HHS Public Access

Author manuscript

Neuroimage. Author manuscript; available in PMC 2017 May 19.

State-space model with deep learning for functional dynamics estimation in resting-state fMRI

Heung-Il Suk[a,*], **Chong-Yaw Wee**[b], **Seong-Whan Lee**[a], and **Dinggang Shen**[a,c]

[a]Department of Brain and Cognitive Engineering, Korea University, Republic of Korea

[b]Department of Biomedical Engineering, National University of Singapore, Singapore

[c]Biomedical Research Imaging Center, Department of Radiology, University of North Carolina at Chapel Hill, USA

Abstract

Studies on resting-state functional Magnetic Resonance Imaging (rs-fMRI) have shown that different brain regions still actively interact with each other while a subject is at rest, and such functional interaction is not stationary but changes over time. In terms of a large-scale brain network, in this paper, we focus on time-varying patterns of functional networks, i.e., functional dynamics, inherent in rs-fMRI, which is one of the emerging issues along with the network modelling. Specifically, we propose a novel methodological architecture that combines deep learning and state-space modelling, and apply it to rs-fMRI based Mild Cognitive Impairment (MCI) diagnosis. We first devise a Deep Auto-Encoder (DAE) to discover hierarchical non-linear functional relations among regions, by which we transform the regional features into an embedding space, whose bases are complex functional networks. Given the embedded functional features, we then use a Hidden Markov Model (HMM) to estimate dynamic characteristics of functional networks inherent in rs-fMRI via internal states, which are unobservable but can be inferred from observations statistically. By building a generative model with an HMM, we estimate the likelihood of the input features of rs-fMRI as belonging to the corresponding status, i.e., MCI or normal healthy control, based on which we identify the clinical label of a testing subject. In order to validate the effectiveness of the proposed method, we performed experiments on two different datasets and compared with state-of-the-art methods in the literature. We also analyzed the functional networks learned by DAE, estimated the functional connectivities by decoding hidden states in HMM, and investigated the estimated functional connectivities by means of a graph-theoretic approach.

Keywords

---

*Corresponding author. hisuk@korea.ac.kr (H.-I. Suk).

Introduction

A human brain can be understood as a complex system with different structural regions dedicated for different functions, which are locally segregated but globally integrated to process various types of information. Over the last decades, it has been one of the major concerns to investigate the underlying functional mechanism of a human brain in the fields of basic and clinical neuroscience. The functional Magnetic Resonance Imaging (fMRI) that measures the changes of Blood Oxygen Level-Dependent (BOLD) signal in a non-invasive manner has become one of the most successful investigative tools to explore the functional characteristics or properties of a brain.

In the meantime, ever since Biswal et al. (1995) discovered that different brain regions still actively interact with each other while a subject is at rest, i.e., not performing any cognitive task, resting-state fMRI (rs-fMRI) has been widely used as one of the major tools to investigate regional associations or brain networks (Rombouts et al., 2005; Fox et al., 2005; Buckner et al., 2008). The rs-fMRI provides insights to explore the brain's functional organization and to examine altered or aberrant functional networks possibly caused by brain disorders such as Alzheimer's Disease (AD) (Greicius et al., 2004; Li et al., 2002), Mild Cognitive Impairment (MCI) (Rombouts et al., 2005; Sorg et al., 2007; Wang et al., 2007; Zhang et al., 2012; Chase, 2014), autism spectrum disorder (Monk et al., 2009; Khan et al., 2013), schizophrenia (Liang et al., 2006; Zhou et al., 2007; Garrity et al., 2007; Lynall et al., 2010), and depression (Anand et al., 2005; Greicius et al., 2007; Craddock et al., 2009). In this work, we focus on the early diagnosis of MCI based on the computational analysis of rs-fMRI. Due to a high rate of progression from MCI to AD in one year, approximately 10 to 15% according to Alzheimer's Association's's (2012)), it has been of great importance for early detection or diagnosis of MCI and seeking a proper treatment to prevent from progressing to AD. From a clinical point of view, it is advantageous to use rs-fMRI to investigate functional characteristics in the rs-fMRI of patients, who may not be able to perform complicated cognitive tasks during scanning. In these regards, the analysis of functional characteristics inherent in rs-fMRI is playing a core role for brain disease diagnosis or prognosis (Handwerker et al., 2012; Li et al., 2012; Leonardi et al., 2013; Hjelm et al., 2014; Wee et al., 2014; Suk et al., 2015c).

To date, the functional characteristics in a brain have been studied in two different approaches. The *effective connectivity* (Friston et al., 1993) investigates the causal relations between regions, e.g., one region exerts over another. The *functional connectivity* (Van Dijk et al., 2010), the other type of approach, measures functional associations between regions by means of temporal coherence or correlation. In this paper, we mainly consider the functional connectivity, which is computationally less intensive for whole-brain network analysis. It is worth noting that recent studies investigating the complex brain functions have observed the phenomenon that functional connectivity spontaneously changes over time (Chang and Glover, 2010; Bassett et al., 2011; Hutchison et al., 2013), i.e., dynamic rather than stationary. Functional dynamics include changes in the strength of connection between regions, and also the number of connections linked to regions. Motivated by those studies, there have been efforts to estimate temporal changes in functional connectivities and then

use functional properties extracted from the estimated dynamic functional connectivities for disease diagnosis.

To our best knowledge, many existing methods for MCI diagnosis with rs-fMRI typically assumed stationarity on a functional network over time and explicitly modelled it by different methods such as Pearson's correlation, partial correlation (Liang et al., 2012), independent component analysis (Jafri et al., 2008; Li et al., 2012), and sparse linear regression (Wee et al., 2014). Recently, (Faisan et al., 2007; Hutchinson et al., 2009), and (Janoos et al., 2011) independently devised different types of state-space models to explore the dynamic characteristics of functional activation and applied for event-related fMRI data analysis. Due to the use of variables related to external stimulus, i.e., event, those models are not suitable for rs-fMRI data analysis. Meanwhile, Leonardi et al. devised an Eigen-decomposition based method to model functional dynamics with a sliding-window technique (Leonardi et al., 2013) and Eavani et al. proposed to model sparse basis learning within a Hidden Markov Model (HMM) framework (Eavani et al., 2013).

In this paper, we propose a novel method of modelling functional dynamics inherent in rs-fMRI by means of probabilistic models. Specifically, rather than computing correlation matrices and extracting graph-theoretic features (Rubinov and Sporns, 2010) such as clustering coefficients and modularity as commonly performed in the literature, we explicitly model dynamic changes of functional characteristics obtained from regional mean time series of rs-fMRI. In a testing phase, our model estimates the likelihood of a testing sample as MCI and Normal healthy Control (NC), based on which we diagnose MCI. Note that, compared to Eavani et al.'s work, where they utilized the original high-dimensional features, in our method, we devise a Deep Auto-Encoder (DAE) that hierarchically discovers non-linear relations among regional features and helps circumvent the problem of high dimensionality, a common in neuroimaging analysis, and then train a dynamic state-space model, i.e., HMM. While Leonardi et al.'s method fails to reflect the spontaneous changes due to the use of a sliding window strategy, the proposed method probabilistically determines the spontaneous changes based on an observation.

It should be noted that the preliminary version of this work was presented in (Suk et al., 2015a). Compared to the preliminary version of this manuscript, we have extended our work by: 1) carrying out more extensive experiments with an additional dataset from the ADNI2 cohort and 2) analyzing the learned models and the estimated functional connectivities in various perspectives. Although, in this paper, we deal with MCI data only, our method can be also used to understand the functional characteristics of other diseases such as autism, schizophrenia, and depression. In addition, thanks to its capability of estimating dynamic functional connectivities from rs-fMRI, our method can be used for neuroscientific studies on functional organization in a brain.

Materials and preprocessing

In this work, we use two independent rs-fMRI datasets, namely, an ADNI2 dataset publicly available online[1] and an in-house dataset.

ADNI2 cohort

We used a cohort of 31 early MCI subjects (14F/17M) and 31 age-matched NC subjects (17F/14M) from ADNI2[2]. The mean ages of MCI and NC groups are 73.9 ± 4.9 and 73.8 ± 5.5, respectively. All subjects were scanned at different centers using 3.0 T Philips Achieva scanners with the same scanning protocol and parameters of Repetition Time (TR) = 3000 ms, Echo Time (TE) = 30 ms, flip angle = 80°, acquisition matrix size = 64 × 64, 48 slices, 140 volumes, and a voxel thickness = 3.3 mm.

In-house cohort

It is recruited for 37 participants of 12 MCI subjects (6F/6M) and socio-demographically matched 25 NC subjects (16F/9M). The mean ages of MCI and NC are 75.0 ± 8.0 and 72.9 ± 7.9, respectively. The data were acquired on a 3.0 T GE scanner (Signa EXCITE, GE Healthcare) using a SENSE inverse-spiral pulse sequence with the parameters of TR = 2000 ms, TE = 32 ms, flip angle = 77°, acquisition matrix size = 64 × 64, 34 slices, 180 volumes, and a voxel thickness = 4 mm.

Preprocessing

The prevalent preprocessing procedure for rs-fMRI was performed using the SPM8 software package[3]. Specifically, we discarded the first 10 rs-fMRI volumes of each subject prior to further processing to ensure magnetization equilibrium. The remaining volumes were then corrected for the staggered order of slice acquisition that was used during echo-planar scanning so that the data on each slice correspond to the same point in time. The images were realigned with the image at the time point of TR/2 as reference to minimize relative errors across each TR. After correcting acquisition time delay, the rs-fMRI volumes of each subject were realigned by means of a least squares technique and a rigid body spatial transformation. The first volume of each subject was used as the reference to which all subsequent volumes were realigned for the purpose of head-motion artifact removal in the rs-fMRI time-series. We assessed the rotation and translation of every subject and found that all the participants showed no excessive head motion with a displacement of less than 1.5 mm or an angular rotation of less than 1.5° in any direction. There were no significant group differences in head-motion for all subjects[4]. To further minimize the effects of head motion, we also applied Friston 24-parameter model (6 head motion parameters, 6 head motion parameters from the previous time point, and 12 corresponding squared items). After realignment, the volumes were resliced such that they match the first volume voxel-by-voxel. We then normalized rs-fMRI images to the MNI space with a voxel size of 3 × 3 × 3 mm^3.

To further reduce the effects of nuisance signals and focus on the signals of gray matter, we regressed out ventricle and white matter signals as well as six head-motion profiles based on

---

[1]Available at 'http://www.loni.ucla.edu/ADNI'.
[2]There are 3 different subtypes of MCI subjects in ADNI2 dataset, i.e., early MCI, normal MCI, and late MCI. In this study, to minimize the effect of different image sizes and resolutions, we selected images from early MCI and healthy normal subjects with the same image dimension and image resolution.
[3]Available at 'http://www.fil.ion.ucl.ac.uk/spm/software/spm8/'.
[4]The smallest p-value for 6 head motion parameters between patient and healthy subjects was 0.218.

(Van Dijk et al., 2010). Due to the ongoing controversy of removing the global signal in the processing of rs-fMRI data (Fox et al., 2009; Murphy et al., 2009), we omitted the process of global signal regression (Supekar et al., 2008; Lynall et al., 2010). The regressed rs-fMRI images were then parcellated into 116 Regions-Of-Interest (ROIs) according to the Automated Anatomical Labeling (AAL) template (Tzourio-Mazoyer et al., 2002), 45 ROIs per hemisphere and 26 cerebellar ROIs (refer to the B1 for details), and calculated the representative mean time series of each ROI by averaging the intensities of all voxels within an ROI. Therefore, we had a set of time series $\mathbf{I} \in \{\mathbf{I}^{(n)} \in \mathbb{R}^{R \times T}\}_{n=1}^{N}$, where N is the number of samples[5] and R and T denote, respectively, the number of ROIs (=116) and the number of volumes.

It has been shown that a frequency range between 0.025 and 0.06 or 0.07 is reliable for test–retest experiment (Malinen et al., 2010). In this regard, we bandpass-filtered the mean time series $\mathbf{I}$ with a frequency band of 0.01 f 0.08 Hz to better utilize the characteristics of low frequency fluctuation in rs-fMRI (Biswal et al., 1995). The bandpass-filtered regional rs-fMRI time series $\mathbf{G} \in \{\mathbf{G}^{(n)} = [\mathbf{g}_1^{(n)}, \ldots, \mathbf{g}_T^{(-n)}] \in \mathbb{R}^{R \times T}\}_{n=1}^{N}$ were used for the following procedures that are circumstantiated below.

Proposed method

In this section, we propose a novel probabilistic method of modelling functional dynamics inherent in rs-fMRI and estimating the data likelihood of a test subject as NC and MCI for diagnosis. Fig. 1 illustrates the overall framework of our method for MCI identification. Given rs-fMRI images, we first extract mean time series of ROIs, preceded by preprocessing images as described in Preprocessing section. We then train a Deep Auto-Encoder (DAE), which takes the mean intensities of ROIs in a volume at one time point as input and finds the non-linear associations among ROIs in an unsupervised and hierarchical manner. Structurally, our DAE is composed of two parts, namely, *deep encoder* and *deep decoder*. After learning the optimal parameters of both encoder and decoder, we use only the lower part of the trained DAE, i.e., deep encoder, to encode the mean intensities of ROIs in a volume at each time point into an embedding low-dimensional space. Thus, we obtain an encoded time series of the rs-fMRI. Based on these encoded time series of ROIs, we then train two state-space models with HMM for functional dynamics in rs-fMRI of NC and MCI, respectively. In a testing phase, we perform the procedures of preprocessing, computing mean time series of ROIs, encoding mean time series of ROIs via deep encoder, and finally estimating the likelihood as NC and MCI with the trained HMMs, based on which we make a clinical decision.

Deep Auto-Encoder (DAE)

Recently, Hjelm et al. demonstrated that Restricted Boltzmann Machines (RBMs) can be used to identify functional networks from fMRI (Hjelm et al., 2014) and supported its use as a building block for deeper network models in neuroimaging researches. Justified by their

---

[5]In this work, each sample is from one subject.

work along with recent successes in various application domains including neuroimaging analysis (Hjelm et al., 2014; Plis et al., 2014; Wu et al., 2015; Suk et al., 2014, 2015b; Kim et al., 2016), we devise a deep architecture by stacking multiple RBMs to transform the input features of an fMRI at a time point, i.e., mean intensities of ROIs in a volume in our case, into an embedding low-dimensional space by discovering non-linear relations among ROIs, i.e., functional patterns inherent in rs-fMRI.

Restricted Boltzmann Machine (RBM)

An RBM is an undirected probabilistic graphical model, composed of a number D of units in a visible layer and a number F of units in a hidden layer. It can be fully specified with a parameter set Θ of inter-layer connection weights $\mathbf{W}=[W_{ij}] \in \mathbb{R}^{D \times F}$, a visible layer's bias $\mathbf{z}=[z_i] \in \mathbb{R}^D$, and a hidden layer's bias $\mathbf{q}=[q_j] \in \mathbb{R}^F$, i.e., $\Theta=\{\mathbf{W},\mathbf{z},\mathbf{q}\}$. In a model architecture, it assumes symmetric inter-layer connections, but no intra-layer connections as illustrated in Fig. 2, where the one on the left presents a fully specified architecture of an RBM and the other one on the right presents its simplified graphical representation. For uncluttered illustration, the parameters of $\mathbf{z}$ and $\mathbf{q}$ are omitted in Fig. 2.

In an RBM, according to the value types of visible units and hidden units, two different energy functions can be considered. Specifically, when both visible and hidden units are binary random variables, it is common to utilize a Bernoulli–Bernoulli energy function defined as follows:

$$\mathscr{E}\left(\mathbf{v}, \mathbf{h};\Theta\right) = \sum_{i=1}^{D}\sum_{j=1}^{F} v_i W_{ij} h_j - \sum_{i=1}^{D} z_i v_i - \sum_{j=1}^{F} q_i h_j. \tag{1}$$

Meanwhile, if visible units are continuous-valued and hidden units are binary-valued, we can use a Gaussian-Bernoulli energy function (Hinton et al., 2006) defined as follows:

$$\mathscr{E}\left(\mathbf{v}, \mathbf{h};\Theta\right) = -\sum_{i=1}^{D}\sum_{j=1}^{F} \frac{v_i}{\sigma_i} W_{ij} h_j + \sum_{i=1}^{D} \frac{(v_i - z_i)^2}{2\sigma_i^2} - \sum_{j=1}^{F} q_j h_j. \tag{2}$$

where σ_i denotes a standard deviation of the i-th visible unit.

The conditional distribution of binary hidden variables given the binary visible variables and the conditional distribution of binary visible variables given the binary hidden variables are, respectively, computed by utilizing a logistic sigmoid function sigm($\cdot$) as follows:

$$P\left(h_j=1 | \mathbf{v};\Theta\right) = \text{sigm}\left(q_j + \sum_{i=1}^{D} W_{ij} v_j\right) \tag{3}$$

$$P\left(v_j{=}1\middle|\mathbf{h};\boldsymbol{\Theta}\right) = \mathrm{sigm}\left(z_j + \sum_{j=1}^{F} W_{ij} h_j\right).$$

(4)

When the visible units take continuous values, the conditional distribution of hidden units and visible units are, respectively, computed as follows:

$$P\left(h_j{=}1\middle|\mathbf{v};\boldsymbol{\Theta}\right) = \mathrm{sigm}\left(q_j + \sum_{i=1}^{D} \frac{W_{ij} v_i}{\sigma_i}\right)$$

(5)

$$P\left(v_i\middle|\mathbf{h};\boldsymbol{\Theta}\right) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{\left(v_i - z_i - \sum_{j=1}^{F} h_j W_{ij}\right)^2}{2\sigma_i^2}\right).$$

(6)

In both cases, i.e., binary and continuous visible units, due to the unobservable hidden variables, the objective function is defined as the marginal distribution of the visible variables by summing out hidden variable as follows:

$$P\left(\mathbf{v};\boldsymbol{\Theta}\right) = \frac{1}{Z\left(\boldsymbol{\Theta}\right)} \sum_{\mathrm{h}} \exp\left(-\varepsilon\left(\mathbf{v},\mathbf{h};\boldsymbol{\Theta}\right)\right)$$

(7)

where $Z(\Theta)$ is a normalizing factor. In order to learn the parameters $\Theta = \{\mathbf{W}, \mathbf{q}, \mathbf{z}\}$ of an RBM, we can use a contrastive divergence algorithm (Hinton, 2002), which maximizes the log-likelihood of the marginal distribution of the visible variables, i.e., $\Theta = \mathrm{argmax}_\Theta P(v; \Theta)$.

Deep architecture by stacking multiple RBMs

In our method, the visible units of an RBM take the mean intensities of ROIs in an fMRI volume as input and each hidden unit models different relations among ROIs in a probabilistic manner. Hence, the connection weights between visible units and hidden units can be regarded as functional networks (Hjelm et al., 2014). In other words, an RBM finds relational information among input variables, i.e., mean value of ROIs, and the relational information among ROIs can be interpreted as functional networks as Hjeml et al. did.

Further, due to the structural characteristics of an RBM, i.e., symmetric connection between visible and hidden layers and the probabilistic nature of the units in both layers, given the values of input units, i.e., mean intensities of ROIs, we can infer the values of hidden units via Eq. (3) and vice versa. In this regard, an RBM is also called as an auto-encoder (Hinton and Salakhutdinov, 2006) and thanks to this favourable characteristic, the parameter learning can be performed in an unsupervised manner, allowing to use a large number of unlabelled

samples. Thus, the values of hidden units can be used as a different representation of an observation.

Interestingly, when the number of hidden units is smaller than the number of visible units, the representation with the values of hidden units has the effect of reducing the dimensionality of the input, but still preserving the original information. In this work, we utilize this favorable characteristic of an RBM to find functional patterns inherent in rs-fMRI. However, due to the shallow architectural nature of an RBM, it discovers only the simple associations among ROIs. To circumvent this limitation, we design a deep architecture by stacking multiple RBMs to discover more complex relations among ROIs, by which we can better investigate the functional characteristics in rs-fMRI.

Specifically, we stack the number L of RBMs in a hierarchical manner by setting the outputs of the hidden units in a lower RBM as the input to the neighbouring upper RBM in hierarchy, computed as follows:

$$P\left(h^{(l)}\big|h^{(l-1)}, \boldsymbol{\Theta}^{(l)}\right) = \text{sigm}\left(q_j^{(l)} + \sum_i \frac{W_{ij}^{(l)} h_i^{(l-1)}}{\sigma_i^{(l-1)}}\right) \tag{8}$$

where $\mathbf{h}^{(0)}$=v and $\sigma_i^{(l-1)}$=1 for a binary random vector $\mathbf{h}^{(l-1)}$. Thus, our model has one visible layer $\mathbf{v}$ and a series of hidden layers $\mathbf{h}^{(1)}, \cdots, \mathbf{h}^{(L)}$ as shown in Fig. 3. Between any two consecutive layers, we have the corresponding RBM parameters, denoted by $\left\{\boldsymbol{\Theta}^{(1)}\right\}_{l=1}^L$. However, once we stack multiple RBMs, then the resulting model loses a generative characteristic, becoming a deep neural network (Hinton and Salakhutdinov, 2006), which means that the deep model is not an auto-encoder any longer. In this regard, we build our model deeper by unfolding the stacked RBMs (Fig. 3) to be structured with deep encoder and deep decoder, each of which has a number (L+1) of layers. By training the structurally augmented model with a backpropagation algorithm (LeCun et al., 1998), we finally construct a 'Deep Auto-Encoder' (DAE) (Fig. 3). Note that the lower part of DAE, i.e., deep encoder, is used to find an embedded representation of an observation, i.e., mean intensities of ROIs in an rs-fMRI volume.

In our work, the units of the bottom input layer, i.e., $\mathbf{v}$ in Fig. 3, is modelled with a Gaussian function, while the units of hidden layers remain binary except for those of the L-th hidden layer $\mathbf{h}^{(L)}$, whose output becomes an encoded representation of an input. Specifically, we use linear continuous units for the L-th hidden layer with Gaussian noises to obtain continuous values for better representational power of the encoded representations (Hinton and Salakhutdinov, 2006).

In regard to learning the parameter sets $\{\Theta^{(1)},\ldots,\Theta^{(2L)}\}$, we perform the following three steps sequentially with a set of mean time series $\mathbf{G}$ as training samples:

1. Pre-train the parameters $\left\{\Theta^{(1)}\right\}_{l=1}^{L}$ of a deep encoder as shown in Fig. 3, in a greedy layer-wise manner via a contrastive divergence algorithm (Hinton et al., 2006). Note that the mean intensities of ROIs in the t-th fMRI volume of a subject n, i.e., $\mathbf{g}_t^{(n)}$, becomes the input to v.

2. Unfold the pre-trained deep encoder to build a deeper network of encoder and decoder as shown in Fig. 3. For the decoder, we initially use the pre-trained weights of the encoder in the first step, i.e., $\Theta^{(L+k)} \leftarrow \Theta^{(L-k+1)}$, $k=1,\ldots,L$.

3. Fine-tune the parameter sets $\{\Theta^{(1)},\ldots,\Theta^{(2L)}\}$ of the whole deep neural network, i.e., DAE in Fig. 3, jointly by using a back-propagation algorithm (Hinton and Salakhutdinov, 2006) with the inputs $\mathbf{v}$ and target outputs $\mathbf{v}'$ kept identical.

Hereafter, we omit the subject index $^{(n)}$ for uncluttered. After training our DAE, we use the deep encoder, i.e., the lower half of our DAE, to transform the rs-fMRI feature vectors, i.e., $G = [\mathbf{g}_1,\cdots,\mathbf{g}_t,\cdots,\mathbf{g}_T]$ into encoded representations $X = [\mathbf{x}_1,\cdots,\mathbf{x}_t,\cdots,\mathbf{x}_T]$, which are further fed into HMMs to estimate the data likelihood of a subject as NC and MCI to identify a clinical status. Note that by setting the number of hidden units in the top hidden layer $\mathbf{h}^{(L)}$ of a deep encoder smaller than the dimension of the original input, i.e., D, we can naturally embed into a low-dimensional space, but still having rich information necessary to reproduce the input in a non-linear way through the deep decoder. However, a DAE finds non-linear relations among ROIs at one time point, without considering the temporal changes of mean intensities of ROIs, which is also important in discriminating MCI from NC. We handle such temporal information or dynamics in rs-fMRI with HMM as described below.

Hidden Markov models (HMMs)

Recent studies that reported network changes in neurological and psychiatric diseases (Handwerker et al., 2012), (Leonardi et al., 2013), (Stam, 2014) inspired us to hypothesize that MCIs may exhibit different dynamic functional characteristics from NCs (Greicius, 2008), depending on the unobservable functional states that spontaneously change over time. It would be reasonable to model different functional dynamics for different clinical labels, i.e., NC and MCI. In this paper, we model such dynamics inherent in rs-fMRI by the first-order Markov chain in HMMs (Rabiner, 1989) for NC and MCI, separately.

An HMM is a state-space model with doubly embedded stochastic models, namely, 1) a hidden process that is unobservable but can be estimated from observations indirectly and 2) an observable process. Let s_t and $\mathbf{o}_t$ denote random variables of a hidden state and an observation at time t, respectively. Then, the hidden and observable processes over T time-length can be denoted by $\{s_1,\cdots,s_t,\cdots,s_T\}$ and $\{\mathbf{o}_1,\cdots,\mathbf{o}_t,\cdots,\mathbf{o}_T\}$, respectively. As for the hidden process, we model it with a first-order Markov chain, while for the observable process, we define a state-dependent probability distribution. Fig. 4 illustrates the graphical representation of an HMM, where the arrows denote dependencies between variables attached in both ends. Specifically, the hidden state at time t, i.e., s_t is dependent on the hidden state at time $(t-1)$, i.e., s_{t-1}, because of the assumption of a first-order Markov chain.

Meanwhile, the observable variable at time t, i.e., $\mathbf{o}_t$, is dependent on the hidden state at the same time, i.e., s_t.

Formally, when we consider a number K of hidden states, a hidden process is defined by two probability distributions, namely, a state transition probability $A=[a_{ij}]_{i,j=\{1,...,K\}}$, where $a_{ij}\equiv P(s_t=j|s_{t-1}=i)$ denotes the probability of changing from one hidden state ($s_{t-1}=i$) to another hidden state ($s_t=j$), and an initial state probability $\Pi=[\pi_i]_{i=\{1,...,K\}}$, where $\pi_i\equiv P(s_1=i)$ models the probability of starting from a specific hidden state at time $t=1$. As for the state transition, we define an ergodic topology as illustrated in Fig. 4, where four different states are considered as an example, allowing a functional network to change from one to any other state. The observable process is depicted by so-called an emission probability density function (pdf) $B=\{b_i\}_{i=\{1,...,K\}}$, where $b_i\equiv p(\mathbf{o}_t=\mathbf{x}_t|s_t=i)$ denotes the likelihood of observing the specific observation $\mathbf{x}_t$ when residing at the hidden state of i. In this work, we use a mixture of Gaussians for an emission pdf b_i. Thus, an HMM is completely defined by the parameter set of $\lambda=(A,B,\Pi)$. For simplicity, we denote, hereafter, HMMs for NC and MCI with λ_{NC} and λ_{MCI}, respectively.

In this work, a sequence of the encoded representations of the mean intensities of ROIs becomes the observations for HMMs, i.e., $[\mathbf{o}_1=\mathbf{x}_1,\cdots,\mathbf{o}_t=\mathbf{x}_t,\cdots,\mathbf{o}_T=\mathbf{x}_T]$, and the unobservable inherent dynamics in rs-fMRI is modelled by the hidden process, i.e., a state transition probability A and an initial state probability Π. After learning the parameters for each class, i.e., MCI and NC, the likelihoods from HMMs can be used as the evaluation of the goodness of fit of models on a testing sample.

By training HMMs with a Baum–Welch algorithm (Rabiner, 1989) for NC and MCI individually, they can be used as a way to represent the functional characteristics of the respective groups in a probabilistic manner. In other words, given a sequence of functional features, i.e., encoded representation $X=[\mathbf{x}_1,\cdots,\mathbf{x}_t,\cdots,\mathbf{x}_T]$ in our work, we infer that how likely the sequence of functional features X is generated from HMMs of NC (λ_{NC}) and MCI (λ_{MCI}), respectively, as follows:

$$p\left(X|\lambda_c\right)=\sum_s p\left(X|S,\lambda_c\right)P\left(S|\lambda_c\right) \tag{9}$$

where $c\in\{NC,MCI\}$, $S=[s_1,\cdots,s_t,\cdots,s_T]$, $s_t=\{1,...,K\}$, and $t=\{1,\cdots,T\}$. Eq. (9) can be efficiently computed by the forward algorithm (Rabiner, 1989). Based on the data likelihoods, we identify the clinical label of an input rs-fMRI with the higher likelihood.

Experimental settings and results

In this section, we validate the effectiveness of the proposed method for MCI diagnosis with rs-fMRI by comparing with the state-of-the-art methods in the literature, which assume stationarity or non-stationarity (i.e., dynamic) in functional connectivity. As described in Materials and preprocessing section, we conducted experiments on two different datasets, i.e., ADNI2 and in-house cohorts.

Experimental settings

With regard to the structure of a deep model, i.e., the number of layers and the number of units in each layer, there are no golden rules or general guidelines. In this work, we set four hidden layers, i.e., $L=4$, by following (Hinton and Salakhutdinov, 2006) and performed a grid search for the number of hidden units in the space of $|\mathbf{h}^{(1)}| \in \{100,200,300\}$–$|\mathbf{h}^{(2)}| \in \{100,150,200\}$–$|\mathbf{h}^{(3)}| \in \{10,30,50\}$–$|\mathbf{h}^{(4)}| \in \{2,3\}$, where $|\cdot|$ denotes a cardinality operator. Empirically, the architecture with $200(|\mathbf{h}^{(1)}|)$–$100(|\mathbf{h}^{(2)}|)$–$50(|\mathbf{h}^{(3)}|)$–$2(|\mathbf{h}^{(4)}|)$ achieved the best performance in both datasets. Hence, our deep decoder was structured as $2(|\mathbf{h}^{(4)}|)$–$50(|\mathbf{h}^{(5)}|)$ –$100(|\mathbf{h}^{(6)}|)$–$200(|\mathbf{h}^{(7)}|)$–$116(|\mathbf{v}'|)$.[6]

A Gaussian-Bernoulli energy function was used for the visible units $\mathbf{v}$ and the units in the first hidden layer $\mathbf{h}^{(1)}$ in Fig. 1, while a Bernoulli–Bernoulli energy function was exploited for the other hidden layers by taking the outputs of the lower layer as inputs. But the units of the top hidden layer in the encoder, i.e., $\mathbf{h}^{(4)}$ in Fig. 1, had stochastic real-values, allowing the low-dimensional codes to distribute in a continuous feature space (Hinton and Salakhutdinov, 2006).

It should be noted that in our DAE learning, it takes features of one fMRI volume as input, rather than the sequence of volumes. Although our dataset is composed of rs-fMRIs from 62 subjects in the ADNI2 cohort, note that there are 130 volumes per subject (since we discarded the first 10 volumes for magnetization equilibrium). By regarding ROI features of a volume at one time point as one sample, in our experiments with leave-one-subject-out cross-validation, we could utilize $62 \times 130 = 8060$ samples to learn the model parameters, thus with no concern of over-fitting. Due to high computational complexity and long training time with a Gaussian-Bernoulli energy function, we normalized each feature of the data to have zero mean and unit variance by subtracting the mean of the training samples and dividing with the respective variance, and then used noise free reconstructions (Hinton, 2012) by setting the variance parameters set to 1, i.e., $\{\sigma_i=1\}_{i=1}^{D}$ in Eq. (8).

For the HMMs of NC and MCI classes, we varied the number of hidden states K from 2 to 7 with different numbers of Gaussians for emission pdfs, varying between 1 and 3. An ergodic topology in Fig. 4, where the transition from one functional state to any other functional state is allowed, was used for state transition modelling. We set the initial values of the parameters (A, B, Π) in random and learned them via a Baum–Welch algorithm by using a BNT toolbox.[7]

To validate the effectiveness of the proposed method, we compared our method with a baseline method with Pearson correlation and also three competing methods in the literature, namely, group Sparse Representation (gSR) (Wee et al., 2014), joint framework of HMM and Sparse Dictionary Learning (HMM + SDL) (Eavani et al., 2013), and sliding window-based Dynamic Functional Network (sDFN) (Leonardi et al., 2013):

---

[6]Therefore, the complete architecture of our DAE was 116-200-100-50-2-50-100-200-116.
[7]Available at 'https://github.com/bayesnet/bnt'.

- Baseline: We first computed the Pearson correlation for inter-regional connectivity and then applied Fisher's z-transformation, which produces a so-called z-map, to improve the normality of correlation coefficients. It is common to convert the z-map into a graph by considering each ROI as a node and the signed correlation coefficient as an edge weight. We extracted features utilizing a graph theory. Specifically, the weighted local clustering coefficient (ω), which quantifies the cliqueness of the nodes in a graph, was considered.

$$\omega\left(r\right) = 2 \times \frac{\sum_{q \in \mathbb{S}_r} e_{r,q}}{c_r\left(c_r - 1\right)} \quad (10)$$

  where c_r is the number of ROIs connected to the r-th ROI in a z-map, $\mathbb{S}_r$ is a sub-network composed of nodes directly connected to the r-th ROI, and $e_{r,q}$ is the connection coefficient between the r-th and the q-th ROIs. A feature vector was then constructed by concatenating the clustering coefficients over all ROIs, i.e., $\mathbf{f} = [\omega(1), \cdots, \omega(R)]^T$.

- gSR: This method models functional connectivity by means of sparse linear regression with a group sparsity regularization. For the regularization control parameter, we applied a grid search technique in the space of $\{0.01, 0.05, 0.1, 0.15, 0.2, 0.5\}$. By following Wee et al.'s work and also same as the baseline, we computed weighted clustering coefficients from a z-map of the estimated functional connectivity and used them as features.

- HMM + SDL[8]: Similar to our method, this method models functional dynamics with HMM, but explicitly learns the functional connectivity in each state by considering coefficients of a functional connectivity, i.e., covariance matrices, as parameters. As for the number of hidden states, we varied it from 2 to 7, same as our method. We set the weighting parameters to the priors of the covariance matrices to 1 by following the original work.

- sDFN[9]: This method utilizes a sliding-window to take care of time-varying functional patterns and applies principal component analysis for latent functional information learning. We used a window size of 30 time points and a stride of 5 time points between consecutive windows, which resulted in $m=20$ functional networks from an rs-fMRI sequence of 130 volumes per subject in the case of ADNI2 cohort. Each of those m functional networks were then projected into eigen-networks, the number of which, denoted by ε, was determined based on the eigenvalues. Thus, an ε-dimensional feature vector represented a functional network, and we had m such feature vectors. Therefore, we had a ($\varepsilon \times m$)-dimensional feature vector for an rs-fMRI sequence.

---

[8]The codes were provided by the author of the original paper (Eavani et al., 2013).
[9]The codes are available at 'http://miplab.epfl.ch/leonardi/'.

For Baseline, gSR, and sDFN, we further applied p-value based thresholding[10] and minimum redundancy maximum relevance (Peng et al., 2005) sequentially for feature selection and used a linear support vector machine as classifier. To evaluate the performance, we conducted a leave-one-subject-out cross-valuation technique because of a limited number of available samples.

Performance comparison

Let TP, TN, FP, and FN denote, respectively, True Positive, True Negative, False Positive, and False Negative. For quantitative evaluation and comparison among the competing methods, we considered the metrics of accuracy = (TP + TN)/(TP + TN + FP + FN), sensitivity = TP/(TP + FN), specificity = TN/(TN + FP), Positive Predictive Value (PPV) = TP/(TP + FP), and Negative Predictive Value (NPV) = TN/(TN + FN).

ADNI2 cohort

Table 1 summarizes the performance of all competing methods on the ADNI2 cohort. In terms of the diagnostic accuracy, the proposed method achieved the maximal accuracy of 72.58% with a sensitivity of 70.59% and a specificity of 75%. Compared to the methods of Baseline and gSR, which both assume functional stationarity in rs-fMRI, our method improved by 11.29% (Baseline) and 6.45% (gSR). In the meantime, the other two methods of HMM + SDL and sDFN, which consider functional dynamics in rs-fMRI, achieved 62.90% and 70.97%, respectively, which were lower than the accuracy of our method as much as 9.86% (HMM + SDL) and 1.61% (sDFN).

As for the sensitivity and specificity, the proposed method outperformed the other competing methods with 70.59% of sensitivity and 75.00% of specificity. Specifically, our method improved the sensitivity by 9.3% (vs. Baseline), 8.69% (vs. gSR), 3.92% (vs. HMM + SDL), and 2.02% (vs. sDFN), and the specificity by 15.00% (vs. Baseline), 14.47% (vs. HMM + SDL), and 0.93% (vs. sDFN).

Regarding the predictive values, the proposed method achieved the PPV of 77.42% by showing the highest performance with the improvement of 22.58% (vs. Baseline), 12.9% (vs. gSR), and 25.81 (vs. HMM + SDL). In the mean time, HMM + SDL showed the highest NPV value of 74.19%, while the proposed method achieved the NPV of 67.84%, which is still higher than the other competing methods, i.e., Baseline, gSR, and sDFN.

In-house cohort

The experimental results on the in-house cohort are summarized in Table 2. The proposed method achieved the maximal accuracy of 81.08% with the sensitivity of 77.78% and the specificity of 82.14%. Compared to the competing methods of Baseline, sGR, HMM + SDL, and sDFN, our method enhanced the accuracy by 16.22%, 5.4%, 10.81%, and 5.4%, respectively.

---

[10]In our experiment, the threshold was set to 0.05.

As for the predictive values, the proposed method achieved the highest PPV of 58.33% by improving 25.00% (vs. Baseline), 25.00% (vs. gSR), 50.00% (vs. HMM + SDL), and 16.66% (vs. DFN). Meanwhile, the NPV value of the proposed method was 92.00% and that of the other competing methods were 80.00% (Baseline), 96.00% (gSR), 100.00% (HMM + SDL), and 76.67% (sDFN).

Discussions

Performance interpretation

In our experiments on the ADNI2 and in-house datasets, the proposed method achieved the best diagnostic accuracy of 72.58% (ADNI2) and 81.08% (in-house), respectively.

It is noteworthy that the methods of sDFN and our method considering the potential dynamics inherent in rs-fMRI showed better performance than the methods of Baseline and gSR that ignored the functional dynamics. As for HMM + SDL, it achieved the lowest performance among the competing methods except for Baseline on both datasets. As described above, since HMM + SDL explicitly model the functional connectivities within the model by regarding the connectivity weights as parameters, it may suffer from the lack of training samples during parameter learning.

Since it is important to consider the prevalence of the disease in clinic, we also considered the metrics of PPV and NPV. Statistically, PPV and NPV measure, respectively, the proportion of subjects with MCI who are correctly diagnosed as patients and the proportion of subjects without MCI who are correctly diagnosed as cognitive normal. The proposed method obtained the PPV of 77.42% and the NPV of 67.84% on ADNI2 dataset and the PPV of 58.33% and the NPV of 92.00% on in-house dataset. It is noticeable that our method overwhelmed the competing methods except for HMM + SDL on both datasets in terms of both PPV and NPV. As for HMM + SDL, it reported the abnormally high values of NPV, 100.00% (ADNI2) and 74.19% (in-house), and the abnormally low values of PPV, 8.33% (ADNI2) and 51.61% (in-house), which can be another phenomenon making us suspect the problem of data in-sufficiency for parameter learning in HMM + SDL.

Concretely, from a clinical standpoint, out of the five metrics (i.e., accuracy, sensitivity, specificity, PPV, and NPV) considered in this work, sensitivity and PPV are the most critical ones. Specifically, the higher the sensitivity, the lower the chance of mis-diagnosing MCI patients; also, the higher the PPV, the larger the proportion of correctly diagnosed subjects with MCI. In these regards, we believe that the proposed method is superior to the competing methods.

Effects of the DAE and HMM

Based on the performances summarized in Tables 1 and 2, the proposed method of combining DAE and HMM to model functional dynamics presented better performance than the state-of-the-art methods, considered in this work. Although it would be helpful to conduct a statistical significance test, unfortunately, it is hindered for us from conducting such test due to the use of our leave-one-out cross-validation strategy. However, from a methodological standpoint, since our method combines two machine learning methods of

DAE and HMM, it may wonder whether the performance enhancement is caused by the use of DAE or HMM. In this regard, we additionally performed experiments by replacing each method of DAE and HMM with linear Principal Component Analysis (PCA), kernel PCA (kPCA), and Mixture of Gaussians (MoG), respectively, as follows:

- PCA + HMM: We used a classical linear PCA for dimension reduction and HMM for modelling pattern dynamics. The dimensionality of a new space was determined based on the eigenvalues, i.e., reflecting more than 85% of the total variance.

- kPCA + HMM: In order to see the effect of our deep architecture to discover an embedded space, to which we project functional features in a non-linear way, we also conducted an experiment by replacing the part of DAE with a kernel PCA. In this way, we could also find a non-linear projection function, similar to DAE. As for the kernel, we used a 5th-order polynomial kernel, and, for the dimension selection, we used the same strategy with PCA + HMM.

- DAE + MoG: The mean time series were transformed into the low-dimensional space via our DAE, trained with the same architecture of the proposed method, and then the embedded features were then modelled with MoG, based on which the clinical label of the test rs-fMRI images of a subject was determined. Basically, this method corresponds to a special case of the proposed DAE + HMM with a single hidden state, thus not considering dynamic characteristics inherent in rs-fMRI.

The diagnostic accuracy of PCA + HMM, kPCA + HMM, and DAE + MoG as well as the proposed method is shown in Fig. 5. From the results, it is apparent that the proposed method of combining DAE and HMM outperforms the counterpart methods.

Analysis of DAE weights

Our deep auto-encoder is composed of multiple Restricted Boltzmann Machines (RBMs) in a hierarchical manner, and the lowest two layers in our deep model, i.e., visible layer $\mathbf{v}$ and the first hidden layer $\mathbf{h}^{(1)}$ corresponds to an RBM. Note that an RBM finds relational information among input variables, i.e., mean value of ROIs in this work. While there is no temporal information involved in our DAE, the features of different ROIs are time-locked and our model finds the relation among ROIs based on the time-locked features. Thus, the learned weights or relations can be interpreted as functional networks, similar to the correlation-based methods (Liang et al., 2012). In the meantime, the second hidden layer $\mathbf{h}^{(2)}$ in our deep model finds relations among the functional networks learned in the first hidden layer $\mathbf{h}^{(1)}$. Concretely, upper hidden layers (i.e., beyond the first hidden layer) in our deep architecture learn relational patterns of the variables in the neighboring lower layer, which can thus be interpreted as 'network of networks'.

In order to investigate the role of our DAE, we visualized the learned weights of the units of the first hidden layer,[11] i.e., $\mathbf{W}^{(1)}$, in Figs. 6, 7, and 8, where each row represents a

[11]Although the interpretation of weights in the upper layers is also informative, it is very challenging due to their high complexity. For better use of deep models in neuroimaging data analysis, more efforts should be devoted to this problem in the future.

functional network in three different views (sagittal, axial, and coronal). We used the BrainNet Viewer (Xia et al., 2013) to generate these images. For uncluttered, we showed regions whose absolute weight is larger than 0.3 only.[12] As stated in Deep Auto-Encoder (DAE) section, the main role of our DAE is to find the latent associations among regions under consideration, based on which we can learn an embedding low-dimensional space. Remarkably, our DAE discovered interesting relations, which can be neurologically understood. First, as visualized in Fig. 6, the learned weights decomposed ROIs into regions involved in Default-Mode Network (DMN) (Broyd et al., 2009), Executive Attention Network (EAN) (Rueda et al., 2005), visual cortex, subcortical regions, and cerebellum. Second, it detected regions of different networks as related with subcortical regions in Fig. 7 and with cerebellar regions in Fig. 7. Third, it learned the joint relations among regions of multiple networks in Fig. 8. In a nutshell, the units of the first hidden layer of our DAE discovered the latent relations among regions of the same network (i.e., DMN and EAN), visual cortex, subcortex, or cerebellum and among regions of different networks, cortices, and cerebellum. Based on these relations, the units of the second hidden layer further discover the non-linear relations among these networks discovered in the first hidden layer, and so on. In this manner, our DAE could learn the complicated relations among regions of networks, cortices, or cerebellum, based on which an embedding low-dimensional space was learned at the end.

It is also noteworthy that the fine-tuning step in our DAE training was conducted in a supervised manner for class-discriminative pattern discovery. Hence, the learned patterns, e.g., functional networks presented in Figs. 6, 7, and 8, can be understood to contribute for clinical label discrimination.

Our DAE found the well-known DMN, EAN, visual network, subcortical regions, cerebellar regions, and their inter-relations, solely from data in a 'self-taught learning' manner. While there were many other functional networks (in total 200, one for each hidden unit in the first hidden layer) learned in our DAE, we presented only some of them in Figs. 6, 7, and 8, which show some relatively simple and interpretable relations.

Estimation and analysis of dynamic functional connectivity

In the following, we analyzed our experimental results on the ADNI2 cohort obtained with the number of hidden states equal to 5 for both classes, with which the best classification accuracy of 72.58% was obtained. The visual illustration of our trained HMM is given in Appendix A. Although we achieved a high performance with the proposed method for MCI identification in various metrics, it should be noted that they do not manifest anything about the internal characteristics of dynamic functional connectivity, which is the main concern in neuroscience and clinics. Fortunately, however, we could estimate the latent functional dynamics probabilistically from the trained HMMs by decoding the best state sequence of an observation with a Viterbi algorithm (Rabiner, 1989). Here, the best state sequence corresponds to a complete state sequence, one for each hidden node, from time 1 to T that produces the maximum likelihood of the given encoded representations, i.e., $X=[\mathbf{x}_1, \cdots, \mathbf{x}_t,$

---

[12]The trained weights of the units of the first hidden layer were in the range of −0.79 and 1.04.

···,$\mathbf{x}_T$]. That is, instead of summing over all possible state sequences in Eq. (9), we find a single state sequence that maximizes the likelihood of an encoded representation X by means of dynamic programming. Fig. 9 presents examples of state decoding for rs-fMRIs of NC and MCI. From the figure, we can see that the state transition occurs spontaneously, and thus different samples show different dynamic patterns. It is also noteworthy that not all samples involve all the functional states during functional state changes as exemplified at the bottom of Fig. 9.

Based on the decoded state sequence, we could further investigate the functional characteristics by constructing a functional connectivity (FC) map per state. Specifically, in constructing a functional connectivity map of the rs-fMRI in the k-th state, we used the bandpass-filtered regional rs-fMRI time series G=[$\mathbf{g}_1$, ···,$\mathbf{g}_T$], out of which those were assigned to the state k in our state decoding as follows: $FC_k = Corr (\cup_{\forall t; st=k} \mathbf{g}_t)$.

Where Corr(·) denotes an operator of calculating the Pearson correlation coefficients among the input variables, i.e., ROIs. We showed a set of functional connectivities of NC and MCI in Fig. 10, calculated in a groupwise manner. Note that the number of functional connectivities for each class, i.e., NC and MCI, is equal to the number of states, with which we achieved the maximal diagnostic accuracy on the ADNI2 cohort.

These maps reflect the functional connectivities inherent in rs-fMRI and a subject's functional states spontaneously change among those over time.

It should be emphasized that, compared to the previous work that mostly used a sliding window technique (Leonardi et al., 2013) to estimate dynamic functional connectivity, our method could estimate the functional dynamics in rs-fMRI, where a varying number of time points can be involved in different states, by means of a generative state-space model, i.e., HMM.

One interesting observation in Fig. 10, where the functional connectivities with the absolute correlation values larger than 0.65 are presented only, is that the functional connectivities of NC seems to be sparser than the functional connectivities of MCI. Specifically, we hypothesize the following distributional characteristics in functional connectivities, supporting the previous studies: increased number of connections as well as correlations in temporal lobe (Das et al., 2013), increased correlations within the prefrontal lobe, and parietal lobe (Wang et al., 2007), (Supekar et al., 2008), (Yao et al., 2010). These can be interpreted as the compensation of the cognitive functional losses caused by pathologies of brain disorder (Grady et al., 2003), (Celone et al., 2006), (Bai et al., 2009). It should be also noted that there is no correspondence between states of NC and MCI, i.e., the i-th state in NC is not necessarily corresponding to the i-th state in MCI. In Fig. 11, we also illustrated the number of positive and negative connections and the sum of the absolute magnitude of the connectivities in each lobe, averaged over 5 states. In regards to positive connections, i.e., connections with positive correlation (Fig. 11), there was no significant difference between NC and MCI. Meanwhile, in negative connections, i.e., connections with negative correlation (Fig. 11), compared to NC, MCI showed a tendency of increased connections in

frontal, insula, temporal, parietal, subcortical, cerebral, and vermis, where MCI also reported higher absolute magnitudes than NC (Fig. 11).

To further investigate the functional connectivities estimated by decoding the hidden states, we analyzed them by means of a graph theory, which is the natural framework for the exact mathematical representation of complex networks, like functional connectivities. Formally, a complex network can be represented as a graph by $\mathscr{G}\left(\mathrm{V}, \mathrm{E}\right)$, with V denoting the set of nodes and E the set of edges in the graph $\mathscr{G}$ quantitative interpretation of a graph or network, i.e., functional Gconnectivity, in this work, we considered degree, clustering coefficient, and local efficiency. The degree is a simple measurement for the connectivity of a node with the rest of the nodes in a network. The clustering coefficient is the number of connections between a node's neighbors divided by the number of possible such connections, or the probability that a node's neighbors are also connected to each other, indicating how close they are to forming a clique. The local efficiency is the inverse of the average shortest path, which reflects the level of global integration in the network, connecting all neighbors of a node.

Fig. 12 compares the graph theoretic characteristics of the estimated functional connectivities of NC and MCI. First, we could see an aspect that the degree of regions in insula and subcortical cortex (Fig. 12) are increased, which presumably reflect the increased connectivity to those regions. This can be linked with the previous studies (Grady et al., 2003), (Celone et al., 2006), (Bai et al., 2009), (Das et al., 2013) that suggested the compensation of the cognitive functional losses caused by pathologies of brain disorder. Second, regarding the clustering coefficient, it showed a feature of increase for the regions of insula and subcortical cortex, but decrease for the regions of the limbic system (Fig. 12). Third, in local efficiency, similar to the measurements of degree and clustering coefficient, regions of insula and subcortical cortex also seemed to be increased, while those of temporal cortex and limbic system presented decreased local efficiency (Fig. 12).

Model complexity

As for a model's complexity, our method requires more computational time and resources than the competing methods. However, the computational burden of our method is mostly involved in the computation during a training phase, which can be performed offline. In other words, the high computational burden or complexity affects only the learning step, while the required computation for testing is matrix–vector multiplication and simple non-linear function operations. Hence, from a clinical perspective, we believe that it deserves to sacrifice more computational time and resources (only for a training phase) for higher diagnostic accuracy.

For reference, all our experiments were performed on a desktop with an Intel Core i7 CPU (3.4 GHz) and 16 GB DDR3 memory. It took approximately 4 min for DAE training and given a model parameter pair, i.e., the number of hidden states and the number of Gaussians, and it took less than 1 min for HMM training.

Conclusion

In this paper, we focused on computational modelling of functional dynamics inherent in rs-fMRI. Specifically, we proposed a novel method of combining a deep architecture model, i.e., deep-auto encoder, with a state-space model, i.e., hidden Markov model, in a unified framework to investigate the underlying functional dynamics in rs-fMRI and ultimately to identify subjects with MCI. The rationale of using DAE is to discover the latent non-linear relations among brain regions, which are spatially apart but functionally integrated, in the form of functional networks, and thus to better understand the functional characteristics of NC and MCI subjects. Meanwhile, the HMM estimates spontaneous functional changes over time in a probabilistic manner. In order to validate the effectiveness of the proposed method, we evaluated the performance of the proposed method and compared with the state-of-the-art methods in the literature on two different datasets, namely, ADNI2 cohort and in-house cohort. The proposed method achieved the maximal diagnostic accuracies of 72.58% (ADNI2) and 81.08% (in-house) overwhelming the competing methods. Our analysis on the functional networks learned by DAE and the estimated functional connectivities by decoding hidden states in HMM further validated the effectiveness of the proposed method.

Clinically, it is also important to build a computational model to predict whether a subject with MCI will progress to AD or not. As it is one of the major issues in neuroimaging-based AD diagnosis, the extension of our method for MCI conversion prediction will be our forthcoming research topic. Last but not least, it is noteworthy that although we considered only MCI diagnosis in this work, the proposed method can also be applied to diagnosis of other brain disorders such as autism spectral disorder, schizophrenia, and depression.

Acknowledgments

Appendix A. Visualization of our trained HMMs

We visualized an example of the trained HMM models, i.e., state transition matrices, observation distributions, mean observations, and average state durations, in Fig. A1. For the mean observation per state and the state durations, we first decoded the rs-fMRI time series of the training samples and took the average of the bandpass-filtered regional rs-fMRI time series assigned to each state, and then computed the average duration for each state, respectively.

(a) Normal healthy control

(b) Mild cognitive impairment
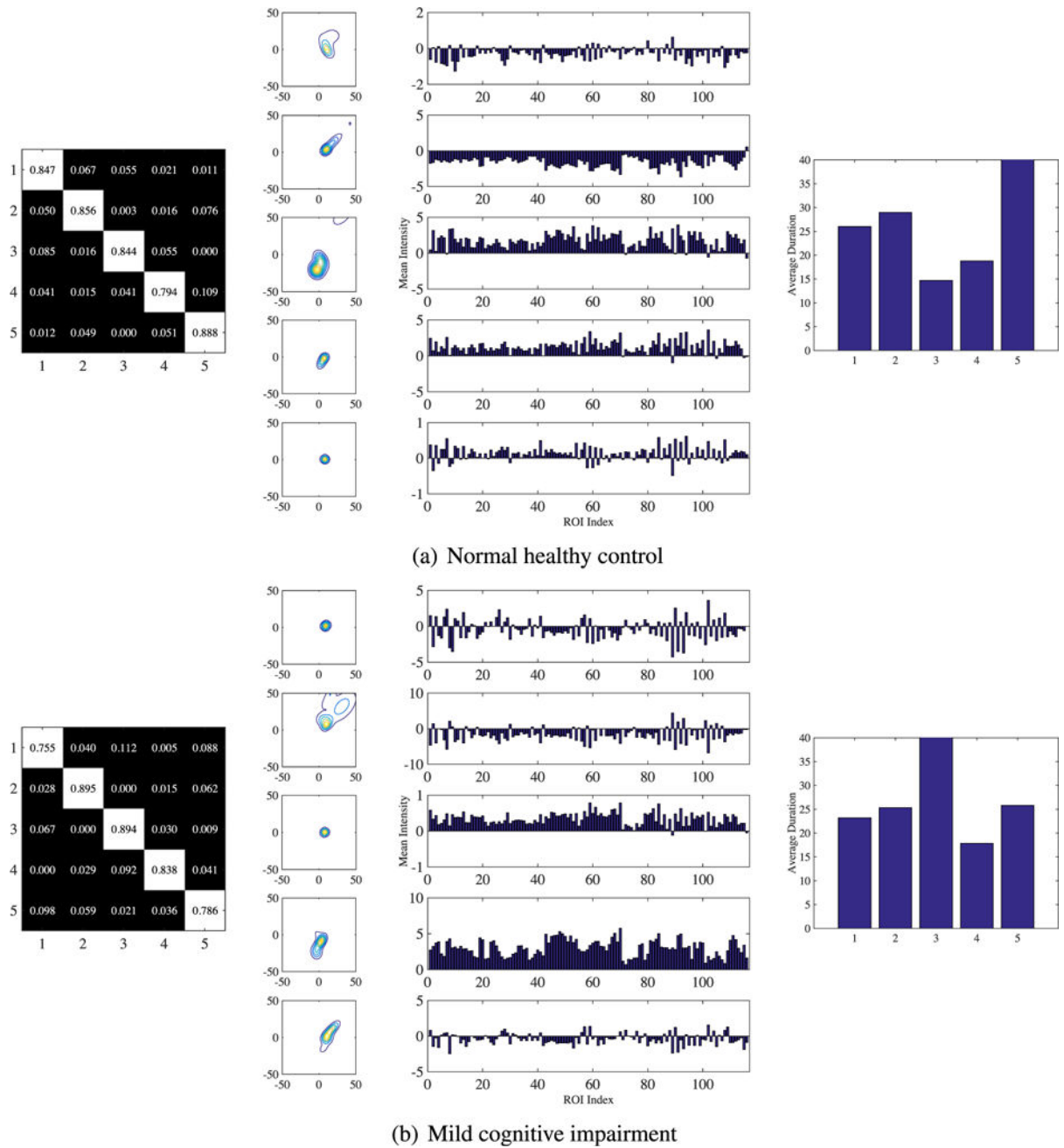
Fig. A1.
Visualization of our trained HMMs: (from left to right) transition matrix, observation distributions, mean observations, and average state durations.

Appendix B. Names of the ROIs in the AAL template

Table B1

The indices and names of the ROIs in the AAL template. The odd and even indices refer, respectively, to the left- and right-hemispheric regions.

Index	ROI label	Index	ROI label
1,2	PReCentral cyrus (PReCG)	3,4	Superior frontal gyrus (dorsal) (SFGdor)
5,6	ORBitofrontal cortex (superior) (ORBsupb)	7,8	Middle frontal gyrus (MFG)
9,10	ORBitofrontal cortext (middle) (ORBmid)	11,12	Inferior frontal gyrus (opercular) (IFGoperc)
13,14	Inferior frontal gyrus (triangular) (IFGtrang)	15,16	ORBitofrontal cortex (inferior) (ORBinf)
17,18	ROLandic operculum (ROL)	19,20	Supplementary motor area (SMA)
21,22	OLFactory (OLF)	23,24	Superior frontal gyrus (middle) (SFGmed)
25,26	ORBitofrontal cortex (medial) (ORBmed)	27,28	RECtus gyrus (REC)
29,30	INSula (INS)	31,32	Anterior cingulate gyrus (ACG)
33,34	Middel cingulate gyrus (MCG)	35,36	Posterior cingulate gyrus (PCG)
37,38	Hippocampus (HIP)	39,40	Parahippocampal gyrus (PHG)
41,42	AMYGdala (AMYG)	43,44	CALcarine cortex (CAL)
45,46	CUNeus (CUN)	47,48	LINGual gyrus (LING)
49,50	Superior occipital gyrus (SOG)	51,52	Middle occipital gyrus (MOG)
53,54	Inferior occipital gyrus (IOG)	55,56	FusiForm gyrus (FFG)
57,58	PostCentral gyrus (PoCG)	59,60	Superior parietal gyrus (SPG)
61,62	Inferior parietal lobule (IPL)	63,64	SupraMarginal gyrus (SMG)
65,66	ANGular gyrus (ANG)	67,68	PreCUNeus (PCUN)
69,70	ParaCentral lobule (PCL)	71,72	CAUdate (CAU)
73,74	PUTamen (PUT)	75,76	PALlidum (PAL)
77,78	THAlamus (THA)	79,80	HEShl gyrus (HES)
81,82	Superior temporal gyrus (STG)	83,84	Temporal POle (superior) (TPOsup)
85,86	Middel temporal gyrus (MTG)	87,88	Temporal POle (middle) (TPOmid)
89,90	Inferior temporal gyrus (ITG)	91–94	Crus I–II of cerebellar hemisphere (Crus)
107–108	Lobule III-X of cerebellar hemisphere (Lobule)	109–116	Lobule I–X of vermis (vermis)

Frontal = (1–16, 19–28, 69–70); insula = (29:30); temporal = (79–90); parietal = (17–18, 57–68); occipital = (43–56); limbic = (31–40); subcortical = (41–42, 71–78); cerebellum = (91–108); vermis = (109–116)

References

Alzheimer's Association. 2012 Alzheimer's disease facts and figures. Alzheimers Dement. 2012; 8:131–168. [PubMed: 22404854]

Anand A, Li Y, Wang Y, Wu J, Gao S, Bukhari L, Mathews VP, Kalnin A, Lowe MJ. Activity and connectivity of brain mood regulating circuit in depression: a functional magnetic resonance study. Biol Psychiatry. 2005; 57:1079–1088. [PubMed: 15866546]

Bai F, Watson DR, Yu H, Shi Y, Yuan Y, Zhang Z. Abnormal resting-state functional connectivity of posterior cingulate cortex in amnestic type mild cognitive impairment. Brain Res. 2009; 1302:167–174. [PubMed: 19765560]

Bassett DS, Wymbs NF, Porter MA, Mucha PJ, Carlson JM, Grafton ST. Dynamic reconfiguration of human brain networks during learning. Proc Natl Acad Sci. 2011; 108:7641–7646. [PubMed: 21502525]

Biswal B, Yetkin FZ, Haughton VM, Hyde JS. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. Magn Reson Med. 1995; 34:537–541. [PubMed: 8524021]

Broyd SJ, Demanuele C, Debener S, Helps SK, James CJ, Sonuga-Barke EJ. Default-mode brain dysfunction in mental disorders: a systematic review. Neurosci Biobehav Rev. 2009; 33:279–296. [PubMed: 18824195]

Buckner RL, Andrews-Hanna JR, Schacter DL. The brain's default network. Ann N Y Acad Sci. 2008; 1124:1–38. [PubMed: 18400922]

Celone KA, Calhoun VD, Dickerson BC, Atri A, Chua EF, Miller SL, DePeau K, Rentz DM, Selkoe DJ, Blacker D, Albert MS, Sperling RA. Alterations in memory networks in mild cognitive impairment and Alzheimer's disease: an independent component analysis. J Neurosci. 2006; 40:10,222–10,231.

Chang C, Glover GH. Time-frequency dynamics of resting-state brain connectivity measured with fMRI. NeuroImage. 2010; 50:81–98. [PubMed: 20006716]

Chase A. Alzheimer disease: Altered functional connectivity in preclinical dementia. Nat Rev Neurol. 2014; 10:609. [PubMed: 25330722]

Craddock RC, Holtzheimer PE III, Hu XP, Mayberg HS. Disease state prediction from resting state functional connectivity. Magn Reson Med. 2009; 62:1619–1628. [PubMed: 19859933]

Das SR, Pluta J, Mancuso L, Kliot D, Orozco S, Dickerson BC, Yushkevich PA, Wolk DA. Increased functional connectivity within medial temporal lobe in mild cognitive impairment. Hippocampus. 2013; 23:1–6. [PubMed: 22815064]

Eavani, H., Satterthwaite, T., Gur, R., Gur, R., Davatzikos, C. Unsupervised learning of functional network dynamics in resting state fMRI. In: Gee, J.Joshi, S.Pohl, K.Wells, W., Zllei, L., editors. Information Processing in Medical Imaging volume 7917 of Lecture Notes in Computer Science. 2013. p. 426-437.

Faisan S, Thoraval L, Armspach JP, Heitz F. Hidden Markov multiple event sequence models: a paradigm for the spatio-temporal analysis of fMRI data. Med Image Anal. 2007; 11:1–20. [PubMed: 17097334]

Fox MD, Snyder AZ, Vincent JL, Corbetta M, Essen DCV, Raichle ME. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. Proc Natl Acad Sci U S A. 2005; 102:9673–9678. [PubMed: 15976020]

Fox MD, Zhang D, Snyder AZ, Raichle ME. The global signal and observed anticorrelated resting state brain networks. J Neurophysiol. 2009; 101:3270–3283. [PubMed: 19339462]

Friston KJ, Frith CD, Liddle PF, Frackowiak RS. Functional connectivity: the principal-component analysis of large (PET) data sets. J Cereb Blood Flow Metab. 1993; 13:5–14. [PubMed: 8417010]

Garrity AG, Pearlson GD, McKiernan K, Lloyd D, Kiehl KA, Calhoun VD. Aberrant "default mode" functional connectivity in schizophrenia. Am J Psychiatr. 2007; 164:450–457. [PubMed: 17329470]

Grady CL, McIntosh AR, Beig S, Keightley ML, Burian H, Black SE. Evidence from functional neuroimaging of a compensatory prefrontal network in Alzheimer's disease. J Neurosci. 2003; 23:986–993. [PubMed: 12574428]

Greicius M. Resting-state functional connectivity in neuropsychiatric disorders. Curr Opin Neurol. 2008; 21:424–430. [PubMed: 18607202]

Greicius MD, Srivastava G, Reiss AL, Menon V. Default-mode network activity distinguishes Alzheimer's disease from healthy aging: evidence from functional MRI. Proc Natl Acad Sci U S A. 2004; 101:4637–4642. [PubMed: 15070770]

Greicius MD, Flores BH, Menon V, Glover GH, Solvason HB, Kenna H, Reiss AL, Schatzberg AF. Resting-state functional connectivity in major depression: abnormally increased contributions from subgenual cingulate cortex and thalamus. Biol Psychiatry. 2007; 62:429–437. [PubMed: 17210143]

Handwerker DA, Roopchansingh V, Gonzalez-Castillo J, Bandettini PA. Periodic changes in fMRI connectivity. NeuroImage. 2012; 63:1712–1719. [PubMed: 22796990]

Hinton GE. Training products of experts by minimizing contrastive divergence. Neural Comput. 2002; 14:1771–1800. [PubMed: 12180402]

Hinton, G. A practical guide to training restricted Boltzmann machines. In: Montavon, G.Orr, G., Mü, KR., editors. Neural Networks: Tricks of the Trade Volume 7700 of Lecture Notes in Computer Science. 2012. p. 599-619.

Hinton GE, Salakhutdinov RR. Reducing the dimensionality of data with neural networks. Science. 2006; 313:504–507. [PubMed: 16873662]

Hinton GE, Osindero S, Teh YW. A fast learning algorithm for deep belief nets. Neural Comput. 2006; 18:1527–1554. [PubMed: 16764513]

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

Hutchinson RA, Niculescu RS, Keller TA, Rustandi I, Mitchell TM. Modeling fMRI data generated by overlapping cognitive processes with unknown onsets using hidden process models. NeuroImage. 2009; 46:87–104. [PubMed: 19457397]

Hutchison RM, Womelsdorf T, Allen EA, Bandettini PA, Calhoun VD, Corbetta M, Penna SD, Duyn JH, Glover GH, Gonzalez-Castillo J, Handwerker DA, Keilholz S, Kiviniemi V, Leopold DA, de Pasquale F, Sporns O, Walter M, Chang C. Dynamic functional connectivity: promise, issues, and interpretations. NeuroImage. 2013; 80:360–378. [PubMed: 23707587]

Janoos F, Machiraju R, Singh S, Morocz I. Spatio-temporal models of mental processes from {fMRI}. NeuroImage. 2011; 57(2):362–377. [PubMed: 21440069]

Jafri MJ, Pearlson GD, Stevens M, Calhoun VD. A method for functional network connectivity among spatially independent resting-state components in schizophrenia. NeuroImage. 2008; 39:1666–1681. [PubMed: 18082428]

Khan S, Gramfort A, Shetty NR, Kitzbichler MG, Ganesan S, Moran JM, Lee SM, Gabrieli JDE, Tager-Flusberg HB, Joseph RM, Herbert MR, Hämäläinen MS, Kenet T. Local and long-range functional connectivity is reduced in concert in autism spectrum disorders. Proc Natl Acad Sci U S A. 2013; 110:3107–3112. [PubMed: 23319621]

Kim J, Calhoun VD, Shim E, Lee JH. Deep neural network with weight sparsity control and pre-training extracts hierarchical features and enhances classification performance: evidence from whole-brain resting-state functional connectivity patterns of schizophrenia. NeuroImage. 2016; 124(Part A):127–146. [PubMed: 25987366]

LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. Proc IEEE. 1998; 86:2278–2324.

Leonardi N, Richiardi J, Gschwind M, Simioni S, Annoni JM, Schluep M, Vuilleumier P, Ville DVD. Principal components of functional connectivity: a new approach to study dynamic brain connectivity during rest. NeuroImage. 2013; 83:937–950. [PubMed: 23872496]

Li SJ, Li Z, Wu G, Zhang MJ, Franczak M, Antuono PG. Alzheimer disease: evaluation of a functional MR imaging index as a marker. Radiology. 2002; 225:253–259. [PubMed: 12355013]

Li Y, Wang Y, Wu G, Shi F, Zhou L, Lin W, Shen D. Discriminant analysis of longitudinal cortical thickness changes in Alzheimer's disease using dynamic and network features. Neurobiol Aging. 2012; 33:427.e15–427.e30.

Liang M, Zhou Y, Jiang T, Liu Z, Tian L, Liu H, Hao Y. Widespread functional disconnectivity in Schizophrenia with resting-state functional magnetic resonance imaging. Neuroreport. 2006; 17:209–213. [PubMed: 16407773]

Liang X, Wang J, Yan C, Shu N, Xu K, Gong G, He Y. Effects of different correlation metrics and preprocessing factors on small-world brain functional networks: a resting-state functional MRI study. PLoS ONE. 2012; 7:e32766. [PubMed: 22412922]

Lynall ME, Bassett DS, Kerwin R, McKenna PJ, Kitzbichler M, Muller U, Bullmore ET. Functional connectivity and brain networks in schizophrenia. J Neurosci. 2010; 30:9477–9487. [PubMed: 20631176]

Malinen S, Vartiainen N, Hlushchuk Y, Koskinen M, Ramkumar P, Forss N, Kalso E, Hari R. Aberrant temporal and spatial brain activity during rest in patients with chronic pain. Proc Natl Acad Sci U S A. 2010; 107:6493–6497. [PubMed: 20308545]

Monk CS, Peltier SJ, Wiggins JL, Weng SJ, Carrasco M, Risi S, Lord C. Abnormalities of intrinsic functional connectivity in autism spectrum disorders. NeuroImage. 2009; 47:764–772. [PubMed: 19409498]

Murphy K, Birn RM, Handwerker DA, Jones TB, Bandettini PA. The impact of global signal regression on resting state correlations: are anti-correlated networks introduced? NeuroImage. 2009; 44:893–905. [PubMed: 18976716]

Peng H, Long F, Ding C. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. IEEE Trans Pattern Anal Mach Intell. 2005; 27:1226–1285. [PubMed: 16119262]

Plis SM, Hjelm D, Salakhutdinov R, Allen EA, Bockholt HJ, Long JD, Johnson HJ, Paulsen J, Turner JA, Calhoun VD. Deep learning for neuroimaging: a validation study. Front Neurosci. 2014; 8

Rabiner L. A tutorial on hidden Markov models and selected applications in speech recognition. Proc IEEE. 1989; 77:257–286.

Rombouts SARB, Barkhof F, Goekoop R, Stam CJ, Scheltens P. Altered resting state networks in mild cognitive impairment and mild Alzheimer's disease: an fMRI study. Hum Brain Mapp. 2005; 26:231–239. [PubMed: 15954139]

Rubinov M, Sporns O. Complex networks measures of brain connectivity: uses and interpretations. NeuroImage. 2010; 52:1059–1069. http://dx.doi.org/10.1016/j.NeuroImage.2009.10.003. [PubMed: 19819337]

Rueda MR, Posner MI, Rothbart MK. The development of executive attention: contributions to the emergence of self-regulation. Dev Neuropsychol. 2005; 28:573–594. [PubMed: 16144428]

Sorg C, Riedl V, Mühlau M, Calhoun VD, Läer TEL, Drzezga A, Kurz HFA, Zimmer C, Wohlschläger AM. Selective changes of resting-state networks in individuals at risk for Alzheimer's disease. Proc Natl Acad Sci U S A. 2007; 104:18,760–18,765.

Stam CJ. Modern network science of neurological disorders. Nat Rev Neurosci. 2014; 15:683–695. [PubMed: 25186238]

Suk HI, Lee SW, Shen D. Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis. NeuroImage. 2014; 101:569–582. [PubMed: 25042445]

Suk, HI., Lee, SW., Shen, D. A hybrid of deep network and hidden Markov model for MCI identification with resting-state fMRI. In: Navab, N.Hornegger, J.Wells, WM., Frangi, AF., editors. Medical Image Computing and Computer-assisted Intervention volume 9349 of Lecture Notes in Computer Science. 2015a. p. 573-580.

Suk HI, Lee SW, Shen D. Latent feature representation with stacked auto-encoder for AD/MCI diagnosis. Brain Struct Funct. 2015b; 220:841–859. [PubMed: 24363140]

Suk HI, Wee CY, Lee SW, Shen D. Supervised discriminative group sparse representation for mild cognitive impairment diagnosis. Neuroinformatics. 2015c; 13:277–295. [PubMed: 25501275]

Supekar K, Menon V, Rubin D, Musen M, Greicius MD. Network analysis of intrinsic functional brain connectivity in Alzheimer's disease. PLoS Comput Biol. 2008; 4:e1000100. [PubMed: 18584043]

Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. NeuroImage. 2002; 15:273–289. [PubMed: 11771995]

Van Dijk KRA, Hedden T, Venkataraman A, Evans KC, Lazar SW, Buckner RL. Intrinsic functional connectivity as a tool for human connectomics: theory, properties, and optimization. J Neurophysiol. 2010; 103:297–321. [PubMed: 19889849]

Wang K, Liang M, Wang L, Tian L, Zhang X, Li K, Jiang T. Altered functional connectivity in early Alzheimer's disease: a resting-state fMRI study. Hum Brain Mapp. 2007; 28:967–978. [PubMed: 17133390]

Wee CY, Yap PT, Zhang D, Wang L, Shen D. Group-constrained sparse fMRI connectivity modeling for mild cognitive impairment identification. Brain Struct Funct. 2014; 219:641–656. [PubMed: 23468090]

Wu G, Kim M, Sanroma G, Wang Q, Munsell BC, Shen D. Hierarchical multi-atlas label fusion with multi-scale feature representation and label-specific patch partition. NeuroImage. 2015; 106:34–46. [PubMed: 25463474]

Xia M, Wang J, He Y. Brainnet viewer: a network visualization tool for human brain connectomics. PLoS ONE. 2013; 8:e68910. [PubMed: 23861951]

Yao Z, Zhang Y, Lin L, Zhou Y, Xu C, Jiang T, The Alzheimer's Disease Neuroimaging Initiative. Abnormal cortical networks in mild cognitive impairment and Alzheimer's disease. PLoS Comput Biol. 2010; 6:e1001006. [PubMed: 21124954]

Zhang D, Shen D, Alzheimer's Disease Neuroimaging, I. Predicting future clinical changes of MCI patients using longitudinal and multimodal biomarkers. PLoS ONE. 2012; 7:e33182. [PubMed: 22457741]

Zhou Y, Liang M, Tian L, Wang K, Hao Y, Liu H, Liu Z, Jiang T. Functional disintegration in paranoid schizophrenia using resting-state fMRI. Schizophr Res. 2007; 97:194–205. [PubMed: 17628434]
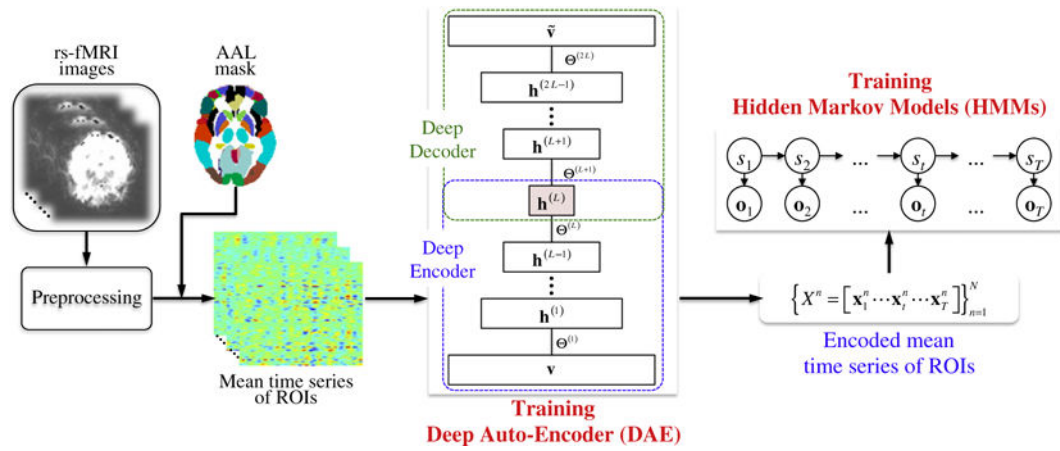
Fig. 1.
Illustration of the proposed method for modelling functional dynamics in rs-fMRI.
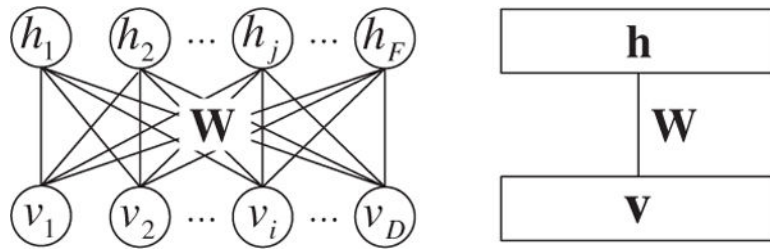
Fig. 2.
Graphical illustration of a restricted Boltzmann machine: (left) inter-layer fully connected bipartite graph, (right) simplified illustration. Here, the bias parameters are omitted for uncluttered.
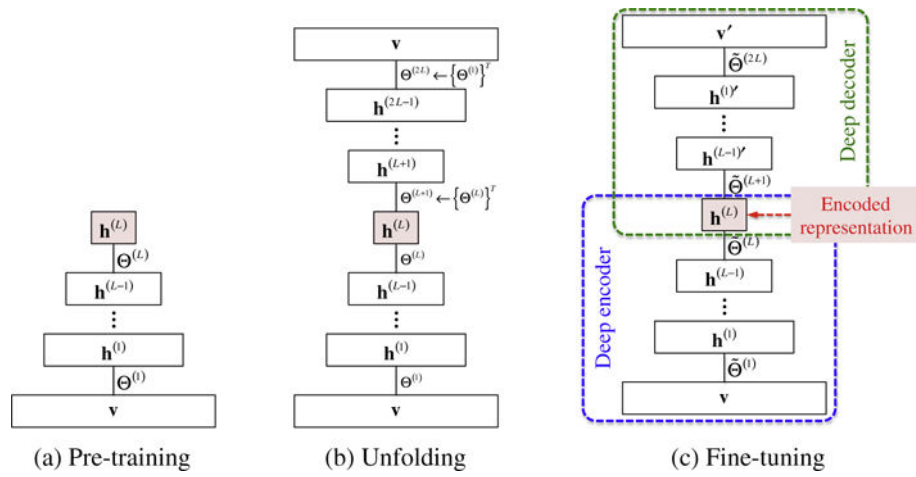
Fig. 3.
Three steps to learn a deep auto-encoder.

(a) Graphical representation of HMM

(b) An ergodic topology in HMM
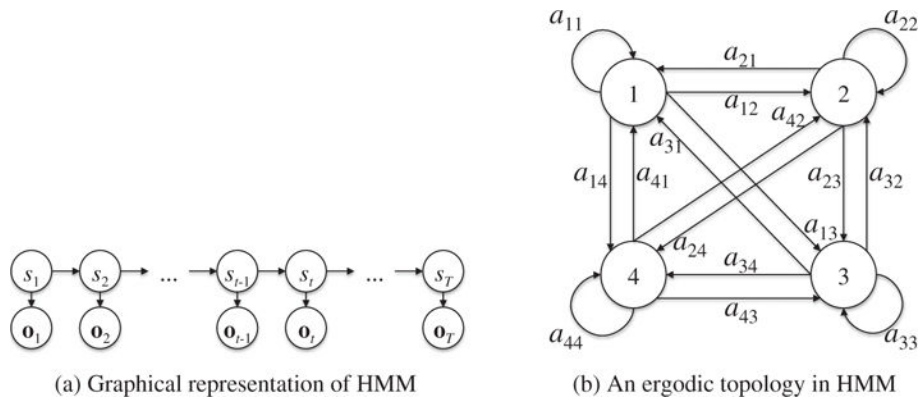
Fig. 4.
Graphical representation of an HMM and an ergodic topology for transition among states, e.g., 4 states in this case, to model functional dynamics in rs-fMRI.
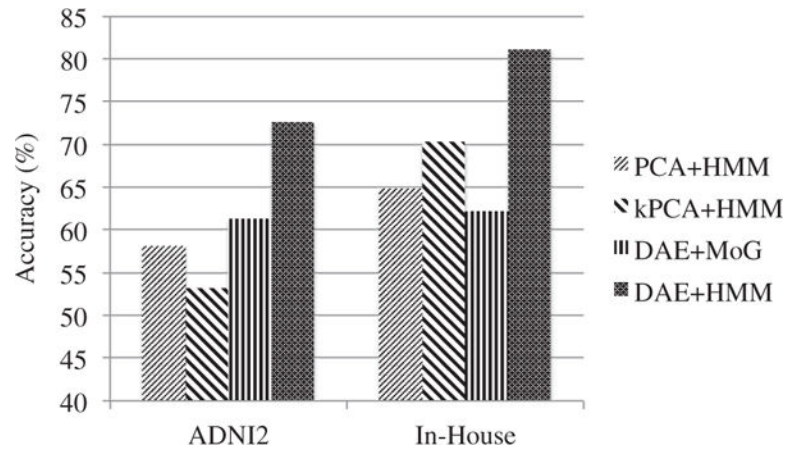
Fig. 5.
Changes of diagnostic accuracy with the replacement of DAE and HMM by their counterparts, i.e., PCA, kPCA, and MoG.
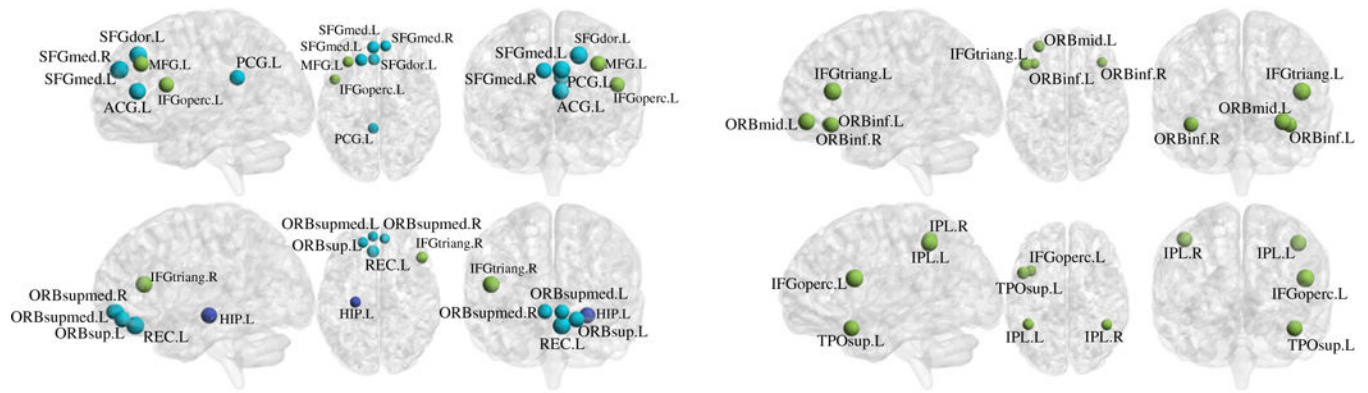
(a) Default-mode network

(b) Executive attention network
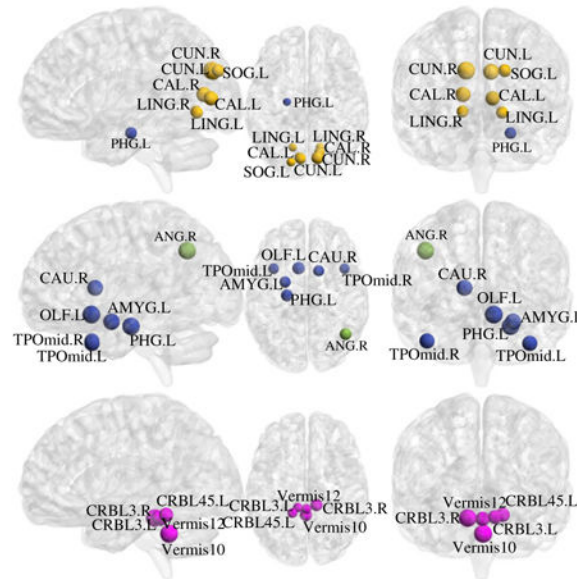
(c) Visual network, subcortical regions, and cerebellum

Fig. 6.
Visualization of the trained weights of units in the first hidden layer of our DAE, i.e., $\mathbf{W}^{(1)}$, which decomposed ROIs into regions involved in different types of functional networks. Here, we denoted only the regions whose absolution value is greater than 0.3. Each network corresponds to weights of one hidden unit. For better legibility, see the figures in an online version.

(a) Relations with subcortical regions

(b) Relations with cerebellar regions
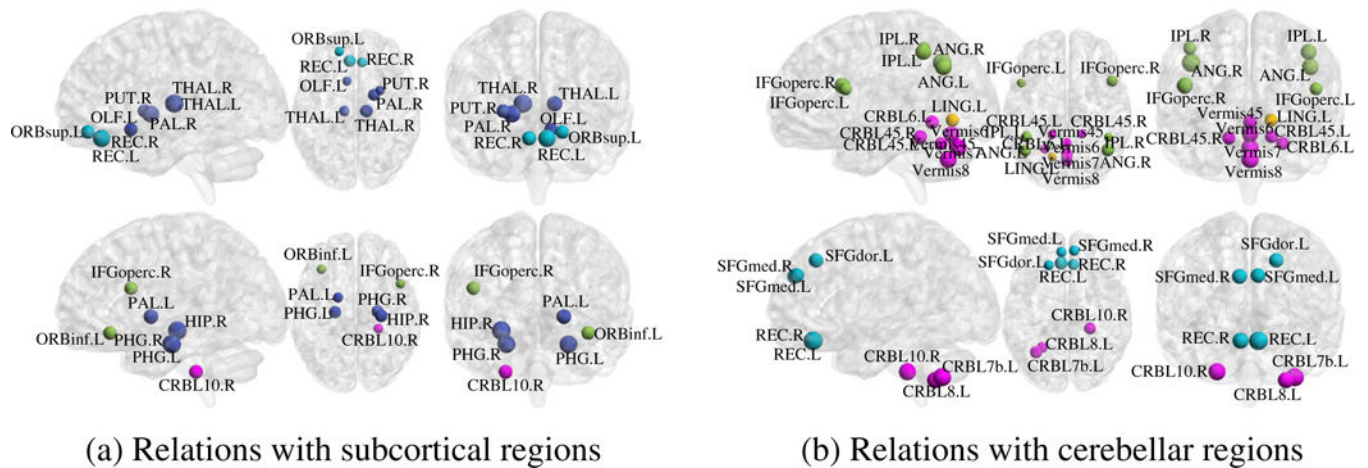
Fig. 7.

Visualization of the learned weights of units in the first hidden layer of our DAE, i.e., $\mathbf{W}^{(1)}$, which found relations with subcortical and cerebellar regions (cyan: default-mode network; blue: regions of the subcortical nuclei; red: regions of the sensorimotor cortex; yellow: regions of the visual cortex; green: regions involved in the executive attention network; pink: regions in the cerebellum). Here, we denoted only the regions whose absolution value is greater than 0.3. Each relational network corresponds to weights of one hidden unit. For better legibility, see the figures in an online version.ω.
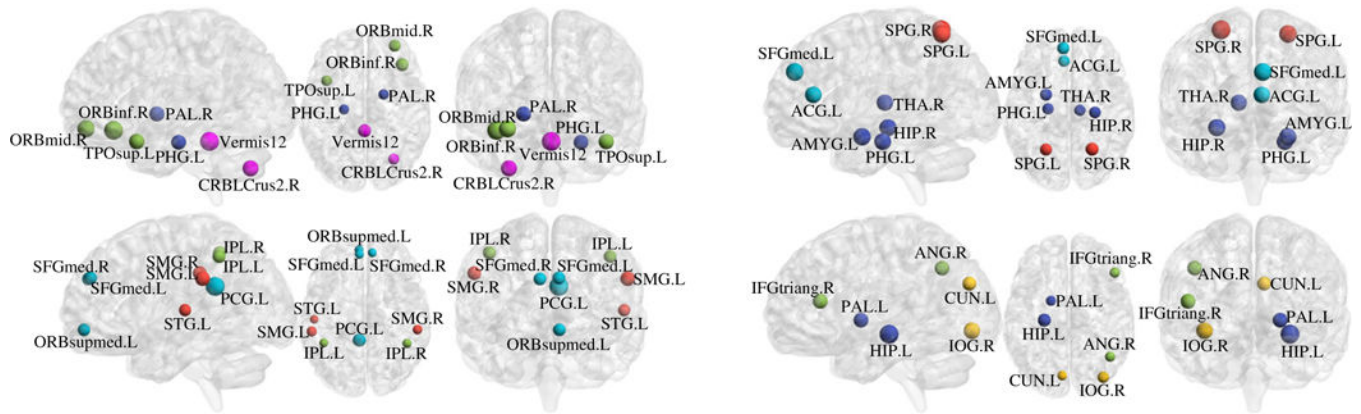
Fig. 8.
Visualization of the learned weights of units in the first hidden layer of our DAE, i.e., $\mathbf{W}^{(1)}$, which found relations among regions of networks, cortices, and cerebellum (cyan: default-mode network; blue: regions of the subcortical nuclei; red: regions of the sensorimotor cortex; yellow: regions of the visual cortex; green: regions involved in the executive attention network; pink: regions in the cerebellum). Here, we denoted only the regions whose absolution value is greater than 0.3. Each relational network corresponds to weights of one hidden unit. For better legibility, see the figures in an online version.
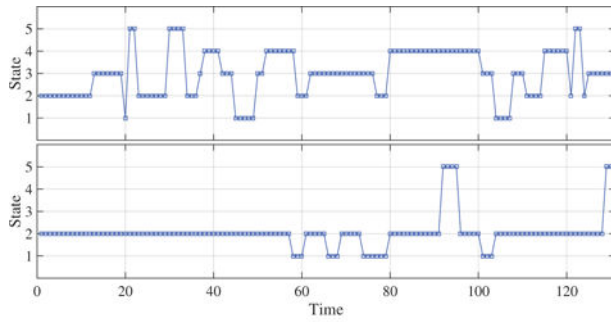
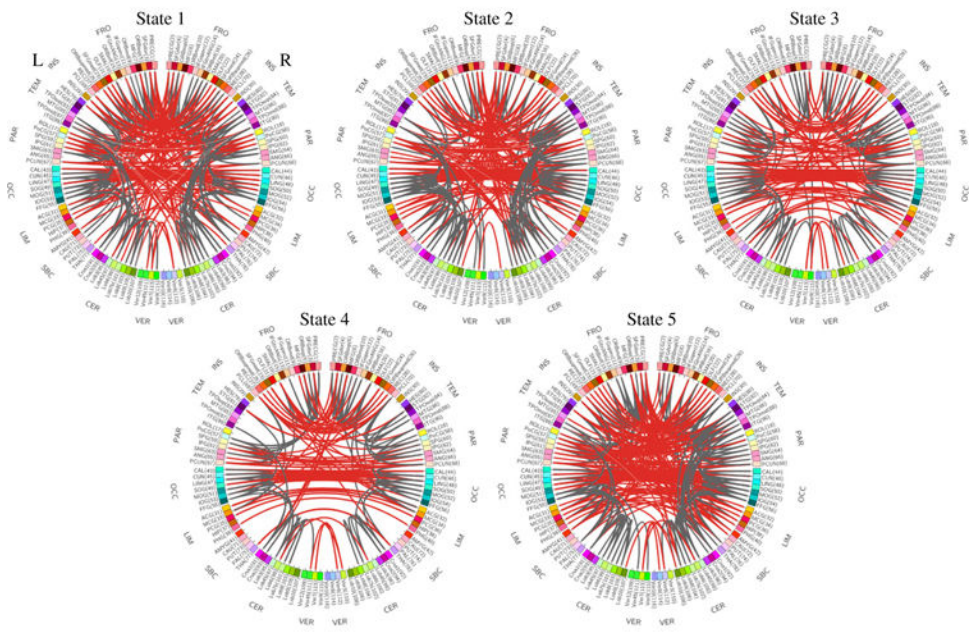(a) Normal healthy control
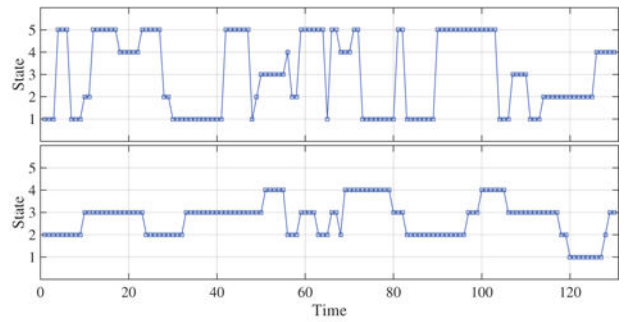
(b) Mild cognitive impairment

Fig. 9.
Examples of state decoding for each class. Note that the state transition occurs spontaneously and not all samples involve all the functional states during functional state changes (e.g., the bottom cases of each class).
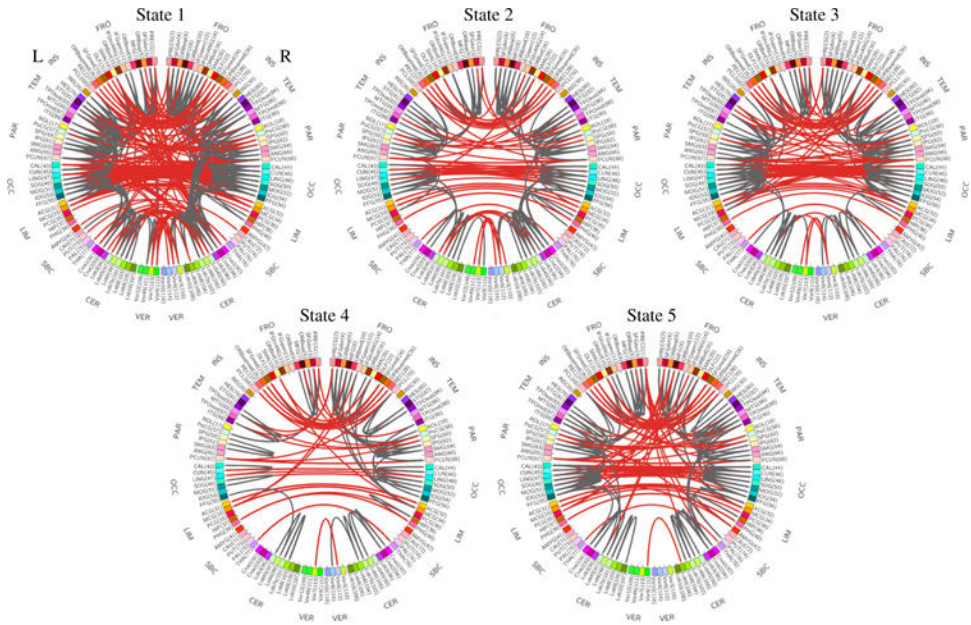
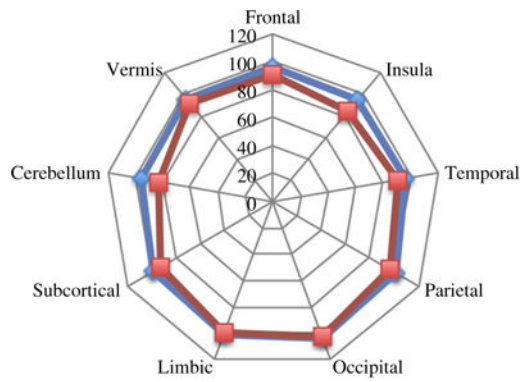(a) Normal healthy control

(b) Mild cognitive impairment

Fig. 10.
Comparison of functional connectivities of NC and MCI on the ADNI2 cohort. Note that connections whose absolute weights are higher than 0.65 are presented only and there is no correspondence between states of NC and MCI, i.e., the *i*-th state in NC is not necessarily related with the *i*-th state in MCI.
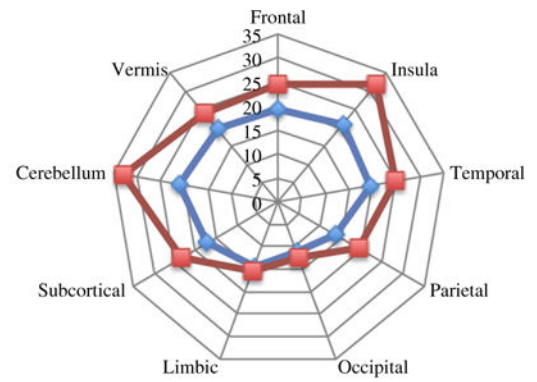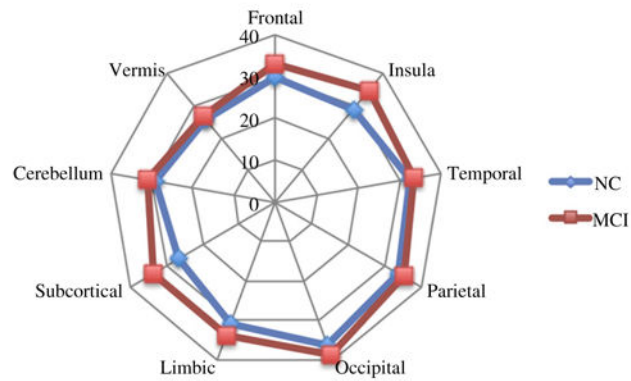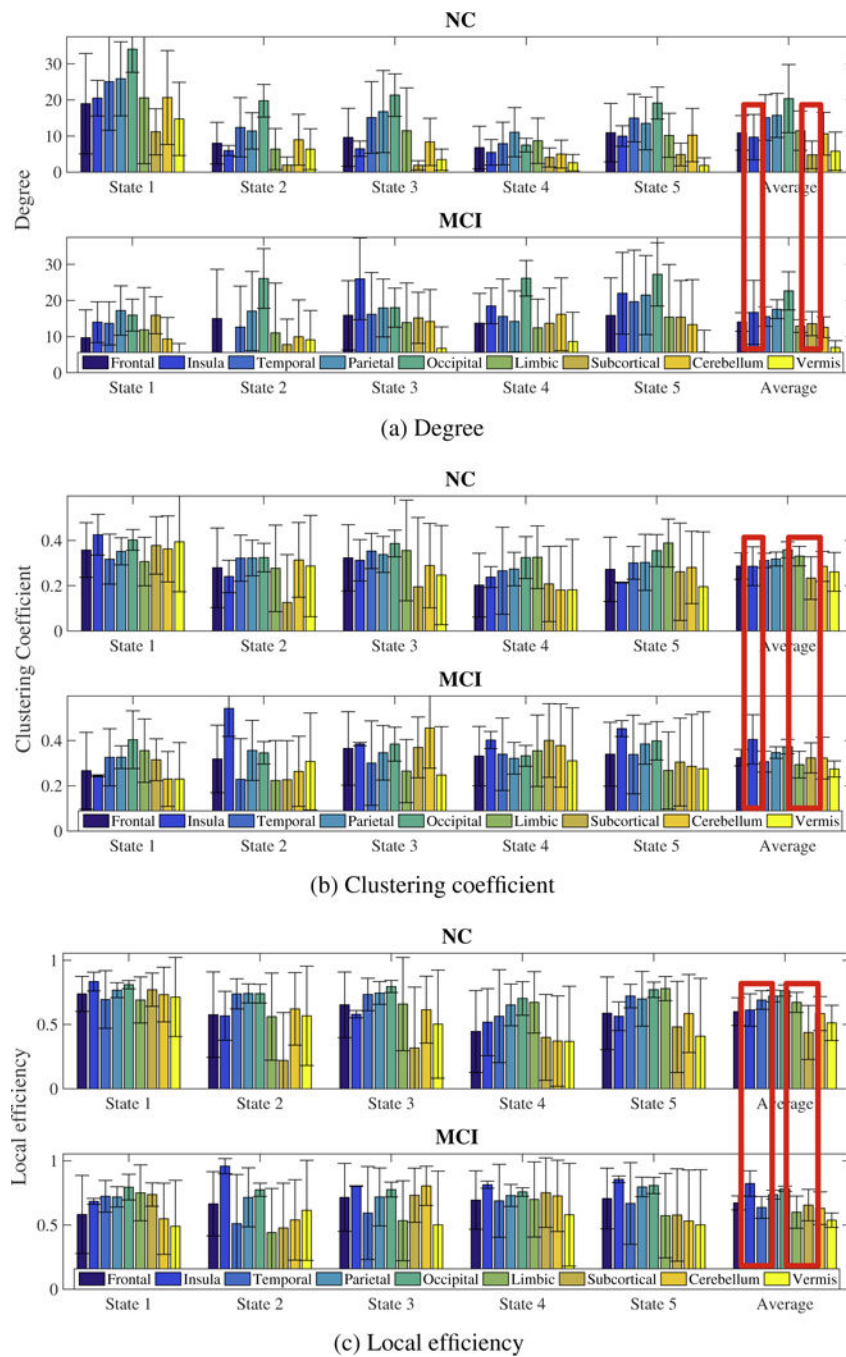
(a) Number of positive connections

(b) Number of negative connections

(c) Sum of absolute magnitudes

Fig. 11.
Comparison of the number of positive and negative connections and the sum of the absolution magnitudes in each lobe, averaged over all the estimated states on the ADNI2 cohort.

(a) Degree

(b) Clustering coefficient

(c) Local efficiency

Fig. 12.
Graph theoretic characteristics of the estimated functional connectivities. Note that there is no correspondence between states of NC and MCI.

Table 1

A summary of the performances of the competing methods on ADNI dataset. The boldface denotes the best performance in each metric. The performance of the proposed method was obtained with 5 hidden states and 3 Gaussians for NC and 5 hidden states and 4 Gaussians for MCI. (ACC: ACCuracy; SEN: SENsitivity; SPEC: SPECificity; PPV: Positive Predictive Value; NPV: Negative Predictive Value).

Methods	ACC (%)	SEN (%)	SPEC (%)	PPV (%)	NPV (%)
Baseline	61.29	62.96	60.00	54.84	67.74
gSR (Wee et al., 2014)	66.13	61.90	**75.00**	64.52	48.39
HMM + SDL (Eavani et al., 2013)	62.90	66.67	60.53	51.61	**74.19**
sDFN (Leonardi et al., 2013)	70.97	68.57	74.07	**77.42**	64.52
Proposed method	**72.58**	**70.59**	**75.00**	**77.42**	67.84

Table 2

A summary of the performances of the competing methods on in-house dataset. The boldface denotes the best performance in each metric. The performance of the proposed method was obtained with 5 hidden states and 3 Gaussians for NC and 5 hidden states and 4 Gaussians for MCI. (ACC: ACCuracy; SEN: SENsitivity; SPEC: SPECificity; PPV: Positive Predictive Value; NPV: Negative Predictive Value).

Methods	ACC (%)	SEN (%)	SPEC (%)	PPV (%)	NPV (%)
Baseline	64.86	44.44	71.43	33.33	80.00
gSR (Wee et al., 2014)	75.68	80.00	75.00	33.33	96.00
HMM + SDL (Eavani et al., 2013)	70.27	**100.0**	69.44	8.33	**100.0**
sDFN (Leonardi et al., 2013)	75.68	71.43	**92.00**	41.67	76.67
Proposed method	**81.08**	77.78	82.14	**58.33**	92.00