2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)
Venice, Italy, April 8-11, 2019

STRUCTURALLY-INFORMED DECONVOLUTION OF FUNCTIONAL MAGNETIC RESONANCE IMAGING DATA

Thomas AW Bolton[⋆†], *Younes Farouj*[⋆†], *Mert Inan*[‡], *Dimitri Van De Ville*[⋆†]

[⋆] Institute of Bioengineering, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland
[†] Department of Radiology and Medical Informatics, University of Geneva (UNIGE), Switzerland
[‡] Department of Biological Sciences, Mellon College of Science, Carnegie Mellon University, PA

ABSTRACT

Neural activity occurs in the shape of spatially organized patterns: networks of brain regions activate in synchrony. Many of these functional networks also happen to be strongly structurally connected. We use this information to revisit the fundamental problem of functional magnetic resonance imaging (fMRI) data deconvolution. Using tools from graph signal processing (GSP), we extend total activation, a spatio-temporal deconvolution technique, to data defined on graph domains. The resulting approach simultaneously cancels out the effect of the haemodynamics, and promotes spatial patterns that are in harmony with predefined structural wirings. More precisely, we minimize a functional involving one data fidelity and two regularization terms. The first regularizer uses the concept of generalized total variation to promote sparsity in the activity transients domain. The second term controls the overall spatial variation over the graph structure. We demonstrate the relevance of this structurally-driven regularization on synthetic and experimental data.

Index Terms— Functional magnetic resonance imaging, Deconvolution, Graph signal processing, Total activation

1. INTRODUCTION

Since its appearance in the early 1990s, blood oxygenated level-dependent (BOLD) functional magnetic resonance imaging (fMRI) [1] has been shedding light on brain responses to diverse tasks, or at rest. The acquired BOLD signal indirectly reflects neuronal activity changes through the dynamics of cerebral blood flow, vascular responses and fluctuations in (de)oxygenated haemoglobin concentration. The overall mechanism linking neuronal events to the BOLD signal output can be, roughly, described as a nonlinear system derived from dynamic models. In practice, it is common to use linearised versions of such systems, leading to linear time-invariant (LTI) systems that are fully characterised by their impulse response, the *haemodynamic response function* (HRF).

In opposition to task-related activity, for which regression techniques can be employed [2], moments of resting-state (RS) activity are more challenging to retrieve. A natural solution to this problem is to view the reconstruction of activity signals as a temporal deconvolution task. In particular, the emergence of ℓ_1-type regularization inspired total activation (TA) [3], a framework which originally combined a sparsity prior on transients in the fMRI signal and a spatial regularization expressed as ℓ_2-smoothness within regions of a predefined brain atlas. It was then shown that controlling the overall spatial total variation of the activity is sufficient to produce well-defined activity maps in data-driven fashion [4].

In the present work, we follow the growing interest in multi-modal approaches in fMRI data mining pipelines [5, 6], and propose a region-level development of the original TA version by providing information about the underlying physical brain wiring (obtained from diffusion-weighted MRI data) to the spatial regularization term.

2. MODELING

2.1. Temporal modeling

Following previous work [3], we consider a matrix $\mathbf{Y} \in \mathbb{R}^{N\times T}$ of fMRI data, with N the number of brain regions and T the number of time points, and consider that each regional time course verifies:

$$\mathbf{Y}(n,\cdot) = \mathbf{X}(n,\cdot) + \varepsilon_n \quad \forall n = 1,\cdots,N, \tag{1}$$

where $\mathbf{X}$ is the matrix of *activity-related signals* and ε_n is a Gaussian noise component. An activity-related time course $\mathbf{X}(n,\cdot)$ is related to its *activity-inducing* variant $\mathbf{S}(n,\cdot)$ as $\mathbf{X}(n,\cdot) = \mathcal{H}\{\mathbf{S}(n,\cdot)\} \quad \forall n = 1,\cdots,N$, where $\mathcal{H}$ is a differential operator that models the effect of the haemodynamics. A crucial assumption is to model $\mathbf{S}(n,\cdot)$ as a boxcar signal composed of a baseline and sustained on and off moments. This implies the sparsity of transients: each activity-inducing time course has a derivative composed of Dirac pulses.

2.2. Spatial modeling

We describe the spatial connections between the N regions as a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$ characterised by a set of vertices

$\{v_n\} \in \mathcal{V}$ with $|\mathcal{V}| = N$ and edges $\{e_{n,m}\} \in \mathcal{E} \subset \mathcal{V} \times \mathcal{V}$, with $1 \leq n < m \leq N$. Each edge has an associated weight $w_{n,m}$, and the matrix $\mathbf{W} \in \mathbb{R}^{|\mathcal{E}| \times |\mathcal{E}|}$ contains the square roots of these weights: $W_{n,m} = \sqrt{w_{n,m}}$. It controls the confidence in a given connection between a pair of regions; *i.e.*, how likely they are supposed to activate together.

3. METHODS

In this paper, we recover $\mathbf{X}$, and consequently $\mathbf{S}$, by minimising a functional of the form

$$\begin{aligned}\widehat{\mathbf{X}} = \arg\min_{\mathbf{X}} \{ & \frac{1}{2}||\mathbf{Y} - \mathbf{X}||_2^2 \\ & + \sum_n \mu_1(n) ||\mathrm{L}_1\{\mathbf{X}(n,\cdot)\}||_1 \\ & + \sum_t \mu_2(t) ||\mathrm{L}_2\{\mathbf{X}(\cdot,t)\}||_1 \},\end{aligned} \tag{2}$$

where L_1 and L_2 are sparsifying operators that respectively act on temporal and spatial variables, and μ_1 and μ_2 are the corresponding regularization parameters. Solving (2) aims at controlling the overall total variation of the activity. We now describe below how to construct L_1 and L_2.

3.1. Hæmodynamic deconvolution

The hæmodynamic deconvolution aims at undoing the observed signals from the effect of $\mathcal{H}$ by means of the operator L_1. Its construction was discussed in previous work [3, 4], which we briefly recall here for the reader's convenience. We denote by D the one-dimensional derivation (finite difference) operator. For $n = 1, \cdots, N$, the starting point is the set of activity-inducing signals $\mathbf{S}(n,\cdot)$, whose derivatives $\mathrm{D}\{\mathbf{S}(n,\cdot)\}$ are sparse. The operator L_1 should be constructed in a way that imposes this sparsity feature on $\mathbf{X}$:

$$\mathrm{L}_1\{\mathbf{X}(n,\cdot)\} = \mathrm{D}\{\mathbf{S}(n,\cdot)\} = \mathrm{D}\{\mathcal{H}^{-1}\{\mathbf{X}(n,\cdot)\}\}, \tag{3}$$

where $\mathcal{H}^{-1}$ is the inverse of $\mathcal{H}$. By identification, we have $\mathrm{L}_1 = \mathrm{D}\{\mathcal{H}^{-1}\}$. When $\mathcal{H}$ can be defined as a transfer function, L_1 can be constructed elegantly by inverting the roles of the poles and zeros and adding a zero at the origin.

3.2. Structurally-informed regularization

The main idea of the structurally-driven regularization is to minimise the differences between connected brain regions using the operator L_2. Similarly to the temporal regularization, this can be done using a spatial total variation over the graph $\mathcal{G}$. For this, we need foundational tools from graph signal processing (GSP). In particular, we want to define the discrete gradient operation for signals defined on $\mathcal{G}$. This can be done through the so-called *incidence matrix* $\mathbf{B} \in \mathbb{R}^{|\mathcal{E}| \times N}$ in its normalised form:

$$B_{e_{n,m},v_k} := \begin{cases} \dfrac{-1}{\sqrt{deg(v_n)}}, & \text{if} \quad n = k \\ \dfrac{1}{\sqrt{deg(v_m)}}, & \text{if} \quad m = k \\ 0 & \text{otherwise,} \end{cases} \tag{4}$$

where $deg(v_n)$ is the degree of the vertex v_n, that is, the number of edges incident to it. It is now possible to define the weighted gradient of column signal from $\mathbf{X}$ using the matrices $\mathbf{W}$ and $\mathbf{B}$. The operator L_2 is defined easily as:

$$\mathrm{L}_2 = \mathbf{WB}, \tag{5}$$

and $||\mathrm{L}_2\{\mathbf{X}(\cdot,t)\}||_1$ is the weighted spatial total variation of the fMRI signal at time step t.

3.3. Minimisation algorithm

A distinctive feature of (2) is the fact that it involves two regularization terms. Fortunately, it is possible to split the problem into two parts using the generalised forward-backward (GFB) proximal splitting method [7]. The corresponding algorithm iterates through the following steps until convergence:

$$\begin{aligned}
&1: \mathbf{Y_t} = \widehat{\mathbf{X}} - \mathbf{X_t} + \mathbf{Y}, \\
&2: \mathbf{X_t} = \arg\min_{\mathbf{X}} \left\{ \tfrac{1}{2}||\mathbf{Y_t} - \mathbf{X}||_2^2 + \textstyle\sum_n \mu_1(n) ||\mathrm{L}_1\{\mathbf{X}(n,\cdot)\}||_1 \right\}, \\
&3: \mathbf{Y_s} = \widehat{\mathbf{X}} - \mathbf{X_s} + \mathbf{Y}, \\
&4: \mathbf{X_s} = \arg\min_{\mathbf{X}} \left\{ \tfrac{1}{2}||\mathbf{Y_s} - \mathbf{X}||_2^2 + \textstyle\sum_t \mu_2(t) ||\mathrm{L}_2\{\mathbf{X}(\cdot,t)\}||_1 \right\}, \\
&5: \widehat{\mathbf{X}} = \omega_t \mathbf{X_t} + \omega_s \mathbf{X_s},
\end{aligned}$$

where ω_t and ω_s are weight parameters to control the trade-off between the temporal and spatial regularizations. Here, all variables $(\mathbf{X_t}, \mathbf{X_s})$, $(\mathbf{Y_t}, \mathbf{Y_s})$ and $\widehat{\mathbf{X}}$ are initialized as zero-valued matrices. Steps 2 and 4 require the minimisation of functionals involving non-differentiable ℓ_1-norm terms. Here, we opted for fast iterative shrinkage thresholding algorithms (FISTA) [8] to estimate $\mathbf{X_t}$ and $\mathbf{X_s}$. Finally, another difficulty is the tuning of the regularization parameter vectors μ_1 and μ_2. For each region n, we estimate $\mu_1(n)$ as the standard deviation (SD) of first-scale coefficients of a Daubechies wavelet decomposition. At this scale, the coefficients are supposed to catch mainly high frequency fluctuations. Thus, their SD is an estimator of the noise component SD. On the other hand, we assume that the spatial regularization parameter $\mu_2(t)$ is the same for all time points: $\mu_2(t) = \mu_2^* \quad \forall t = 1, \cdots, T$. In practice, we find μ_2^* via an oracle search.

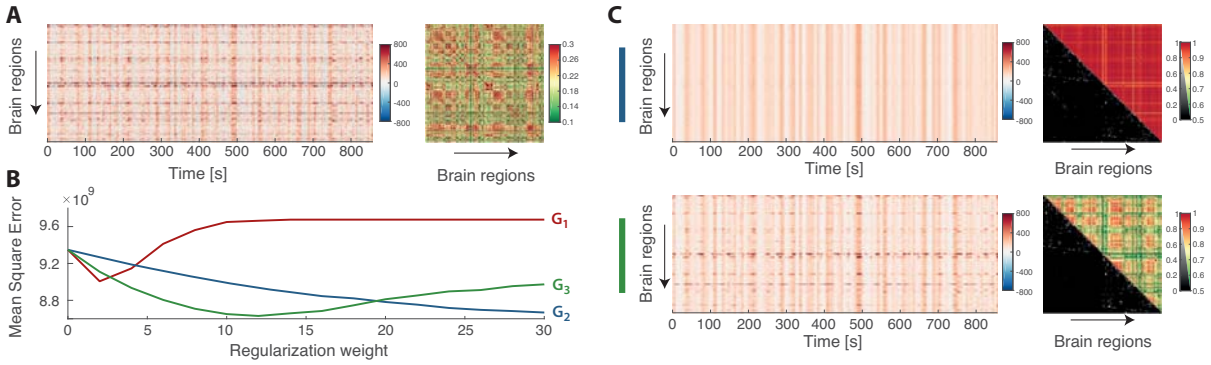

Fig. 1. **(A)** Regional time series (left panel) and correlation pattern (right panel) for example simulated data. **(B)** MSE for G1 (red curve), G2 (blue curve) and G3 (green curve) cases as a function of the spatial regularization parameter. **(C)** Output regional time courses, and correlation patterns (top right matrix halves), for the G2 (top row) and G3 (bottom row) cases. The associated employed structural information is also depicted as the bottom left halves of the matrices.

4. EXPERIMENTS

4.1. Evaluation of the approach

We retrieved a set of 360 considered regions from an atlas previously derived by multimodal imaging analysis [9]. To determine their modular assignment, we employed 14 maps derived in a previous independent component analysis (ICA) study [10]. Each region was assigned to the network with which it shared maximal overlap; this way, 275 areas could be matched and were thus considered in the presented results.

We analysed minimally preprocessed resting-state fMRI data acquired within the Human Connectome Project (HCP) initiative [11]. On top of realignment, co-registration and MNI (Montreal Neurological Institute) space warping, we manually performed detrending of the data, regression of constant, linear, quadratic, cerebrospinal fluid and white matter regressors, atlasing, scrubbing at a threshold of 0.5mm [12], and high-pass filtering at 0.01Hz.

4.1.1. Simulated data

To simulate artificial time courses, we hypothesised that the activity over time of a given brain region would consist in a weighted average of whole-brain spatial patterns (here, the aforementioned ICA maps). First, we converted the maps in the atlas domain by assigning them a value between 0 (a map does not overlap with a region) to 1 (full overlap). We then generated their activation time series by linear regression, and subsequently performed phase randomisation to generate artificial data with identical second-order properties as compared to the real ones. We returned to a regional time course description, using the weighted assignment of each brain region to a map. Finally, Gaussian noise was added, for each regional time course, sampled from a distribution with equal mean and standard deviation as the data.

We considered three constructions of $\mathcal{G}$ to compare, based on different nearness criteria: all nodes connected equivalently together (G1), each node connected to its k nearest neighbours (G2), or $\mathcal{G}$ constructed by structural wiring data obtained from diffusion-weighted MRI (G3). We assessed mean square error (MSE) between the ground truth artificial time courses, and the ones retrieved by our algorithm, for a range of spatial regularizer values $\mu_2 \in [0:2:30]$. We also examined correlation between regional time series at optimal regularization level (minimal MSE), comparing the observed patterns to the ground truth case.

4.1.2. Experimental fMRI data

The analysed experimental data were the regional time courses obtained from a test HCP subject. We compared the fMRI data to the activity-inducing signals obtained by our method (G3), using the optimal regularization level $\mu_2^{(G3)}$ derived from above. In addition, we also generated the activity-inducing signals when solely relying on the temporal or spatial regularization term.

5. RESULTS

5.1. Simulated data

Fig. 1A shows an example simulation of regional activity (left panel), and the associated correlation pattern across brain areas (right panel). Regarding performance of our algorithm (**Fig. 1B**), the G1 approach (red curve) reaches an optimum $MSE^{(G1)} = 9.006 \cdot 10^9$ at $\mu_2^{(G1)} = 2$. There was no optimum within the probed range of μ_2 values for G2 (blue curve), as MSE kept decreasing, highlighting convergence towards an over-regularized, trivial solution (as confirmed in **Fig. 1C**, top row). For G3 (green curve), however, there was also a clear error minimum $MSE^{(G3)} = 8.63 \cdot 10^9$ at $\mu_2^{(G3)} = 12$, and

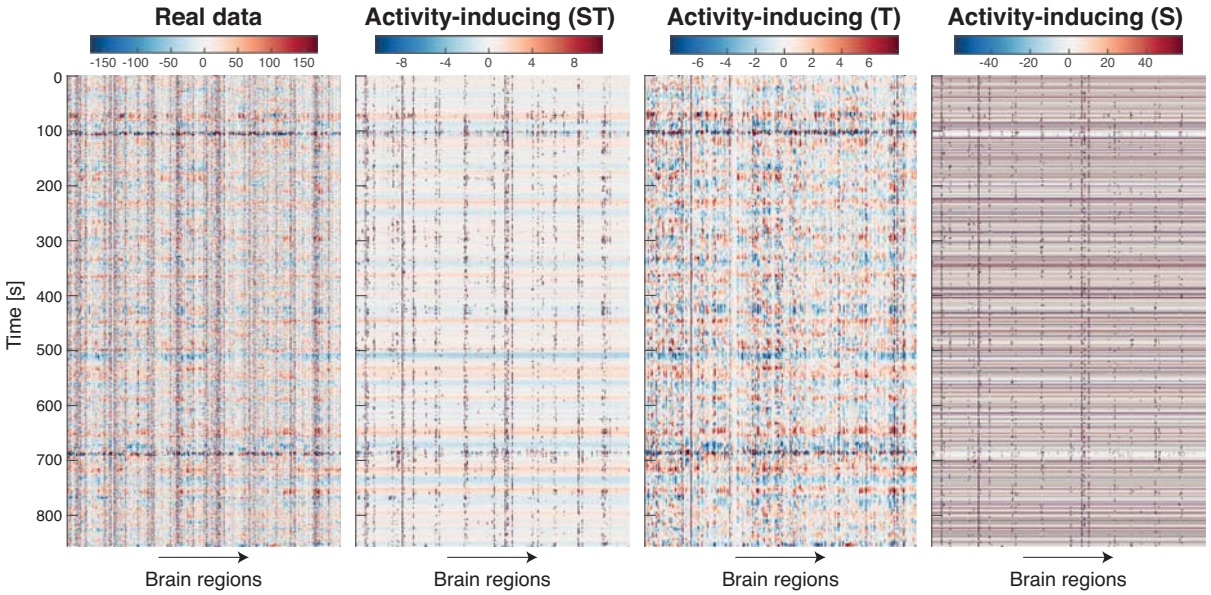

Fig. 2. Example real fMRI time courses (left panel) and recovered activity-inducing signals using the full model (ST panel), only the temporal regularizer (T panel), or only the spatial regularizer (S panel). Note that signal intensities differ across cases since we compare non-deconvolved data (real fMRI panel) to activity-inducing (deconvolved) signals.

it corresponded to retrieved signals close to the ground truth, both in terms of regional activity and correlation pattern (**Fig. 1C**, bottom row).

5.2. Experimental fMRI data

Fig. 2 shows an example output from our method on a test HCP subject, using $\mu_2^{(G3)} = 12$ as regularization level. Compared to the real fMRI time courses (left panel), the recovered activity-inducing signals (ST panel) are cleaner, and still show both whole-brain changes consistent across regions, and more regionally-specific fluctuations in activity. When only the temporal (T) or spatial (S) regularizer was included, the outputs were less accurate.

6. DISCUSSION

6.1. Neurophysiological relevance of the approach

Functional activity as captured by fMRI derives from the structural brain scaffold [13, 14]. Since many of the noise components that corrupt the truly neuronal signal, such as motion-related artefacts [15] or cross-subject vascular specificities [16], do not depend on structural brain architecture, the use of this knowledge to inform algorithms aimed at recovering functional information is expected to be valuable.

In accordance with this postulate, error measurements obtained on our simulated data case using a graph describing white matter architecture (G3) showed an optimal regularization regime, for which the results outperformed other investigated approaches (G1 and G2). This was translated, when applying the method to real fMRI data, into cleaner time courses nonetheless conserving whole-brain and regionally-localised signal changes.

6.2. Methodological comments

Generally speaking, the combination of structural and functional information into a joint analysis has been blossoming, be it using deep learning [17], multimodal parcellation approaches [9] or graph signal processing tools [18]. Our approach exploits this last avenue, as we use a graph description of brain structure to more accurately retrieve the set of functional signals.

Obtained results may be influenced by various methodological factors: first, the quality of tractography (amount of false positives/false negatives, extent of propagation to cortical areas); second, the exact employed preprocessing strategy on RS fMRI data [19]; third, the atlas used for parcellation: indeed, a larger amount of parcels into which to segment brain activity will, on the one hand, enable more accurate results (as long as extra information of relevance is gained by the further splitting), but on the other hand, computational time will be increased. At the limit, we reach the voxel-wise case, for which alternative deconvolution methods imposing piecewise smoothness of the responses over space have previously been suggested [4]. One could envision to push this approach one step further by making use of whole-brain (as opposed to regional as here) white matter information in a graph format.

7. REFERENCES

[1] S. Ogawa, T.-M. Lee, A. R. Kay, and D. W. Tank, "Brain magnetic resonance imaging with contrast dependent on blood oxygenation," *PNAS*, vol. 87, no. 24, pp. 9868–9872, 1990.

[2] K. J. Friston, A. P. Holmes, K. J. Worsley, J.-P. Poline, C. D. Frith, and R. S. Frackowiak, "Statistical parametric maps in functional imaging: a general linear approach," *Human brain mapping*, vol. 2, no. 4, pp. 189–210, 1994.

[3] F. I. Karahanoğlu, C. Caballero-Gaudes, F. Lazeyras, and D. Van De Ville, "Total activation: fMRI deconvolution through spatio-temporal regularization," *Neuroimage*, vol. 73, pp. 121–134, 2013.

[4] Y. Farouj, F. I. Karahanoğlu, and D. Van De Ville, "Regularized spatiotemporal deconvolution of fMRI data using gray-matter constrained total variation," in *Biomedical Imaging (ISBI 2017), 2017 IEEE 14th International Symposium on*. Ieee, 2017, pp. 472–475.

[5] T. A. Bolton, Y. Farouj, S. Obertino, and D. Van De Ville, "Graph slepians to strike a balance between local and global network interactions: Application to functional brain imaging," in *Biomedical Imaging (ISBI 2018), 2018 IEEE 15th International Symposium on*. IEEE, 2018, pp. 1239–1243.

[6] W. Huang, T. A. Bolton, J. D. Medaglia, D. S. Bassett, A. Ribeiro, and D. Van De Ville, "A Graph Signal Processing Perspective on Functional Brain Imaging," *Proceedings of the IEEE*, 2018.

[7] H. Raguet, J. Fadili, and G. Peyré, "A generalized forward-backward splitting," *SIAM Journal on Imaging Sciences*, vol. 6, no. 3, pp. 1199–1226, 2013.

[8] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM journal on imaging sciences*, vol. 2, no. 1, pp. 183–202, 2009.

[9] M. F. Glasser, T. S. Coalson, E. C. Robinson, C. D. Hacker, J. Harwell, E. Yacoub, K. Ugurbil, J. Andersson, C. F. Beckmann, M. Jenkinson *et al.*, "A multi-modal parcellation of human cerebral cortex," *Nature*, vol. 536, no. 7615, pp. 171–178, 2016.

[10] W. Shirer, S. Ryali, E. Rykhlevskaia, V. Menon, and M. D. Greicius, "Decoding subject-driven cognitive states with whole-brain connectivity patterns," *Cerebral cortex*, vol. 22, no. 1, pp. 158–165, 2012.

[11] S. M. Smith, C. F. Beckmann, J. Andersson, E. J. Auerbach, J. Bijsterbosch, G. Douaud, E. Duff, D. A. Feinberg, L. Griffanti, M. P. Harms *et al.*, "Resting-state fMRI in the human connectome project," *Neuroimage*, vol. 80, pp. 144–168, 2013.

[12] J. D. Power, K. A. Barnes, A. Z. Snyder, B. L. Schlaggar, and S. E. Petersen, "Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion," *Neuroimage*, vol. 59, no. 3, pp. 2142–2154, 2012.

[13] C. Honey, O. Sporns, L. Cammoun, X. Gigandet, J.-P. Thiran, R. Meuli, and P. Hagmann, "Predicting human resting-state functional connectivity from structural connectivity," *Proceedings of the National Academy of Sciences*, vol. 106, no. 6, pp. 2035–2040, 2009.

[14] M. P. Van Den Heuvel, R. C. Mandl, R. S. Kahn, and H. E. Hulshoff Pol, "Functionally linked resting-state networks reflect the underlying structural connectivity architecture of the human brain," *Human brain mapping*, vol. 30, no. 10, pp. 3127–3141, 2009.

[15] J. D. Power, B. L. Schlaggar, and S. E. Petersen, "Recent progress and outstanding issues in motion correction in resting state fMRI," *Neuroimage*, vol. 105, pp. 536–551, 2015.

[16] A. M. Golestani, J. B. Kwinta, S. C. Strother, Y. B. Khatamian, and J. J. Chen, "The association between cerebrovascular reactivity and resting-state fMRI functional connectivity in healthy adults: The influence of basal carbon dioxide," *Neuroimage*, vol. 132, pp. 301–313, 2016.

[17] S. M. Plis, M. F. Amin, A. Chekroud, D. Hjelm, E. Damaraju, H. J. Lee, J. R. Bustillo, K. Cho, G. D. Pearlson, and V. D. Calhoun, "Reading the (functional) writing on the (structural) wall: Multimodal fusion of brain structure and function via a deep neural network based translation approach reveals novel impairments in schizophrenia," *NeuroImage*, vol. 181, pp. 734–747, 2018.

[18] J. D. Medaglia, W. Huang, E. A. Karuza, A. Kelkar, S. L. Thompson-Schill, A. Ribeiro, and D. S. Bassett, "Functional alignment with anatomical networks is associated with cognitive flexibility," *Nature Human Behaviour*, vol. 2, no. 2, p. 156, 2018.

[19] W. R. Shirer, H. Jiang, C. M. Price, B. Ng, and M. D. Greicius, "Optimization of rs-fMRI pre-processing for enhanced signal-noise separation, test-retest reliability, and group discrimination," *Neuroimage*, vol. 117, pp. 67–79, 2015.