effective connectivity matrix as the output, while learning the flow of information in the data. Based on the matrices, effective connectivity networks were constructed, and explored the functional activities of EMCI and neurologically healthy control (HC) brains in the resting-state.

Based on the above analysis, GRU_GC was applied to the research of EMCI brain regions, and the EC matrices were used to classify EMCI patients and HC subjects. It provides an auxiliary method for the clinical diagnosis of EMCI. The main organization of this paper is enumerated as follows: The second part introduces the materials and methods. The third part is about the results. The fourth part and the fifth part are the discussion and the conclusion, respectively.

Materials and methods

Firstly, the resting-state fMRI (rs-fMRI) data with multiple time points were preprocessed and the data was used to select the region of interest (ROI) through prior knowledge. Then, the effective connectivity between brain regions was acquired by using the GRU_GC. Next, according to the ranking of feature importance scores, the features were selected and input to the random forest (RF) classifier for classification. After that, the difference network and hub nodes were built through the obvious different edges found by two sample T-test. Finally, we analyzed the EMCI brain region changes according to the results of DN and selected features. The overall experimental process is shown in Figure 1.

Data acquisition and preprocessing

In this paper, two groups of subjects including 89 EMCI patients and 73 neurologically healthy controls were selected from the ADNI database1 (http://adni.loni.usc.edu/) for the experiments. ADNI is the multisite observational study of normal aging, MCI, and AD. MCI subjects are sub-classed in two subtypes, EMCI and LMCI in ADNI, based on the WMS-R Logical Memory II Story A score. The EMCI is considered to reflect those at the earlier point in the clinical

1. Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu) that was launched in 2003 and led by Principal Investigator Michael W. Weiner, MD. For up-to-date information, see www.adni-info.org.

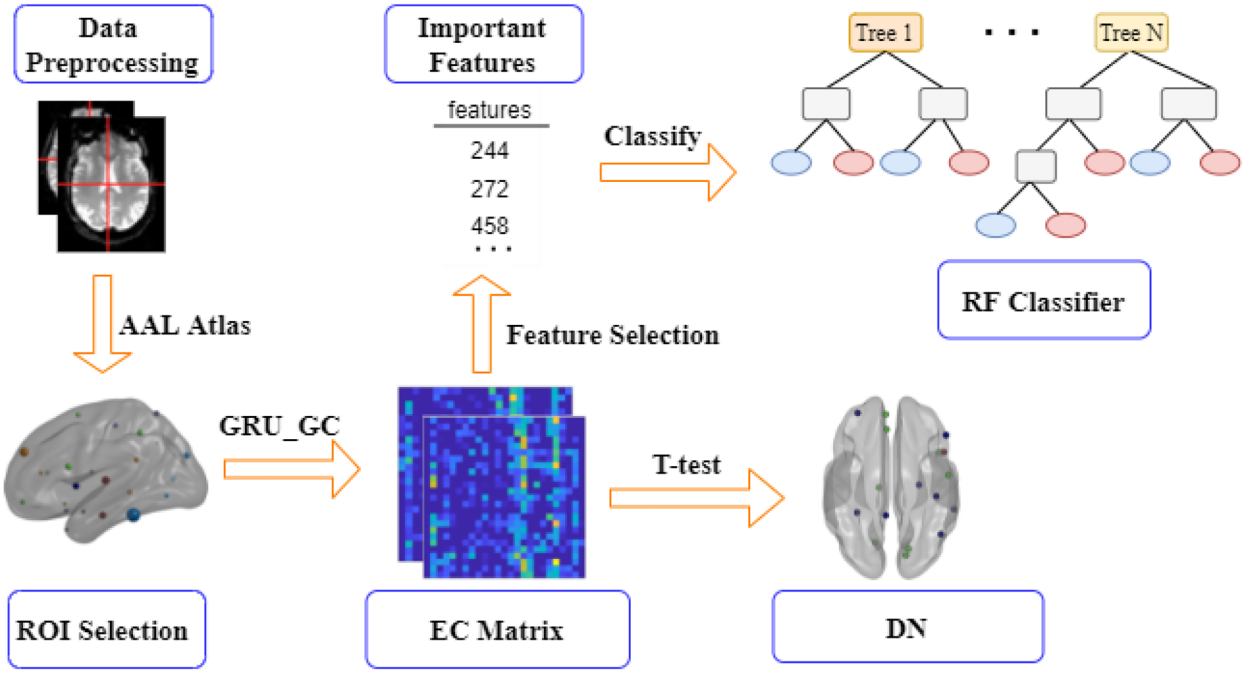


Figure 1. Overview of our proposed framework for EMCI classification and constructed difference network to seek the changes in brain regions between EMCI and healthy control group. (EC Matrix: effective connectivity matrix, DN: difference network, RF classifier: random forest classifier, Feature Selection: tree-based feature selection.).

spectrum, while LMCI is at the later point to progress to AD (Jitsuishi & Yamaguchi, 2022). According to the ADNI protocol, the data were all scanned by Philips' 3.0 Tesla. During the scanning, the subjects needed to be supine with the whole body relaxed in the nuclear magnetic resonance equipment. The head should be fixed as far as possible to avoid the experimental error.

The fMRI images were acquired using an echo-planar imaging (EPI) sequence with a repetition time (TR) of 3000.0 ms; Echo Time (TE) = 30 ms; Time Points = 140s; Flip Angle (FA) = 80.0 degree; Slices = 48. The Pixel Spacing in the X and Y dimensions was 3.3 mm, and the slice thickness was 3.3 mm.

The Data Processing Assistant for the Resting-State Toolbox DPABI (http://rfmri.org/dpabi) in the MATLAB2018b platform was used to preprocess the fMRI data of EMCI and HC groups (Yan & Zang, 2010). The preprocessing steps were as follows: (1) Convert the data from DICOM to NIFTI. (2) Remove the first 10 time points. (3) Slice timing with slice 47 as the reference. (4) Head motion correction, as well as offset and angular motion correction. (5) Spatial normalization using EPI templates and reslicing to 3 mm × 3 mm × 3 mm voxels. (6) Smoothing with a Gaussian kernel function (FWHM = 4 mm) to reduce the noise of the different signals. (7) Detrending. (8) Nuisance covariates regression. (9) Time series of regions of interest was extracted.

ROIs selection

The definition of ROIs adopted in this paper was derived from the 25 brain regions with abnormal functions in the AAL template related to EMCI, LMCI and AD proposed in the reference (Guo et al., 2017), as shown in Table 1.

Granger causality

Granger causality analysis method is a common method for effective connectivity judgment in multivariate time series (Goebel et al., 2003; Roebroeck et al., 2005). The Granger causality assumes that the time series $X=\{X_1,\ldots,X_t,\ldots\}, Y=\{Y_1,\ldots,Y_t,\ldots\}, Z=\{Z_1,\ldots,Z_t,\ldots\}$ are stable, if the time series X can improve the accuracy of future prediction of X with the addition of time series Y, then Y is called the causality of X. To test whether Y is the cause of X, Granger causality analysis is estimated by fitting a vector autoregressive model to the time series with the time lag of P. Under a given time series Z, it is able to indicate whether the past of the Y signal helps to reduce the variance of the prediction error of X without the influence of Z. In this case of conditional causality analysis, the linear autoregressive equations are as follows:

$$X_t = \sum_{p=1}^{P} a_p X_{p-i} + \sum_{p=1}^{P} c_p Z_{t-p} + \xi_t \quad (1)$$

$$X_t' = \sum_{p=1}^{P} a_p' X_{t-p} + \sum_{p=1}^{P} b_p' Y_{t-p} + \sum_{p=1}^{P} c_p' Z_{t-p} + \xi_t' \quad (2)$$

where a_p, c_p, a_p', c_p', b_p' are the best regression parameters of the model, the model order P can be determined by the Akaike Information Criterion (AIC) or Bayesian Information Criterion (BIC), ξ_t, ξ_t' are the predicted residuals. The quality of the prediction is determined by the variance σ of the error in the prediction. Under the condition of time series Z, $\sigma(\xi_t)$ is the error variance in predicting X when time

Table 1. ROIs selection. This table is the set of selected brain regions and No stands for the ROI serial number in this paper. ABBR. stands for the abbreviation of brain regions, followed by the number in brackets that corresponds to the number in the AAL template. The brain region represents the brain region name.

No	Abbr.	ROI	No	Abbr.	ROI
1	Frontal_Mid_Orb_L(9)	Left Middle frontal gyrus, orbital part	14	Cuneus_R(46)	Right Cuneus
2	Frontal_Inf_Oper_L(11)	Left Inferior frontal gyrus, opercular part	15	Lingual_L(47)	Left Lingual gyrus
3	Frontal_Inf_Tri_R(14)	Right Inferior frontal gyrus, triangular part	16	Occipital_Sup_L(49)	Left Superior occipital gyrus
4	Rolandic_Oper_R(18)	Right Rolandic operculum	17	Fusiform_L(55)	Left Fusiform gyrus
5	Supp_Motor_Area_L(19)	Left Supplementary motor area	18	Postcentral_R(58)	Right Postcentral gyrus
6	Olfactory_R(22)	Right Olfactory cortex	19	Parietal_Sup_R(60)	Right Superior parietal gyrus
7	Frontal_Sup_Medial_L(23)	Left Superior frontal gyrus, medial	20	Parietal_Inf_L(61)	Left Inferior parietal, but supramarginal and angular gyri
8	Insula_L(29)	Left Insula	21	Angular_R(66)	Right Angular gyrus
9	Cingulum_Ant_L(31)	Left Anterior cingulate and paracingulate gyri	22	Thalamus_L(77)	Left Thalamus
10	Cingulum_Post_L(35)	Left Posterior cingulate gyrus	23	Temporal_Pole_Sup_R(84)	Right Temporal pole: superior temporal gyrus
11	ParaHippocampal_R(40)	Right Parahippocampal gyrus	24	Temporal_Mid_R(86)	Right Middle temporal gyrus
12	Amygdala_R(42)	Right Amygdala	25	Temporal_Mid_R(88)	Right Temporal pole: middle temporal gyrus
13	Calcarine_R(44)	Right Calcarine fissure and surrounding cortex			

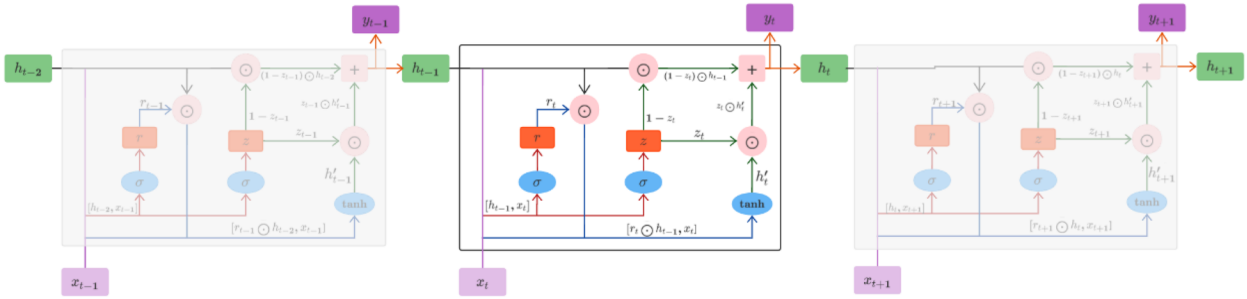


Figure 2. GRU unit structure diagram. h_{t-1}, x_t are the input of unit structure, h_t, y_t are the output, and h_t as input to the next unit. r, z are the reset gate and update gate, respectively. The symbols $\odot$ inside the pink circle represents the "Hadamard product" operation, and + is "plus". σ, tanh in the blue ellipses indicate functions, σ is the sigmoid function.

series X is used, whereas $\sigma(\xi_t')$ the error variance in predicting X when time series X and Y are used. Therefore, the intensity of Granger causality Y → X under the condition of Z can be defined as:

$$F_{Y\to X|Z} = \ln \frac{\sigma(\xi)}{\sigma(\xi')} \tag{3}$$

If there is no direct causality between X and Y, but indirect causality between them is caused by Z, then $\sigma(\xi_t) = \sigma(\xi_t')$, thus $F_{Y\to X|Z} = 0$. This indicates that under the condition of Z, the addition of Y does not improve the prediction accuracy of X, meaning there is no causality between Y and X.

GRU_GC model

The traditional neural network cannot use the previous information to intervene in the subsequent prediction, while the RNN can solve this problem by considering the output of the previous time step in the transmission. RNN allows the continuous transmission of information, in other words, RNN is a neural network that reuses structural units (Graves, 2012) and this continuous transmission is the power of this network. As a member of RNN, GRU holds the superiority of the time dimension. At the same time, GRU can reduce the calculation based on maintaining the advantage of LSTM in solving the gradient descent problem in backpropagation (Sutskever et al., 2014) which is more convenient.

In Figure 2, the GRU unit structure diagram is shown, inputs of the cell are current time information x_t and historical information h_{t-1}. Intuitively, the reset gate r determines how the current time series x_t combined with previous memory h_{t-1}, and update gate z defines the amount of h_{t-1} saved to the current time step. After two gates, new information was calculated and passed to the next cell. Unlike the LSTM, there are only two gates in this unit, while LSTM has three, including forget gate. Importantly, since $(1-z)$ has the same effect as the forget gate in LSTM, GRU can achieve same even better performance. σ is the sigmoid function that transforms information into a value in the range of $0 \sim 1$, tanh function

converts information to $-1 \sim 1$. The red lines represent the inputs to the r,z at the current moment. The blue and green lines express the selection of the current time series and the previous information, and finally, the orange ones indicate the data passed to the next state and the output at that moment.

The GRU with the ability to remember long or short historical information is endowed by the structure of the gates. Therefore, using GRU-based GC to model connectivity in given time series of different transmission lags, which can be well adapted to the problem of brain connectivity estimation. In the process of using GRU_GC to calculate the brain effective connectivity, the time series is predicted by GRU and then calculating the causality with GC. The GRU_GC model can be defined as:

$$F_{Y\to X|Z} = \ln \frac{\sigma(X_t - GRU(X_{t-1} + Z_{t-1}))}{\sigma(X_t - GRU(X_{t-1} + Y_{t-1} + Z_{t-1}))} \quad (4)$$

where GRU is the Gate Recurrent Unit. $GRU(X_{t-1} + Z_{t-1})$ and $GRU(X_{t-1} + Y_{t-1} + Z_{t-1})$ used time series X, Z, and X, Y, Z as inputs in GRU, respectively. The time series prediction of brain regions is performed by GRU neural network, followed by calculation of the predicted error variance to obtain the EC matrices between brain regions. GRU_GC has shown noteworthy advantages in detecting brain connectivity problems. First, compared with vector autoregressive models, nonlinear factors can be considered. Then compared with neural networks, this method considers the time dimension more comprehensively and retains historical time information. Maciej Rosoł et al. have proved the model was able to detect non-linear causality, make accurate forecasting, and not indicate false causality. The neural network-based approach is a suitable method that allows the detection of a nonlinear causal relationship, which cannot be detected by the classical Granger method. The GRU-based method overcomes this problem by cycling time series and can predict more accurately.

Features selection and classifiers

In this paper, two methods were used to select features and classify two groups of subjects. Two ways including random forest and support vector machine (SVM) were used to clear up the impact of the classifier and go a step further to verify the advantages and necessity of the proposed model.

Random forest

Before using effective connectivity matrices of 162 subjects to classify, the data is divided into 10 parts, among which eight parts are training sets for feature selection by sorted according to the feature importance and the remaining two are testing sets for classification.

The RF with the characteristics of high accuracy, strong robustness and convenient of use. The random forest was used for the classification test, and the feature attributes were added in order of feature importance score from the largest to the smallest for classification. Selected the features with the highest classification accuracy as candidates and retained the highest accuracy by comparing the classification results. The features importance according to the Gini index. It is assumed that the proportion of the k-th sample in the current set D is listed as p_k (k = 1,2, …, K), then the purity of D can be measured by Gini value:

$$Gini(p) = 1 - \sum_{k=1}^{K} p_k^{\,2} \quad (5)$$

The Gini index reflects the probability that two samples randomly selected in set D have inconsistent category markers. Therefore, a set more purity, higher Gini index it is. The features importance means the sum of the Gini index decline. In general, the importance score measures the value of the attribute in the construction of the boosting decision tree in the model. The importance of an attribute is relative, as each attribute will be calculated and ordered. It means that the more an attribute is used in constructing the decision tree of the model, the higher its importance is. For the classification task, performance on the testing set was assessed using the accuracy of classification and the area under the receiver operating characteristic curve (AUC), in which an AUC of 1 indicates a perfect classifier while an AUC of 0.5 indicates a classifier that performs no better than random chance.

Support vector machine

Each subject brain effective connectivity matrix according to the GRU_GC model was arranged and reorganized in rows after removing the diagonal. It means that 600 feature attributes were concatenated into a single feature vector after removing diagonals, and each of them is connectivity strength between various brain regions. Then, the vector in which the lowest features' weight was cyclically deleted was used as the input features for the SVM classifier. To train and test the model, a 10-fold iterative cross-validation scheme was employed, in which the data were split into a training set (80%) and a testing set (20%) the same as the split in RF. An SVM was generated using the training set and its performance was assessed on the testing set. The output was the classification of each subject as EMCI or HC or the prediction for each subject.

Evaluation

To ensure that the addition of GRU can play a positive role in classifying EMCI and HC. The feature importance scores were calculated, and the RF algorithm was used for classification. Furthermore, we repeated this process 20 times to measure the effectiveness of the proposed method and reported the averaged performance with standard deviation. For quantitative measurement, accuracy, sensitivity and specificity are used to test the stability and robustness of the classification, the formulas are as follows:

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP}$$
$$Sensitivity = \frac{TP}{TP + FN} \quad (6)$$
$$Specificity = \frac{TN}{TN + FP}$$

where TP means true positive which is the number of EMCI classified correctly, TN denotes true negative, which is the number of HC correctly classified, FP is false positive which means the number of HC classified as EMCI, and FN denotes false negative which is the number of EMCI classified as HC. Besides, the receiver operating curve (ROC) of classification also can be plotted and AUC for evaluation can be calculated, which is a better performance indicator than accuracy (Fawcett, 2005).

Results

Results of classification

Random forest

The effective connectivity matrices as feature sets and effective edges as features were repeated 20 times for feature selection and classification. Finally, four metrics were obtained: accuracy, sensitivity, specificity, and AUC. The classification performance of comparing the MVGC, LSTM_GC and GRU_GC by using an RF classifier, is shown in Table 2. Figure 3 shows a boxplot of 20 times classification accuracy, specificity, and sensitivity. In the boxplot, the mean, variance, and other statistics of the three indicators of the experiments are shown. It can be seen the highest accuracy is 88% in the figure.

The reason for feature selection is that redundancy will affect the accuracy of the classification and increase the consumption of time. In this paper, the tree-based feature selection algorithm was used to determine the contribution of the features to the classification by calculating their importance, so as to determine the marked features.

According to the maximal accuracy of 87.88% (88% in Figure 3) and the importance score, we fixed the selected 36 features. Remarkable and stable classification accuracy (87.88%) can be reached using the selected features and the corresponding sensitivity and specificity were 80% and 94.4%, respectively. The corresponding serial numbers of the top 10 features, which are brain regions' effective edges as shown in Figure 4(a). As shown in Figure 4(b), The ROC curve is very close to the upper left corner, with an area of 0.88, close to 1.

Table 2. Classification performance of EC matrices by using GRU_GC and MVGC after repeated 20 times.

Method	AUC	Accuracy (%)	Sensitivity (%)	Specificity (%)
GRU_GC	0.807	79.70	85.28	73
LSTM_GC	0.573	62.73	73.61	49.67
MVGC	0.525	61.52	61.80	44

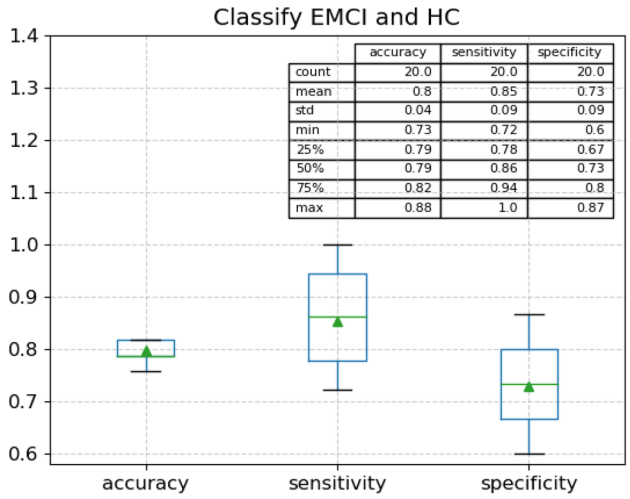
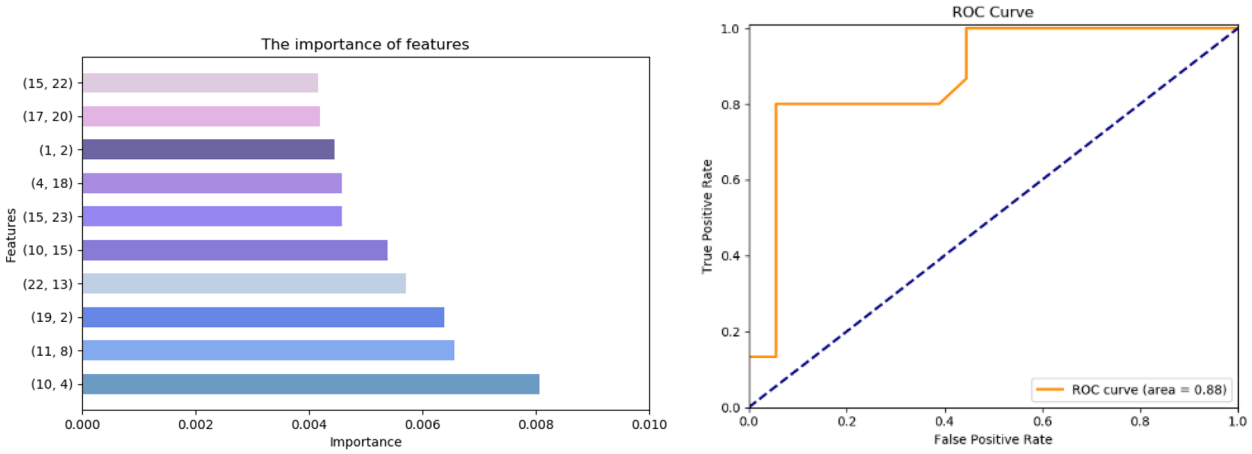


Figure 3. Classification performance. Demonstrate the classification performance after 20 repetitions. Box plot on classification accuracy, specificity, and sensitivity by GRU_GC model.



(a) Feature importance score

(b) ROC curve

Figure 4. (a) Top 10 important features in the importance ranking of 36 feature attributes. The horizontal axis represents the importance of features, and the vertical axis represents the EC between brain regions, such as (10, 4) means the feature from No.10 to No.4 in the EC matrices. (b) ROC curve of classification with 36 feature attributes using random forest.

Support vector machine

In addition to the random forest, there is also a support vector machine to classify the two groups of subjects. The EC matrices obtained from MVGC, LSTM_GC and GRU_GC model are also used as the input of SVM. The results of the classification are shown in Figure 5. It can be clearly observed from the figure that GRU_GC model (blue in the figure) has the highest classification accuracy, followed by LSTM_GC (orange) and MVGC (yellow). The largest value of blue is 71.88%, while orange and yellow are 65.63% and 68.75%, respectively. The higher accuracy of MVGC is stable at 62.5%. The blue line fluctuates around 0.6, while the yellow lines fluctuate around 0.5. The maximum classification accuracy of LSTM_GC and MVGC are less than that of the GRU-GC.

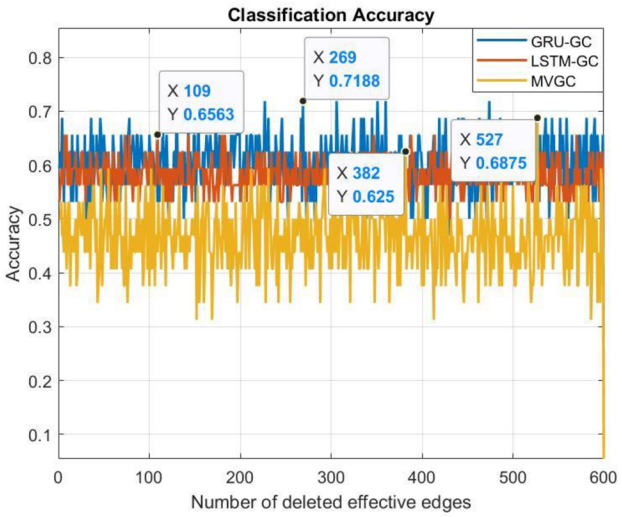


Figure 5. Classification accuracy. Compare the accuracy of classification using the three models: GRU_GC, LSTM_GC and MVGC. The blue, orange and yellow lines represent the classification accuracy of effective connectivity matrices using GRU-GC, LSTM_GC and MVGC, respectively.

Difference network

The relationship between each pair of brain regions in the causality matrix of HC and EMCI was examined to discover the connectivity with significant differences. Because of 162 subjects in the experiment, the causality between each pair of brain regions contained data from 162 samples. The two sample T-test was performed on the brain regions corresponding to EMCI and HC, and $\alpha = 0.05$ was the standard. Calculated the DN with 21 causalities in significant differences in the final. Figure 6(a) represents the DN in which EMCI has significantly enhanced causality compared to HC, including 18 differences in brain connectivity. Figure 6(b) indicates the DN where the causality of EMCI is significantly weaker than that of HC, including 3 connectivity between different brain regions.

In the significant DN of EMCI, six brain regions showed significant influence at the level of information inflow, including left supplementary motor area (19), right cuneus (46), left lingual gyrus (47), left superior occipital gyrus (49), right superior parietal gyrus (60). There are five brain regions including the left supplementary motor area, right cuneus, left lingual gyrus, and left superior occipital gyrus have increased significantly in the information inflow on the physiological activities of the brain in EMCI. The brain regions in noteworthy differences have no fully weakened EC in the information inflow level, such as some of the connectivity about the right superior parietal gyrus was enhanced while some of them were decreased. Fifteen brain regions with prominent changes at the information outflow level: left middle frontal gyrus (orbital part) (9), left supplementary motor area (19), left posterior cingulate gyrus (35), right parahippocampal gyrus (40), right calcarine fissure and surrounding cortex (44), left lingual gyrus (47), left superior occipital gyrus (49), right superior parietal gyrus (60) and right temporal pole: superior temporal gyrus (84) et al. The brain regions with significant differences in the level of

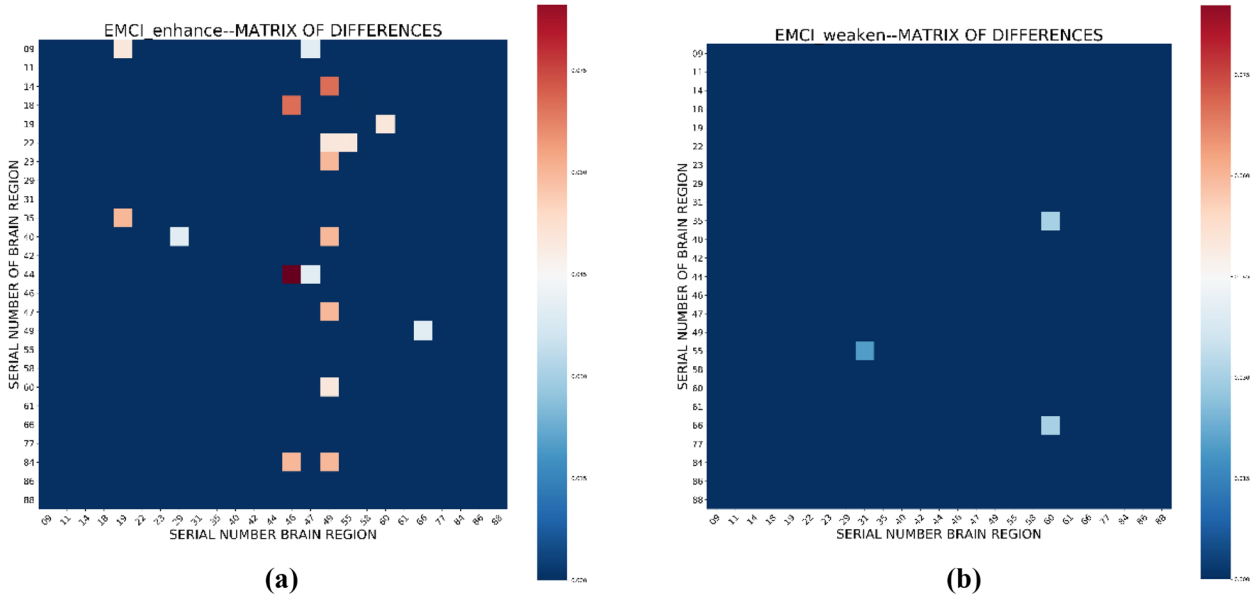


Figure 6. (a) and (b) show the difference networks of significantly enhanced and decreased brain connectivity in EMCI compared with HC, respectively.

information inflow and outflow include the left supplementary motor area (19), left lingual gyrus (47), left superior occipital gyrus (49), and right superior parietal gyrus (60).

To further explore the influence of EMCI on brain regions, the in-degree and out-degree of the brain regions corresponding to the selected effective edges were analyzed. The hub nodes in the information of inflow level and outflow level as shown in Table 3. After calculating the degree of each node in the DN, the hub nodes can be determined. If the in-degree of a node is greater than the mean of the in-degree of all nodes plus 1/4 variance (i.e., mean + 1/4*variance), then the node is a hub node. The hub nodes of out-degree are obtained in the same way. The standard used to measure whether the node is pivotal is to indicate the significance of the node change. Figure 7 displays the brain regions corresponding to hub nodes.

Features and difference network

The two sample T-test is performed on HC and EMCI and obtained the edges $E_T = 21$ with the obvious difference. The important features $E_F = 36$ were obtained when the accuracy of the above classification reached the highest. And then take the intersection of them to get the edges $E = E_T \cap E_F = 9$ with an obvious difference and important effect on classification, which are [84, 113, 233, 247, 254, 300, 392, 498, 541]. The positions of these features on the average matrices of EMCI and HC are shown in Figure 8. The 9 features in figure respectively represent the effective connectivity strength between ROI, including a total of 13 related brain regions as shown in Table 4. Figure 9 shows the distribution of brain regions that distinguish EMCI from HC.

Discussion

As a neurological disease, the number of patients with Alzheimer's disease has generally increased and attracted more and more attention (Alzheimer's Association, 2018; Mckhann, 2011). Since MCI is an intermediate state between HC and AD, compared with AD patients and HC patients, the neural network of MCI patients may be more delicate and converted to AD to a large extent, which makes research on EMCI extensive (Khazaee et al., 2017; Xiaobo et al., 2016; Yang et al., 2019). MCI can be divided into EMCI and LMCI. The subtle neural network relationship between EMCI and HC makes it challenging to classify HC and EMCI effectively and accurately. The abnormal brain activity of AD and MCI in 25 brain regions divided according to the AAL template has been proved by many scholars using various methods. This study selected these brain regions as ROIs. Simultaneously, Given the complexity of the human brain, the level of non-linearity and the related delays in rs-fMRI data, the GRU and Granger causality are combined and improved to find the more complex nonlinear system with delay correlation as much as possible. The GRU_GC method was applied to the classification of EMCI patients and HC. At the same time, the brain region connectivity between EMCI and HC was researched and analyzed to explore their differences.

The classification accuracy of EMCI and HC were compared to verify that the GC method added to GRU can play a positive role in the diagnosis of EMCI to a certain extent. As shown in Table 2, compared with MVGC and LSTM_GC, GRU_GC can improve the accuracy of classification, as well as the sensitivity and specificity. The results showed that GRU_GC was superior to MVGC and LSTM_GC in identifying patients with EMCI based on rs-fMRI data. It is vital for the identification of the EMCI stage and the subsequent development of the disease and failure to identify and classify correctly can lead to serious consequences, such as the delay of critical treatment periods. The ROC curve according to Figure 4 (b) shows that the GRU_GC has great performance for the classification of EMCI patients and normal subjects. Since GRU is a successful variant model that inherits most characteristics of RNN and LSTM models, it has the advantage of dealing with problems that are highly related to time series. At the same time, it solves the vanishing gradient problem caused by the gradual reduction in the gradient direction propagation process. GRU_GC can learn from different lengths of time delays and simulate the logical development of human behavior and the cognitive process of neural organization realistically. In addition, it does not rely on the independent variable regression model, so the time series no longer requires random and generalized stability. It can adapt to linear and non-linear nervous systems while reducing the problem of more potential variables. Besides, GRU-GC only requires a small set of parameters due to the parameters sharing structure of the recurrent neural network. In the process of constructing the EC matrices, GRU exerts its advantages in sequence modeling: passing information selectively, adding or removing information transmitted to the cell state in the time sequences, and long-term memory. The ability of GRU_GC to more accurately describe the flow of information between brain regions. The gate mechanism not only with a positive effect on information selection but also can solve gradient explosion and gradient vanishing to a certain extent problem. This kind of globalization of time series data can improve the localization in input problems that may lose information and cause the model to be less accurate. In summary, the method utilizes the powerful self-learning ability of neural networks to fit and predict time series. The gate mechanism in GRU plays an important role in controlling global

Table 3. The in-degree and out-degree of the DN and found in the corresponding level of hub nodes. The red represents pivotal nodes between in-degree and out-degree.

Brain regions		9	11	14	18	19	22	23	29	31	35	40	42	44	46	47	49	55	58	60	61	66	77	84	86	88
DN	In	0	0	1	0	2	0	0	1	1	0	0	0	0	2	2	7	1	0	3	0	1	0	0	0	0
	Out	2	0	1	1	1	2	1	0	0	2	2	0	2	0	1	1	1	0	1	0	1	0	2	0	0
In-degree: 19,46,47,49,60															Out-degree: 9,14,18,19,22,23,35,40,44,47,49,55,60,66,84											

information in the model. The reset gate of GRU resets the transmission of the previous stage state, and the update gate forgets and selects the updated data. Using two gate structures achieves a better effect than LSTM and dependence on hardware computing power. The GRU-based GC method can effectively characterize the information flow between brain regions, which can play a role in the diagnosis of EMCI.

To make the conclusion more convincing, the SVM classifier was also used for classifying HC and EMCI with three methods of constructing the brain effective connectivity

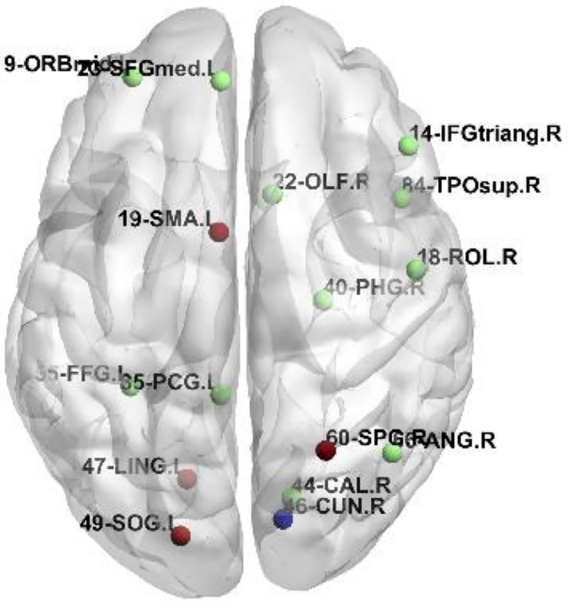
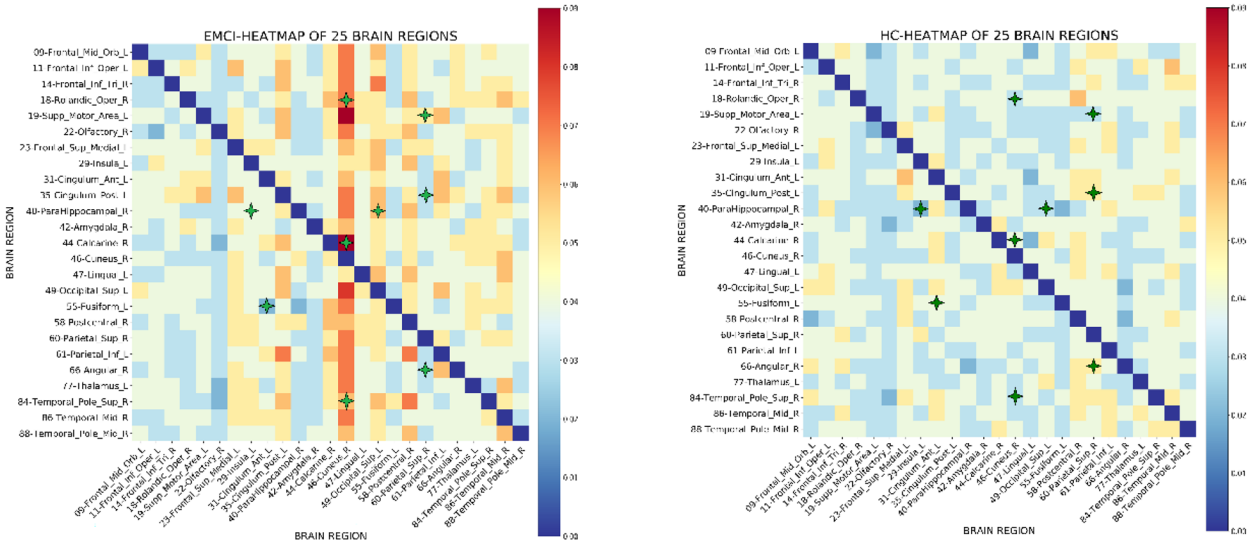


Figure 7. There are 16 brain nodes in the figure. The blue, green and red nodes represent the brain nodes with significant differences in the information inflow level, information outflow level and the two levels of the EMCI difference network, respectively.

besides the RF classifier. Two methods were used to eliminate the influence of classifiers. Combined with the classification results in Table 2 and Figure 5, the classification accuracy of both the RF classifier and the SVM classifier has remarkably improved when the results of GRU_GC were used. It indicated that the effect of GRU_GC is better than LSTM_GC and MVGC, which reveals that the Granger causality analysis method with GRU has played a positive role in the diagnosis of EMCI. Feature selection is necessary because feature redundancy will affect the accuracy of the classification and increase time consumption. In this study, we determined the features according to the importance score calculated by the tree-based features selection. It can be seen from Figure 4 that the accuracy of the 36 features by feature selection can reach 87.88%, which shows that the method of LSTM_GC can effectively and accurately distinguish between EMCI and HC.

The DN constructed by the two sample T-test on the causality between brain regions shows. 21 connectivity in the brains of EMCI and HC have noteworthy differences (Figure 6). 22 causalities between EMCI brain regions were obviously enhanced (Figure 6(a)) and the remaining 3 connectivity were significantly decreased (Figure 6(b)). According to the DN, the brain region nodes in the brain network were found (Figure 7), temporal pole: superior temporal gyrus, and superior parietal gyrus respectively belong to the temporal lobe and the parietal lobe. The frontal lobe includes the middle frontal gyrus (orbital part), supplementary motor area and olfactory cortex, the limbic lobe contains the parahippocampal gyrus and posterior cingulate gyrus. The calcarine fissure and surrounding cortex, cuneus, lingual gyrus and superior occipital gyrus belong to the occipital lobe.

In order to further explore the brain function differences between EMCI and HC, the intersection of the 36 important features and 21 significantly different causality was conducted through two sample T-tests to obtain 9 important causality



(a) **(b)**

Figure 8. Both axes are 25 brain regions, (a) and (b) represent the average effective connectivity matrices of 89 EMCI subjects and 73 HC subjects, respectively, the green stars in the figure indicate features that are significantly different and play an important role in classification.

relationships. The map of EC about two groups is shown in Figure 8, it can be seen from the figure that most of the connectivity of EMCI were stronger than HC while only four connectivity of EMCI were weaker than HC. Figure 9 shows the 9 connectivity between different brain regions. It is seen that the great part of the enhanced connectivity occurs in or is associated with the right brain, the connectivity includes from the right calcarine fissure and surrounding cortex to the right cuneus, from the right parahippocampal gyrus to the left superior occipital gyrus, from right parahippocampal gyrus to left insula, from left supplementary motor area to right superior parietal gyrus, from right temporal pole: superior temporal gyrus to right cuneus and from right Rolandic operculum to the right cuneus. Compared with the HC group, the connectivity of EMCI was generally stronger, especially in the right brain and the most remarkable changes were related to connectivity in the parahippocampal gyrus. In these connectivities, the causality related to the parahippocampal gyrus showed a trend of enhancement. The parahippocampal gyrus is involved in associative learning and episodic memory (Maddock et al., 2001). In EMCI, EC enhancement may be due to the presence of signal transmission compensation problems in brain regions. As the activity of some brain regions is reduced, the signal transmission between other brain regions is strengthened. As an early stage of MCI, EMCI has similarities with AD, and the connection between the parahippocampal gyrus has significant changes. As shown in Table 4, the Rolandic operculum and supplementary motor area belong to the frontal cortex. Language, numeracy skills (Kimberg & Farah, 1993) and decision-making (Yang et al., 2017) which are all controlled by the frontal lobe. Dopamine neurons are mostly located in the midbrain, and midbrain cortical projections affect cognitive functions during planning, short-term memory tasks, execution, and attention mechanisms of frontal cortical activity. The superior temporal gyrus belongs to the temporal pole, which is mainly related to memories and emotions.

The connectivity of EMCI from left posterior cingulate gyrus to right superior parietal gyrus, from left fusiform gyrus to left anterior cingulate and paracingulate gyri and from right angular gyrus to right superior parietal gyrus to left fusiform gyrus are weakened. The fusiform gyrus (55) plays an important role in advanced visual processing and recognition and its functions may include processing color information and recognizing and classifying the face, body and font. The posterior cingulate gyrus (PCC) is the cortical part of the limbic system. Structural and functional abnormalities in the PCC result in a range of neurological and psychiatric disorders, including Alzheimer's disease (Scheff et al., 2015), autism, hyperactivity disorder, major depression, traumatic brain injury, and anxiety disorders

Table 4. The numbers behind each brain region are the serial numbers of the corresponding AAL template. The first column contains the number of ROIs in the brackets.

	Brain regions
Frontal lobe (2)	Rolandic operculum (18), Supplementary motor area (19)
Temporal lobe (2)	Fusiform gyrus (55), Temporal pole: superior temporal gyrus(84)
Parietal lobe (2)	Superior parietal gyrus (60)
Occipital lobe (3)	Calcarine fissure and surrounding cortex (44), Cuneus (46), Superior occipital gyrus (49)
Limbic lobe (3)	Anterior cingulate and paracingulate gyri (31), Posterior cingulate gyrus (35), Parahippocampal gyrus (40)
Else (2)	Insula (29), Angular gyrus (66)

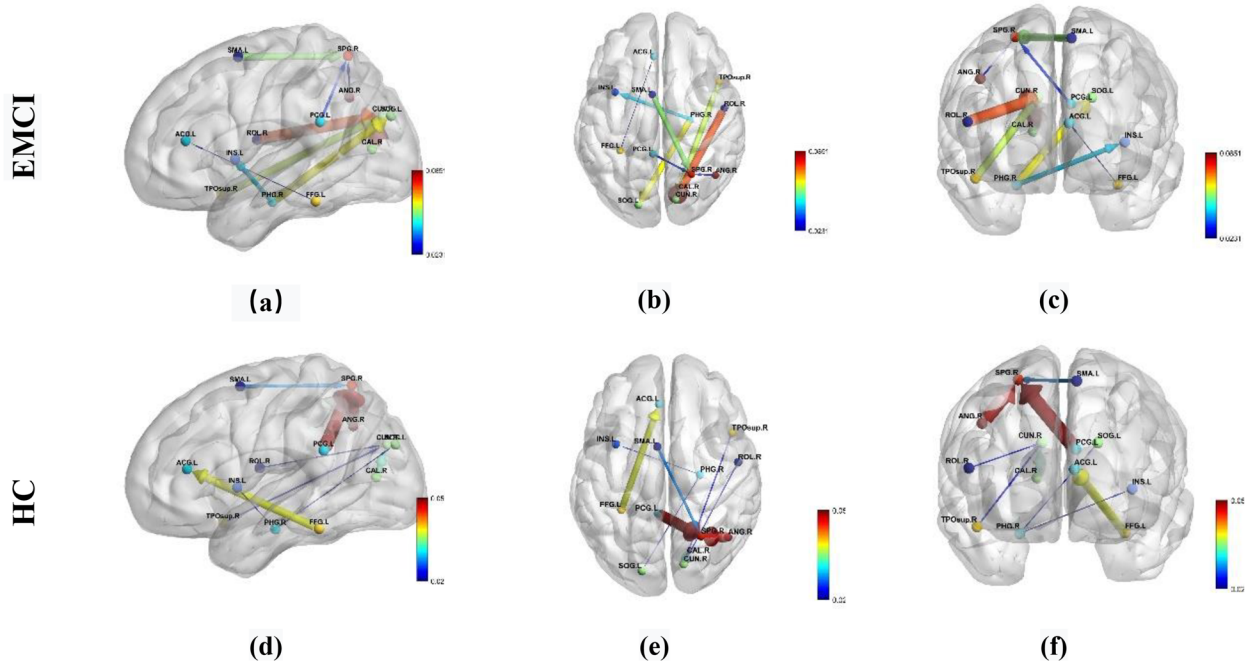


Figure 9. The figure shows the strength of 9 effective connectivity between different brain regions. The nodes in the figure represent different brain regions. The edges between brain regions represent effective connectivity and the arrows on the edges represent from cause to effect. The thickness and color of the line represent the strength of the connectivity.

(Leech & Sharp, 2013; Oblak et al., 2011). The PCC likely integrates and mediates information in the brain. Therefore, functional abnormalities of the PCC might be an accumulation of remote and widespread damage in the brain. The parahippocampal gyrus and posterior cingulate gyrus are all part of the limbic system. The connectivity between the three brain regions showed a trend of weakening and change obviously, which indicated that EMCI might have changes of significantly weakening of EC in the limbic system. From the comparison between EMCI and HC in Figure 9, it can be directly observed that the lines of HC are generally thin and cool in color, which tends to be blue. While compared with HC, most of the connection lines in the EMCI group are generally thick and bright colors, the color tending to yellow or red. By comparison, it was found that the difference in lines between the two groups of EMCI and HC in Figure 9. The lines intuitively reflected the noteworthy difference in the effective connectivity between the 25 brain regions, indicating that the EMCI patients may have changes in these regions. In this work, we examined the effectiveness of the GRU_GC model in diagnosing EMCI patients. The results indicated that the GRU_GC model can better investigate the EC of EMCI as compared to other models (e.g., LSTM_GC and MVGC) as shown in Table 2. We hope that our method can be applied to more fields of psychiatric disorders. In the comparison experiments, we have not yet validated more effective connectivity methods, such as dynamic causal modeling, which can be further validated in the follow-up work. In addition, we introduced a recurrent neural network in the Granger causality model to obtain the relationship between brain regions, and the effective connectivity method and model framework based on GC can be further optimized in the future. We can try to introduce newer structures or neural networks such as graph neural networks to improve the identification of effective connectivity between brain regions.

Conclusion

EMCI and HC can be effectively classified by constructing effective connectivity for the subjects with GRU_GC. The GRU_GC model can acquire non-linear connectivity with different time delays in the signal because of the powerful time series modeling capabilities in the GRU model. Compared with LSTM_GC and MVGC, this model shows its significant superiority. Furthermore, according to EC matrices, the difference network of causality between the two groups of subjects is constructed. The results showed that there were noteworthy differences between EMCI and HC in parahippocampal gyrus and posterior cingulate gyrus. In conclusion, it shows that GRU_GC can play an important role in classification and brain region diagnostic analysis. The model can be used as a practical significant diagnostic method for mild cognitive impairment diseases.Notes

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the National Natural Science Foundation of China (Grant No. 31870979), National Natural Science Foundation of China (Grant No. 61906117), Pudong New Area Science and Technology Development Fund (Grant No. PKJ2014-Y08), Shanghai Sailing Program (Grant No. 19YF1419000).

References

Alzheimer's Association. (2018). 2018 Alzheimer's disease facts and figures. *Alzheimers & Dementia*, *14*(3), 367–429.

Bressler, S. L., & Seth, A. K. (2011). Wiener–Granger causality: A well established methodology. *NeuroImage*, *58*(2), 323–329. https://doi.org/10.1016/j.neuroimage.2010.02.059

Cho, K., Merrienboer, B. V., Gulcehre, C., et al. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*(pp. 1724–1734).

Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. In *NIPS 2014 Workshop on Deep Learning, December 2014*.

Deng, J., Kong, W., Mou, X., & Wang, S. (2018). Pathway crosstalk analysis based on signaling pathway impact analysis in Alzheimer's disease. *Current Proteomics*, *15*(2), 142–150. https://doi.org/10.2174/1570164614666171030162949

Deshpande, G., Laconte, S., James, G. A., Peltier, S., & Hu, X. (2010). Multivariate granger causality analysis of fMRI data. *Human Brain Mapping*, *30*(4), 1361–1373. https://doi.org/10.1002/hbm.20606

Dimitriadis, S. I., Laskaris, N. A., Tsirka, V., Vourkas, M., & Micheloyannis, S. (2012). An EEG study of brain connectivity dynamics at the resting state. *Nonlinear Dynamics Psychology & Life Sciences*, *16*(1), 5.

Farokhzadi, M., Hossein-Zadeh, G.-A., & Soltanian-Zadeh, H. (2018). Nonlinear effective connectivity measure based on adaptive neuro fuzzy inference system and Granger causality. *NeuroImage*, *181*, 382–394. https://doi.org/10.1016/j.neuroimage.2018.07.024

Fawcett, T. (2005). An introduction to ROC analysis. *Pattern Recognition Letters*, *27*(8), 861–874. https://doi.org/10.1016/j.patrec.2005.10.010

Friston, K. J. (1995). Functional and effective connectivity in neuroimaging: A synthesis. *Human Brain Mapping*, *2*(1–2), 56–78. https://doi.org/10.1002/hbm.460020107

Friston, K. J., Holmes, A. P., Poline, J. B., Grasby, P. J., Williams, S. C., Frackowiak, R. S., & Turner, R. (1995). Analysis of fMRI time-series revisited. *Neuroimage*, *2*(1), 45–53. https://doi.org/10.1006/nimg.1995.1007

Friston, K., Moran, R., & Seth, A. K. (2013). Analysing connectivity with granger causality and dynamic causal modeling. *Current Opinion in Neurobiology*, *23*(2), 172–178. https://doi.org/10.1016/j.conb.2012.11.010

Goebel, R., Roebroeck, A., Kim, D.-S., & Formisano, E. (2003). Investigating directed cortical interactions in time-resolved fMRI data using vector autoregressive modeling and Granger causality mapping. *Magnetic Resonance Imaging*, *21*(10), 1251–1261. https://doi.org/10.1016/j.mri.2003.08.026

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, *37*(3), 424. https://doi.org/10.2307/1912791

Graves, A. (2012). Supervised sequence labeling with recurrent neural networks. *Studies in Computational Intelligence* (Vol. 385). Springer.

Greff, K. (2016). LSTM: A search space odyssey. *IEEE transactions on neural networks and learning systems*, *28*(10), 2222–2232.

Group, Nlc. (2017). R-NET: Machine reading comprehension with self-matching networks.

Guo, S., Seth, A. K., Kendrick, K. M., Zhou, C., & Feng, J. (2008). Partial Granger causality—Eliminating exogenous inputs and latent variables. *Journal of Neuroscience Methods*, *172*(1), 79–93. https://doi.org/10.1016/j.jneumeth.2008.04.011

Guo, H., Zeng, W., Shi, Y., Deng, J., & Zhao, L. (2020). Kernel Granger causality based on back propagation neural network fuzzy inference

system on fMRI data. *IEEE Transactions on Neural Systems and Rehabilitation Engineering: A Publication of the IEEE Engineering in Medicine and Biology Society*, *28*(5), 1049–1058. https://doi.org/10.1109/TNSRE.2020.2984519

Guo, H., Zhang, F., Chen, J., Xu, Y., & Xiang, J. (2017). Machine learning classification combining multiple features of a hyper-network of fMRI data in Alzheimer's disease. *Frontiers in Neuroscience*, *11*, 615. https://doi.org/10.3389/fnins.2017.00615

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Jitsuishi, T., & Yamaguchi, A. (2022). Searching for optimal machine learning model to classify mild cognitive impairment (MCI) subtypes using multimodal MRI data. *Scientific Reports*, *12*(1), 4284. https://doi.org/10.1038/s41598-022-08231-y

Khadem, A., & Hossein-Zadeh, G.-A. (2014). Estimation of direct nonlinear effective connectivity using information theory and multilayer perceptron. *Journal of Neuroscience Methods*, *229*, 53–67. https://doi.org/10.1016/j.jneumeth.2014.04.008

Khazaee, A., Ebrahimzadeh, A., & Babajani-Feremi, A, Alzheimer's Disease Neuroimaging Initiative. (2017). Classification of patients with MCI and AD from healthy controls using directed graph measures of resting-state fMRI. *Behavioural Brain Research*, *322*(Pt B), 339–350. https://doi.org/10.1016/j.bbr.2016.06.043

Kimberg, D. Y., & Farah, M. J. (1993). A unified account of cognitive impairments following frontal lobe damage: the role of working memory in complex, organized behavior. *Journal of Experimental Psychology. General*, *122*(4), 411–428. https://doi.org/10.1037/0096-3445.122.4.411

Leech, R., & Sharp, D. J. (2013). The role of the posterior cingulate cortex in cognition and disease. *Brain*, *137*(1), 12–32. https://doi.org/10.1093/brain/awt162

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature*, *412*(6843), 150–157. https://doi.org/10.1038/35084005

Maddock, R. J., Garrett, A. S., & Buonocore, M. H. (2001). Remembering familiar people: The posterior cingulate cortex and autobiographical memory retrieval. *Neuroscience*, *104*(3), 667–676. https://doi.org/10.1016/s0306-4522(01)00108-7

Mckhann, G. (2011). The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's Dement*, *7*(3), 263–269.

Montalto, A., Stramaglia, S., Faes, L., Tessitore, G., Prevete, R., & Marinazzo, D. (2015). Neural networks with non-uniform embedding and explicit validation phase to assess granger causality. *Neural Networks: The Official Journal of the International Neural Network Society*, *71*, 159–171. https://doi.org/10.1016/j.neunet.2015.08.003

Oblak, A. L., Gibbs, T. T., & Blatt, G. J. (2011). Reduced GABAA receptors and benzodiazepine binding sites in the posterior cingulate cortex and fusiform gyrus in autism. *Brain Research*, *1380*, 218–228. https://doi.org/10.1016/j.brainres.2010.09.021

Pereda, E., Quiroga, R. Q., & Bhattacharya, J. (2005). Nonlinear multivariate analysis of neurophysiological signals. *Progress in Neurobiology*, *77*(1–2), 1–37. https://doi.org/10.1016/j.pneurobio.2005.10.003

Porta, A., & Faes, L. (2016). Wiener–Granger causality in network physiology with applications to cardiovascular control and neuroscience. *Proceedings of the IEEE*, *104*(2), 282–309. https://doi.org/10.1109/JPROC.2015.2476824

Riley, J. D., Chen, E. E., Winsell, J., Davis, E. P., Glynn, L. M., Baram, T. Z., Sandman, C. A., Small, S. L., & Solodkin, A. (2018). Network specialization during adolescence: Hippocampal effective connectivity in boys and girls. *NeuroImage*, *175*, 402–412. https://doi.org/10.1016/j.neuroimage.2018.04.013

Roebroeck, A., Formisano, E., & Goebel, R. (2005). Mapping directed influence over the brain using Granger causality and fMRI. *NeuroImage*, *25*(1), 230–242. https://doi.org/10.1016/j.neuroimage.2004.11.017

Rosoł, M., Młyńczak, M., & Cybulski, G. (2022). Granger causality test with nonlinear neural-network-based methods: Python package and simulation study. *Computer Methods and Programs in Biomedicine*, *216*, 106669. https://doi.org/10.1016/j.cmpb.2022.106669

Scheff, S. W., Price, D. A., Ansari, M. A., Roberts, K. N., Schmitt, F. A., Ikonomovic, M. D., & Mufson, E. J. (2015). Synaptic change in the posterior cingulate gyrus in the progression of Alzheimer's disease. *Journal of Alzheimer's Disease: JAD*, *43*(3), 1073–1090. https://doi.org/10.3233/JAD-141518

Schlösser, R. G. M., Wagner, G., Koch, K., Dahnke, R., Reichenbach, J. R., & Sauer, H. (2008). Fronto-cingulate effective connectivity in major depression: A study with fMRI and dynamic causal modeling. *NeuroImage*, *43*(3), 645–655. https://doi.org/10.1016/j.neuroimage.2008.08.002

Seth, A. K., Chorley, P., & Barnett, L. C. (2013). Granger causality analysis of fMRI BOLD signals is invariant to hemodynamic convolution but not downsampling. *NeuroImage*, *65*, 540–555. https://doi.org/10.1016/j.neuroimage.2012.09.049

Shi, Y. H., Zeng, W. M., Nie, W. F., & Yang, J. J. (2020). Multi-channel hierarchy functional integration analysis between large-scale brain networks for migraine: An fMRI study. *NeuroImage: Clinical*, *28*, 102462. https://doi.org/10.1016/j.nicl.2020.102462

Shi, Y., Zeng, W., Wang, N., & Chen, D. (2015). A novel FMRI group data analysis method based on data-driven reference extracting from group subjects. *Computer Methods and Programs in Biomedicine*, *122*(3), 362–371. https://doi.org/10.1016/j.cmpb.2015.09.002

Shi, Y., Zeng, W., Wang, N., & Zhao, L. (2018). A new constrained spatiotemporal ICA method based on multi-objective optimization for fMRI data analysis. *IEEE Transactions on Neural Systems and Rehabilitation Engineering: a Publication of the IEEE Engineering in Medicine and Biology Society*, *26*(9), 1690–1699. https://doi.org/10.1109/TNSRE.2018.2857501

Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to sequence learning with neural networks. *Advances in Neural Information Processing Systems*, *4*, 3104–3112.

Wang, Y., Lin, K., Qi, Y., Lian, Q., Feng, S., Wu, Z., & Pan, G. (2018). Estimating brain connectivity with varying-length time lags using a recurrent neural network. *IEEE Transactions on Bio-Medical Engineering*, *65*(9), 1953–1963. https://doi.org/10.1109/TBME.2018.2842769

Chen, X., Zhang, H., Gao, Y., Wee, C. Y., Li, G., Shen, D., & Alzheimer's Disease Neuroimaging Initiative. (2016). High-order resting-state functional connectivity network for MCI classification. *Human Brain Mapping*, *37*(9), 3282–3296.

Xin, D., & Biswal, B. B. (2014). Identifying the default mode network structure using dynamic causal modeling on resting-state functional magnetic resonance imaging. *NeuroImage*, *86*(2), 53–59. https://doi.org/10.1016/j.neuroimage.2013.07.071

Yan, C. C., & Zang, Y. F. (2010). DPARSF: A MATLAB toolbox for "pipeline" data analysis of resting-state fMRI. *Frontiers in Systems Neuroscience*, *4*(13), 13.

Yang, P., Zhou, F., Ni, D., Xu, Y., Chen, S., Wang, T., & Lei, B. (2019). Fused sparse network learning for longitudinal analysis of mild cognitive impairment. *IEEE Transactions on Cybernetics*, *51*(1), 233–246..

Yang, X., Gao, M., Shi, J., Ye, H., & Chen, S. (2017). Modulating the activity of the DLPFC and OFC has distinct effects on risk and ambiguity decision-making: A tDCS study. *Frontiers in Psychology*, *8*, 1417. https://doi.org/10.3389/fpsyg.2017.01417

Zhao, L., Zeng, W., Shi, Y., Nie, W., & Yang, J. (2020). Dynamic visual cortical connectivity analysis based on functional magnetic resonance imaging. *Brain and Behavior*, *10*(7), e01698. https://doi.org/10.1002/brb3.1698