HHS Public Access

Author manuscript

Med Image Anal. Author manuscript; available in PMC 2024 April 01.

Computing personalized brain functional networks from fMRI using self-supervised deep learning

Hongming Li[a,b], **Dhivya Srinivasan**[a,b], **Chuanjun Zhuo**[c], **Zaixu Cui**[d], **Raquel E. Gur**[a,b,f,h,i], **Ruben C. Gur**[a,f,h,i], **Desmond J. Oathes**[e,f], **Christos Davatzikos**[a,b], **Theodore D. Satterthwaite**[f,g,h], **Yong Fan**[a,b,*]

[a]Center for Biomedical Image Computation and Analytics, University of Pennsylvania, Philadelphia, PA 19104, USA

[b]Department of Radiology, University of Pennsylvania, Philadelphia, PA 19104, USA

[c]Key Laboratory of Brain Circuit Real Time Tracing (BCRTT-Lab), Beijing, 102206, China

[d]Tianjin University Affiliated Tianjin Fourth Center Hospital, Department of Psychiatry, Tianjin Medical University, Tianjin, China Chinese Institute for Brain Research, Beijing, 102206, China

[e]Center for Neuromodulation in Depression and Stress, University of Pennsylvania, Philadelphia, PA 19104, USA

[f]Department of Psychiatry, University of Pennsylvania, Philadelphia, PA 19104, USA

[g]Penn Lifespan Informatics and Neuroimaging Center, University of Pennsylvania, Philadelphia, PA 19104, USA

[h]Penn/CHOP Lifespan Brain Institute, University of Pennsylvania, Philadelphia, PA 19104, USA

[i]Department of Neurology, University of Pennsylvania, Philadelphia, PA 19104, USA

Abstract

A novel self-supervised deep learning (DL) method is developed to compute personalized brain functional networks (FNs) for characterizing brain functional neuroanatomy based on functional MRI (fMRI). Specifically, a DL model of convolutional neural networks with an encoder-decoder architecture is developed to compute personalized FNs directly from fMRI data. The DL model is trained to optimize functional homogeneity of personalized FNs without utilizing any external supervision in an end-to-end fashion. We demonstrate that a DL model trained on fMRI scans from the Human Connectome Project can identify personalized FNs and generalizes well across four different datasets. We further demonstrate that the identified personalized FNs are informative for predicting individual differences in behavior, brain development, and schizophrenia status.
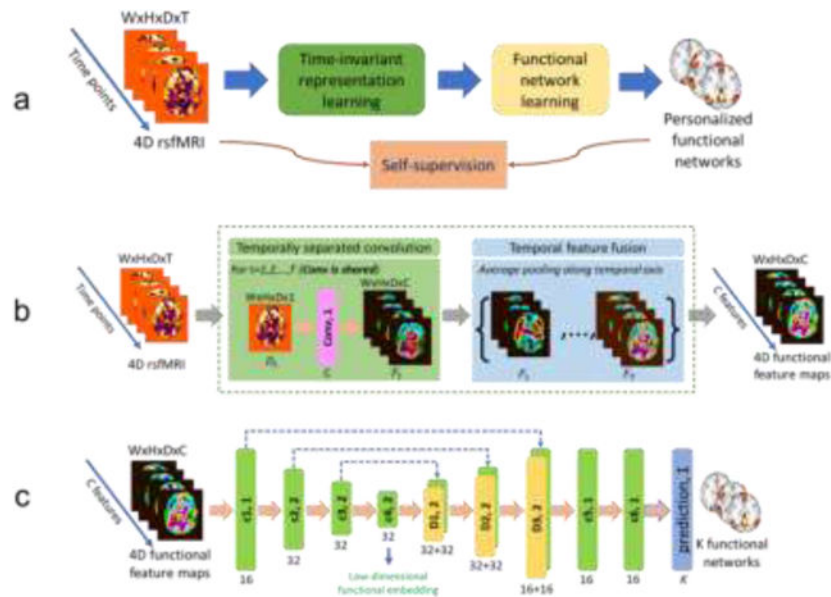
*Correspondence: Yong.Fan@pennmedicine.upenn.edu.

Taken together, the self-supervised DL allows for rapid, generalizable computation of personalized FNs.

Graphical Abstract

Keywords

Brain functional networks; personalized; self-supervised learning

INTRODUCTION

Functional magnetic resonance imaging (fMRI) is a powerful tool to study functional networks (FNs) in the human brain. Accumulating evidence demonstrates that FNs undergo predictable normative development in youth and abnormal development is associated with diverse psychopathology (Grayson and Fair, 2017). However, most of our knowledge of FNs is drawn from studies that used "one size fits all" group atlases that are not tailored to the individual. Recent convergent evidence from multiple independent efforts has established that FNs are indeed person-specific. These studies emphasize that there is marked inter-individual variability in the *functional topography* of FNs, which is defined as the location and spatial arrangement of FNs on the anatomic cortex (Bijsterbosch et al., 2019; Bijsterbosch et al., 2018; Cui et al., 2020; Glasser et al., 2016; Gordon et al., 2017; Kong et al., 2018; Laumann et al., 2015; Li et al., 2017; Li et al., 2019; Tavor et al., 2016). Notably, it has been shown that the greatest variability in functional topography is present in association networks such as the fronto-parietal control network (FPN), the default mode network (DMN), and the ventral attention network (VAN). Recent studies have also revealed that personalized FNs are refined in adolescence and predict executive function (Cui et al., 2020), individual differences in functional network topography in association networks are related to overall psychopathology in youth (Cui et al., 2022), the normative brain development is associated with substantial sex differences in the

functional topography of personalized association networks (Shanmugan et al., 2021), and the development of personalized FNs aligns with and refines a hierarchy linked to cognition (Pines et al., 2022). It has also been shown that personalized FNs facilitate effective targeting in brain stimulation studies (Ren et al., 2022; Wang et al., 2015). All these findings have demonstrated that personalized FNs provide a promising way to investigate functional organization of the human brain and its associations with behavior and brain disorders. Group-level definitions of FNs used in the vast majority of translational studies fail to account for such inter-individual variability of FNs, inevitably mixing signals from disparate FNs, and reducing sensitivity in clinical applications.

An alternative to using group-level atlases is to create personalized FNs (Abraham et al., 2013; Beckmann et al., 2005; Calhoun et al., 2001; Daubechies et al., 2009; Du and Fan, 2011, 2013; Harrison et al., 2020; Harrison et al., 2015; Kong et al., 2018; Lee et al., 2011; Lee et al., 2008; Li et al., 2017; Yeo et al., 2014). To enforce correspondence of FNs across different individuals, these methods typically compute personalized FNs under certain constraints to be similar to group level FNs, such as those built upon independent component analysis (ICA) (Beckmann et al., 2005; Calhoun et al., 2001) or based on assumptions that loadings of corresponding FNs of different individuals follow certain distributions (Abraham et al., 2013; Harrison et al., 2020; Harrison et al., 2015; Kong et al., 2018). However, such techniques may yield biased results since little is known about FNs' statistical distribution. Such limitations have been partially overcome using spatially-regularized non-negative matrix factorization (Li et al., 2017). Recent studies have also demonstrated that deep learning methods, such as deep autoencoders, can be used to learn low-dimensional representations of fMRI data (Brown et al., 2020; Hjelm et al., 2014; Hu et al., 2018; Jang et al., 2017; Kim et al., 2021; Plis et al., 2014). However, deep learning models built by these methods do not generate personalized FNs in that their outputs are optimized to approximate their inputs. Supervised deep learning methods have been developed to predict personalized FNs, such as DMN (Zhao et al., 2020). However, CNNs in the supervised learning setting have to be trained with personalized FNs obtained by conventional brain decomposition methods as ground truth.

Since personalized FNs account for inter-individual differences in functional neuroanatomy (Satterthwaite and Davatzikos, 2015), we developed a novel self-supervised deep learning (DL) method for computing personalized FNs from fMRI data while maintaining inter-individual correspondence. We hypothesized that learning an effective intrinsic representation of FNs in a low-dimensional latent space would facilitate the computation of bias-free personalized FNs with improved functional homogeneity. We adopted convolutional neural networks (CNNs) with an encoder-decoder architecture to identify individual-specific FNs in an end-to-end learning fashion. In this framework, the encoder is used to learn an effective intrinsic representation of FNs in a low-dimensional latent space and the decoder is used to directly generate personalized FNs, instead of approximations of input data (Brown et al., 2020; Hjelm et al., 2014; Hu et al., 2018; Jang et al., 2017; Kim et al., 2021; Plis et al., 2014), from the low-dimensional representation. We trained the encoder-decoder CNNs using self-supervised information including functional homogeneity and spatial sparsity (Li et al., 2017); this trained model could then be applied to fMRI data of an unseen individual's data to identify FNs in one forward-pass computation. We have

trained a DL model on resting-state fMRI (rsfMRI) data of individuals from the Human Connectome Project (HCP) (Van Essen et al., 2013), and applied it to rsfMRI data of new individuals from different datasets including healthy individuals with different age ranges and patients with schizophrenia. Results reveal that the DL model could obtain personalized FNs which coincide with well-established large-scale FNs. Furthermore, these personalized FNs were informative for predicting behavioral measures and brain age of the healthy individuals, and also could distinguish schizophrenia patients from healthy individuals. As described below, these results indicate that the personalized FNs identified by our DL method accurately capture the variability of personalized functional neuroanatomy across diverse datasets in a way that meaningfully captured individual differences in behavior and brain health. As far as we know, this proposed method is the first study demonstrating self-supervised deep learning is capable of identifying personalized FNs from resting-state fMRI data, a shift from iterative optimization or supervised learning based methodologies.

METHODS

Our self-supervised DL framework includes a representation learning module and a functional network learning module for identifying personalized FNs (Fig. 1a). The representation learning module (Fig. 1b) learns time-invariant feature maps from the input fMRI, and the functional network learning module (Fig. 1c) with an Encoder-Decoder architecture computes personalized FNs from these feature maps by optimizing self-supervision measures including functional homogeneity and spatial sparsity so that the network is trained and optimized based on the input fMRI data without any external supervision.

Functional brain decomposition based on matrix factorization.

Given rsfMRI data was measured with partial correlation coefficient $X^i \in R^{T \times S}$ of individual i, consisting of S voxels and T time points, a matrix factorization based decomposition model tries to identify K FNs $V^i = \left(V_{k,s}^i\right) \in R^{K \times S}$ and their corresponding time courses $U^i = \left(U_{t,k}^i\right) \in R^{T \times K}$, so that the original rsfMRI data can be approximated by two low-rank matrices $X^i \approx U^i V^i$. To identify FNs that do not contain anti-correlated functional units and are spatially compact and overlapped, non-negative constraint and spatial sparsity regularization are usually applied to V^i (Li et al., 2017), and the FNs can be obtained by optimizing

$$\min_{U^i, V^i} \left\| X^i - U^i V^i \right\|_F^2 + \lambda \sum_{j=1}^K \frac{\sum_{s=1}^S \left| v_{j,s}^i \right|}{\sqrt{\sum_{s=1}^S \left(V_{j,s}^i\right)^2}}, \qquad s.t. \ V^i \in R_+, \tag{1}$$

where the first term is a data fitting term, the second term is the Hoyer regularization for spatial sparsity (Hoyer, 2004), and λ is the trade-off parameter.

An alternative optimization method is usually adopted to optimize Eq. (1) (Ding et al., 2008). When V^i is fixed, U^i can be calculated analytically as $U^i = X^i \left(V^i\right)^T \left(V^i \left(V^i\right)^T\right)^{-1}$. Substituting this expression for U^i in Eq. (1), we have

$$\min_{V^i} \left\| X^i - X^i(V^i)^T \left(V^i(V^i)^T\right)^{-1} V^i \right\|_F^2 + \lambda \sum_{j=1}^{K} \frac{\sum_{s=1}^{S} |V_{j,s}^i|}{\sqrt{\sum_{s=1}^{S}(V_{j,s}^i)^2}}, \qquad s.t. \; V^i \in R_+. \qquad (2)$$

Instead of optimizing V^i by the multiplicative weight update method (Ding et al., 2008), we propose to train a deep convolutional neural networks (CNNs) to estimate V^i directly from the input rsfMRI data as shown in Fig. 1a.

Self-supervised deep learning of personalized FNs.

Personalized FNs are computed from rsfMRI using CNNs with an Encoder-Decoder architecture under a self-supervised deep learning framework by optimizing data fitting and sparsity regularization terms that are commonly used in brain decomposition models (Li et al., 2017; Li et al., 2018). A trained model can be applied to rsfMRI data of a new individual to identify personalized FNs in one forward-pass computation. As illustrated in Fig. 1, we adopt an Encoder-Decoder U-Net (Ronneberger et al., 2015) to identify personalized FNs from rsfMRI data by optimizing data fitting and spatial sparsity regularization terms (Li et al., 2017; Li et al., 2018) so that the network is trained and optimized based on the input rsfMRI data without any external supervision. The Encoder-Decoder CNNs facilitate the inter-individual comparability of FNs implicitly by an effective representation of intrinsic FNs in a low-dimensional latent space characterized by the Encoder-Decoder's bottleneck layer.

Given a group of n individuals, each having 4D rsfMRI data $I^i \in R^{W \times H \times D \times T}$, $i = 1, \ldots, n$, where W, H, and D are width, height, and depth of each 3D volume respectively and T is the number of time points, we train a deep network of 3D CNNs $M_{\theta_c}(I^i) = F^i$ with convolutional parameters θ_c, which takes the 4D fMRI data I^i as input and identifies its corresponding FNs $F^i \in R^{W \times H \times D \times K}$ as output, where K is the number of FNs. We use rsfMRI data from n individuals as training data to minimize the loss function in Eq. (2) for learning the convolutional parameters θ_c. When calculate the loss function, I^i is flattened across the spatial dimension, transposed, and reshaped as a matrix X^i with size $T \times S$, where $S = W \times H \times D$. The same transformation is applied to F^i to get the reshaped matrix V^i with size $K \times S$. As FNs are optimized as formulated by Eq. (2), which does not rely on any external supervision, the deep learning model is optimized based on input data alone, referred to as self-supervised. Once θ_c is optimized, the deep learning model can be used to predict personalized FNs for different individuals. It is worth noting that our deep learning model is optimized to directly generate personalized FNs, instead of approximations of the input data.

The deep network includes a representation learning module and a functional network learning (U-Net (Ronneberger et al., 2015) like) module for identifying FNs, as illustrated in Fig. 1a. The representation learning module (with details in next paragraph) extracts time-invariant feature maps from the input rsfMRI using 3D CNNs, and the feature maps are used as input to the functional network learning module. As shown in Fig. 1c, the functional network learning module consists of one convolutional layer with 16 filters, three

convolutional layers with 32 filters and a stride of 2, three deconvolutional layers with 32, 32, and 16 filters and a stride of 2, and finally 2 convolutional layers with 16 filters. Leaky ReLU activation and batch normalization (Ioffe and Szegedy, 2015) are used for all the convolutional and deconvolutional layers. One output convolutional layer is used to predict personalized FNs and the number of output channels is K. Rectified linear activation function (ReLU) activation is used for the output layer so that the output is non-negative. Each output channel is linearly scaled so that its maximum equals to 1 before calculating the loss function. The kernel size in all layers is set to $3 \times 3 \times 3$.

Inspired by a prior study (Liu and Duyn, 2013) which demonstrated that replication of FNs can be obtained through averaging of selected fMRI time frames, we propose to learn features from each individual time point of rsfMRI scans and fuse the features of all time points for learning personalized FNs by the following functional network learning module. As shown in Fig. 1b, the representation learning module consists of two parts referred as temporally separated convolution and temporal feature fusion. The temporally separated convolution part is a 3D convolutional layer with C filters and a stride of 1, which is applied to each time point (a 3D volume with size of $W \times H \times D$) of the rsfMRI scan and output C feature maps (with size of $W \times H \times D \times C$) for each time point. The temporal feature fusion part is an element-wise average pooling layer which outputs C average feature maps of all the time points. In the present study, C is set to 16 and the kernel size of the convolutional layer is set to $3 \times 3 \times 3$, and Leaky ReLU activation and batch normalization are also adopted. While the convolution learning part is expected to capture the co-activation patterns of different brain regions by different feature channels at each time point, the feature fusion part summarizes the patterns obtained from all the time points. This module is plugged into the model just before the functional network learning module, and the whole model can be optimized in an end-to-end fashion.

In the present study, the number of FNs is set to 17 in order to facilitate a direct comparison with the well-established FNs (Yeo et al., 2011). Our model is implemented using Tensorflow. Adam optimizer is adopted to optimize the network, the learning rate is set to 1×10^{-4}, the batch size is set to 1, and the number of iterations is set to 30000 during training. One NVIDIA TITAN Xp GPU is used for training and testing. We set $\lambda = 10$ empirically in our experiments.

EXPERIMENTS AND RESULTS

Imaging data.

One simulated functional dataset and four resting-state fMRI (rsfMRI) datasets from different cohorts were used to develop and validate the proposed method in the current study, including data from the Human Connectome Project (HCP), the Philadelphia Neurodevelopmental Cohort (PNC), the Pediatric Imaging, Neurocognition, and Genetics (PING) Data Repository, and one cohort of healthy controls and schizophrenia patients (SCZ).

Simulated functional dataset.—The simulated functional dataset was generated using the SimTB toolbox (Erhardt et al., 2012). In particular, the dataset consisted of simulated data of 100 individuals, and every individual had 120 2D images with 128×128 pixels, generated by linear combinations of 20 distinct FNs. Spatial variability in translation, rotation, and spread of FNs was adopted during the data generation as inter-subject variability. Rician noise was added to the simulated images with random contrast-to-noise ratios (CNR) ranging from 0.65 to 1.0 for different subjects. This simulated dataset provided ground-truth FNs that can be compared with those computed by different methods for evaluating their modeling accuracy.

HCP dataset.—This dataset consisted of 1078 healthy individuals from the HCP cohort (Van Essen et al., 2013). For each individual, two rsfMRI sessions (left-to-right direction phase encoding) were included. Details of the scanning protocols of the rsfMRI data has been published previously (Glasser et al., 2013). Frame-wise displacement (FD) was adopted to measure volume-to-volume changes in head position and participants with large head-movements (mean FD > 0.25 mm) were excluded. The ICA-FIX denoised rsfMRI in volumetric space were used, and the denoised rsfMRI data was spatially smoothed with a 6-mm full-width half-maximum (FWHM) kernel and downsampled at a spatial resolution of $3 \times 3 \times 3$ mm^3. The image data was downsampled to make its spatial resolution consistent with that of fMRI data from other 3 cohorts.

PNC.—This dataset consisted of 969 young individuals (ages from 8 to 23 years, 535 females) from the PNC (Satterthwaite et al., 2014). For each individual, a rsfMRI scan was acquired using 3T Siemens Tim Trio whole-body scanner with single-shot, interleaved multi-slice, gradient-echo, echo planar imaging (GE-EPI) sequence sensitive to BOLD contrast (TR = 3000 ms, TE = 32 ms; flip angle = 90°, FOV = 192×192 mm^2, matrix = 64×64, 46 slices, slice thickness/gap = 3/0 mm, effective voxel resolution = $3 \times 3 \times 3$ mm^3, 124 volumes/TRs). The rsfMRI data were preprocessed using an optimized procedure, including removal of the initial volumes, slice timing correction, 36-parameter confound regression, and band-pass filtering (Ciric et al., 2017). Participants with large head-movements (mean FD > 0.25 mm) were excluded.

PING dataset.—This dataset consisted of 210 young individuals whose MRI scans were acquired using 3T Siemens scanners (ages from 3 to 20 years, 106 females) from the PING cohort (Jernigan et al., 2016). For each individual, a rsfMRI scan with varied effective voxel resolutions of $4 \times 4 \times 4$ mm^3 or $3.75 \times 3.75 \times 3.5$ mm^3 was acquired using imaging sequence parameters optimized for equivalence in contrast properties and consistency in image-derived quantitative measures with integrated distortion correction. The rsfMRI data were preprocessed using the same procedure as the PNC dataset. Participants with large head-movements (mean FD > 0.25 mm) were excluded.

SCZ dataset.—This cohort consisted of 195 individuals (101 healthy controls, and 94 SCZ patients) from previous studies (Jing et al., 2019; Zhuo et al., 2017). All the subjects underwent structural MRI and rsfMRI scanning with a 3.0-Tesla MR system (Discovery MR750, General Electric, Milwaukee, WI, USA). Tight but comfortable foam padding was

used to minimize head motion, and earplugs were used to reduce scanner noise. Sagittal 3D T1-weighted images were acquired using a brain volume sequence (TR/TE/TI=8.2/3.2/450 ms; FA=12°; FOV=256×256 mm; matrix=256×256; slice thickness=1 mm, no gap; and 188 sagittal slices). A gradient-echo single-short echo planar imaging sequence was used to collect rsfMRI data (TR/TE=2000/45 ms; FOV=220×220 mm; matrix=64×64; FA=90°; slice thickness=4mm; gap=0.5mm; 32 interleaved transverse slices; and 180 volumes). All subjects were instructed to keep their eyes closed, relax, move as little as possible, think of nothing in particular, and not fall asleep during the rsfMRI scanning. All images were visually checked for artefacts, structural abnormalities, and pathologies by neuroradiologists. The rsfMRI scans were preprocessed using fMRIPrep (Esteban et al., 2019) (version: 20.2.1, https://github.com/nipreps/fmriprep/archive/20.2.1.tar.gz) and head motion artifacts were removed using ICA-AROMA (Pruim et al., 2015). Participants with large head-movements (mean FD > 0.25 mm) were excluded. The final dataset comprised data from 101 healthy controls (HC, age: 33.86 ± 10.77 years, 58 females) and 94 schizophrenia (SCZ, age: 33.23 ± 8.03 years, 41 females) subjects. There is no significant difference between the healthy controls and schizophrenia subjects regarding age ($p = 0.647$, two-sample T-test), gender ($p = 0.0539$, Chi-squared test), or head motion ($p = 0.3069$, two-sample T-test).

All the preprocessed rsfMRI data across datasets were in the MNI space and with spatial resolution of $3 \times 3 \times 3$ mm^3 so that the DL model worked with reception field of the same size (in convolutional operations) on all datasets.

Functional homogeneity.

For each FN, a homogeneity measure is calculated as the weighted sum of the correlation coefficients between the time courses of all the voxels within the FN and its centroid time course, which is a weighted mean time course within the FN and the FN's voxelwise loadings are used weights. The median value of FN-wise homogeneity measures is used to gauge functional homogeneity of FNs for each subject.

Sanity testing for personalized FNs computed by the proposed DL model.

Sanity testing was performed for quality assurance of the personalized FNs in terms of both functional homogeneity and spatial correspondence. First, the personalized FNs was compared with their corresponding group average FNs in terms of within-network functional homogeneity. Specifically, given a subject with preprocessed fMRI data, functional homogeneity measures are computed based on its personalized FNs and the group average FNs. If a subject's personalized FNs have higher functional homogeneity than the group average FNs, then the sanity testing is passed in terms of the functional homogeneity. Second, the personalized FNs' spatial correspondence across individuals is tested. Instead of comparing FNs across individuals directly, each individual subject's FNs are compared with group average FNs based on their spatial correlation coefficients. Specifically, one personalized FN is deemed to maintain correspondence with its corresponding group level FN if $\Delta Sim_i = Corr(FN_i^p, FN_i^g) - \max_{j, j \neq i} Corr(FN_i^p, FN_j^g) > 0$, where FN_i^p denotes the i-th personalized FN, FN_i^g denotes its corresponding group average FN, FN_j^g denotes other group average FNs, and $Corr(\cdot, \cdot)$ is spatial correlation coefficient between two FNs. If a subject's all personalized FNs maintain the correspondence with their group average FNs, then the

sanity testing is passed in terms of the spatial correspondence. The group average FNs are computed by averaging the FNs across subjects without any other operations, such as thresholding or binarization, and the group-level FNs were only used for the sanity testing.

Ablation experiments.

We compared the proposed DL model with its degraded version without the time-invariant representation learning module, which does not explicitly take into consideration the correspondence of resting-state functional dynamics across subjects. These two DL models shared the same Encoder-Decoder network architecture of the functional network learning module of the proposed model (Fig. 1c), but the degraded version's input was the original fMRI (4D volume data). We compared these two models based on both simulated functional dataset and HCP dataset.

Self-supervised DL model could identify FNs on simulated data with high accuracy.

We trained the proposed DL model (Fig. 1) and its degraded version using simulated functional data of 80 randomly selected individuals and evaluate the DL models using data of the remaining 20 individuals. Pearson correlation coefficient between a computed FN and its corresponding ground truth FN was adopted to gauge its spatial accuracy, and an overall spatial accuracy for each individual was computed as an average value of correlation coefficients of all FNs. The FNs computed by the DL models were matched with the ground truth FNs using the Hungarian algorithm (Carpaneto and Toth, 1980). Personalized FNs of three randomly selected testing individuals from the simulated dataset are illustrated in Fig. 2. The proposed DL model could successfully capture the differences in location, rotation, and spatial extent of different FNs across subjects, with an average spatial correlation of 0.959 ± 0.008, indicating that it could effectively learn the representation that informative to predict personalized FNs. The degraded DL model without time-invariant learning was not able to identify the FNs accurately, with an average spatial correlation of 0.118 ± 0.073. The inaccurate result was mainly due to that the temporal dimensions of resting-state functional dynamics across subjects were not inherently aligned as task-evoked fMRI, and the degraded model without time-invariant learning could not generalize well to unseen data. This demonstrated the effectiveness of our proposed time-invariant representation learning module.

Self-supervised DL model could identify FNs which coincide with well-established FNs.

We trained the proposed DL model (Fig. 1) on rsfMRI data of 400 individuals randomly selected from the HCP cohort. The number of FNs was set to 17 in order to facilitate a direct comparison with well-established FNs (Yeo et al., 2011). The trained DL model was applied to another 678 HCP individuals served as testing data for evaluation. As two 3T rsfMRI sessions were acquired for each individual, we have adopted the first session as the primary dataset (HCP REST1) and used the second session for replication (HCP REST2). The personalized FNs of all testing individuals computed by the DL model passed the sanity testing with respect to both functional homogeneity and spatial correspondence. Particularly, the personalized FNs had higher within-network functional homogeneity than the group average FNs of all testing individuals and maintained spatial correspondence with their

corresponding group average FNs (Fig. S1). The average FNs of the 678 testing individuals identified by our DL model on the REST1 dataset are illustrated in Fig. 3. It can be observed that the DL model successfully identified both spatially localized and distributed FNs, including visual networks, somatomotor networks, dorsal attention networks (DAN), ventral attention networks (VAN), fronto-parietal networks (FPN), default mode networks (DMN). These identified FNs show high correspondence to the well-established functional atlases (Yeo et al., 2011). Using a conservative spatial permutation test based on spatial-correlation-preserving surrogate maps generated by BrainSMASH (Burt et al., 2020), we found a significant alignment ($p < 0.001$) between the FNs identified by the DL model and the canonical networks from the Yeo atlas (Yeo et al., 2011).

Self-supervised DL model could identify personalized FNs with improved functional homogeneity.

Personalized FNs of three randomly selected testing individuals from HCP REST1 dataset are illustrated in Fig. 4a. Inter-individual differences can be clearly observed in the personalized FNs, illustrating that the proposed method captured the inherent differences of individualized functional neuroanatomy. We also compared the personalized FNs computed using the proposed DL model and a top-performing spatially-regularized NMF method (Li et al., 2017) in terms of their within-network functional homogeneity (Li et al., 2017). The personalized FNs obtained by the DL model had significantly higher functional homogeneity than those computed using the spatially-regularized NMF (Fig. 4b, $p < 10^{-5}$, Wilcoxon signed rank test). This increase in functional homogeneity may happen because no specific regularization was adopted to enforce the FNs to follow a prior distribution – facilitating better characterization of inter-individual variability in functional neuroanatomy.

Personalized FNs from the DL model are highly reproducible.

To evaluate the reproducibility of personalized FNs computed by the DL model, we compared personalized FNs of the same testing subjects, computed from their rsfMRI scans collected in different imaging sessions, including the HCP REST1 (R1) and REST2 (R2) datasets. Default mode network and fronto-parietal network of three randomly selected testing individuals are shown in Fig. 5a, illustrating that the intra-subject (inter-session) differences between corresponding FNs are visually smaller than the inter-subject counterpart. To evaluate this quantitatively, we further tested if the personalized FNs computed from two imaging sessions (R1 and R2) of the same subjects are more similar than those of different subjects so that FNs can be used to identify subjects. As illustrated in Fig. 5b, each subject in R1 was compared with all subjects in R2 to identify the individual with the maximally similar FNs. Average spatial correlation coefficient across FNs was used as the similarity metric. The subject identification was considered as correct if the maximally similar FNs between the two sessions were from rsfMRI scans of the same subject. The identification rate was calculated as the ratio between the number of correctly identified target individuals and the total number of target individuals. The identification procedures were carried out in two directions, i.e., identification of subjects of R1 from R2 (R1->R2) and vice versa (R2->R1). The identification rate was 93.6% and 97.4%, respectively for the procedures of R1->R2 and R2->R1 when all 17 FNs were used for the identification (Fig. 5c), which was similar to that obtained by the functional connectome

fingerprinting study (Finn et al., 2015). Since the inter-subject functional topography variability is maximal in association networks (Cui et al., 2020; Kong et al., 2018; Li et al., 2019; Wang et al., 2015), the same subject identification procedure was also carried out based on the fronto-parietal FNs alone (as illustrated in Fig. 3), and improved identification rates were obtained, with 98.3% and 98.8% respectively for the procedures of R1->R2 and R2->R1 (Fig. 5c). The subject identification performance is on par with that obtained on the HCP subjects in the functional connectome fingerprinting study (Finn et al., 2015), although our study used a much larger testing dataset. These results indicate that the DL model could identify reproducible and subject-specific FNs that capture variations of inter-subject functional neuroanatomy.

We further compared the proposed DL model with its degraded version for subject identification, and a higher identification rate is expected if a DL model can better identify the personalized FNs. The degraded DL model was trained using the same setting as our proposed DL model. The DL models were trained on rsfMRI data of the same set of 400 randomly selected individuals and evaluated on the remaining 678 individuals. Similar to that on the simulated dataset, the degraded DL model was not able to accurately identify personalized FNs across subjects, with an identification rate of 0.006 (R1->R2) and 0.002 (R2->R1).

Moreover, we investigated the performance of the personalized functional embedding (FE) for subject identification, where the personalized FE was obtained as the feature maps output at the bottleneck layer in the proposed DL model (as illustrated in Fig. 1c). Particular, the personalized FE was flattened into a vector, and the Pearson correlation coefficient between FEs were used as inter-subject similarity metric. The same identification procedure as Fig. 5b was adopted. The identification rate was 88.3% and 88.3%, respectively for the procedures of R1->R2 and R2->R1, demonstrating that the Encoder-Decoder CNNs effectively encodes the functional data into a low-dimensional latent space that captures inter-subject variability, which facilitate the decoder to identify the inter-subject variations in FNs. The identification rate based on FEs was lower than that based FNs, which might be due to that the bottleneck feature maps contain less detailed information than the FNs for characterizing the personalized functional neuroanatomy.

Personalized FNs from DL model are associated with behavior.

To determine whether the individual variations in FNs are behaviorally meaningful, we investigated whether personalized FNs identified by the DL model could be used to predict individuals' behavior and cognitive performance, with a focus on 13 cognitive measures highlighted in the HCP data dictionary (Kong et al., 2018). The loading coefficients of all 17 FNs were concatenated into a feature vector for each individual, and ridge regression with 2-fold cross-validation was used to predict the behavior measures respectively. The regularization parameter in ridge regression was optimized in the range $[2^{-10}, 2^{-9}, \ldots, 2^4, 2^5]$ using nested 2-fold cross-validation within the training cohort. Partial correlation coefficient between the predicted and real measures was adopted to evaluate the prediction accuracy, with age, sex, and in-scanner motion as co-variates. The prediction for each behavior measure was repeated 100 times. As shown in Fig. 5d, the cognitive measures were

predicted with an average accuracy of 0.1092 ± 0.0097 (7 out of 13 were significant, $p < 0.05$, permutation test), in par with the prediction performance obtained in a previous study (Kong et al., 2018). These results indicated that the DL model was able to identify the variability of personalized FNs that meaningfully captured individual differences in behavior.

Self-supervised DL model robustly generalizes to new datasets.

To determine whether the DL model was generalizable to individuals whose fMRI scans were collected with imaging protocols/parameters different from the fMRI data used to train the DL model, we applied the model built on the HCP training cohort to fMRI data from two external datasets. Notably, both datasets were different from the HCP cohort in terms of age, scan length, and scanning protocol. The first cohort contained 969 individuals (ages from 8 to 23) from the PNC (Satterthwaite et al., 2014), and the second cohort included 210 individuals (ages from 3 to 20) from the PING cohort (Jernigan et al., 2016). The personalized FNs of all those individuals computed by the DL model passed the sanity testing with respect to both functional homogeneity and spatial correspondence (Fig. S1). Five representative personalized FNs of three randomly selected individuals from the PNC and PING cohort are shown in Fig. 6a and Fig. S2a, respectively. It can be observed that the DL model generalized well to individuals from different cohorts, with both localized and distributed FNs identified successfully. Differences in FNs could also be clearly observed across individuals for both cohorts. The FNs obtained by the DL model had significantly higher functional homogeneity than those computed using the spatially-regularized NMF (Fig. 6b and Fig. S2b, $p < 10^{-5}$ Wilcoxon signed rank test).

To further determine whether these personalized FNs are biologically meaningful, we investigated whether personalized FNs identified by the DL model could be used to predict individuals' age, as a recent study demonstrated that functional topography could encode development (Cui et al., 2020). The same strategy as the prediction of behavior phenotypes on HCP cohort was adopted for the age prediction. The prediction accuracy was measured with partial correlation coefficient between the predicted brain age and chronological age with sex and in-scanner motion as co-variates, in addition to mean absolute error (MAE). The prediction results shown in Fig. 6c and Fig. S2c revealed that the personalized FNs identified by the DL model could predict age accurately on both cohorts, with an average partial correlation coefficient of 0.6014 ± 0.0153 ($p < 0.001$, permutation test) and $MAE = 2.099 \pm 0.0324$ ($p < 0.001$, permutation test) on the PNC (Fig. 6d and Fig. 6e), and an average partial correlation coefficient of 0.7838 ± 0.0168 ($p < 0.001$, permutation test) and $MAE = 2.9201 \pm 0.0861$ ($p < 0.001$, permutation test) on the PING cohort (Fig. S2d and Fig. S2e). Together, these results indicate that the DL model generalizes well to new datasets.

Self-supervised DL model could capture the differences of FNs between healthy control and schizophrenia patients.

The DL model trained on HCP cohort generalized well to the PNC and PING cohort, despite they contain subjects of quite different age ranges. To determine whether the DL model trained on healthy individuals (HCP cohort) is applicable to individuals with neuropsychiatric illness, we further applied it to one additional dataset, consisting of fMRI data from 101 healthy controls (HC) and 94 schizophrenia (SCZ) patients. The

personalized FNs of all those individuals computed by the DL model passed the sanity testing with respect to both functional homogeneity and spatial correspondence (Fig. S1). Five representative personalized FNs of three randomly selected individuals (2 HC and 1 SCZ) are shown in Fig. 7a, suggesting that the DL model could identify personalized FNs for both healthy controls and patients, with significantly higher functional homogeneity than those computed using the spatially-regularized NMF (Fig. 7b, $p < 10^{-5}$, Wilcoxon signed rank test).

To further determine whether the differences of FNs between HC and SCZ patients can identify features relevant to neuropsychiatric illness, we investigated whether the personalized FNs identified by the DL model could be used to distinguish HCs from patients with SCZ in a pattern classification setting. Support vector machine (SVM) with linear kernel was used for the classification with the loading coefficients of all 17 FNs used as features. The classification was carried out under a 2-fold cross-validation setting for training and testing, and this procedure was repeated 100 times. The regularization parameter C in SVM was optimized in the range $[2^{-10}, 2^{-9}, ..., 2^4, 2^5]$ using nested 2-fold cross-validation within the training cohort. The prediction performance was evaluated with classification accuracy and area under the Receiver operating characteristic (ROC) curve (AUC). The classification based on personalized FNs distinguished HCs from SCZ patients (Fig. 7c), with an average classification rate of 0.7503 ± 0.0283 ($p < 0.001$, permutation test) and an average AUC value of 0.8245 ± 0.0204 ($p < 0.001$, permutation test) over 100 classification runs (Fig. 7d and Fig. 7e). These results reinforce that the DL model generalizes well and is sensitive to individual differences in functional neuroanatomy associated with schizophrenia.

Self-supervised DL model could identify personalized FNs rapidly.

The computational time to compute FNs using our trained DL model was proportional to the scan length of the fMRI scan. On average, it took 22.5 ± 3.1, 4.8 ± 0.2, 7.14 ± 2.61, and 4.26 ± 0.645 seconds to compute 17 FNs for each testing individual subject of the HCP, PNC, PING, and SCZ datasets respectively, highlighting the DL model's computational efficiency.

DISCUSSION

A rapidly accruing body of evidence has demonstrated that the spatial distribution of FNs on the anatomic cortex differ substantially across individuals, and that personalized FNs are necessary to account for individual variation in functional neuroanatomy (Bijsterbosch et al., 2019; Bijsterbosch et al., 2018; Cui et al., 2020; Glasser et al., 2016; Gordon et al., 2017; Kong et al., 2018; Laumann et al., 2015; Li et al., 2017; Li et al., 2019; Satterthwaite and Davatzikos, 2015; Tavor et al., 2016). Here we present a novel method for identifying the personalized FNs through self-supervised deep learning in an end-to-end fashion. Our self-supervised deep learning method is capable of learning an effective intrinsic representation of FNs in a low-dimensional latent space, which captures the differences in functional topography across individuals and therefore facilitates the computation of personalized FNs by optimizing functional homogeneity of the personalized FNs (Li et al., 2017; Li et al., 2018). We demonstrated that our deep learning method could identify well-established FNs

at an individual level with high reproducibility, and that the individual variation of FNs is relevant to individual behavior, brain development, and neuropsychiatric illness such as schizophrenia.

The traditional paradigm of identifying personalized FNs is to model them as latent factors of rsfMRI data in an unsupervised learning framework using matrix decomposition techniques or variants of autoencoders. Most of the existing methods adopt a two-step inference strategy that first estimates group FNs based on fMRI data from a group of individuals and then use the group FNs to infer personalized FNs under different assumptions (Calhoun et al., 2001; Dong et al., 2020; Du and Fan, 2013; Hjelm et al., 2014; Kong et al., 2018; Nickerson et al., 2017). As a result, these methods inevitably introduce bias to the group FNs and do not necessarily yield an optimal solution since their personalized FNs are not optimized in the same way as their group level counterparts. Several alternative methods have been developed to identify personalized FNs of different individuals jointly and enforce inter-individual correspondence of FNs using spatial group sparsity regularization (Li et al., 2017) or regularizations based on certain assumptions about statistical distribution of loadings of corresponding FNs of different individuals (Abraham et al., 2013; Harrison et al., 2015). Deep belief networks (DBNs) have also been utilized to identify FNs of multiple individuals jointly (Zhang et al., 2019). However, all these methods cannot directly make inference for new individuals and are computationally expensive. Variants of autoencoders have been adopted to learn low-dimensional representations of fMRI data (Brown et al., 2020; Hjelm et al., 2014; Hu et al., 2018; Jang et al., 2017; Kim et al., 2021; Plis et al., 2014). Since these methods built deep learning models to reconstruct the original fMRI as outputs, they do not compute personalized FNs directly. CNNs have also been used to predict individual-specific DMN in a supervised learning setting (Zhao et al., 2020), with DMNs obtained by conventional brain decomposition methods as ground truth. However, CNNs in the supervised learning setting may be limited by the silver standard ground truth adopted.

Our method is different from the traditional paradigm mainly in two aspects. First, our method is under the paradigm of predictive modeling, which learns the personalized FNs from rsfMRI data in an end-to-end self-supervised deep learning fashion. Our deep learning model is trained based on self-supervised information including functional homogeneity and spatial sparsity of FNs, which do not require external guidance such as FNs computed using conventional techniques. Therefore, any rsfMRI data could be used to train the model. Once the training is completed, the model can be used to identify personalized FNs from a new, unseen individual using one forward-pass computation without extra optimization. This feature makes our method very efficient: it takes about 5 seconds to compute the FNs for an fMRI scan with spatial resolution of 3mm and scan length of 120 time points (TR=2 seconds) on one modern NVIDIA GPU. Such computational efficiency facilitates its use in large-scale studies. Second, the inter-individual correspondence of FNs is enforced implicitly though the encoder-decoder architecture of our deep learning model. An effective intrinsic representation of FNs in a low-dimensional latent space is extracted by the encoder from the original functional signal, from which the decoder can infer personalized FNs with inter-individual correspondence. From a perspective of predictive modeling, voxels with similar functional profiles will be assigned to the same FN by the deep learning model.

While different output channels of the deep learning model capture different FNs, one output channel will identify FNs across individuals with functional correspondence. Compared with explicit spatial regularization, this kind of implicit enforcement is more flexible to characterize the inter-individual variation and therefore facilitates the identification of FNs with improved functional homogeneity. While the improved functional homogeneity indicated that the personalized FNs captured the functional neuroanatomy better to a certain extent, it should be interpreted with caution in downstream analysis tasks, as the personalized analysis might be more susceptible to confounding factors, such as motion and image noises.

The proposed deep learning model generalizes well from the training dataset to different external datasets without fine-tuning, even though the external datasets used had notably different sample characteristics (i.e., adults vs. children) and different acquisition protocols. This generalization may be attributed to the self-supervised nature of our method, which is optimized to learn effective representation of the intrinsic functional brain organization from input data without external guidance or prior assumptions. The robust generalization emphasizes that this approach is likely to be widely applicable to heterogeneous data in diverse applications.

Though the deep learning model is promising to learn highly reproducible and functionally homogenous FNs at individual level, there are still several potential limitations which should be noted. First, based on prior work, our method currently generates 17 networks. Future work will be devoted to evaluation of deep learning models for identifying FNs with different spatial resolutions. Second, as the human brain is a multi-scale system with hierarchical organization (Bassett and Siebenhühner, 2013; Baum et al., 2017; Betzel and Bassett, 2016; De Domenico, 2017; Doucet et al., 2011; Meunier et al., 2010; Okamoto, 2015; Park and Friston, 2013; Zhou et al., 2006), modeling personalized FNs at multiple scales with a hierarchical structure simultaneously could allow us to better account for the finding that hetero-modal association cortex in humans has multiple cognitive functions (Barde and Thompson-Schill, 2002; Dosenbach et al., 2007). Third, while we currently focus on identifying FNs for volumetric imaging data only, our method is directly appliable to cortical surface data by projecting the surface data to a 2D space (Cheng et al., 2020). Fourth, the functional network learning module adopts a U-Net architecture, which can be optimized either manually or using neural architecture search (NAS) techniques (Elsken et al., 2019). Evolutionary algorithm based NAS and differentiable architecture search (DARTS) have been successfully applied to recurrent neural network (RNN) based brain decomposition models (Li et al., 2021; Li et al., 2022), and therefore it is highly likely that they can also be used to optimize our network architecture. Finally, while the proposed method can generalize to different datasets well even with slight differences in the image preprocessing pipelines, it merits further investigation of their effects on the computation of FNs and also the effects of data harmonization for fMRI scans of different sites.

Notwithstanding these limitations, we developed a self-supervised deep learning method to identify personalized functional networks accurately and efficiently. The evaluation of our method on multiple datasets has demonstrated that the deep learning model built using our method could identify functional networks in healthy adults, children, and

patients with mental illness. Notably, personalized functional networks effectively captured individual differences in functional topography that could be used to predict development, cognitive performance, and schizophrenia status. Moving forward, this approach may be particularly useful for both large-scale functional imaging studies as well as smaller studies that target neuromodulatory interventions (e.g., TMS) using individual-specific functional neuroanatomy.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

REFERENCES

Abraham A, Dohmatob E, Thirion B, Samaras D, Varoquaux G, 2013. Extracting brain regions from rest fMRI with total-variation constrained dictionary learning. Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention 16, 607–615.

Barde LH, Thompson-Schill SL, 2002. Models of functional organization of the lateral prefrontal cortex in verbal working memory: evidence in favor of the process model. Journal of cognitive neuroscience 14, 1054–1063. [PubMed: 12419128]

Bassett DS, Siebenhühner F, 2013. Multiscale Network Organization in the Human Brain, Multiscale Analysis and Nonlinear Dynamics. Wiley-VCH Verlag GmbH & Co. KGaA, pp. 179–204.

Baum GL, Ciric R, Roalf DR, Betzel RF, Moore TM, Shinohara RT, Kahn AE, Vandekar SN, Rupert PE, Quarmley M, Cook PA, Elliott MA, Ruparel K, Gur RE, Gur RC, Bassett DS, Satterthwaite TD, 2017. Modular Segregation of Structural Brain Networks Supports the Development of Executive Function in Youth. Current Biology 27, 1561–1572.e1568. [PubMed: 28552358]

Beckmann CF, DeLuca M, Devlin JT, Smith SM, 2005. Investigations into resting-state connectivity using independent component analysis.

Betzel RF, Bassett DS, 2016. Multi-scale brain networks. NeuroImage.

Bijsterbosch JD, Beckmann CF, Woolrich MW, Smith SM, Harrison SJ, 2019. The relationship between spatial configuration and functional connectivity of brain regions revisited. eLife 8, e44890. [PubMed: 31066676]

Bijsterbosch JD, Woolrich MW, Glasser MF, Robinson EC, Beckmann CF, Van Essen DC, Harrison SJ, Smith SM, 2018. The relationship between spatial configuration and functional connectivity of brain regions. eLife 7, e32992. [PubMed: 29451491]

Brown JA, Lee AJ, Pasquini L, Seeley WW, 2020. Intrinsic brain activity gradients dynamically coordinate functional connectivity states. bioRxiv, 2020.2008.2012.248112.

Burt JB, Helmer M, Shinn M, Anticevic A, Murray JD, 2020. Generative modeling of brain maps with spatial autocorrelation. NeuroImage 220, 117038. [PubMed: 32585343]

Calhoun VD, Adali T, Pearlson GD, Pekar JJ, 2001. A method for making group inferences from functional mri data using independent component analysis. Human Brain Mapping 14, 140–151. [PubMed: 11559959]

Carpaneto G, Toth P, 1980. Algorithm 548: Solution of the assignment problem [H]. ACM Transactions on Mathematical Software (TOMS) 6, 104–111.

Cheng J, Dalca AV, Fischl B, Zöllei L, 2020. Cortical surface registration using unsupervised learning. NeuroImage 221, 117161. [PubMed: 32702486]

Ciric R, Wolf DH, Power JD, Roalf DR, Baum GL, Ruparel K, Shinohara RT, Elliott MA, Eickhoff SB, Davatzikos C, Gur RC, Gur RE, Bassett DS, Satterthwaite TD, 2017. Benchmarking of participant-level confound regression strategies for the control of motion artifact in studies of functional connectivity. NeuroImage 154, 174–187. [PubMed: 28302591]

Cui Z, Li H, Xia CH, Bart Larsen, Adebimpe A, Baum GL, Cieslak M, Gur RE, Gur RC, Moore TM, Oathes DJ, Raznahani A, Roalf DR, Shinohara RT, Wolf DH, Fairk DA, Bassett DS, Davatzikos C, Fan Y, Satterthwaite TD, 2020. Individual Variation in Functional Topography of Association Networks in Youth. Neuron 106, 340–353. [PubMed: 32078800]

Cui Z, Pines AR, Larsen B, Sydnor VJ, Li H, Adebimpe A, Alexander-Bloch AF, Bassett DS, Bertolero M, Calkins ME, Davatzikos C, Fair DA, Gur RC, Gur RE, Moore TM, Shanmugan S, Shinohara RT, Vogel JW, Xia CH, Fan Y, Satterthwaite TD, 2022. Linking Individual Differences in Personalized Functional Network Topography to Psychopathology in Youth. Biological Psychiatry.

Daubechies I, Roussos E, Takerkart S, Benharrosh M, Golden C, D'Ardenne K, Richter W, Cohen JD, Haxby J, 2009. Independent component analysis for brain fMRI does not select for independence. PNAS 106, 10414–10422.

De Domenico M, 2017. Multilayer modeling and analysis of human brain networks. GigaScience 6, 1–8.

Ding CH, Li T, Jordan MI, 2008. Convex and semi-nonnegative matrix factorizations. IEEE transactions on pattern analysis and machine intelligence 32, 45–55.

Dong Q, Ge F, Ning Q, Zhao Y, Lv J, Huang H, Yuan J, Jiang X, Shen D, Liu T, 2020. Modeling Hierarchical Brain Networks via Volumetric Sparse Deep Belief Network. IEEE Trans Biomed Eng 67, 1739–1748. [PubMed: 31647417]

Dosenbach NU, Fair DA, Miezin FM, Cohen AL, Wenger KK, Dosenbach RA, Fox MD, Snyder AZ, Vincent JL, Raichle ME, Schlaggar BL, Petersen SE, 2007. Distinct brain networks for adaptive and stable task control in humans. Proc Natl Acad Sci U S A 104, 11073–11078. [PubMed: 17576922]

Doucet G, Naveau M, Petit L, Delcroix N, Zago L, Crivello F, Jobard G, Tzourio-Mazoyer N, Mazoyer B, Mellet E, Joliot M, 2011. Brain activity at rest: a multiscale hierarchical functional organization. Journal of neurophysiology 105, 2753–2763. [PubMed: 21430278]

Du Y, Fan Y, 2011. Group information guided ICA for analysis of multi-subject fMRI data, 17th Annual Meeting of the Organization for Human Brain Mapping, Quebec City, Canada.

Du Y, Fan Y, 2013. Group information guided ICA for fMRI data analysis. NeuroImage 69, 157–197. [PubMed: 23194820]

Elsken T, Metzen JH, Hutter F, 2019. Neural architecture search: A survey. The Journal of Machine Learning Research 20, 1997–2017.

Erhardt EB, Allen EA, Wei Y, Eichele T, Calhoun VD, 2012. SimTB, a simulation toolbox for fMRI data under a model of spatiotemporal separability. NeuroImage 59, 4160–4167. [PubMed: 22178299]

Esteban O, Markiewicz CJ, Blair RW, Moodie CA, Isik AI, Erramuzpe A, Kent JD, Goncalves M, DuPre E, Snyder M, Oya H, Ghosh SS, Wright J, Durnez J, Poldrack RA, Gorgolewski KJ, 2019. fMRIPrep: a robust preprocessing pipeline for functional MRI. Nat Methods 16, 111–116. [PubMed: 30532080]

Finn ES, Shen X, Scheinost D, Rosenberg MD, Huang J, Chun MM, Papademetris X, Constable RT, 2015. Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. Nat Neurosci 18, 1664–1671. [PubMed: 26457551]

Glasser MF, Coalson TS, Robinson EC, Hacker CD, Harwell J, Yacoub E, Ugurbil K, Andersson J, Beckmann CF, Jenkinson M, Smith SM, Van Essen DC, 2016. A multi-modal parcellation of human cerebral cortex. Nature 536, 171–178. [PubMed: 27437579]

Glasser MF, Sotiropoulos SN, Wilson JA, Coalson TS, Fischl B, Andersson JL, Xu J, Jbabdi S, Webster M, Polimeni JR, Van Essen DC, Jenkinson M, Consortium WU-MH, 2013. The minimal preprocessing pipelines for the Human Connectome Project. NeuroImage 80, 105–124. [PubMed: 23668970]

Gordon EM, Laumann TO, Gilmore AW, Newbold DJ, Greene DJ, Berg JJ, Ortega M, Hoyt-Drazen C, Gratton C, Sun H, Hampton JM, Coalson RS, Nguyen AL, McDermott KB, Shimony JS, Snyder AZ, Schlaggar BL, Petersen SE, Nelson SM, Dosenbach NUF, 2017. Precision Functional Mapping of Individual Human Brains. Neuron 95, 791–807 e797. [PubMed: 28757305]

Grayson DS, Fair DA, 2017. Development of large-scale functional networks from birth to adulthood: A guide to the neuroimaging literature. NeuroImage 160, 15–31. [PubMed: 28161313]

Harrison SJ, Bijsterbosch JD, Segerdahl AR, Fitzgibbon SP, Farahibozorg S-R, Duff EP, Smith SM, Woolrich MW, 2020. Modelling subject variability in the spatial and temporal characteristics of functional modes. NeuroImage 222, 117226. [PubMed: 32771617]

Harrison SJ, Woolrich MW, Robinson EC, Glasser MF, Beckmann CF, Jenkinson M, Smith SM, 2015. Large-scale probabilistic functional modes from resting state fMRI. NeuroImage 109, 217–231. [PubMed: 25598050]

Hjelm RD, Calhoun VD, Salakhutdinov R, Allen EA, Adali T, Plis SM, 2014. Restricted Boltzmann machines for neuroimaging: An application in identifying intrinsic networks. NeuroImage 96, 245–260. [PubMed: 24680869]

Hoyer PO, 2004. Non-negative matrix factorization with sparseness constraints. Journal of machine learning research 5.

Hu X, Huang H, Peng B, Han J, Liu N, Lv J, Guo L, Guo C, Liu T, 2018. Latent source mining in FMRI via restricted Boltzmann machine. Human Brain Mapping 39, 2368–2380. [PubMed: 29457314]

Ioffe S, Szegedy C, 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167.

Jang H, Plis SM, Calhoun VD, Lee J-H, 2017. Task-specific feature extraction and classification of fMRI volumes using a deep neural network initialized with a deep belief network: Evaluation using sensorimotor tasks. NeuroImage 145, 314–328. [PubMed: 27079534]

Jernigan TL, Brown TT, Hagler DJ Jr., Akshoomoff N, Bartsch H, Newman E, Thompson WK, Bloss CS, Murray SS, Schork N, Kennedy DN, Kuperman JM, McCabe C, Chung Y, Libiger O, Maddox M, Casey BJ, Chang L, Ernst TM, Frazier JA, Gruen JR, Sowell ER, Kenet T, Kaufmann WE, Mostofsky S, Amaral DG, Dale AM, Pediatric Imaging N, Genetics S, 2016. The Pediatric Imaging, Neurocognition, and Genetics (PING) Data Repository. NeuroImage 124, 1149–1154. [PubMed: 25937488]

Jing R, Li P, Ding Z, Lin X, Zhao R, Shi L, Yan H, Liao J, Zhuo C, Lu L, Fan Y, 2019. Machine learning identifies unaffected first-degree relatives with functional network patterns and cognitive impairment similar to those of schizophrenia patients. Hum Brain Mapp 40, 3930–3939. [PubMed: 31148311]

Kim JH, Zhang Y, Han K, Wen Z, Choi M, Liu Z, 2021. Representation learning of resting state fMRI with variational autoencoder. NeuroImage 241, 118423. [PubMed: 34303794]

Kong R, Li J, Orban C, Sabuncu MR, Liu H, Schaefer A, Sun N, Zuo X-N, Holmes AJ, Eickhoff SB, Yeo BTT, 2018. Spatial Topography of Individual-Specific Cortical Networks Predicts Human Cognition, Personality, and Emotion. Cerebral Cortex 29, 2533–2551.

Laumann TO, Gordon EM, Adeyemo B, Snyder AZ, Joo SJ, Chen MY, Gilmore AW, McDermott KB, Nelson SM, Dosenbach NU, Schlaggar BL, Mumford JA, Poldrack RA, Petersen SE, 2015. Functional System and Areal Organization of a Highly Sampled Individual Human Brain. Neuron 87, 657–670. [PubMed: 26212711]

Lee JH, Hashimoto R, Wible CG, Yoo SS, 2011. Investigation of Spectrally Coherent Resting-State Networks Using Non-Negative Matrix Factorization for Functional MRI Data. Int J Imag Syst Tech 21, 211–222.

Lee JH, Lee TW, Jolesz FA, Yoo SS, 2008. Independent vector analysis (IVA): Multivariate approach for fMRI group study. NeuroImage 40, 86–109. [PubMed: 18165105]

Li H, Satterthwaite TD, Fan Y, 2017. Large-scale sparse functional networks from resting state fMRI. NeuroImage 156, 1–13. [PubMed: 28483721]

Li H, Zhu X, Fan Y, 2018. Identification of multi-scale hierarchical brain functional networks using deep matrix factorization. the 21st International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI 2018) LNCS 11072, 223–231.

Li ML, Wang DH, Ren JX, Langs G, Stoecklein S, Brennan BP, Lu J, Chen HF, Liu HS, 2019. Performing group-level functional image analyses based on homologous functional regions mapped in individuals. Plos Biol 17.

Li Q, Wu X, Liu T, 2021. Differentiable neural architecture search for optimal spatial/temporal brain function network decomposition. Med Image Anal 69, 101974. [PubMed: 33588118]

Li Q, Zhang W, Zhao L, Wu X, Liu T, 2022. Evolutional Neural Architecture Search for Optimization of Spatiotemporal Brain Network Decomposition. IEEE Trans Biomed Eng 69, 624–634. [PubMed: 34357861]

Liu X, Duyn JH, 2013. Time-varying functional network information extracted from brief instances of spontaneous brain activity. Proc Natl Acad Sci U S A 110, 4392–4397. [PubMed: 23440216]

Meunier D, Lambiotte R, Bullmore ET, 2010. Modular and hierarchically modular organization of brain networks. Front Neurosci 4.

Nickerson LD, Smith SM, Ongur D, Beckmann CF, 2017. Using Dual Regression to Investigate Network Shape and Amplitude in Functional Connectivity Analyses. Front Neurosci 11, 115. [PubMed: 28348512]

Okamoto H, 2015. Hierarchical organization of multiscale communities in brain networks is non-tree structured. BMC Neuroscience 16, P187.

Park HJ, Friston KJ, 2013. Structural and Functional Brain Networks: From Connections to Cognition. Science 342, 579-+.

Pines AR, Larsen B, Cui Z, Sydnor VJ, Bertolero MA, Adebimpe A, Alexander-Bloch AF, Davatzikos C, Fair DA, Gur RC, Gur RE, Li H, Milham MP, Moore TM, Murtha K, Parkes L, Thompson-Schill SL, Shanmugan S, Shinohara RT, Weinstein SM, Bassett DS, Fan Y, Satterthwaite TD, 2022. Dissociable multi-scale patterns of development in personalized brain networks. Nat Commun 13, 2647. [PubMed: 35551181]

Plis SM, Hjelm DR, Salakhutdinov R, Allen EA, Bockholt HJ, Long JD, Johnson HJ, Paulsen JS, Turner JA, Calhoun VD, 2014. Deep learning for neuroimaging: a validation study. Frontiers in Neuroscience 8.

Pruim RHR, Mennes M, van Rooij D, Llera A, Buitelaar JK, Beckmann CF, 2015. ICA-AROMA: A robust ICA-based strategy for removing motion artifacts from fMRI data. NeuroImage 112, 267–277. [PubMed: 25770991]

Ren J, Chi Q, Hubbard CS, Cui W, Wang D, Li L, Zhang H, Liu H, 2022. Personalized functional imaging identifies brain stimulation target for a patient with trauma-induced functional disruption. Brain Stimul 15, 53–56. [PubMed: 34749006]

Ronneberger O, Fischer P, Brox T, 2015. U-net: Convolutional networks for biomedical image segmentation, International Conference on Medical image computing and computer-assisted intervention. Springer, pp. 234–241.

Satterthwaite TD, Davatzikos C, 2015. Towards an Individualized Delineation of Functional Neuroanatomy. Neuron 87, 471–473. [PubMed: 26247857]

Satterthwaite TD, Elliott MA, Ruparel K, Loughead J, Prabhakaran K, Calkins ME, Hopson R, Jackson C, Keefe J, Riley M, Mentch FD, Sleiman P, Verma R, Davatzikos C, Hakonarson H, Gur RC, Gur RE, 2014. Neuroimaging of the Philadelphia neurodevelopmental cohort. NeuroImage 86, 544–553. [PubMed: 23921101]

Shanmugan S, Seidlitz J, Cui Z, Adebimpe A, Bassett DS, Bertolero MA, Davatzikos C, Fair DA, Gur RE, Gur RC, Larsen B, Li H, Pines A, Raznahan A, Roalf DR, Shinohara RT, Vogel J, Wolf DH, Fan Y, Alexander-Bloch A, Satterthwaite TD, 2021. Sex Differences in Functional Topography of Association Networks. bioRxiv, 2021.2005.2025.445671.

Tavor I, Jones OP, Mars RB, Smith SM, Behrens TE, Jbabdi S, 2016. Task-free MRI predicts individual differences in brain activity during task performance. Science 352, 216–220. [PubMed: 27124457]

Van Essen DC, Smith SM, Barch DM, Behrens TE, Yacoub E, Ugurbil K, Consortium WU-MH, 2013. The WU-Minn Human Connectome Project: an overview. NeuroImage 80, 62–79. [PubMed: 23684880]

Wang D, Buckner RL, Fox MD, Holt DJ, Holmes AJ, Stoecklein S, Langs G, Pan R, Qian T, Li K, Baker JT, Stufflebeam SM, Wang K, Wang X, Hong B, Liu H, 2015. Parcellating cortical functional networks in individuals. Nat Neurosci 18, 1853–1860. [PubMed: 26551545]

Yeo BT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, Roffman JL, Smoller JW, Zollei L, Polimeni JR, Fischl B, Liu H, Buckner RL, 2011. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. Journal of neurophysiology 106, 1125–1165. [PubMed: 21653723]

Yeo BTT, Krienen FM, Chee MWL, Buckner RL, 2014. Estimates of segregation and overlap of functional connectivity networks in the human cerebral cortex. NeuroImage 88, 212–227. [PubMed: 24185018]

Zhang S, Dong Q, Zhang W, Huang H, Zhu D, Liu T, 2019. Discovering hierarchical common brain networks via multimodal deep belief network. Med Image Anal 54, 238–252. [PubMed: 30954851]

Zhao Y, Li X, Huang H, Zhang W, Zhao S, Makkie M, Zhang M, Li Q, Liu T, 2020. 4D Modeling of fMRI Data via Spatio-Temporal Convolutional Neural Networks (ST-CNN). IEEE Trans Cogn Dev Syst 12, 451–460. [PubMed: 33748420]

Zhou CS, Zemanova L, Zamora G, Hilgetag CC, Kurths J, 2006. Hierarchical organization unveiled by functional connectivity in complex brain networks. Phys Rev Lett 97.

Zhuo C, Zhu J, Wang C, Qu H, Ma X, Tian H, Liu M, Qin W, 2017. Brain structural and functional dissociated patterns in schizophrenia. BMC Psychiatry 17, 45. [PubMed: 28143464]
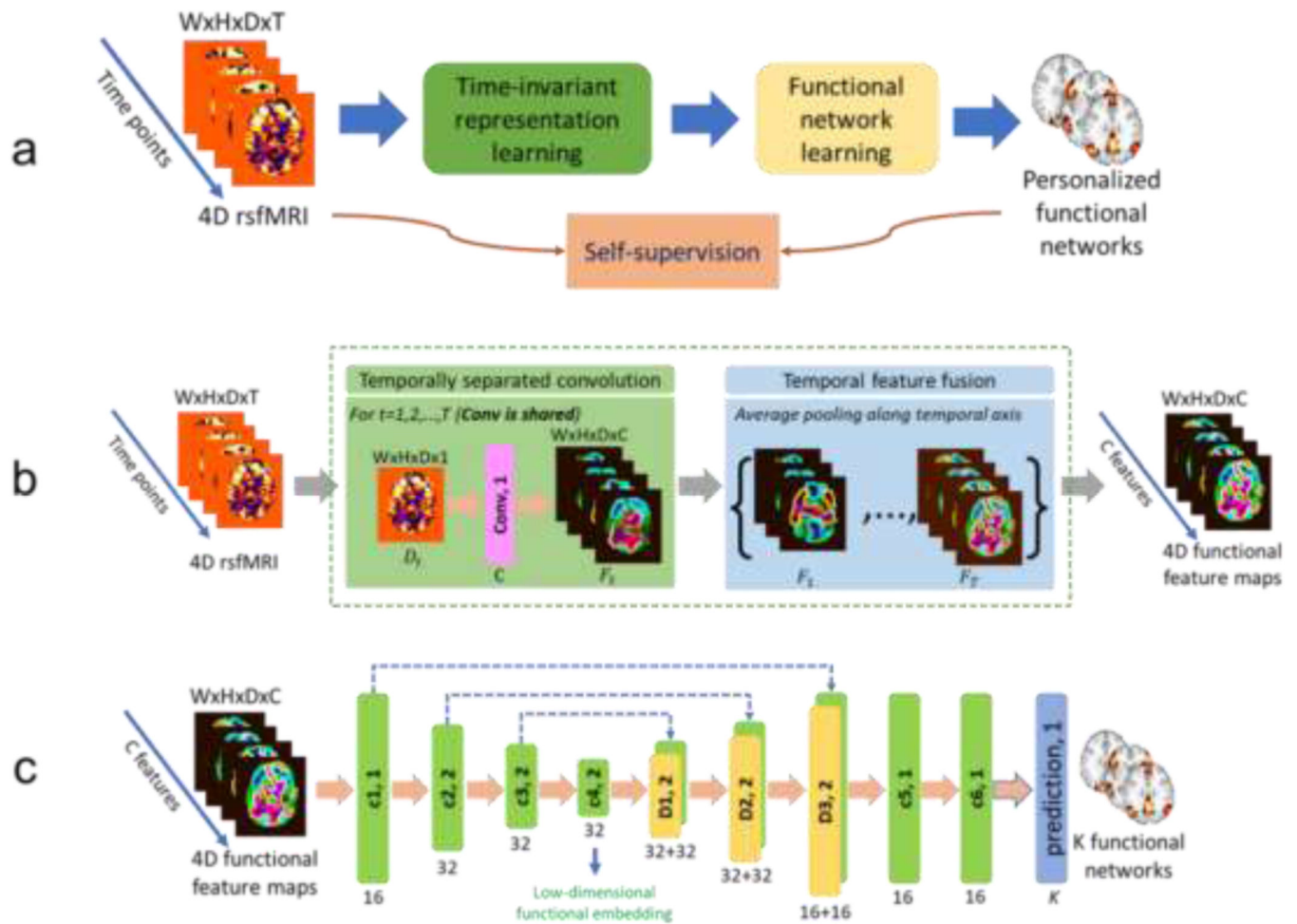
Fig. 1.

A deep learning framework for computing personalized functional networks. **(a)** Schematic diagram of the self-supervised deep learning model for identifying personalized functional networks (FNs), consisting of a time-invariant representation learning module (green) and a functional network learning module (yellow). **(b)** Network architecture of the time-invariant representation learning module to account for temporal misalignment of resting-state fMRI data across different scans. **(c)** Network architecture of the functional network learning module for the prediction of personalized FNs. The numbers underneath convolutional (c1, c2, c3, c4, c5, and c6, green blocks) and deconvolutional (d1, d2, and d3, yellow blocks) layers indicate their corresponding numbers of kernels with a stride of 1 or 2. The kernel size in all layers is set to 3×3×3. Different colormaps are used to differentiate the input fMRI and feature maps learned by the deep learning model. The colors are for illustrations only.
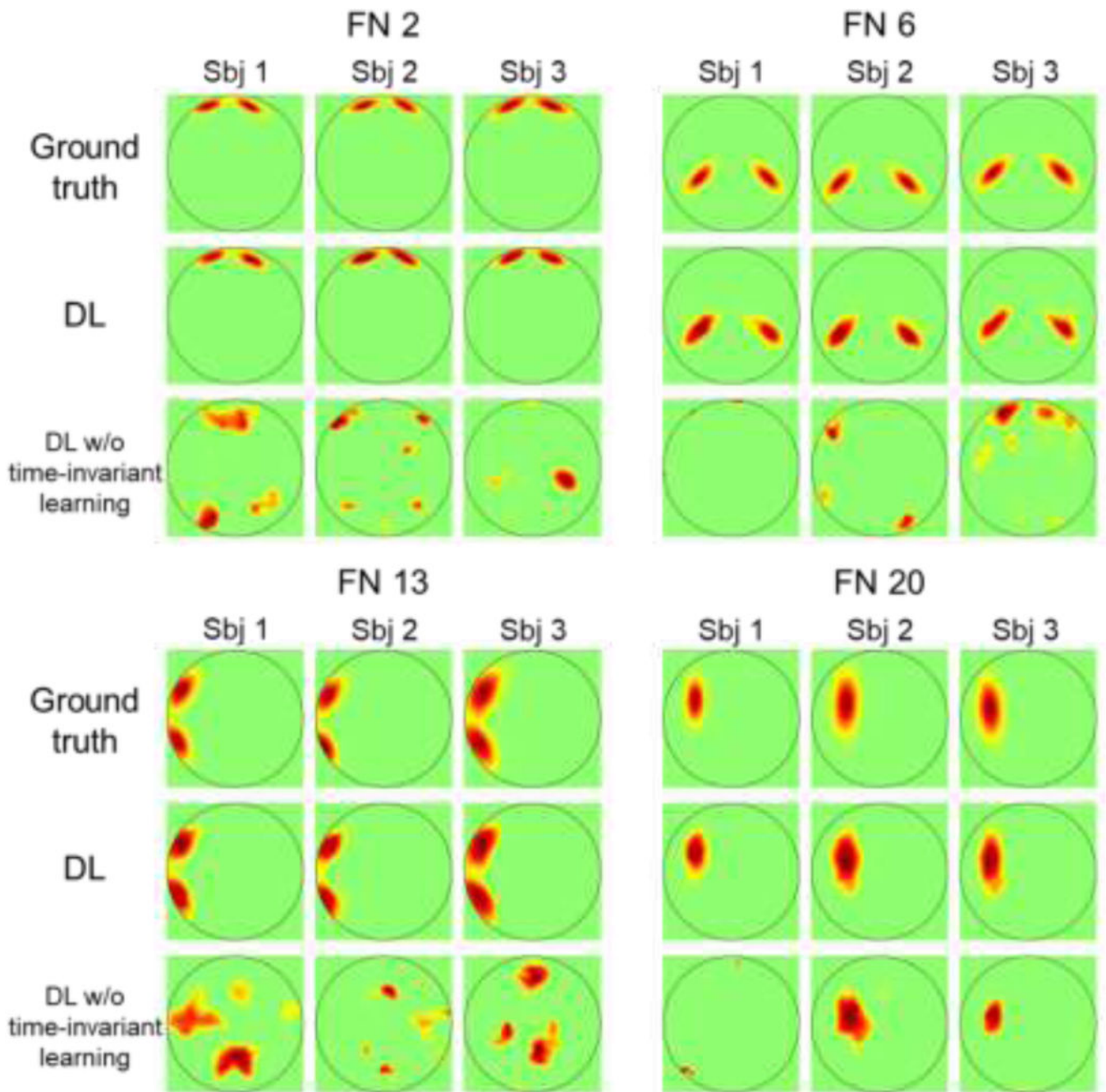
Fig. 2.
Personalized functional networks (FNs) identified by the proposed DL model and the DL model without time-invariant representation learning on the simulated dataset. Four FNs of three randomly selected individuals are shown.
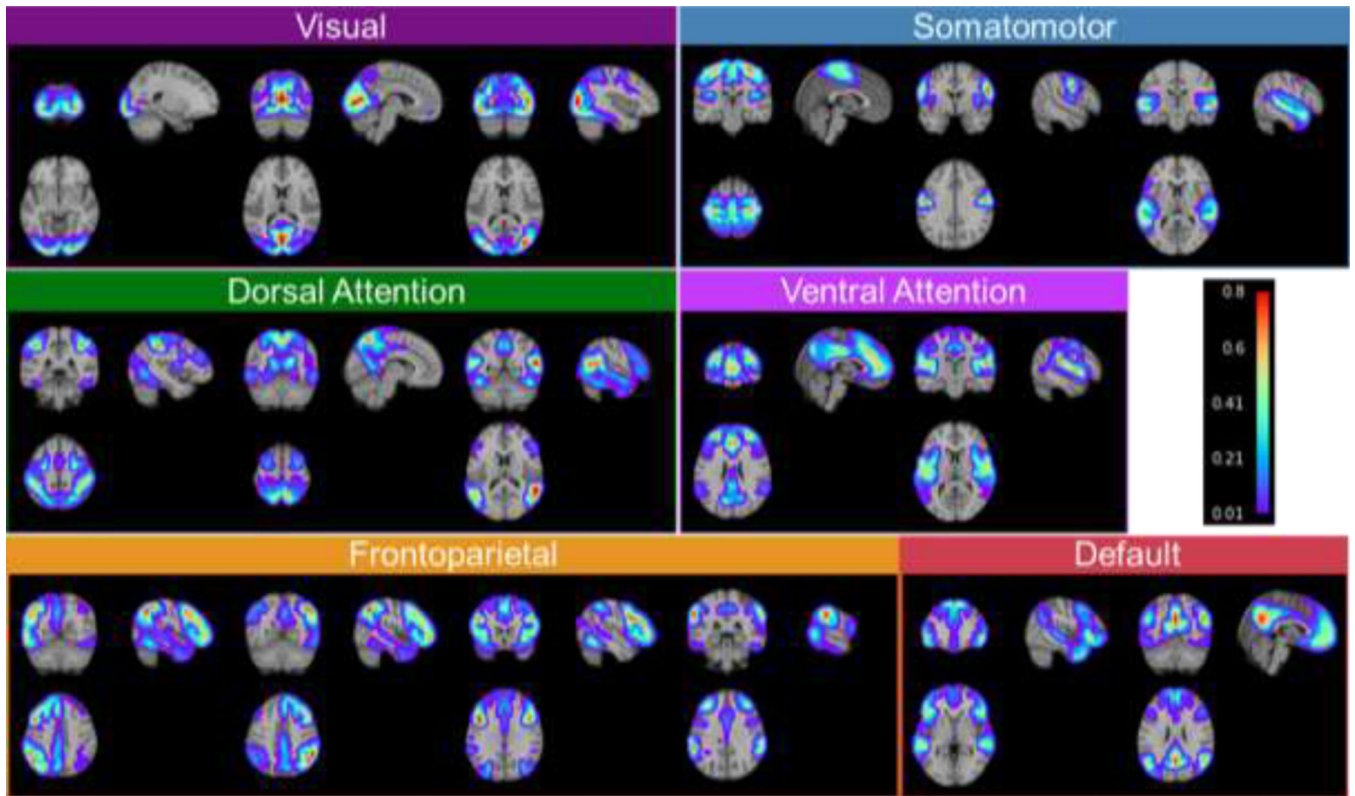
Fig. 3.
Group average functional networks identified by the DL model on the testing individuals of HCP REST1 dataset. The 17 networks have been categorized into visual networks, somatomotor networks, dorsal attention networks, ventral attention networks, fronto-parietal networks, and default mode networks, based on their spatial overlap with the Yeo atlas.
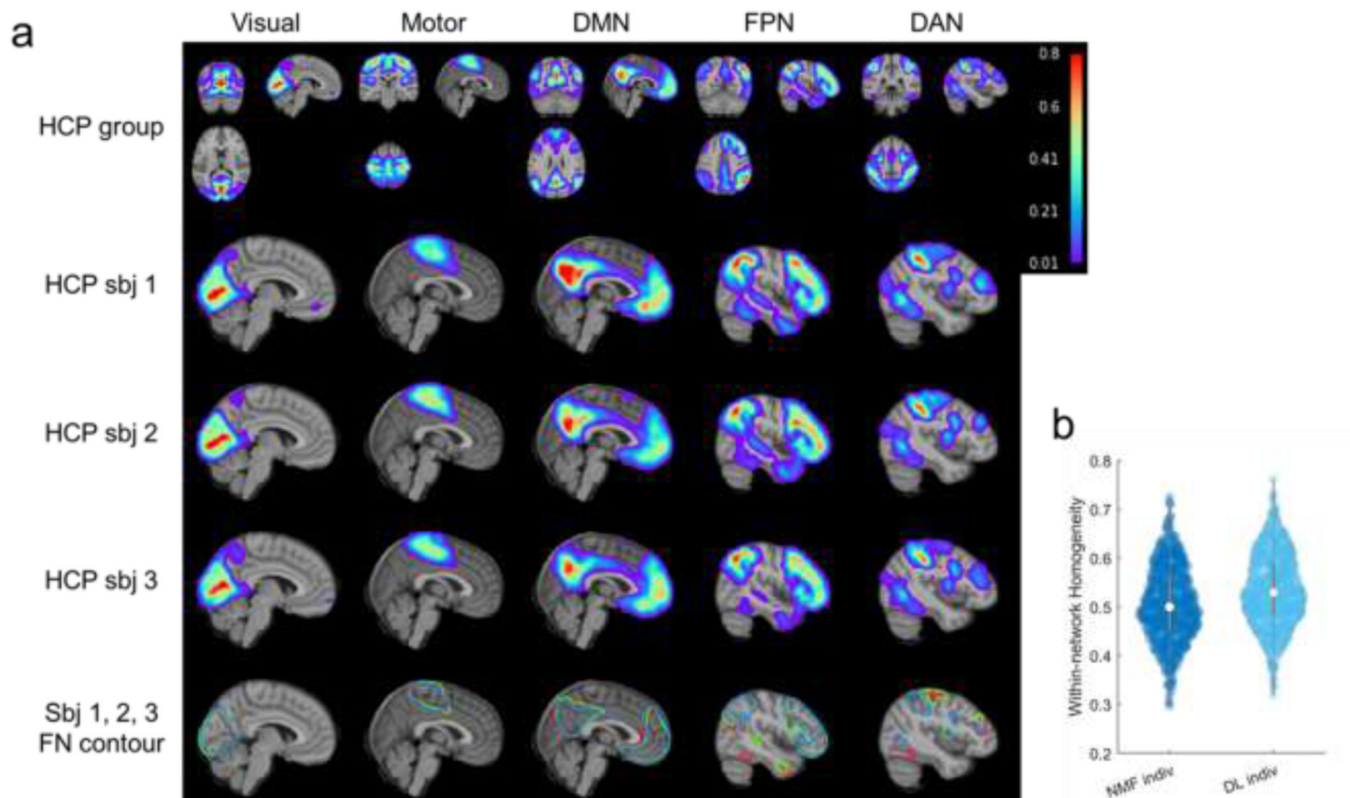
Fig. 4.
Personalized functional networks identified by the DL model for three randomly selected individuals from the testing cohort of HCP REST1 dataset, including visual network, somatomotor network, default mode network (DMN), fronto-parietal control network (FPN), and dorsal attention network (DAN). (**a**) Average FNs of all testing HCP subjects (the first row), three individual subjects' personalized FNs in sagittal view (the second, third, and fourth rows), and isolines of the FNs at a value of 0.15 in different colors (the fifth row). (**b**) Personalized FNs computed using the DL model had significantly higher functional within-network homogeneity than those computed using the spatially-regularized NMF ($p < 10^{-5}$, Wilcoxon signed rank test). Each datapoint in the plot denotes the average functional homogeneity across all FNs of one subject.
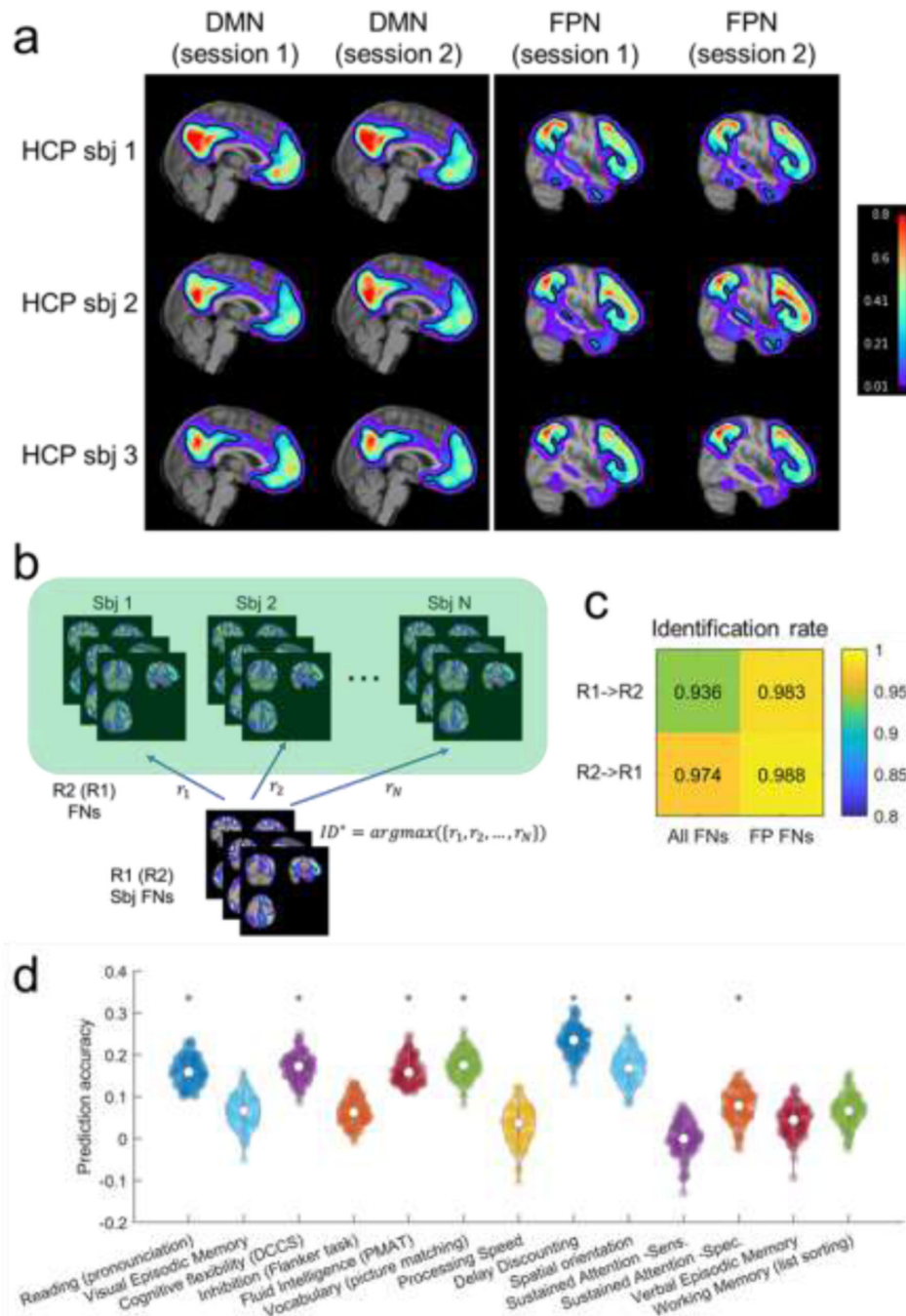
Fig. 5.

Personalized functional networks (FNs) identified by the DL model are with high reproducibility and associated with behavior. **(a)** Personal FNs including DMN and FPN of three randomly selected individuals using HCP REST1 and REST2 rsfMRI data. The isoline of value 0.15 in each FN is demonstrated by the black contour to facilitate the visual comparison across sessions and subjects. **(b)** FN based identification procedure. Given a query set of FNs from one target subject (sbj), we computed the similarity r_i between its FNs and all the sets of FNs in the Database. The matched identity ID^* is the one with the highest

similarity. **(c)** The identification rate when 17 FNs (All FNs) or combined fronto-parietal FNs (FP FNs) were used for the identification. **(d)** Prediction accuracy of 13 cognitive measures based on personalized FNs computed by the proposed DL model. Violin plots show the distribution of prediction accuracy of 100 repetitions for each measure. Asterisks denote the cognitive measures that are significantly associated with personalized FNs.
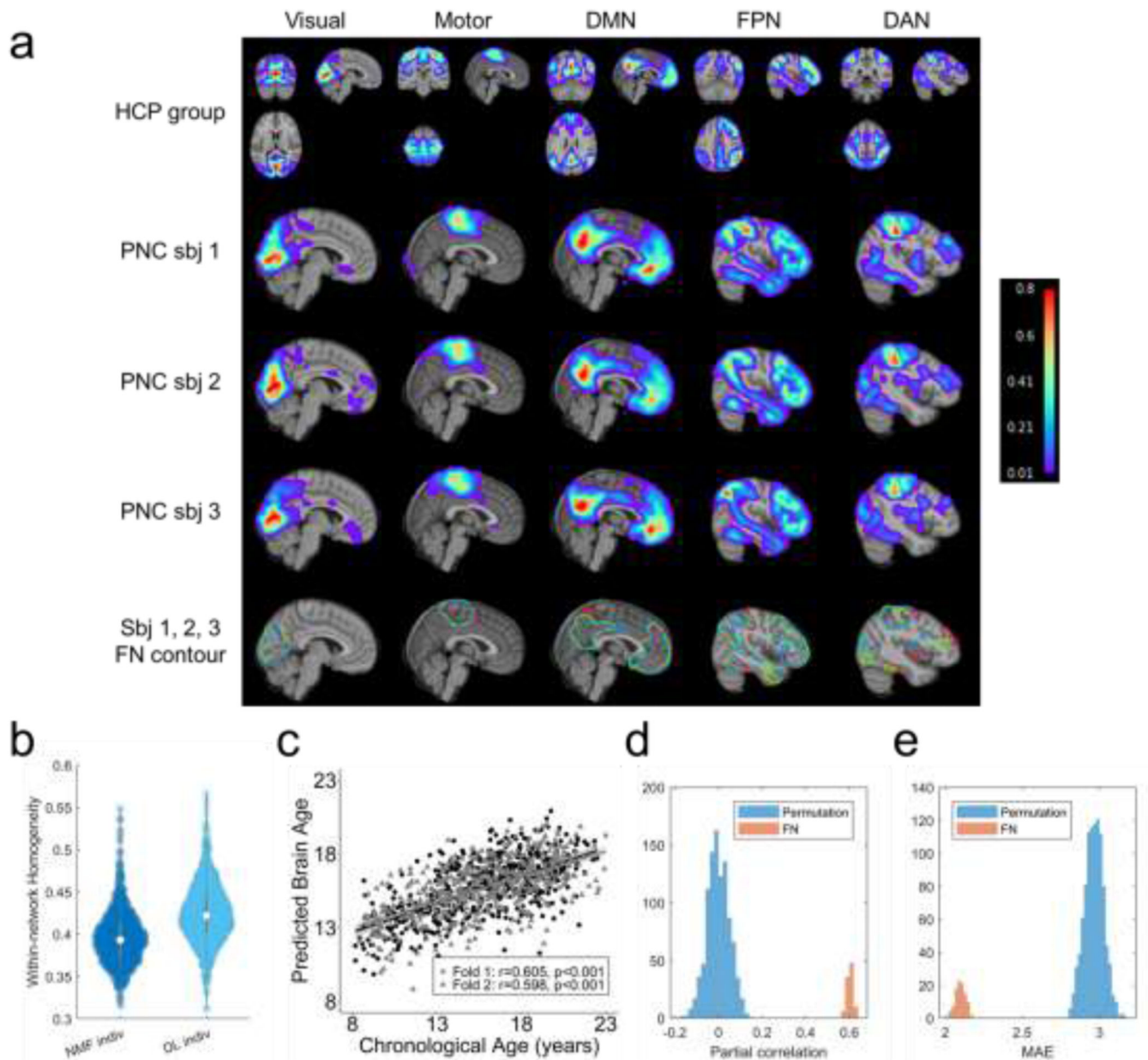
Fig. 6.

The DL model trained on the HCP cohort generalizes well to the PNC. **(a)** Personalized FNs of three randomly selected PNC subjects, identified by the DL model trained on HCP cohort, in sagittal view (the second, third, and fourth rows), isolines of the FNs at a value of 0.15 in different colors (the fifth row), and their corresponding average FNs of all testing HCP subjects (the first row). **(b)** Personalized FNs computed using the DL model had significantly higher functional within-network homogeneity than those computed using the spatially-regularized NMF ($p < 10^{-5}$, Wilcoxon signed rank test). **(c)** Age prediction performance (partial correlation coefficients) of the personalized FNs, obtained with one run of the 2-fold cross-validation. Data points in different colors represent the prediction results of different folds in the 2-fold cross-validation. **(d, e)** Prediction accuracy measured

with partial correlation coefficients and MAE of 100 runs of the 2-fold cross-validation and the null distribution of prediction accuracy from permutation test (1000 runs of 2-fold cross-validation using permuted data).
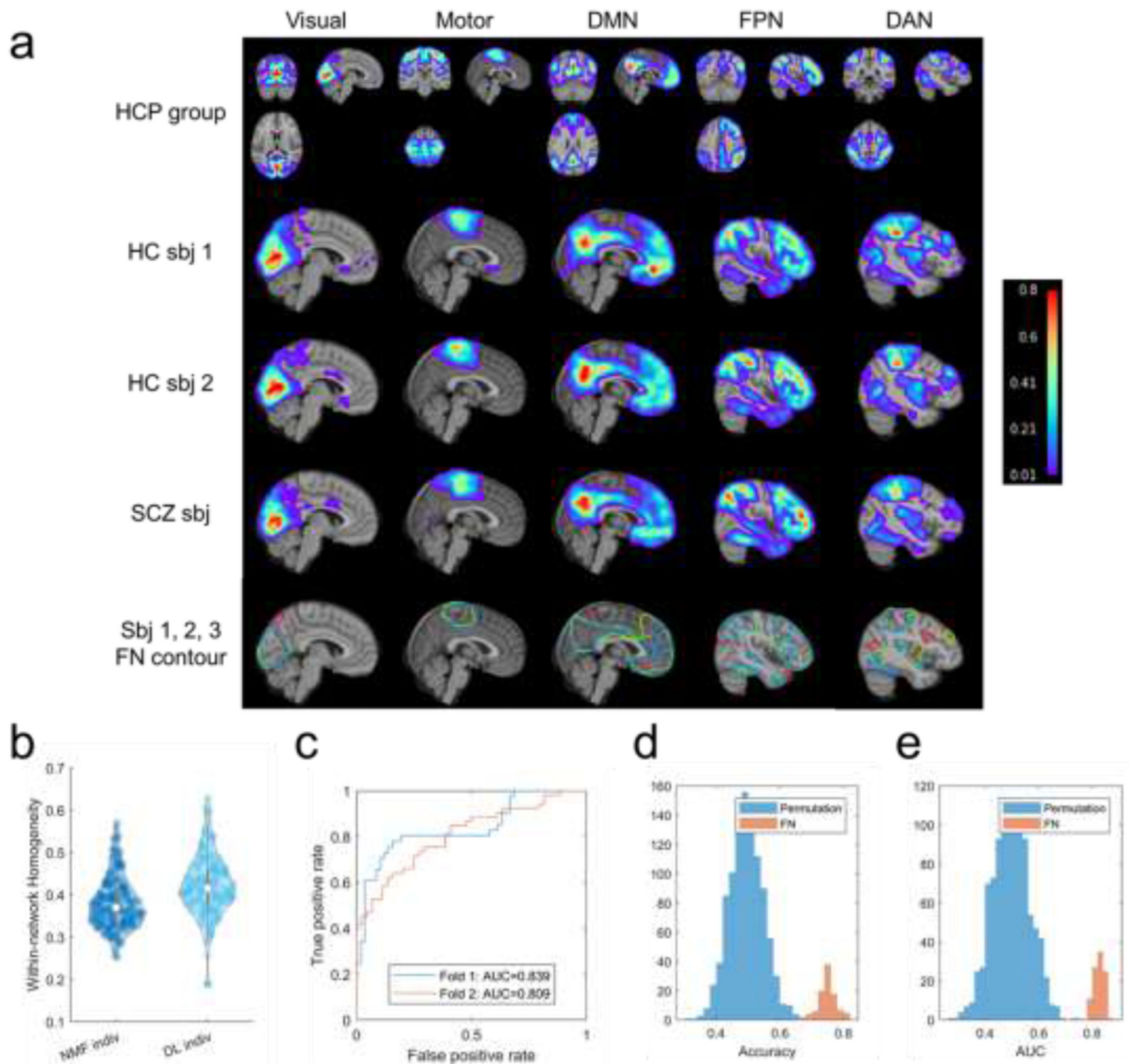
Fig. 7.
The DL model trained on HCP cohort could capture the differences of FNs between healthy controls (HC) and schizophrenia (SCZ) patients. **(a)** Personalized FNs of two HCs (the second and third rows) and one SCZ patient (the fourth row), identified by the DL model trained on HCP cohort, isolines of the FNs at a value of 0.15 in different colors (the fifth row), and their corresponding average FNs of all testing HCP subjects (the first row). **(b)** Personalized FNs computed using the DL model had significantly higher functional within-network homogeneity than those computed using the spatially-regularized NMF (bottom right, $p < 10^{-5}$, Wilcoxon signed rank test). **(c)** Receiver operating characteristic (ROC) curves of classification models built on the personalized FNs of one run of the 2-fold cross-validation. **(d, e)** Classification rates and AUC values of 100 runs of 2-fold cross-validation,

and the null distribution of classification rates and AUC values from permutation test (1000 runs of 2-fold cross-validation using permuted data).