A Deep Network Model on Dynamic Functional Connectivity With Applications to Gender Classification and Intelligence Prediction

Liangwei Fan, Jianpo Su, Jian Qin, Dewen Hu and Hui Shen*

College of Intelligence Science and Technology, National University of Defense Technology, Changsha, China

Increasing evidence has suggested that the dynamic properties of functional brain networks are related to individual behaviors and cognition traits. However, current fMRI-based approaches mostly focus on statistical characteristics of the windowed correlation time course, potentially overlooking subtle time-varying patterns in dynamic functional connectivity (dFC). Here, we proposed the use of an end-to-end deep learning model that combines the convolutional neural network (CNN) and long short-term memory (LSTM) network to capture temporal and spatial features of functional connectivity sequences simultaneously. The results on a large cohort (Human Connectome Project, $n = 1,050$) demonstrated that our model could achieve a high classification accuracy of about 93% in a gender classification task and prediction accuracies of 0.31 and 0.49 (Pearson's correlation coefficient) in fluid and crystallized intelligence prediction tasks, significantly outperforming previously reported models. Furthermore, we demonstrated that our model could effectively learn spatiotemporal dynamics underlying dFC with high statistical significance based on the null hypothesis estimated using surrogate data. Overall, this study suggests the advantages of a deep learning model in making full use of dynamic information in resting-state functional connectivity, and highlights the potential of time-varying connectivity patterns in improving the prediction of individualized characterization of demographics and cognition traits.

Keywords: dynamic functional connectivity (dFC), deep learning, gender classification, intelligence prediction, resting-state functional magnetic resonance imaging

INTRODUCTION

Functional organization principles in the human brain derived from resting-state functional MRI (rs-fMRI) data have been found to improve our understanding of individual cognition and behavioral differences greatly. Functional connectivity (FC) analysis based on rs-fMRI is often applied to quantify the statistical dependencies across different brain regions which are correlated over time (Friston, 2011). Initially, FC was assessed with the assumption that the connections remain unchanged at rest (Hutchison et al., 2013). Moreover, recent studies have found that FC is extremely useful in gender classification (Zhang et al., 2018) and individual prediction of cognition traits such as fluid intelligence (Finn et al., 2015).

However, increasing evidence has suggested that the statistical properties of FC change over different time scales across task states (Elton and Gao, 2015; Fatima et al., 2016) and during periods of unconstrained rest (Chang and Glover, 2010; Allen et al., 2014), i.e., dynamic functional connectivity (dFC). So far, spatiotemporal patterns and dynamics of functional networks derived from rs-fMRI data have been widely studied. The evidence suggested that dynamic interactions of different functional networks were relate to specific tasks (Yuan et al., 2018a; Li et al., 2019) and brain disorders such as Alzheimer's Disease (Huang et al., 2017), which motivated us to explore the spatiotemporal characteristics of functional networks based on dFC. Moreover, dFC (Sakoğlu et al., 2010) has been successfully applied to characterizing neuropsychiatric disorders like schizophrenia (Damaraju et al., 2014; Du et al., 2017), autism spectrum disorder (Zhu et al., 2016) and depression (Demirtaş et al., 2016). Importantly, dFC is linked to individual characteristics such as cognitive flexibility (Douw et al., 2016), emotions (Tobia et al., 2017), and age (Qin et al., 2015; Davison et al., 2016), and even the variability of dFC can be used to predict high-level cognitive performance (Liu et al., 2018).

Although these studies discriminated individual cognitive ability or demographic characteristics based on dFC, most of them employed statistical characteristics of the original dynamic correlation series as features to achieve classification or prediction. Nevertheless, these manually selected features neglected the time-varying patterns in correlation timecourse of FC through folding temporal dimension of dFC, likely causing the loss of some useful information related to individual cognition traits. Actually, the temporal evolution of connectome-scale brain network interactions has been observed to fit well the task-fMRI data (Yuan et al., 2018a), suggesting the neurophysiological significance of spatiotemporal structures in functional network dynamics. Some recent studies also demonstrated the potential of temporal sequence of windowed functional correlations in predicting individual cognitive traits and demographics. For example, the time series of dFC has long-range sequential correlations that vary across the human adult lifespan (Battaglia et al., 2020) and specific temporal structures of several FC microstates have been reported to be subject-specific and heritable, and significantly linked to individual cognitive traits (Vidaurre et al., 2017). Moreover, network switching in dFC is related to task performance and sleep (Pedersen et al., 2018), attention (Madhyastha et al., 2015), schizophrenia (Damaraju et al., 2014), and depression (Zheng et al., 2017). In particular, a growing body of researches links observed patterns of non-stationary switching between FC states with aspects of the underlying neural dynamics (Hansen et al., 2015; Thompson, 2018), indicating short-term alteration in FC time series along with shifting in cognitive states. However, given that the mechanism underlying spontaneous fluctuation in resting-state dFC has not been fully understood, explicitly modeling the sequence of time-resolved FC is still a challenging task.

Currently, deep learning has undergone unprecedented development in neuroscience. The advantage of an end-to-end model like deep learning is automatically extracting abstract spatial-temporal structures from neuroimaging data (Huang et al., 2017; Li et al., 2019), which has been used for discriminating complex mental disorders (Yan et al., 2017; Zeng et al., 2018) and gender classification (Yuan et al., 2018b). In particular, RNN models, such as long short-term memory (LSTM), can effectively extract complex and non-linear time-varying patterns underlying signals' fluctuations in a data-driven way due to its advantages of "deep in time," and have been successfully used in identifying autism (Dvornek et al., 2017) and discriminating schizophrenia (Yan et al., 2019) from healthy controls.

In this work, we ask to what extent the dFC can be used to predict individual cognition traits including gender classification and individual prediction of fluid and crystallized intelligence, using a deep network model on the dFC time course. We proposed the use of a convolutional neural network (CNN) and LSTM framework to identify the spatial and temporal structures underlying the spontaneous fluctuation of dFC. We assumed this framework directly working on the temporal series of dFC, could avoid information reduction and make full use of high-level spatiotemporal information of dFC. This model consists of two parts. The first part includes a multi-scale 1D convolutional layer, concatenation layer, and max-pooling layer, which were designed for spatial feature extraction. The second part involves two stacked LSTM layers, which were used to detect temporal dynamics (by learning short-term sequential switching and unknown long-term/non-linear patterns) in time series of windowed correlations. Finally, the average outputs of the LSTM layer were put as the input for the final fully connected layer. We further assessed the capacity of the CNN + LSTM model on one gender classification task and two prediction tasks of individual intelligence, with comparison to support vector machine (SVM) and support vector regression (SVR) models that usually use statistical characteristics of dFC as features (Liu et al., 2018). Furthermore, we conducted deconvolutional computation to visualize the learned features corresponding to task-related connections in three tasks for demonstrating the validity of the network in learning dynamics of time-resolved FC.

MATERIALS AND METHODS

Participants and Data Acquisition

The dataset was selected from the publicly available S1200 release of the Human Connectome Project dataset (HCP), including 1,206 subjects (age 22–35). The rs-fMRI data of all subjects was scanned in two sessions on two different days. Each session contains a right-left (RL) phase-encoding run and a left-right (LR) phase-encoding run. In addition, the HCP collected many behavioral measures, such as fluid intelligence and crystallized intelligence, which allows us to investigate the relationship between individual traits and their neuroimaging data. 1,050 subjects (gender: 569 females and 481 males) were left for this study by restricting subjects with at least one run and two intelligence measures.

Data was acquired using a 3.0 T Siemens scanner at Washington University at St. Louis. The data acquisition parameters were as follows: repetition time (TR) = 720 ms, time echo (TE) = 33.1 ms, flip angle (FA) = 52°, resolution = 2.0 mm,

field of view (FOV) = 208 × 180 mm (RO × PE), matrix size = 104 × 90 (RO × PE), slices = 72, and volumes = 1,200. More details are available in the previous literature (Van Essen et al., 2013).

Data Preprocessing

The HCP dataset was first preprocessed under the HCP minimal preprocessing pipeline, which mainly includes distortion and spatial artifact removal, motion correction, within-subject cross-modal registration, and cross-subject registration to a standard space (Glasser et al., 2013). In addition, standard preprocessing procedures for resting-state connectivity analysis were performed on the HCP using SPM8[1]. For each subject, the fMRI images were resampled to 3 × 3 × 3 mm isotropic voxels. Next, a Gaussian filter kernel of 6 mm full width at half maximum (FWHM) was used to smooth the images spatially. Then, the images were temporally bandpass filtered from 0.01 to 0.08 Hz. Finally, to further denoise rs-fMRI data, we regressed the white matter (WM) signal, head motion, cerebrospinal fluid (CSF) signal, and their first-order deviations.

Dynamic Functional Connectivity

To obtain regions of interest (ROI) based bold signals, we averaged preprocessed rs-fMRI time courses of voxels within each gray matter region according to a 268-node functional atlas (Shen et al., 2013). Then, a sliding time window approach (Calhoun et al., 2014; Qin et al., 2015) was used to divide the ROI-based brain signals into temporal segments with a window size of 39.6 s (55 TRs), and the dFC of each region pair was calculated using Pearson correlation coefficients. As a result, a series of 268 × 268 correlation matrices were generated (**Figure 1A**). Additionally, to normalize the coefficient values of correlation matrices, we applied Fisher's z-transformation to each correlation matrix. Next, the upper triangular part of each correlation matrix was reshaped into a vector for the following analysis. Ultimately, for each run of each subject, we obtained a sliding-window correlation matrix of 230 (windows) × 35,778 (connections).

dFC-Based Prediction Setup

Gender and behavioral data of intelligence on high-level cognition from HCP protocol were selected, including fluid intelligence (Penn Progressive Matrices, HCP: PMAT24_A_CR) and crystallized intelligence which was measured using NIH toolbox composite scores (Crystallized Cognition Composite), combining Picture Vocabulary Test and Oral Reading Recognition Test into one score (CompGcScore) (Akshoomoff et al., 2013). Next, the gender and intelligence measures of subjects were predicted using the CNN + LSTM model based on dFC matrices.

CNN + LSTM Model

As shown in **Figure 1B**, the used CNN + LSTM network is composed of three 1D convolutional layers which have three filters of different size (4, 8, and 16 time windows), one

concatenation layer combining features from three convolutional layers, one max-pooling layer which was designed to down-sample, two LSTM network layers, and a fully connected layer (Yan et al., 2019). The model was built for both the gender classification task and cognitive performance prediction tasks. Specifically, the obtained dFC matrices were fed into the CNN + LSTM model as inputs for parameter optimization. Then, the optimized model was saved for testing and comparison.
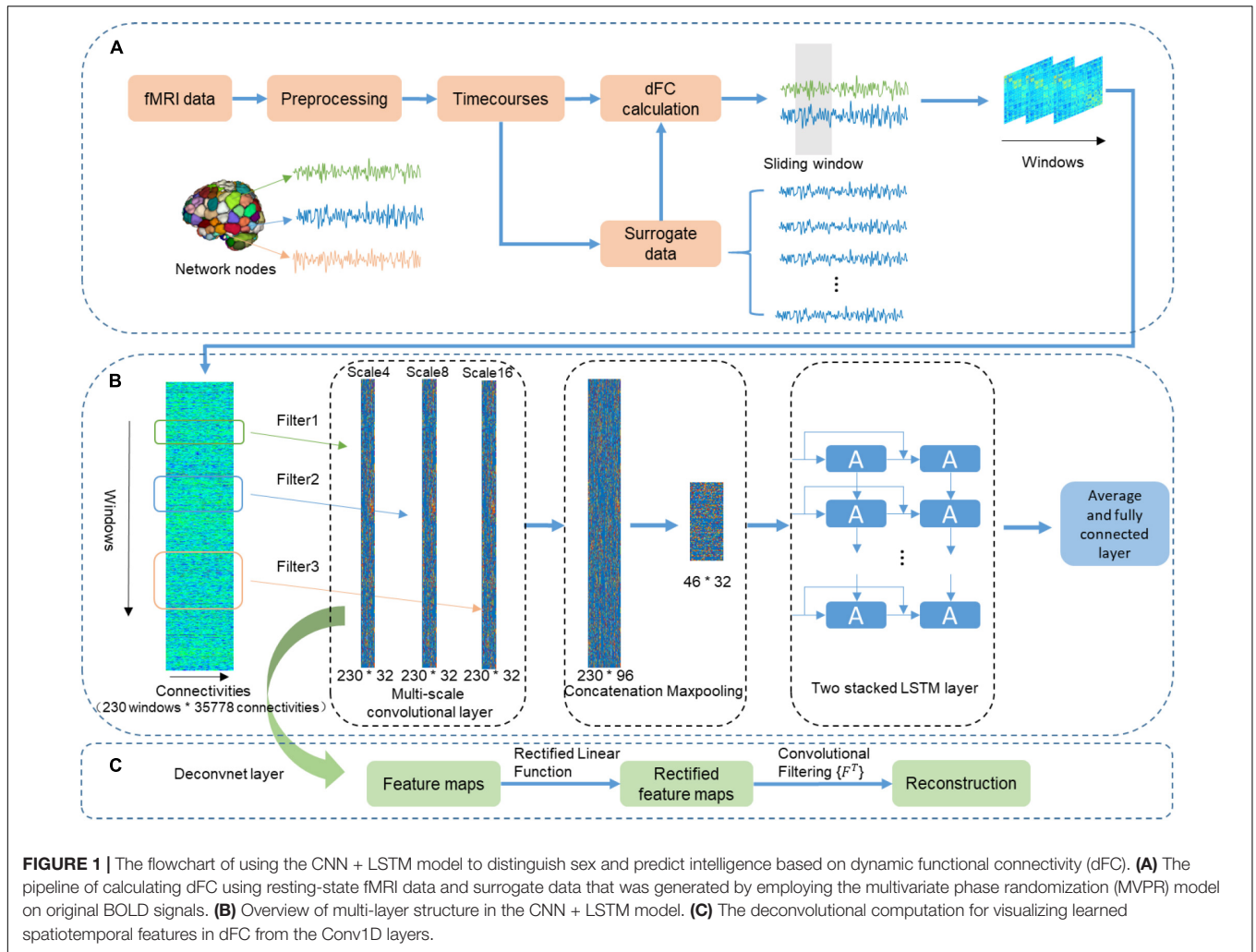
Multi-Scale Convolutional Layer

Although LSTM is powerful for handling temporal correlation, the major drawback of LSTM in handling spatiotemporal data was the redundancy of high-dimensional data. In addition, in the previous CNN-RNN architecture, the capability of each Conv1D layer determined by a single fixed filter size is extracting local information at only one time scale. Moreover, to adapt specific tasks, the filter length of 1D convolutional layer should be hand-picked. Therefore, multi-scale 1D convolutional layers were proposed for feature extraction because it can not only reduce spatial dimension but also account for different scales of brain activity (Roy et al., 2019; Yan et al., 2019). The architecture of multi-scale 1D convolutional layers includes multiple filters with diverse sizes in each convolutional layer. Filter lengths of 1D convolutional layers changed exponentially rather than linearly. Our experiment demonstrated that this architecture led to better performance by using different filter lengths (4, 8, and 16 time windows) in the 1D convolutional layers. Therefore, the size of the convolutional filters of three different scales 1D convolutional layers are 35,778 × 4 × 32, 35,778 × 8 × 32, 35,778 × 16 × 32, respectively. Here, we fed the dFC matrices into the multi-scale convolutional layer. As a result, outputs are three features whose sizes are 230 (time windows) × 32 (feature dimension). Then, a concatenation layer is designed to concatenate the output features in the time dimension, resulting in a feature map with a size of 230 × 96. Furthermore, a downsampling operation is performed using the max-pooling layer along the time dimension with kernel size 5 × 1, resulting in 46 × 96 features as an input of the following LSTM layers.

We performed deconvolutional computation for each convolutional layer (**Figure 1C**) (Zeiler and Fergus, 2014) to obtain the distribution of those connections with relatively high weights in discrimination. The deconvolutional results of one 1D convolutional layer were obtained for 100 randomly selected subjects. The group-level statistical strength of each functional connection was generated by computing the absolute value of its average strength across 100 randomly selected subjects (Zhang et al., 2018).

Two-Layer Stacked LSTM Layer

Long short-term memory is one kind of RNN models, which is different from CNN because of its consideration of the temporal information. LSTM consists of an input gate, an output gate, a forget gate and a cell. The advantage of LSTM is that it has sufficient ability to solve long-term dependencies because of the interactive operation among these three gates, in contrast to general RNNs. In addition, LSTM is designed to combat

FIGURE 1 | The flowchart using the CNN + LSTM model to distinguish sex and predict intelligence based on dynamic functional connectivity (dFC). **(A)** The pipeline of calculating dFC using resting-state fMRI data and surrogate data that was generated by employing the multivariate phase randomization (MVPR) model on original BOLD signals. **(B)** Overview of multi-layer structure in the CNN + LSTM model. **(C)** The deconvolutional computation for visualizing learned spatiotemporal features in dFC from the Conv1D layers.

vanishing/exploding gradients by using a gating mechanism. The LSTM model can be presented in the following form:

$$Forget\,gate : f_t = \text{Sigmoid}\left(W_f x_t + U_f h_{t-1} + b_f\right) \quad (1)$$

$$Input\,gate : i_t = \text{Sigmoid}\left(W_i x_t + U_i h_{t-1} + b_i\right) \quad (2)$$

$$Output\,gate : o_t = \text{Sigmoid}\left(W_o x_t + U_o h_{t-1} + b_o\right) \quad (3)$$

$$Estimated\,current\,cell\,state : \tilde{C}_t = Tanh\left(W_c x_t + U_c h_{t-1} + b_c\right) \quad (4)$$

$$Cell\,state : C_t = i_t \odot \tilde{C}_t + f_t \odot C_{t-1} \quad (5)$$

$$Hidden\,state : h_t = o_t \odot \tanh\left(C_t\right) \quad (6)$$

Where W_f, W_i, W_o, and W_c are input weights; U_f, U_i, U_o, and U_c are recurrent weights; b_f, b_i, b_o, and b_c are bias weights, and $\odot$ is the Hadamard product.

Here, we choose two densely connected LSTM layers because it may better capture latent dynamic information than one LSTM layer. It is worth noting that densely connected LSTM layers may mitigate the vanishing or exploding gradients problem (Roy et al., 2019; Yan et al., 2019). The size of the hidden state was set as 32. Also, we averaged all of the LSTM outputs to combine all fMRI steps (Dvornek et al., 2017). In this way, better classification performance could be obtained through leveraging all brain activities during scanning. Then the learned features were passed to the fully connected layer. The fully connected layer can be expressed as:

$$h^l = b^l + h^{l-1} \times w^l \quad (7)$$

Where w^l and b^l are input weight and bias weight, respectively. When the model is trained for intelligence prediction, the output of the fully connected layer is the predicted intelligence scores. However, for gender classification, another operation of Softmax was added as the last operation of this architecture.

Training, Validation, and Testing

A 10-fold cross-validation procedure was used for evaluating prediction performance. The HCP data was randomly split

into training, validation, and testing sets. More specifically, we divided the 1,050 subjects into ten folds. Note that multiple runs belonging to the same subject were not split across folds (He et al., 2020). For every test fold, the remaining nine folds were used for training and validation. Furthermore, it has been found that head motion was correlated with behavioral measures such as fluid intelligence (Siegel et al., 2017). Therefore, we regressed sex, age, and motion (frame-wise displacement FD) from the intelligence. In each test fold, we estimated the regression coefficients from the training set and applied them into the test fold (Kong et al., 2019; He et al., 2020).

The CNN + LSTM model was coded based on the platform of Pytorch (Paszke et al., 2017), and optimized with Adam optimizer to minimize the loss (Yan et al., 2019). The value of batch size was set as 64. The initial learning rate was 0.0001. We decreased the learning rate with weight decay rate of 10^{-1} after each epoch. To avoid overfitting and achieve higher generalization performance, we used dropout (dropout rate = 0.5) for regulating the model parameters and early stopping to stop training according to the prediction condition of the validation data. Briefly, when training loss continued to decrease but validation loss increased, this means that the training was already overfitting, and we should stop training.

Evaluation of Model Ability to Capture Dynamics of FC With Surrogate Data

It was unclear whether the CNN + LSTM model concentrated on the dFC containing the sequential temporal dynamics, or just captured the static statistics of dFC. To validate the validity of the model in capture dynamic interaction information, we used initial BOLD signals to generate surrogate data as null hypothesis. The aim of surrogate data is to generate the same time probability distribution, while preserving all statistical properties of the observed data like stationary cross correlation, but to destroy the dynamics in FC time courses (Schreiber and Schmitz, 2000; Pereda et al., 2005; Savva et al., 2019).

In this study, a multivariate phase randomization (MVPR) model (Prichard and Theiler, 1994) was applied by randomly shuffling the Fourier phases of the original BOLD signals such that their static FC structure could be preserved (Hindriks et al., 2016; Liégeois et al., 2017; Savva et al., 2019).

$$\hat{X}_k = X_k e^{i\varphi}, k = 1, 2, \cdots, n \qquad (8)$$

Where $X = [X_1, X_2, \cdots, X_n]$ is the discrete Fourier transformations of original time series. $n = 268$ refers to the number of brain regions. $\varphi = [\varphi_1, \varphi_2, \cdots, \varphi_T]$ is a uniformly distributed random phase in the range of $[0, 2\pi]$.

Subsequently, the inverse Fourier transform is applied to $\hat{X}_1, \hat{X}_2, \cdots, \hat{X}_n$ to generate one randomized copy $\hat{x}$. We repeated the procedure and then generated 100 surrogate copies for each subject. In addition, dFC matrices were calculated in the same way, excluding that surrogate copies instead of BOLD signals were used. For each surrogate copy of all subjects, the resulted dFC matrices were used as input of the model for gender classification with 10-fold cross-validation strategies. As a result,

100 surrogate copies were performed to estimate a distribution of accuracies under the null hypothesis of dFC.

Furthermore, a statistical method was proposed to assess the existence of dFC between a pair of ROIs (Savva et al., 2019). Then, this statistical framework was applied to test if the CNN + LSTM model could concentrate on the ROI pairs with statistically significant dFC. For details, 250 surrogate copies of one randomly selected participant were generated with the aforementioned method to formulate the null hypothesis. This null hypothesis can be rejected when any given FC time-series exhibits dFC.

Evaluation of Model Performance in Gender Classification and Intelligence Prediction Tasks

As commonly seen in the recent intelligence prediction studies based on rs-fMRI data (Finn et al., 2015; Liu et al., 2018; He et al., 2020), the Pearson's correlation between predicted and observed intelligence scores of all subjects across all folds was used for assessing the model performance of intelligence prediction tasks. In this study, we reported the Pearson's correlation and mean absolute error (MAE) to evaluate our model prediction performance. In the case of gender classification, the model was evaluated with the classification accuracy. Additionally, the area under ROC curve (AUC) which is a very widely used measure of performance for classification was also reported. Furthermore, it has been revealed that chronnectome (Calhoun et al., 2014) could be used to identify individuals and predict individual higher cognitive performance (Liu et al., 2018). For comparing with the CNN + LSTM model, we used linear SVM and linear epsilon SVR models (LIBSVM toolbox in Matlab[2]) based on dynamic characteristic of dFC (dFC-Str, which refers to the overall strength of dFC) (Liu et al., 2018), to achieve gender classification and intelligence prediction with the same 10-fold cross-validation strategies. The hyperparameter of linear SVM and linear epsilon-SVR includes the regularization parameter C, which was optimized for achieving its best performance.

Control Analysis

We further examined the effects of parcellation schemes and sliding window sizes on model performance. The selection of window size was controversial in previous studies (Kiviniemi et al., 2011; Jones et al., 2012; Hutchison et al., 2013; Allen et al., 2014; Qin et al., 2015). A smaller window size can better detect the potential low-frequency fluctuations in FC (Sakoğlu et al., 2010; Hutchison et al., 2012). It has been suggested that the window sizes from around 30–60 s can achieve the best classification and prediction performances (Shirer et al., 2012; Qin et al., 2015). In this work, three additional window sizes (60, 80, and 100 s) were used to investigate the potential impact of window width on gender classification and intelligence prediction performance. Additionally, to evaluate the effects caused by different brain parcellations, two additional parcellations were used for generating dFC matrices, including Power-264 consisting of 264 ROIs (Power et al., 2011),

---

and functional brain atlas of 160 ROIs (Dosenbach et al., 2010). Then, we re-performed the gender classification and intelligence prediction analyses using these dFC matrices based on these parcellations.

RESULTS

Performance of Gender Classification and Intelligence Prediction Tasks

Gender Classification Results

Figure 2A shows the gender classification performance of the CNN + LSTM and SVM model. The accuracy of $93.05 \pm 1.91\%$ was obtained by using the CNN + LSTM model, which is significantly higher than that obtained using SVM ($p < 0.001$, two-sample t-tests). Moreover, their ROC curves are shown in **Figure 2B**. The CNN + LSTM model achieved an AUC of 0.9805, while the SVM achieved an AUC of 0.9195. Note that the CNN + LSTM model achieved better performance than SVM by integrating the advantages of CNN and LSTM.

We conducted 100 times of repeated tests using 100 surrogate copies to testify whether the trained deep network discriminates genders based on dynamics in FC rather than other features such as static functional correlation across ROIs. The results are shown in **Figure 2C**. As expected, empirical distributions of the accuracies scattered around 57%, suggesting the classifier performance was just better than random guessing for the surrogate copies. All the accuracies of surrogate copies fell behind that of real BOLD signals, demonstrating that the statistical significance of gender classification based on temporal dynamics in FC was high ($p < 0.001$), so that the null hypothesis that the deep network failed in capturing spatiotemporal features of dFC could be rejected. The learning curves while training the CNN+LSTM model was shown in **Figure 2D**.

Intelligence Prediction

Figure 3A depicts the prediction accuracy of fluid and crystallized intelligence (Pearson's correlation coefficient) in a 10-fold cross-validation test. The CNN + LSTM model achieved higher prediction accuracy in both tasks than the SVR model, with Pearson's correlation $r = 0.3129$ for fluid intelligence and $r = 0.4946$ for crystallized intelligence, respectively, in contrast to SVR's prediction accuracy of $r = 0.2245$ for fluid intelligence and $r = 0.3889$ for crystallized intelligence. The conclusions of MAE are similar, as illustrated in **Figure 3B**. The MAE of 11.9561 ± 0.7412 (for crystallized intelligence) and 3.7287 ± 0.2673 (for fluid intelligence) in the CNN + LSTM model are significantly lower than those obtained using SVR ($p < 0.05$, two-sample t-tests).

Estimating the Most Discriminative Connection Features

To explore the ability of the CNN + LSTM model in extracting features related to the three discriminating tasks, we used deconvolution computations to show the important functional connection features characterized with high weights. **Figure 4A**

shows the distributions of FC with feature weights above the threshold, which was set for an important contribution at the 75th percentile of all feature weights. Important FC features are widespread across the brain for all three discriminating tasks. However, as for gender classification, a large number of FC features above the threshold are distributed in inter-network and intra-network of frontoparietal (networks 1 and 2), default mode (network 3), and motor (network 5), especially for the default mode and motor networks. While FC features with high weights are mostly in the inter-network and intra-network of frontoparietal for fluid and crystallized intelligence tasks.
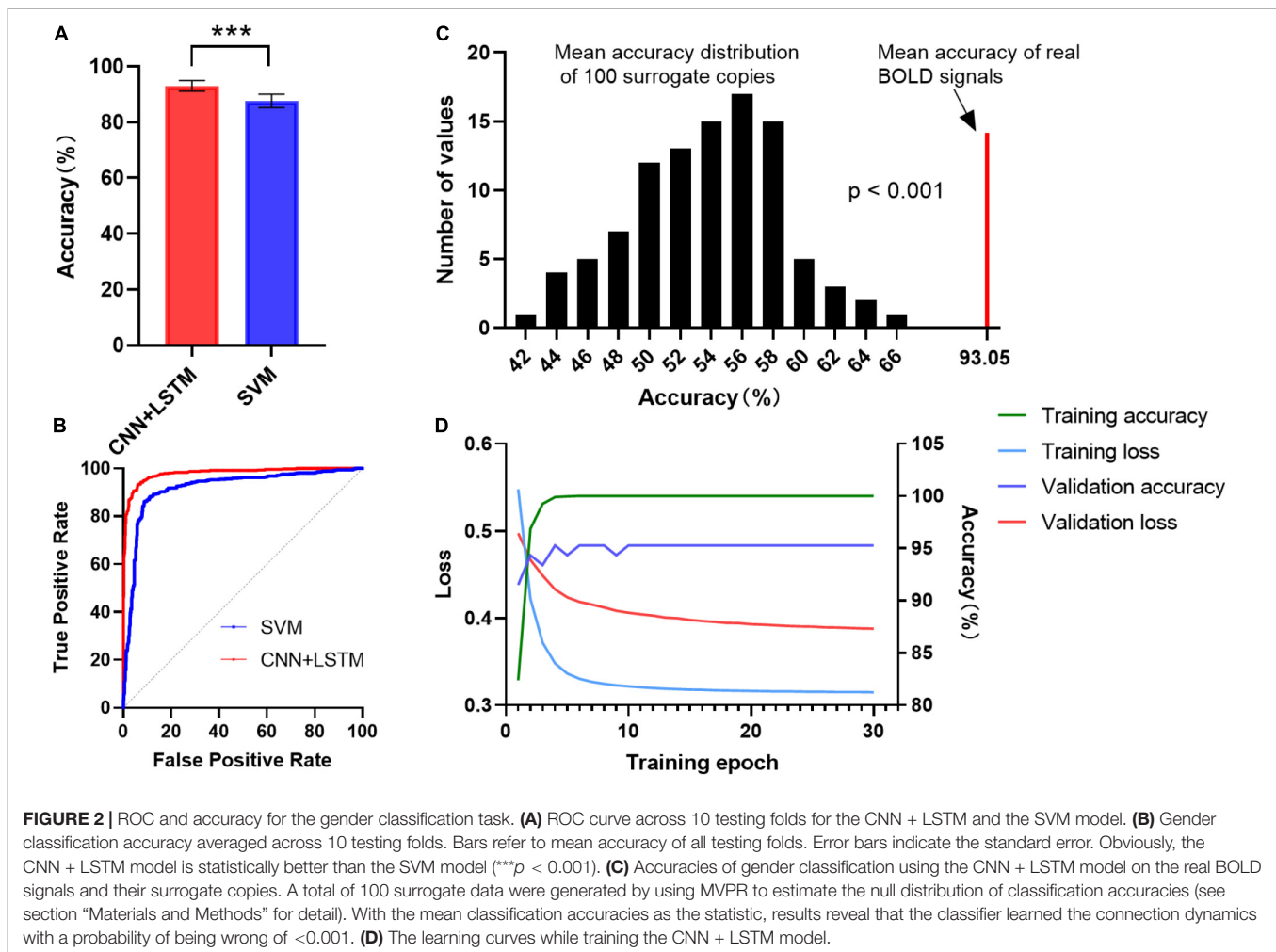
To explore which networks had the most predictive power, average intra-network and inter-network feature weights of eight networks were calculated without thresholding. As illustrated in **Figure 4B**, the default mode network (DMN) has the highest inter-network and intra-network feature weights of gender classification, especially for the inter-network. The motor and frontoparietal networks follow the default mode. The other networks have slightly low weights of intra-network and inter-network features. Additionally, patterns from fluid intelligence and crystallized intelligence predictions are very similar, i.e., the frontoparietal network has slightly higher weights than the others. However, the visual I network (network 6) has low predictive power for predicting intelligence. Intra-network and inter-network feature weights are comparable in all the networks.

Identifying Significant Dynamic Connections With Surrogate Data

The null hypothesis for dFC using the surrogate data was tested on one randomly selected subject, with the results indicated in **Figure 5A**. The red points represent those connections exhibiting statistically significant dFC. Note that significant connections were widely distributed across the brain. To further identify if the CNN + LSTM model could capture those regions with statistically significant dynamics, we computed the distributions of averaged deconvolutional weights for all significantly dynamic and non-significantly dynamic connections. As shown in **Figures 5B,C**, both the weight distributions of ROI pairs follow the approximately normal distribution. Moreover, the features with high weights cover nearly half of those connections with significant dFC. In other words, the important regions extracted by using the CNN + LSTM model carry sufficient dynamic connectivity information.

DISCUSSION

In this work, we used a CNN + LSTM model combining the advantages of CNN and LSTM to learn spatiotemporal information in rest-state dFC. This was the first attempt to capture the dynamic interaction information of dFC series using deep learning, which avoids information reduction and takes advantage of time-varying spatiotemporal information. More importantly, we showed that the CNN + LSTM model not only successfully achieved a high accuracy of gender classification but could significantly predict individual intelligence, including fluid and crystallized intelligence across a large-scale dataset totaling

FIGURE 2 | ROC and accuracy for the gender classification task. **(A)** ROC curve across 10 testing folds for the CNN + LSTM and the SVM model. **(B)** Gender classification accuracy averaged across 10 testing folds. Bars refer to mean accuracy of all testing folds. Error bars indicate the standard error. Obviously, the CNN + LSTM model is statistically better than the SVM model (***$p < 0.001$). **(C)** Accuracies of gender classification using the CNN + LSTM model on the real BOLD signals and their surrogate copies. A total of 100 surrogate data were generated by using MVPR to estimate the null distribution of classification accuracies (see section "Materials and Methods" for detail). With the mean classification accuracies as the statistic, results reveal that the classifier learned the connection dynamics with a probability of being wrong of <0.001. **(D)** The learning curves while training the CNN + LSTM model.
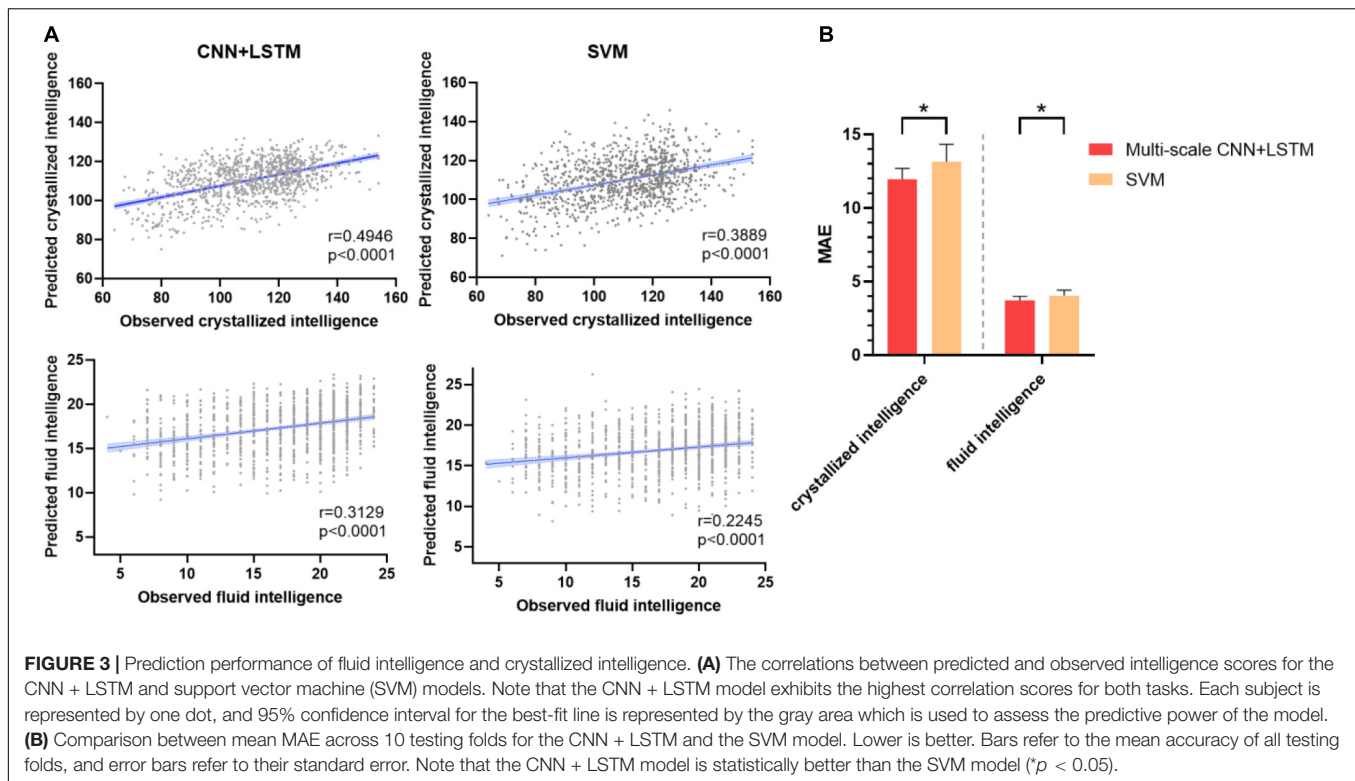
1,050 participants. Moreover, the results of deconvolutional computation provided interpretation for extracted features. Our results suggest that the CNN + LSTM model can simultaneously learn temporal and spatial information of dFC series instead of dFC's statistical characteristics, which could significantly improve the prediction power of individual cognition traits.

Gender Classification and Intelligence Prediction Performance

Table 1 summarizes the rs-fMRI based model performance of gender classification and intelligence prediction from some recent studies. Zhang et al. (2018) showed that with three different methods combining FC features of multiple runs from the HCP (820 subjects), the optimal classification accuracy and AUC were 85% and 0.93, respectively. Ktena et al. (2018) achieved a gender classification accuracy of 80% with static functional connectivity (sFC) from UK Biobank (2,500 subjects). In addition, He et al. (2020) reported a sex prediction accuracy of 91.6% with four different models in the UK Biobank (8,868 subjects). Our model based on dFC time series could achieve a higher accuracy of gender classification than those using static FC feature.

In the case of behavioral measures such as fluid and crystallized intelligence, combining Picture Vocabulary Test (vocabulary) and Oral Reading Recognition Test (reading) into one score, previous studies (Finn et al., 2015; Liu et al., 2018) reporting high fluid intelligence prediction accuracy were both using a leave-one-subject-out cross-validation for small samples from the HCP dataset (126 subjects and 105 subjects). Recent studies reported lower accuracies when the number of samples increased. For example, He et al. (2020) reported prediction accuracies (Pearson's correlation) ranging from 0.257 to 0.297 for fluid intelligence, about from 0.25 to 0.4 for reading and approximately from 0.2 to 0.4 for vocabulary, with four different models in the HCP dataset (953 subjects). Dubois et al. (2018) successfully predicted fluid intelligence ($r = 0.27$) using linear models in the HCP dataset (884 subjects). Kashyap et al. (2019) reported prediction accuracies of 58 behavioral measures, including fluid intelligence ($r = 0.25$), reading (about $r = 0.37$) and vocabulary (about $r = 0.39$) in the HCP dataset (803 subjects).

It is worthy to note that all these methods are based on the sFC, assuming that the underlying connections remain unchanged over time. Here, we applied a CNN + LSTM model to learn the spatiotemporal information of the original dFC series.

FIGURE 3 | Prediction performance of fluid intelligence and crystallized intelligence. **(A)** The correlations between predicted and observed intelligence scores for the CNN + LSTM and support vector machine (SVM) models. Note that the CNN + LSTM model exhibits the highest correlation scores for both tasks. Each subject is represented by one dot, and 95% confidence interval for the best-fit line is represented by the gray area which is used to assess the predictive power of the model. **(B)** Comparison between mean MAE across 10 testing folds for the CNN + LSTM and the SVM model. Lower is better. Bars refer to the mean accuracy of all testing folds, and error bars refer to their standard error. Note that the CNN + LSTM model is statistically better than the SVM model (*$p < 0.05$).

Importantly, combining the multiple 1D convolutional layers of different filters can learn the spatial associations of dFC at different time scales. And the stacked LSTM layers can capture latent temporal dynamics. As mentioned earlier, we achieved a gender classification accuracy of 93.05% and prediction accuracies (Pearson's correlation) of 0.3129 and 0.4946 for fluid intelligence and crystallized intelligence, respectively. Overall, this indicates that our model prediction performance is among the best in the literature based on the same size dataset of rs-fMRI. Further, we employed an ablation control analysis to highlight necessity of the use of the LSTM module, i.e., we only retained the multi-scale 1D convolutional layer, one concatenation layer and one fully connected layer to reperform the intelligence prediction tasks. Other parameters are consistent with the CNN + LSTM model. As shown in **Supplementary Figure S1**, the multi-scale CNN model only achieved a significantly low prediction accuracy of crystallized intelligence (Pearson's correlation $r = 0.1866$) and fluid intelligence ($r = 0.1277$), suggesting that the LSTM greatly contributed to the prediction of individual intelligence by capturing the temporal information of dFC time series.

Effectiveness Analysis Using Surrogate Data

So far, many dFC studies have employed MVPR model to generate surrogate data for detecting significance of dynamics in resting-state FC (Hindriks et al., 2016; Savva et al., 2019) by preserving statistical properties of the initial data such as stationary cross-correlation, i.e., static FC structure. Given that

the dynamic information is removed in the surrogate data, the null hypothesis is that the model can achieve high classification or prediction accuracy in the absence of the dFC. However, our results demonstrate that the classification performances of surrogate copies are just better than random guessing, so that the null hypothesis should be rejected ($p < 0.0001$). Moreover, the results from the feature analysis in the model also provide further evidence supporting this deduction, i.e., the features with high weights from the 1D convolution layer covered nearly half of those connections with significant dFC, especially those connections with the highest weights. These outcomes from the surrogate data suggest that the CNN + LSTM model can sufficiently learn temporal dynamics rather than only static structure in the FC.

## Important Networks of Gender and Intelligence Discrimination

As shown in **Figure 4**, the FC features within the DMN', frontoparietal and motor networks had a great contribution to gender classification, especially, the DMN has the highest inter-network and intra-network feature weights, which is generally consistent with previous structural and functional MRI studies (Zhang et al., 2018; Luo et al., 2019). However, the subcortical/cerebellum and visual networks are the majority of the least discriminative functional networks.

Previous researches have shown that the DMN is related to many different functions like social understanding (Li et al., 2014), social cognitive abilities (Schilbach et al., 2008; Mars et al., 2012), and episodic memory (Sestieri et al., 2011). While
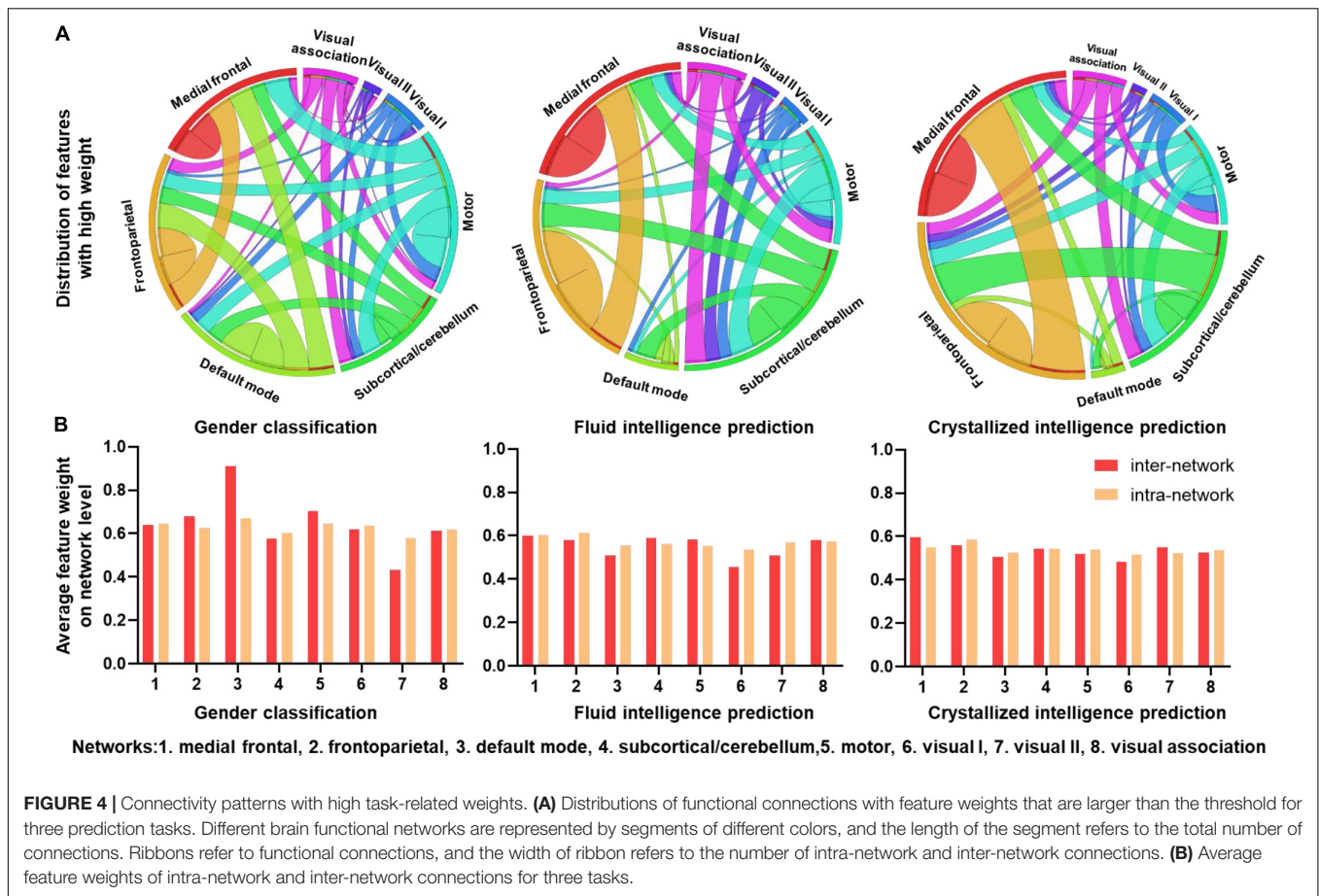
FIGURE 4 | Connectivity patterns with high task-related weights. **(A)** Distributions of functional connections with feature weights that are larger than the threshold for three prediction tasks. Different brain functional networks are represented by segments of different colors, and the length of the segment refers to the total number of connections. Ribbons refer to functional connections, and the width of ribbon refers to the number of intra-network and inter-network connections. **(B)** Average feature weights of intra-network and inter-network connections for three tasks.
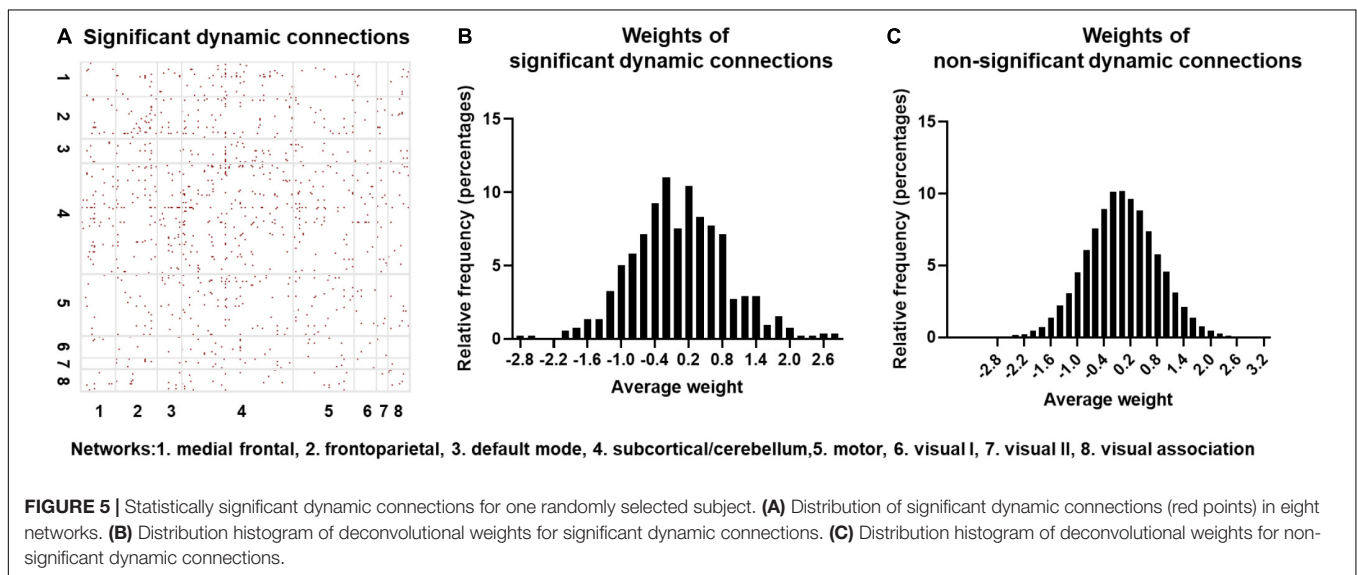
FIGURE 5 | Statistically significant dynamic connections for one randomly selected subject. **(A)** Distribution of significant dynamic connections (red points) in eight networks. **(B)** Distribution histogram of deconvolutional weights for significant dynamic connections. **(C)** Distribution histogram of deconvolutional weights for non-significant dynamic connections.
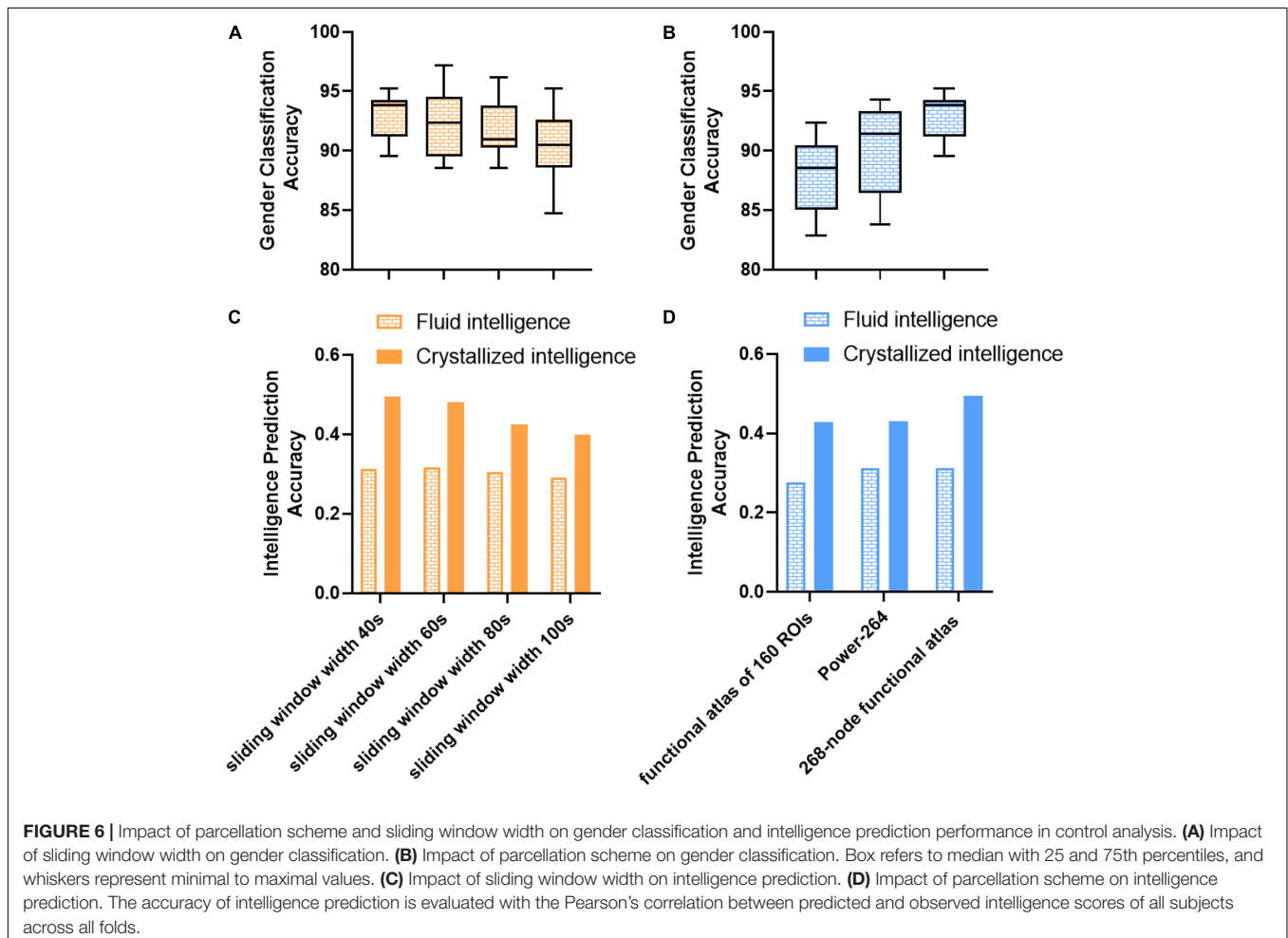
many studies reported sex differences in behavioral measures. For example, women perform better than men on memory tasks (Hedges and Nowell, 1995) as well as measures of social cognition (Gur et al., 2010, 2012; Moore et al., 2015). These results revealed the underlying associations between the DMN and gender.

Besides the DMN, there were also other networks with high contributions to the gender classification. For instance, features with high weights for gender classification are prominent in the frontoparietal and motor networks. Also, the network-level average feature weight was illustrated in **Figure 5B**,

TABLE 1 | Model performance of rs-fMRI based gender classification and intelligence prediction tasks in some recent studies.

Author	Dataset (subs)	Methods	Types of data	Tasks	Accuracy
Chao Zhang	HCP (820)	PLSR	sFC	Gender	87% (10-fold)
Sofia Ira Ktena	UK Biobank (2,500)	GCN	sFC	Gender	80% (5-fold)
Tong He	UK Biobank (8,868)	TML	sFC	Gender	91.6% (9-fold)
Casanova R	FCP (148)	Lasso regression	sFC	Gender	62%
Stephen M. Smith	HCP (131)	Linear discriminant analysis	sFC	Gender	87% (LOOCV)
Susanne Weis	HCP (434)	SVM	sFC	Gender	75.1% (10-fold)
Tong He	HCP (953)	FNN	sFC	Fluid intelligence	0.297 (20-fold)
Julien Dubois	HCP(884)	Linear models	sFC	Fluid intelligence	0.27 (leave-one family out)
Rajan Kashyap	HCP (803)	COBE	sFC	Fluid intelligence	0.25 (20-fold)
Ru Kong	HCP (881)	MS-HBM	sFC	Fluid intelligence	0.22 (20-fold)
Abigail S. Greene	HCP (515)	CPM	sFC	Fluid intelligence	0.196 (LOOCV)
Jin Liu	HCP (105)	SVR	dFC	Fluid intelligence	0.418 (LOOCV)
Emily S. Finn	HCP (142)	Linear regression	sFC	Fluid intelligence	0.5 (LOOCV)

All results are the best performance mentioned in the literature. CNN, convolutional neural network; GCN, graph convolutional network; PLSR, partial least squares regression; COBE, common and orthogonal basis extraction; MS-HBM, multi-session hierarchical Bayesian model; CPM, connectome-based predictive modeling; LOOCV, leave-one-subject-out cross-validation.

FIGURE 6 | Impact of parcellation scheme and sliding window width on gender classification and intelligence prediction performance in control analysis. **(A)** Impact of sliding window width on gender classification. **(B)** Impact of parcellation scheme on gender classification. Box refers to median with 25 and 75th percentiles, and whiskers represent minimal to maximal values. **(C)** Impact of sliding window width on intelligence prediction. **(D)** Impact of parcellation scheme on intelligence prediction. The accuracy of intelligence prediction is evaluated with the Pearson's correlation between predicted and observed intelligence scores of all subjects across all folds.

indicating that the intra-network and inter-network weights of the motor and frontoparietal regions are slightly higher than other networks except for the DMN. Recent studies have

also reported that most FC features within frontalparietal and sensorimotor networks are associated with gender differences (Zhang et al., 2016, 2018). Additionally, the reliable gender

difference in FC has been reported for the sensorimotor network (Weis et al., 2019).

In the case of fluid intelligence and crystallized intelligence, we found stronger FC features are mostly associated with the frontoparietal network. This suggests that individual variability in intelligence is related to higher order systems that reflect individual cognitive ability than those of the primary systems. Similarly, Liu et al. (2018) found that dFC features of default mode, frontoparietal, and dorsal attention networks contributed predominantly to fluid intelligence and executive function. The frontoparietal network was also reported to be the most predictive of fluid intelligence (Finn et al., 2015). Moreover, many previous studies also reported that intelligence is related to the structural and functional properties of these networks, which are consistent with ours (Kanai and Rees, 2011; Cole et al., 2012). On the other hand, for the intelligence prediction tasks, all the networks have comparable average weights, suggesting strong evidence that more than one network accounts for the intelligence, consistent with the previous report that intelligence is related to functional coupling and structural connectivity across widespread brain regions (Choi et al., 2008).

Effect of Parcellation Scheme and Sliding Window Width

Several confounding factors, such as sliding window width and parcellation scheme, might influence prediction performance. As shown in **Figure 6**, we investigated the impact of the parcellation scheme and sliding window width and found that most of the results remained robust. All the sliding window widths from 40 to 100 s could achieve a robust prediction performance, thus providing indirect evidence of dFC's presence at different timescales (Savva et al., 2019). Intriguingly, the window width of 40 s achieved the best performance in gender classification and crystallized intelligence prediction tasks. The window width of 60 s in fluid intelligence prediction task achieved a slightly higher performance ($r = 0.3172$) than the window width of 40 s ($r = 0.3129$). These results suggested that the duration between 40 and 60 s could better capture dynamic information of fluctuations and decode the variability of individuals, in line with the previous findings (Shirer et al., 2012; Kucyi and Davis, 2014; Qin et al., 2015). However, the mean accuracies of gender classification based on the functional brain atlas of 160 ROIs and 264 ROIs were about 88 and 90.3%. The prediction accuracies of intelligence prediction tasks based on the functional brain atlas of 160 ROIs and 264 ROIs were 0.2758 and 0.3120 for fluid intelligence, 0.4283 and 0.4308 for crystallized intelligence, respectively. Reduction in gender classification and intelligence prediction accuracies compared to the 268-node functional parcellation suggests that a finer parcellation may detect more subtle individual variability and dynamics.

Limitation and Conclusion

Some limitations should be considered in the presented approach. Firstly, We used the gender classification task to demonstrate the effectiveness of the deep learning model, considering that gender is an accurate label and many recent studies have focused on gender classification (Smith et al., 2013; Ktena et al., 2018; Zhang et al., 2018). However, compared to SVM, the improvement (about 5%) of gender classification is considerate but still limited which may be caused by the ceiling effect of gender classification. A better way of evaluating the model is to extend the model to applications of individual diagnosis for typical mental disorders such as schizophrenia and depression. Many recent studies aimed to find neuroimaging-based biomarkers for disease diagnosis using deep learning techniques. For example, In our previous study (Zeng et al., 2018), the deep autoencoder network achieved an improvement ($>5.0\%$) in diagnosing schizophrenia across multiple imaging sites compared to other multi-site studies such as using linear classifiers (Skåtun et al., 2017). Furthermore, LSTM also shows potential capabilities in diagnosing disease. Dvornek et al. (2017) achieved a classification accuracy of 68.5% in discriminating autism spectrum disorders with LSTMs, which demonstrated a higher classification accuracy compared to previous studies (Ghiassian et al., 2016; Abraham et al., 2017). Overall, based on the findings of abnormal functional connectivities in psychiatric disorders (Damaraju et al., 2014; Yan et al., 2017; Zeng et al., 2018) and the potential abilities of deep learning in disease diagnosis, the CNN + LSTM model may have excellent prospects in assistant diagnosis of some mental disorders.

Secondly, we applied the CNN + LSTM model to explore brain network dynamics only with rs-fMRI data. While multi-modalities fusion (e.g., combining electroencephalography measures) will be helpful for generating a more accurate model due to the higher temporal resolution of electroencephalography data relative to fMRI signals. Thirdly, interpretation of the LSTM model remains unclear but a critical field of research. We will further aim to interpret deep learning models in the future.

In summary, this is the first time that original dFC series are successfully applied to discriminating individual cognitive ability or demographic characteristics such as sex, fluid, and crystallized intelligence. Furthermore, the high accuracies in these tasks indicated the effectiveness of the used model owing to the full use of the high-level spatiotemporal information of dFC. Also, the deconvolutional computation provides an interpretation of the deep learning methods. Our findings highlight that deep-chronnectome can capture the inherent dynamical interaction information of functional networks and provide the potentials for predicting individualized characterization of demographics and cognition traits.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: https://www.humanconnectome.org.

AUTHOR CONTRIBUTIONS

DH and HS designed the study. JQ and LF performed the experiments. LF, JS, and HS discussed the results and contributed

to the final manuscript. All authors contributed to the article and approved the submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins. 2020.00881/full#supplementary-material

REFERENCES

Abraham, A., Milham, M. P., Di Martino, A., Craddock, R. C., Samaras, D., Thirion, B., et al. (2017). Deriving reproducible biomarkers from multi-site resting-state data: an autism-based example. *NeuroImage* 147, 736–745. doi: 10.1016/j.neuroimage.2016.10.045

Akshoomoff, N., Beaumont, J. L., Bauer, P. J., Dikmen, S. S., Gershon, R. C., Mungas, D., et al. (2013). VIII. NIH Toolbox Cognition Battery (CB): composite scores of crystallized, fluid, and overall cognition. *Monogr. Soc. Res Child Dev.* 78, 119–132. doi: 10.1111/mono.12038

Allen, E. A., Damaraju, E., Plis, S. M., Erhardt, E. B., Eichele, T., and Calhoun, V. D. (2014). Tracking whole-brain connectivity dynamics in the resting state. *Cereb. Cortex* 24, 663–676. doi: 10.1093/cercor/bhs352

Battaglia, D., Boudou, T., Hansen, E. C., Lombardo, D., Chettouf, S., Daffertshofer, A., et al. (2020). Dynamic functional connectivity between order and randomness and its evolution across the human adult lifespan. *bioRxiv* [Preprint]. doi: 10.1101/107243

Calhoun, V. D., Miller, R., Pearlson, G., and Adalı, T. (2014). The chronnectome: time-varying connectivity networks as the next frontier in fMRI data discovery. *Neuron* 84, 262–274. doi: 10.1016/j.neuron.2014.10.015

Chang, C., and Glover, G. H. (2010). Time–frequency dynamics of resting-state brain connectivity measured with fMRI. *Neuroimage* 50, 81–98. doi: 10.1016/j. neuroimage.2009.12.011

Choi, Y. Y., Shamosh, N. A., Cho, S. H., DeYoung, C. G., Lee, M. J., Lee, J.-M., et al. (2008). Multiple bases of human intelligence revealed by cortical thickness and neural activation. *J. Neurosci.* 28, 10323–10329. doi: 10.1523/jneurosci.3259-08. 2008

Cole, M. W., Yarkoni, T., Repovš, G., Anticevic, A., and Braver, T. S. (2012). Global connectivity of prefrontal cortex predicts cognitive control and intelligence. *J. Neurosci.* 32, 8988–8999. doi: 10.1523/jneurosci.0536-12.2012

Damaraju, E., Allen, E., Belger, A., Ford, J., McEwen, S., Mathalon, D., et al. (2014). Dynamic functional connectivity analysis reveals transient states of dysconnectivity in schizophrenia. *Neuroimage Clin.* 5, 298–308. doi: 10.1016/j. nicl.2014.07.003

Davison, E. N., Turner, B. O., Schlesinger, K. J., Miller, M. B., Grafton, S. T., Bassett, D. S., et al. (2016). Individual differences in dynamic functional brain connectivity across the human lifespan. *PLoS Comput. biol.* 12:e1005178. doi: 10.1371/journal.pcbi.1005178

Demirtaş, M., Tornador, C., Falcón, C., López-Solà, M., Hernández-Ribas, R., Pujol, J., et al. (2016). Dynamic functional connectivity reveals altered variability in functional connectivity among patients with major depressive disorder. *Hum Brain Map.* 37, 2918–2930. doi: 10.1002/hbm. 23215

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

Douw, L., Wakeman, D. G., Tanaka, N., Liu, H., and Stufflebeam, S. M. (2016). State-dependent variability of dynamic functional connectivity between frontoparietal and default networks relates to cognitive flexibility. *Neuroscience* 339, 12–21. doi: 10.1016/j.neuroscience.2016.09.034

Du, Y., Pearlson, G. D., Lin, D., Sui, J., Chen, J., Salman, M., et al. (2017). Identifying dynamic functional connectivity biomarkers using GIG−ICA: application to schizophrenia, schizoaffective disorder, and psychotic bipolar disorder. *Hum. Brain Map.* 38, 2683–2708. doi: 10.1002/hbm.23553

Dubois, J., Galdi, P., Han, Y., Paul, L. K., and Adolphs, R. (2018). Resting-state functional brain connectivity best predicts the personality dimension of openness to experience. *Pers. Neurosci.* 1:e6. doi: 10.1017/pen. 2018.8

Dvornek, N. C., Ventola, P., Pelphrey, K. A., and Duncan, J. S. (2017). *Identifying Autism from Resting-State fMRI Using Long Short-Term Memory Networks*. Berlin: Springer International Publishing, 362–370.

Elton, A., and Gao, W. (2015). Task−related modulation of functional connectivity variability and its behavioral correlations. *Hum. Brain Map.* 36, 3260–3272. doi: 10.1002/hbm.22847

Fatima, Z., Kovacevic, N., Misic, B., and Mcintosh, A. R. (2016). Dynamic functional connectivity shapes individual differences in associative learning. *Hum. Brain Map.* 37, 3911–3928. doi: 10.1002/hbm.23285

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

Friston, K. J. (2011). Functional and effective connectivity: a review. *Brain Connect.* 1, 13–36. doi: 10.1089/brain.2011.0008

Ghiassian, S., Greiner, R., Jin, P., and Brown, M. R. (2016). Using functional or structural magnetic resonance images and personal characteristic data to identify ADHD and autism. *PloS One* 11:e0166934. doi: 10.1371/journal.pone. 0166934

Glasser, M. F., Sotiropoulos, S. N., Wilson, J. A., Coalson, T. S., Fischl, B., Andersson, J. L., et al. (2013). The minimal preprocessing pipelines for the human connectome project. *Neuroimage* 80, 105–124. doi: 10.1016/j. neuroimage.2013.04.127

Gur, R. C., Richard, J., Calkins, M. E., Chiavacci, R., Hansen, J. A., Bilker, W. B., et al. (2012). Age group and sex differences in performance on a computerized neurocognitive battery in children age 8-21. *Neuropsychology* 26, 251–265. doi: 10.1037/a0026712

Gur, R. C., Richard, J., Hughett, P., Calkins, M. E., Macy, L., Bilker, W. B., et al. (2010). A cognitive neuroscience-based computerized battery for efficient measurement of individual differences: standardization and initial construct validation. *J. Neurosci. Methods* 187, 254–262. doi: 10.1016/j.jneumeth.2009.11. 017

Hansen, E. C., Battaglia, D., Spiegler, A., Deco, G., and Jirsa, V. K. (2015). Functional connectivity dynamics: modeling the switching behavior of the resting state. *Neuroimage* 105, 525–535. doi: 10.1016/j.neuroimage.2014.11.001

He, T., Kong, R., Holmes, A. J., Nguyen, M., Sabuncu, M. R., Eickhoff, S. B., et al. (2020). Deep neural networks and kernel regression achieve comparable accuracies for functional connectivity prediction of behavior and demographics. *NeuroImage* 206, 116276. doi: 10.1016/j.neuroimage.2019. 116276

Hedges, L. V., and Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science* 269, 41–45. doi: 10.1126/science.7604277

Hindriks, R., Adhikari, M. H., Murayama, Y., Ganzetti, M., Mantini, D., Logothetis, N. K., et al. (2016). Can sliding-window correlations reveal dynamic functional connectivity in resting-state fMRI? (vol 127, pg 242, 2016). *Neuroimage* 132, 115–115. doi: 10.1016/j.neuroimage.2016.02.007

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

Hutchison, R. M., Womelsdorf, T., Allen, E. A., Bandettini, P. A., Calhoun, V. D., Corbetta, M., et al. (2013). Dynamic functional connectivity: promises, issues, and interpretations. *Neuroimage* 80, 360–378. doi: 10.1016/j.neuroimage.2013. 05.079

Hutchison, R. M., Womelsdorf, T., Gati, J. S., Everling, S., and Menon, R. S. (2012). Resting−state networks show dynamic functional connectivity in awake humans and anesthetized macaques. *Hum. Brain Map.* 34, 2154–2177. doi: 10.1002/hbm.22058

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

Kanai, R., and Rees, G. (2011). The structural basis of inter-individual differences in human behaviour and cognition. *Nat. Rev. Neurosci.* 12, 231–242. doi: 10. 1038/nrn3000

Kashyap, R., Kong, R., Bhattacharjee, S., Li, J., Zhou, J., and Thomas Yeo, B. T. (2019). Individual-specific fMRI-Subspaces improve functional connectivity prediction of behavior. *NeuroImage* 189, 804–812. doi: 10.1016/j.neuroimage. 2019.01.069

Kiviniemi, V., Vire, T., Remes, J., Elseoud, A. A., Starck, T., Tervonen, O., et al. (2011). A sliding time-window ICA reveals spatial variability of the default mode network in time. *Brain Connect.* 1, 339–347. doi: 10.1089/brain.2011. 0036

Kong, R., Li, J., Orban, C., Sabuncu, M. R., Liu, H., Schaefer, A., et al. (2019). Spatial topography of individual-specific cortical networks predicts human cognition, personality, and emotion. *Cereb. Cortex* 29, 2533–2551. doi: 10.1093/cercor/ bhy123

Ktena, S. I., Parisot, S., Ferrante, E., Rajchl, M., Lee, M., Glocker, B., et al. (2018). Metric learning with spectral graph convolutions on brain connectivity networks. *NeuroImage* 169, 431–442. doi: 10.1016/j.neuroimage.2017. 12.052

Kucyi, A., and Davis, K. D. (2014). Dynamic functional connectivity of the default mode network tracks daydreaming. *Neuroimage* 100, 471–480. doi: 10.1016/j. neuroimage.2014.06.044

Li, Q., Dong, Q., Ge, F., Qiang, N., Zhao, Y., Wang, H., et al. (2019). "Simultaneous spatial-temporal decomposition of connectome-scale brain networks by deep sparse recurrent auto-encoders," in *Proceedings of the International Conference on Information Processing in Medical Imaging*, (Berlin: Springer), 579–591. doi: 10.1007/978-3-030-20351-1_45

Li, W., Mai, X., and Liu, C. (2014). The default mode network and social understanding of others: what do brain connectivity studies tell us. *Front. Hum. Neurosci.* 8:74. doi: 10.3389/fnhum.2014.00074

Liégeois, R., Laumann, T. O., Snyder, A. Z., Zhou, J., and Yeo, B. T. T. (2017). Interpreting temporal fluctuations in resting-state functional connectivity MRI. *NeuroImage* 163, 437–455. doi: 10.1016/j.neuroimage.2017. 09.012

Liu, J., Liao, X., Xia, M., and He, Y. (2018). Chronnectome fingerprinting: identifying individuals and predicting higher cognitive functions using dynamic brain connectivity patterns. *Hum. Brain Map.* 39, 902–915. doi: 10.1002/hbm. 23890

Luo, Z., Hou, C., Wang, L., and Hu, D. (2019). Gender identification of human cortical 3-D morphology using hierarchical sparsity. *Front. Hum. Neurosci.* 13:e00029. doi: 10.3389/fnhum.2019.00029

Madhyastha, T. M., Askren, M. K., Boord, P., and Grabowski, T. J. (2015). Dynamic connectivity at rest predicts attention task performance. *Brain Connect.* 5, 45–59. doi: 10.1089/brain.2014.0248

Mars, R., Neubert, F.-X., Noonan, M., Sallet, J., Toni, I., and Rushworth, M. (2012). On the relationship between the "default mode network" and the "social brain". *Front. Hum. Neurosci.* 6:e00189.doi: 10.3389/fnhum.2012. 00189

Moore, T. M., Reise, S. P., Gur, R. E., Hakonarson, H., and Gur, R. C. (2015). Psychometric properties of the penn computerized neurocognitive battery. *Neuropsychology* 29, 235–246. doi: 10.1037/neu0000093

Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., et al. (2017). Automatic differentiation in pytorch. *NIPS-W*.

Pedersen, M., Zalesky, A., Omidvarnia, A., and Jackson, G. D. (2018). Brain connectivity dynamics: multilayer network switching rate predicts brain performance. *bioRxiv* [Preprint]. doi: /10.1101/403105

Pereda, E., Quiroga, R. Q., and Bhattacharya, J. (2005). Nonlinear multivariate analysis of neurophysiological signals. *Prog. Neurobiol.* 77, 1–37. doi: 10.1016/j. pneurobio.2005.10.003

Power, J. D., Cohen, A. L., Nelson, S. M., Wig, G. S., Barnes, K. A., Church, J. A., et al. (2011). Functional network organization of the human brain. *Neuron* 72, 665–678.

Prichard, D., and Theiler, J. (1994). Generating surrogate data for time series with several simultaneously measured variables. *Phys. Rev. Lett.* 73, 951–954. doi: 10.1103/PhysRevLett.73.951

Qin, J., Chen, S.-G., Hu, D., Zeng, L.-L., Fan, Y.-M., Chen, X.-P., et al. (2015). Predicting individual brain maturity using dynamic functional connectivity. *Front. Hum. Neurosci.* 9:e00418. doi: 10.3389/fnhum.2015.00418

Roy, S., Kiral-Kornek, I., and Harrer, S. (2019). *ChronoNet: A Deep Recurrent Neural Network for Abnormal EEG Identification*. Berlin: Springer International Publishing, 47–56.

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

Savva, A. D., Mitsis, G. D., and Matsopoulos, G. K. (2019). Assessment of dynamic functional connectivity in resting-state fMRI using the sliding window technique. *Brain Behav.* 9:e01255. doi: 10.1002/brb3.1255

Schilbach, L., Eickhoff, S. B., Rotarska-Jagiela, A., Fink, G. R., and Vogeley, K. (2008). Minds at rest? Social cognition as the default mode of cognizing and its putative relationship to the "default system" of the brain. *Conscious. Cogn.* 17, 457–467. doi: 10.1016/j.concog.2008.03.013

Schreiber, T., and Schmitz, A. (2000). Surrogate time series. *Physica D* 142, 346–382. doi: 10.1016/S0167-2789(00)00043-49

Sestieri, C., Corbetta, M., Romani, G. L., and Shulman, G. L. (2011). Episodic memory retrieval, parietal cortex, and the default mode network: functional and topographic analyses. *J. Neurosci.* 31, 4407–4420. doi: 10.1523/jneurosci.3335- 10.2011

Shen, X., Tokoglu, F., Papademetris, X., and Constable, R. T. (2013). Groupwise whole-brain parcellation from resting-state fMRI data for network node identification. *NeuroImage* 82, 403–415. doi: 10.1016/j.neuroimage.2013.05.081

Shirer, W., Ryali, S., Rykhlevskaia, E., Menon, V., and Greicius, M. (2012). Decoding subject-driven cognitive states with whole-brain connectivity patterns. *Cereb. Cortex* 22, 158–165. doi: 10.1093/cercor/bhr099

Siegel, J. S., Mitra, A., Laumann, T. O., Seitzman, B. A., Raichle, M., Corbetta, M., et al. (2017). Data quality influences observed links between functional connectivity and behavior. *Cereb. Cortex* 27, 4492–4502. doi: 10.1093/cercor/ bhw253

Skåtun, K. C., Kaufmann, T., Doan, N. T., Alnæs, D., Córdova-Palomera, A., Jönsson, E. G., et al. (2017). Consistent functional connectivity alterations in schizophrenia spectrum disorder: a multisite study. *Schizophr. Bull.* 43, 914–924. doi: 10.1093/schbul/sbw145

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

Thompson, G. J. (2018). Neural and metabolic basis of dynamic resting state fMRI. *Neuroimage* 180, 448–462. doi: 10.1016/j.neuroimage.2017.09.010

Tobia, M. J., Hayashi, K., Ballard, G., Gotlib, I. H., and Waugh, C. E. (2017). Dynamic functional connectivity and individual differences in emotions during social stress. *Hum. Brain Map.* 38, 6185–6205. doi: 10.1002/hbm.23821

Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E., Yacoub, E., Ugurbil, K., et al. (2013). The WU-Minn human connectome project: an overview. *Neuroimage* 80, 62–79. doi: 10.1016/j.neuroimage.2013.05.041

Vidaurre, D., Smith, S. M., and Woolrich, M. W. (2017). Brain network dynamics are hierarchically organized in time. *Proc. Natl. Acad. Sci. U.S.A.* 114, 12827– 12832. doi: 10.1073/pnas.1705120114

Weis, S., Hodgetts, S., and Hausmann, M. (2019). Sex differences and menstrual cycle effects in cognitive and sensory resting state networks. *Brain Cogn.* 131, 66–73. doi: 10.1016/j.bandc.2017.09.003

Yan, W., Calhoun, V., Song, M., Cui, Y., Yan, H., Liu, S., et al. (2019). Discriminating schizophrenia using recurrent neural network applied on time courses of multi-site FMRI data. *EBioMedicine* 47, 543–552. doi: 10.1016/j. ebiom.2019.08.023

Yan, W., Plis, S., Calhoun, V. D., Liu, S., Jiang, R., Jiang, T., et al. (2017). "Discriminating schizophrenia from normal controls using resting state functional network connectivity: a deep neural network and layer-wise relevance propagation method," in *Proceedings of the 2017 IEEE 27th International Workshop on Machine Learning for Signal Processing*, (Piscataway, NJ: IEEE), 1–6.

Yuan, J., Li, X., Zhang, J., Luo, L., Dong, Q., Lv, J., et al. (2018a). Spatio-temporal modeling of connectome-scale brain network interactions via time-evolving graphs. *Neuroimage* 180, 350–369. doi: 10.1016/j.neuroimage.2017.10.067

Yuan, L., Wei, X., Shen, H., Zeng, L., and Hu, D. (2018b). Multi-Center Brain Imaging Classification Using a Novel 3D CNN Approach. *IEEE Access* 6, 49925–49934. doi: 10.1109/ACCESS.2018.2868813

Zeiler, M. D., and Fergus, R. (2014). *Visualizing and Understanding Convolutional Networks*. Berlin: Springer International Publishing, 818–833.

Zeng, L.-L., Wang, H., Hu, P., Yang, B., Pu, W., Shen, H., et al. (2018). Multi-site diagnostic classification of schizophrenia using discriminant deep learning with functional connectivity MRI. *EBioMedicine* 30, 74–85. doi: 10.1016/j.ebiom.2018.03.017

Zhang, C., Cahill, N. D., Arbabshirani, M. R., White, T., Baum, S. A., and Michael, A. M. (2016). Sex and age effects of functional connectivity in early adulthood. *Brain Connect.* 6, 700–713. doi: 10.1089/brain.2016.0429

Zhang, C., Dougherty, C. C., Baum, S. A., White, T., and Michael, A. M. (2018). Functional connectivity predicts gender: evidence for gender differences in resting brain connectivity. *Hum. Brain Mapp.* 39, 1765–1776. doi: 10.1002/hbm.23950

Zheng, H., Li, F., Bo, Q., Li, X., and Wu, X. (2017). The dynamic characteristics of the anterior cingulate cortex in resting-state fMRI of patients with depression. *J. Affect Disord* 227, 391–397. doi: 10.1016/j.jad.2017.11.026

Zhu, Y., Zhu, X., Zhang, H., Gao, W., Shen, D., and Wu, G. (2016). "Reveal consistent spatial-temporal patterns from dynamic functional connectivity for autism spectrum disorder identification," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, (Berlin: Springer), 106–114. doi: 10.1007/978-3-319-46720-7_13