

Recovering Brain Structural Connectivity from Functional Connectivity via Multi-GCN based Generative Adversarial Network

Lu Zhang^{1(✉)}, Li Wang^{1,2}, and Dajiang Zhu¹

¹ Computer Science and Engineering, University of Texas at Arlington, Arlington, TX, USA

lu.zhang2@mavs.uta.edu

² Mathematics, University of Texas at Arlington, Arlington, TX, USA

Abstract. Understanding brain structure-function relationship, e.g., the relations between brain structural connectivity (SC) and functional connectivity (FC), is critical for revealing organizational principles of human brain. However, brain's many-to-one function-structure mode, i.e., diverse functional patterns may be associated with the same SC, and the complex direct/indirect interactions in both structural and functional connectivity make it challenge to infer a reliable relationship between SC and FC. Benefiting from the advances in deep neural networks, many deep learning based approaches are developed to model the complex and non-linear relations that can be overlooked by traditional shallow methods. In this work, we proposed a multi-GCN based generative adversarial network (MGCN-GAN) to infer individual SC based on corresponding FC. The generator of MGCN-GAN is composed by multiple multi-layer graph convolution networks (GCNs) which have the capability to model complex indirect connections in brain connectivity. The discriminator of MGCN-GAN is a single multi-layer GCN which aims to distinguish predicted SC from real SC. To overcome the inherent unstable behavior of GAN, we designed a new structure-preserving (SP) loss function to guide the generator to learn the intrinsic SC patterns more effectively. We tested our model on Human Connectome Project (HCP) dataset and the proposed MGCN-GAN model can generate reliable individual SC based on FC. This result implies that there may exist a common regulation between specific brain structural and functional architectures across different individuals.

Keywords: structural connectivity · functional connectivity · graph convolution networks · generative adversarial network.

1 Introduction

One of the major challenges in modern neuroscience is to understand brain structure-function relationship [1], such as the relations between brain structural connectivity (SC) [2] and functional connectivity (FC) [3]. Brain connectivity can be represented using a graph, comprising the nodes (e.g., brain regions)

and the connecting edges. For SC, the edges are often represented as the count of diffusion MRI derived fibers connecting to the regions. FC can be defined via Blood-Oxygen-Level-Dependent (BOLD) signal correlations using functional MRI (fMRI). Many studies have been published to investigate the potential relationship between brain structure and function, specifically, how SC and FC influence each other. Koch et al. [4] directly compared SC and FC and found a positive correlation between them in regions along the central sulcus. Greicius et al. [5] studied the relations between SC and four default mode network (DMN) related regions and found that strong functional connectivity can exist without direct structural connections. Skudlarski et al. [6] reported a significant overall agreement between SC and FC. Other studies [7–9] also reported similar results: strong structural connections are accompanied with strong functional connections. But functional connections can be observed between regions with little or no direct structural connections, which indicates functional connection may be mediated by indirect structural connections. In general, how to jointly represent and analyze these two types of brain networks is still challenging; because of brain’s many-to-one function-structure mode, traditional regression methods cannot be directly used to explore the relationship between SC and FC. Moreover, the individual variability and the non-linearity of SC and FC need to be considered simultaneously.

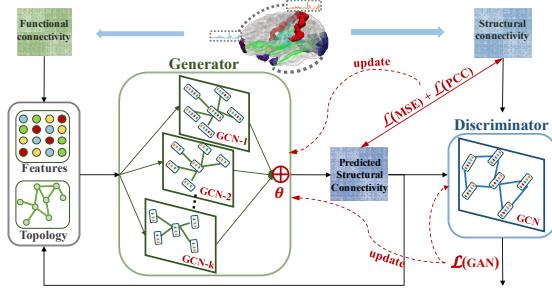


Fig. 1: An illustration of the proposed multi-GCN based generative adversarial network (MGCN-GAN). Firstly, by using Destrieux atlas [10] along with diffusion MRI and rs-fMRI data, we extracted the averaged BOLD signal of each brain region (148 regions in total). Then we created FC by Pearson’s correlation coefficient and constructed SC by ratio of number of fibers connecting two regions to the total number of fibers. SC was used as real samples to train the discriminator. FC was used as: 1) features associated with the nodes and 2) initialization of the GCM topology. The features and topology were fed into generator to predict SC. The predicted SC were used to 1) iteratively update the GCM topology and 2) train discriminator as fake samples. The generator is updated based on structure preserving (SP) loss function composed of MSE loss, PCC loss and GAN loss.

To tackle the abovementioned difficulties and motivated by the recent development of deep neural network based methods, we proposed a multi-GCN based generative adversarial network (MGCN-GAN) (**Fig. 1**) to generate individual

SC based on corresponding individual FC. Specifically, we adopted generative adversarial network [11, 12] to handle brain’s many-to-one function-structure mode. Moreover, in order to capture the complex relationship buried in both direct and indirect brain connections, we designed the generator and discriminator of MGCGN-GAN using graph convolution networks (GCN) [13, 14]. Compared to traditional CNN-based GAN that can only operate on regular, Euclidean data, our model can handle interrelated and hidden structures that beyond the grid neighbors, such as brain connectivity. In addition, to overcome the inherent unstable behavior of GAN, we proposed a novel structure-preserving (SP) loss function to guide the generator to learn the intrinsic SC patterns more effectively. We tested our method on Human Connectome Project (HCP) [15] dataset and the proposed MGCGN-GAN can generate reliable individual SC based on FC. More importantly, our results imply that there may exist a common regulation between specific brain structural and functional architectures across individuals.

2 Methods

2.1 Data Collection and Preprocessing

In this work, we used diffusion magnetic resonance imaging (diffusion MRI) and resting state functional magnetic resonance imaging (rs-fMRI) of 300 subjects in Human Connectome Project (HCP) [15] dataset. The diffusion MRI data has 111 slices and FOV = 210×180 with 1.25mm isotropic voxels, TE=89.5ms, TR=5.52s, flip angle = 78° . The rs-fMRI data has 72 slices and FOV = 208×180 with 2.0 mm isotropic voxels, TE=33.1ms, TR=0.72s, flip angle = 52° and there are 1200 volumes for each subject.

We applied standard preprocessing procedures including skull removal for both modalities, spatial smoothing, slice time correction, temporal pre-whitening, global drift removal and band pass filtering (0.01-0.1 Hz) for rs-fMRI, eddy current correction and fiber tracking via MedINRIA for diffusion MRI, registering rsfMRI to diffusion MRI space using FLIRT and adopt the Destrieux Atlas [10] for ROI labeling. The brain cortex was partitioned into 148 regions after removing two unknown areas and two empty areas.

2.2 Problem Description

In this work, we proposed a MGCGN-GAN model to generate individual SC based on corresponding FC. Specifically, the whole brain is represented as a network with $N = 148$ nodes (brain regions). We used the ratio of the number of fibers connecting two ROIs to the total number of fibers to create SC, denoted as $\mathbf{A} \in \mathbf{R}^{N \times N}$. As for FC, we used Pearson’s correlation coefficient between two average fMRI signals of two ROIs to create FC, denoted as $\mathbf{P} \in \mathbf{R}^{N \times N}$. The proposed MGCGN-GAN model is built on two components: Multi-GCN based generator and single-GCN based discriminator. The generator takes the following steps to generate an SC for a given FC: (i) FC is used as both features associated with nodes and the initialization of topology of brain network; (ii) based on

current topology, FC is mapped to different feature spaces by each multi-layer GCN component of generator in order to explore the latent relationship between SC and FC, so that multiple output feature matrices can be obtained; (iii) all the output feature matrices are combined by learnable coefficients to generate predicted SC; (iv) the topology is updated by the predicted SC. The discriminator is a classifier to label the input SC as real SC samples or fake samples. Given the training data consisting of FC samples and their corresponding real SC samples, the generator is trained based on SP loss function (Section 2.4) and the discriminator is trained by standard cross-entropy loss [11, 12].

2.3 Multi-GCN Based GAN (MGCN-GAN)

Similar to vanilla GAN [11, 12], MGCN-GAN is composed of two components, i.e., generator and discriminator. The generator is trained to generate real-like individual SC by competing with the discriminator based on an adversarial training scheme. Inspired by the great success of CNN that uses multiple filters to learn features from different feature spaces, the proposed generator consists of multiple multi-layer GCNs. Different GCN components are designed for different feature space and each of them will learn a latent mapping from individual FC to its corresponding SC. Through paralleling multiple GCNs, generator has the capacity to model complex relationship between FC and SC, which will be demonstrated by our experimental results in Section 3. The discriminator is composed by a single multi-layer GCN and two fully-connect layers, which aims to distinguish the predicted SC from real SC.

Generator. The generator is composed by k multi-layer GCNs. It is formulated as:

$$g(\mathbf{T}, \mathbf{P}, \boldsymbol{\theta}) = \boldsymbol{\theta} \oplus (\mathbf{G}_1 || \mathbf{G}_2 || \mathbf{G}_3 || \dots || \mathbf{G}_k), \quad (1)$$

$$\mathbf{SC}^p(\mathbf{P}, \boldsymbol{\theta}) = g(g(\mathbf{T}, \mathbf{P}, \boldsymbol{\theta}), \mathbf{P}, \boldsymbol{\theta}), \quad (2)$$

where \mathbf{G}_i , $i = 1, 2, \dots, k$ represents i^{th} GCN and $||$ denotes parallel operation. Each GCN takes the individual FC samples as input and outputs a predicted SC of the same individual. Then, we used the learnable coefficient $\boldsymbol{\theta}$ to fuse (\oplus) these k predictions to form $g(\mathbf{T}, \mathbf{P}, \boldsymbol{\theta})$, and obtained the final prediction, denoted as \mathbf{SC}^p , by updating current topology \mathbf{T} with the updated $\mathbf{T} = g(\mathbf{T}, \mathbf{P}, \boldsymbol{\theta})$. Each n -layer GCN \mathbf{G}_i is defined as:

$$\mathbf{G}_i(\mathbf{T}, \mathbf{P}) = f(\mathbf{T}\mathbf{H}_i^{n-1}\mathbf{W}_i^n), \quad (3)$$

$$\mathbf{H}_i^n = \begin{cases} f(\mathbf{T}\mathbf{H}_i^{n-1}\mathbf{W}_i^n), & n \geq 1, \\ \mathbf{P}, & n = 0. \end{cases} \quad (4)$$

As shown in **Eq. (3)** and **Eq. (4)**, each GCN has two inputs that represent the features and topology of the graph data. In our work, we used FC matrix \mathbf{P} as features. The topology \mathbf{T} was initialized by \mathbf{P} and iteratively updated by \mathbf{SC}^p . f is the nonlinear activation function and we used *Relu* in our experiments. \mathbf{H}_i^n is the output of n^{th} graph convolution layer of \mathbf{G}_i . $\mathbf{W}_i^n \in \mathbf{R}^{D_i \times D_o}$ is the

weight matrix of n^{th} graph convolution layer of \mathbf{G}_i , D_i is the dimension of input features and D_o is the dimension of output features. Each graph convolution layer selects and combines features from its neighbors based on topology and maps this combination to output feature space based on \mathbf{W}_i^n . By stacking multiple layers, information from high-order neighbors (the nodes that are connected via other nodes) are integrated along the topology \mathbf{SC}^p , which enables the generator to capture complex indirect relationship. After training, each multi-layer GCN defines a mapping from FC to SC, which reflects a latent relationship between FC and SC. In order to enhance the capability of generator, we parallel multiple GCNs to capture the complex interrelated relationships between SC and FC.

Discriminator. In order to distinguish the two sets of graph data (predicted SC and real SC), the discriminator consists of a multi-layer GCN, $\mathbf{G}_d(\mathbf{SC}, \mathbf{I})$, and two fully-connect layers. The input \mathbf{SC} represents the real SC matrix – \mathbf{A} , derived from diffusion MRI and predicted SC matrix – \mathbf{SC}^p , from generator. They are treated as real and fake samples for the discriminator training. Different from generator, we used identity matrix as input feature matrix for discriminator. This is because discriminator aims to learn the rules that can be used to decide whether the input connectivity matrix is a valid SC matrix or not, any external knowledge should be excluded.

2.4 Structure-Preserving (SP) Loss Function

The generator is optimized according to the feedback of discriminator. However, in the SC prediction task, the discriminator is much easier to train than the generator. The discriminator may easily classify real SC from predicted SC after a few training iterations and the generative adversarial loss would be close to 0, resulting in zero back-propagated gradients in generator. In such case, the generator cannot be optimized and will keep generating invalid SC. To break this dilemma, maintaining the balance between generator and discriminator regarding the optimization capability during the entire training process is important. We designed a new structure-preserving (SP) loss function to train discriminator and generator, which is a combination of three loss functions: mean squared error (MSE) loss, Pearson’s correlation coefficient (PCC) loss and GAN loss. SP loss function is formulated by Eq. (5).

$$SPLoss = GAN + \alpha MSE + \beta PCC, \quad (5)$$

where the regularization parameters α and β are initialized by 1 and will keep decreasing during the training process. MSE loss aims to force the predicted SC to be the same scale with real SC at element-wise level. PCC loss try to maximize the similarity of overall pattern between predicted SC and real SC. It consists of two components: 1) brain-level PCC loss and 2) region-level PCC loss. Brain-level PCC loss calculates the PCC between predicted SC matrix and real SC matrix, which measures the overall correlation between the predicted SC and real SC. Region-level PCC loss calculates the correlation between predicted SC and real SC of each brain region (each row/column of the connectivity matrix).

3 Results

3.1 Predicted Structural Connectivity

In our experiment, we used 180 subjects as training dataset and 120 subjects as testing dataset. All results showed in this section are from testing dataset. For detailed network architecture, three two-layer GCNs are paralleled in generator and the discriminator is composed of one three-layer GCN followed by two fully-connect layers. The feature dimensions of GCNs in generator are: $G_1 = (74, 148)$, $G_2 = (148, 148)$, $G_3 = (296, 148)$. $G_i = (F_1, F_2, \dots, F_n)$ represents an n -layer GCN and output feature dimension of layer i is F_i . Similarly, feature dimensions of discriminator are : $G_d = (148, 296, 148)$, which is followed by two fully connected layers with output feature dimensions 1024 and 2, respectively.

Fig. 2 (a) shows the predicted SC and real SC of 5 randomly selected subjects. As shown in **Fig. 2(a)**, the overall similarity between the real SC and the corresponding predicted SC is very high. To better demonstrate the details of the prediction result, we extracted two patches at the same location of predicted SC and real SC for all 5 subjects. These patches are enlarged and showed in the middle of **Fig. 2 (a)**. From the enlarged patches, we can see that different individuals possess different SC though their overall patterns are similar. Our model can not only generate the similar patterns across individuals, but also predict the subtle individual differences. All these predictions are based on individual FC.

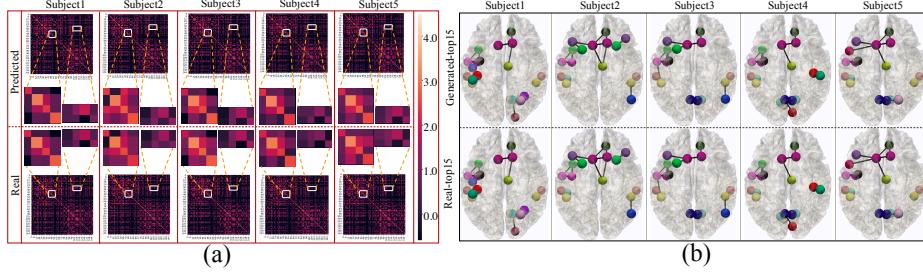


Fig. 2: Illustration of predicted SC and real SC of 5 randomly selected subjects. **(a)** shows the predicted SC matrices (the first row) and real SC matrices (the second row) derived from diffusion MRI. Each column belongs to different individuals. Two patches of the matrices are extracted from the same location and their enlarged patches are showed in the middle. **(b)** shows the top 15 strongest connectivity of real SC and predicted SC in the brain space. The colorful bubbles and links represent different brain regions and structural connectivity, respectively. The colors used in this figure are the same in Destrieux atlas in FreeSurfer. The five subjects showed in **(b)** are the same ones in **(a)**.

Fig. 2 (b) shows the top 15 strongest connectivity of real SC and predicted SC for the same 5 subjects in the brain space. Because of individual variability, the top 15 connectivity of different subjects are different. But our model can

accurately predict these differences as well as individual patterns based on individual FC only. This result further confirms that our method can effectively generate accurate individual SC instead of generating common SC patterns at population level. Note that we used individual SC and FC as a pair-input for training and testing, which means the learned model represents a common mapping between individual SC and FC. This result suggests that there may exist a common regulation between specific brain structural and functional architectures across individuals.

3.2 Model Comparison

As mentioned before, the generator of MGCN-GAN is composed by multiple GCNs. In order to verify the necessity of multi-GCN architecture, we conducted comparison experiments with different generator architectures and evaluated our results by three measures: 1) MSE (real, gen), 2) MSE (other-reals, gen) and 3) MSE (other-reals, gen)-MSE (real, gen). MSE (real, gen) is the average MSE between real SC and predicted SC from the same subject, which measures the similarity between the real SC and the corresponding predicted SC. Smaller MSE (real, gen) indicates higher similarity. Thus, to generate reliable SC, the MSE (real, gen) should keep decreasing before converged. MSE (other-reals, gen) is the average MSE between predicted SC and real SC of other subjects. A reliable predicted SC should avoid to be “trapped” in common SC patterns at population level. Therefore, we expect MSE (other-reals, gen) to keep increasing during training process. MSE (other-reals, gen)-MSE (real, gen) is the difference of MSE (real, gen) and MSE (other-reals, gen) and an increasing value is expected. **Fig. 3** shows the results of different architectures, which are evaluated by the three measures. It is obvious that the predicted SCs generated from multi-GCN generator can maintain the individual differences in SCs, while single-GCN generator shows obvious worse performance.

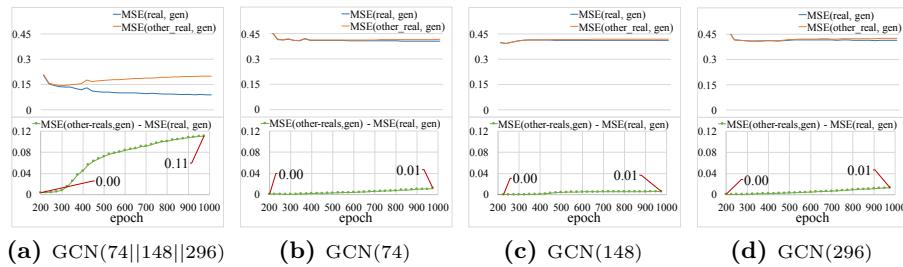


Fig. 3: Results of different generator architectures. The $\text{GCN}(D_1||D_2||\dots||D_k)$ represents the architecture of generator, the generator is composed of k two-layer GCNs, and the output feature dimension of the first layer of i^{th} GCN is D_i . The results of different generator architectures are evaluated by MSE (real, gen), MSE (other-reals, gen) and MSE (other-reals, gen)-MSE (real, gen).

The multiple GCNs in the generator are combined with learnable coefficients. In our experiments, we initialized the coefficients with different values and found that the coefficients with different initializations will converge to a consistent ratio that all the GCN components seem to contribute equally to the results and each of them is indispensable. One explanation is that similar to the filters in CNN, multiple GCNs with different size of output features may be more flexible and efficient for characterizing the complex SC-FC mapping.

3.3 Loss Function Comparison

To demonstrate the superiority of the proposed SP loss function, we trained our network with different loss function and the results of comparison are showed in **Fig. 4**. From the results, we can see that our SP loss function outperforms other loss function. The reason is that MSE only focuses on the element-wise similarity within the connectivity and ignores the overall patterns. Though PCC have better performance to describe the overall connectivity patterns, it may also overlook the connection magnitude across different connectivity. However, both of them are important in our designed SP loss to capture the subtle differences between real and predicted SC.

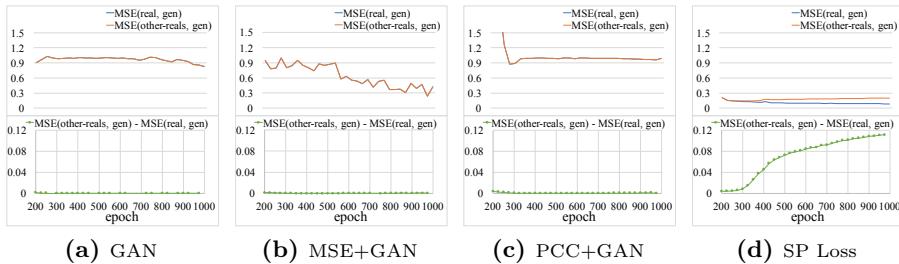


Fig.4: Results of MGCGAN with different loss functions. We used single GAN loss, the combination of GAN loss and MSE loss, the combination of GAN loss and PCC loss and the proposed SP loss to conduct experiments and the results are evaluated based on the three measures discussed in Section 3.2.

4 Conclusion and Discussion

In this work, we proposed a novel multi-GCN based generative adversarial network (MGCGAN) to generate individual SC based on corresponding FC. We adopted GCN based generator and discriminator to model the interrelated hidden structures of brain network and used multi-GCN architecture to capture the complex relationship between SC and FC. Moreover, we designed a new structure-preserving loss function to make generator more effective when differentiating real and predicted SCs. Using HCP dataset as a testbed, our MGCGAN can not only predict reliable individual SC, but also capture the subtle difference across individuals.

References

1. Park, H.J., Friston, K.: Structural and functional brain networks: from connections to cognition. *Science* **342**(6158), 1238411 (2013)
2. Yeh, F.C., Vettel, J.M., Singh, A., Poczos, B., Grafton, S.T., Erickson, K.I., Tseng, W.Y.I., Verstynen, T.D.: Quantifying differences and similarities in whole-brain white matter architecture using local connectome fingerprints. *PLoS computational biology* **12**(11) (2016)
3. Gratton, C., Laumann, T.O., Nielsen, A.N., Greene, D.J., Gordon, E.M., Gilmore, A.W., Nelson, S.M., Coalson, R.S., Snyder, A.Z., Schlaggar, B.L., et al.: Functional brain networks are dominated by stable group and individual factors, not cognitive or daily variation. *Neuron* **98**(2), 439–452 (2018)
4. Koch, M.A., Norris, D.G., Hund-Georgiadis, M.: An investigation of functional and anatomical connectivity using magnetic resonance imaging. *Neuroimage* **16**(1), 241–250 (2002)
5. Greicius, M.D., Supekar, K., Menon, V., Dougherty, R.F.: Resting-state functional connectivity reflects structural connectivity in the default mode network. *Cerebral cortex* **19**(1), 72–78 (2009)
6. Skudlarski, P., Jagannathan, K., Calhoun, V.D., Hampson, M., Skudlarska, B.A., Pearlson, G.: Measuring brain connectivity: diffusion tensor imaging validates resting state temporal correlations. *Neuroimage* **43**(3), 554–561 (2008)
7. Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C.J., Wedeen, V.J., Sporns, O.: Mapping the structural core of human cerebral cortex. *PLoS biology* **6**(7) (2008)
8. Honey, C., Sporns, O., Cammoun, L., Gigandet, X., Thiran, J.P., Meuli, R., Hagmann, P.: Predicting human resting-state functional connectivity from structural connectivity. *Proceedings of the National Academy of Sciences* **106**(6), 2035–2040 (2009)
9. Teipel, S.J., Pogarell, O., Meindl, T., Dietrich, O., Sydkova, D., Hunklinger, U., Georgii, B., Mulert, C., Reiser, M.F., Möller, H.J., et al.: Regional networks underlying interhemispheric connectivity: an eeg and dti study in healthy ageing and amnestic mild cognitive impairment. *Human brain mapping* **30**(7), 2098–2119 (2009)
10. Destrieux, C., Fischl, B., Dale, A., Halgren, E.: Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *Neuroimage* **53**(1), 1–15 (2010)
11. Goodfellow, I.: Nips 2016 tutorial: Generative adversarial networks. arXiv preprint arXiv:1701.00160 (2016)
12. Hong, Y., Hwang, U., Yoo, J., Yoon, S.: How generative adversarial networks and its variants work: an overview of gan. arXiv preprint arXiv:1711.05914 (2017)
13. Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., Yu, P.S.: A comprehensive survey on graph neural networks. arXiv preprint arXiv:1901.00596 (2019)
14. Zhou, J., Cui, G., Zhang, Z., Yang, C., Liu, Z., Wang, L., Li, C., Sun, M.: Graph neural networks: A review of methods and applications. arXiv preprint arXiv:1812.08434 (2018)
15. Van Essen, D.C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T., Bucholz, R., Chang, A., Chen, L., Corbetta, M., Curtiss, S.W., et al.: The human connectome project: a data acquisition perspective. *Neuroimage* **62**(4), 2222–2231 (2012)