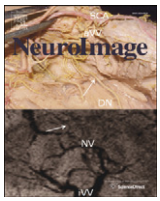
NeuroImage 60 (2012) 1970–1981

Contents lists available at SciVerse ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg



Enhancing reproducibility of fMRI statistical maps using generalized canonical correlation analysis in NPAIRS framework

Babak Afshin-Pour [a,b], Gholam-Ali Hossein-Zadeh [a,c], Stephen C. Strother [b,d], Hamid Soltanian-Zadeh [a,c,e,*]

[a] Control and Intelligent Processing Center of Excellence, School of Electrical and Computer Engineering, College of Engineering, University of Tehran, Tehran, Iran
[b] Rotman Research Institute, Baycrest, Toronto, Ontario, Canada
[c] School of Cognitive Sciences, Institute for Research in Fundamental Sciences (IPM), Tehran, Iran
[d] Department of Medical BioPhysics, University of Toronto, Ontario, Canada
[e] Image Analysis Lab., Radiology Department, Henry Ford Health System, Detroit, Michigan, USA

ARTICLE INFO

Article history:
Accepted 28 January 2012
Available online 14 February 2012

Keywords:
Canonical variant analysis (CVA)
Functional magnetic resonance imaging (fMRI)
Generalized canonical correlation analysis (gCCA)
Multivariate analysis techniques
Reproducibility

ABSTRACT

Common fMRI data processing techniques usually minimize a temporal cost function or fit a temporal model to extract an activity map. Here, we focus on extracting a highly, spatially reproducible statistical parametric map (SPM) from fMRI data using a cost function that does not depend on a model of the subjects' temporal response. Based on a generalized version of canonical correlation analysis (gCCA), we propose a method to extract a highly reproducible map by maximizing the sum of pair-wise correlations between some maps. In a group analysis, each map is calculated from a linear combination of fMRI scans of a subset of subjects under study. The proposed method is applied to BOLD fMRI datasets without any spatial smoothing from 10 subjects performing a simple reaction time (RT) task. Using the NPAIRS split-half resampling framework with a reproducibility measure based on SPM correlations, we compare the proposed approach with canonical variate analysis (CVA) and a simple general linear model (GLM). gCCA provides statistical parametric maps with higher reproducibility than CVA and GLM with correlation reproducibilities across independent split-half SPMs of 0.78, 0.46, and 0.41, respectively. Our results show that gCCA is an efficient approach for extracting the default mode network, assessing brain connectivity, and processing event-related and resting-state datasets in which the temporal BOLD signal varies from subject to subject.

© 2012 Elsevier Inc. All rights reserved.

Introduction

Within the growing number of applications of functional magnetic resonance imaging (fMRI) in neuroscience and in the clinical domain for evaluation of individual treatments, there is a need to enhance the reproducibility of the results. As a key aspect of an fMRI experiment, reproducibility requires that the same local activation maps are likely to be observed in experimental replications. Many researchers have addressed the fact that fMRI experiments will produce more reliable results if one searches for reproducible spatial patterns rather than only significant *p*-values for activity detection (Branch, 1999; Carver, 1993; Genovese et al., 1997; Liou et al., 2009; Nickerson, 2000; Smith et al., 2000; Strother et al., 1997; Tegeler et al., 1999). Along with the above fact, Strother et al. (2002) and Yourganov et al. (2011) derived a relationship between a measure of reproducibility and the global signal-to-noise ratio (SNR) of a statistical parametric map (SPM). As well, many researchers attempt to enhance the reproducibility of a SPM and to extract more spatially-reproducible maps from fMRI data. As there are many steps in the pre-processing and analysis of fMRI data and several definitions for reproducibility, multiple frameworks have been proposed in the literature to enhance a measure of reproducibility for pre-processing and/or analysis steps.

Some methods focus on enhancing the reproducibility of detection of active voxels. Salli et al. (2001) improved a reproducibility measure defined in Genovese et al. (1997) by using contextual information in thresholding of z-maps. Tegeler et al. (1999) and Swallow et al. (2003) showed that using the pooled threshold across subjects negatively affects the reproducibility of brain activation maps. Swallow et al. (2003) suggested selecting a threshold by maximizing the overall reproducibility of active/inactive voxels along subjects. A structural analysis of fMRI data is proposed in Thirion et al. (2007) based on a reproducibility measure to derive a group SPM of the subjects. They proposed to detect activation regions from each subject's data and then, search for correspondences across subjects to find the most reproducible activation regions in the study group.

In univariate processing methods, Liou et al. (2006) showed that using an empirical Bayes method in a general linear model (GLM) provides a way of borrowing information across different runs of an experiment to improve the parameter estimates in each individual

* Corresponding author at: Radiology Image Analysis Lab., Henry Ford Health System, One Ford Place, 2F, Detroit, MI 48202, USA. Fax: +1 313 874 4494.
E-mail address: hamids@rad.hfh.edu (H. Soltanian-Zadeh).

1053-8119/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2012.01.137

run and to enhance the reproducibility of the final SPM. Liou et al. (2009) also argued that there might be functionally active regions that are frequently observed in different experimental sessions, but are modest in response magnitude. They attempt to find reproducible patterns by maximizing the between-run reproducibility via a random effects model.

In multivariate processing approaches and in associated dimension reduction steps, some researchers have tried to choose the best subspace of fMRI data which enhances the reproducibility. Selecting an optimal subspace to maximize reproducibility is fundamental to the NPAIRS approach of Strother et al. (1997, 2002). Yourganov et al. (2011) have shown that for principal components (PC) ranked from 1 to p with decreasing variance, selecting p to optimize reproducibility captures an optimal PCA subspace for activation signal detection associated with fMRI classification. Yang et al. (2008) introduced the RAICAR (Ranking and Averaging Independent Component Analysis by Reproducibility) framework to rank independent components by their reproducibility. They focused on repeated measurements of independent components achieved by running the independent component analysis (ICA) algorithm with different initial conditions (see also Zeng et al., 2009). Wang and Peterson (2009) proposed a method to identify within- or between-subject reproducible independent components that may contain signal from other spurious components in an ICA based fMRI data analysis framework.

There are also some efforts in the literature to use a reproducibility metric to compare the performance of different preprocessing methods and/or choose the best strategy for preprocessing. Ardekani et al. (2004) investigated the impact of inter-subject registration algorithm on the resulting activation maps using reproducibility and sensitivity measures derived from a non-parametric statistical analysis of the data. Strother et al. (2002, 2004), LaConte et al. (2003), and Kjems et al. (2002) proposed NPAIRS (nonparametric prediction, activation, influence, and reproducibility resampling) framework for quantitative evaluation of a processing pipeline based on the reproducibility and prediction accuracy. In NPAIRS, the acquired dataset of an experiment is split into two independent split-half sets, the analysis process is repeated for each set separately and the two statistical maps are calculated with repeated split-half resampling. The Pearson correlation coefficient between these two maps is used as a measure of reproducibility, and provides a global spatial SNR of the fMRI experiment and the whole processing pipeline (Strother et al., 2010). Shaw et al. (2003) proposed to choose the best preprocessing strategy for each subject in the group study based on reproducibility and prediction accuracy calculated in the NPAIRS framework. Moreover, Strother et al. (2004, 2010), LaConte et al. (2003), and Zhang et al. (2008, 2009) evaluated the impact of several preprocessing steps based on the reproducibility and prediction accuracy performance metrics calculated using the NPAIRS framework. Assuming that multiple subjects share a common but unknown temporal response to a stimulation, Evans et al. (2010) used an agnostic canonical variate analysis (aCVA) (Kjems et al., 2002; Strother et al., 2004) in the NPAIRS framework to detect artifacts and choose the best motion correction strategy for different age groups based on reproducibility.

In this paper, we propose a multivariate method based on generalized canonical correlation analysis (gCCA) (Kettenring 1971) to maximize SPM reproducibility without adopting any model for the hemodynamic response or other temporal brain responses. Li et al. (2009) used gCCA to separate different temporal sources in fMRI data. They assumed that there are some common temporal responses to external stimulation in the subjects being studied (as in aCVA), and showed that these may be explored using gCCA. In contrast, in this paper, the underlying assumption is that there are multiple subjects that share an unknown spatial response (or spatial map) to the common experimental manipulation but may show different temporal responses to external stimulation. For each subject, our gCCA explores a broad, potentially inconsistent range of temporal responses in fMRI time-series space while maximizing the mean of correlation coefficients between the pair-wise spatial maps of the subjects.

In addition, Varoquaux et al. (2010) used gCCA as a preprocessing step in an ICA based approach to extract independent and reproducible brain maps. They used gCCA to extract reproducible images from fMRI scans to make ICA more robust to noise and artifacts. They compared their method to common group-ICA methods of fMRI. The different group models used for ICA fMRI are concatenation and tensorial group ICA approaches on fMRI resting-state datasets and fMRI datasets with a stimuli paradigm. It is shown that this approach is more reproducible than traditional ICA approaches to extract activation maps. They assess reproducibility using the approach introduced in the NPAIRS framework (Strother et al., 2004). To this end, they split the subjects into two groups and calculate two thresholded maps for each group. The relative number of voxels declared as active in the two maps is considered as the reproducibility measure.

In the proposed gCCA based framework for group fMRI data analysis, the data of each of the M subjects is put in a (voxel by time) matrix $\boldsymbol{Y}_k$, $1 \le k \le M$. Subjects are then partitioned into two independent split-half sets (i.e., $N = M/2$ for M even, and $N = M/2 \pm 1/2$ for M odd) several times (split-half resampling) and the dimensions of each subject's data matrix are reduced using a dimension reduction technique like PCA ($\boldsymbol{X}_k = \boldsymbol{Y}_k \boldsymbol{W}_k$). $\boldsymbol{W}_k$ is the dimension reduction matrix of the kth subject with columns containing some of the principal components of $\boldsymbol{Y}_k$. For PCA dimension reduction, we use a nested leave-one-out cross-validation loop within each split-half to choose reproducible principal components for each subject. A spatial map for each of the M subjects ($\boldsymbol{z}_k$, $1 \le k \le M$) may be achieved by projecting its data ($\boldsymbol{X}_k$) to a direction $\boldsymbol{a}_k$ ($\boldsymbol{z}_k = \boldsymbol{X}_k \boldsymbol{a}_k$).

Using gCCA for each split-half, the directions ($\boldsymbol{a}_k$, $1 \le k \le N$) are estimated such that the average of the correlation coefficients between the pair-wise spatial maps of the subjects ($\boldsymbol{z}_k$, $1 \le k \le N$) is maximized. The spatial maps of the subjects in a split-half may also be averaged to obtain a split-half spatial map ($\boldsymbol{z} = \frac{1}{N} \sum_{k=1}^{N} \boldsymbol{z}_k$). The reproducibility is denoted by the correlation coefficient between two independent split-half spatial maps. We will show in the results that reproducibility provided by our proposed method is significantly higher than two other common analysis techniques.

As we show, maximizing the mean of correlation coefficients between the pair-wise spatial maps of the subjects empirically maximizes spatial reproducibility calculated in the NPAIRS framework. The results also show that this method offers an effective analysis for fMRI data in which there is a large between (or within)-subject variability in the time-domain hemodynamic responses and provides a stable spatial map of activity across subjects. Such an approach is supported by the recent evidence that optimizing individual subjects' preprocessing pipelines enhances spatial activation overlap between subjects while increasing between-subject signal variability (Churchill et al., 2011).

We show that the gCCA based approach is a flexible framework that can be used to process fMRI data with different degrees of between subject variability in the time-domain responses. For example, in a resting-state experiment where each subject has a different time-domain BOLD response, the gCCA based approach finds functional networks that are common in the subjects. In experiments with stimuli, subjects respond to the same stimuli, but their responses might be slightly different. The responses may be modeled by a set of basis functions. These basis functions can be acquired by modeling of the hemodynamic response with a few unknown parameters (Hossein-Zadeh et al., 2003) or by choosing time-series of a few seed voxels. We show that the gCCA based approach finds an optimal combination of basis functions by maximizing spatial reproducibility. Therefore, the gCCA based approach is a flexible approach that can be used on datasets with high variability in the BOLD response (e.g., resting state datasets) as well as datasets with low variability in the

hemodynamic response where the temporal responses are modeled by a few basis functions.

First, we introduce CCA and its generalized version for assessing a spatially reproducible map. Then we describe the proposed framework for applying gCCA to fMRI data. Next, we present the results on a group of 10 young subjects (20–30 years old) performing a simple reaction time (RT) task and compare the gCCA results on spatially unsmoothed data with those from CVA implemented in the NPAIRS software package and from a general linear model (GLM).

Materials and methods

Multi-task fMRI data

The data were acquired using a 3.0 T Siemens Trio MRI scanner. Subjects were responding to several visual stimuli during the experiment by pressing a button. The visual stimuli were band-pass filtered white noise patches with different center frequencies. During the scans there were blocks of five task conditions: 1) fixation (FIX); 2) simple reaction time (RT); 3) perceptual matching (PMT); 4) attentional cueing (ATT); and 5) delayed match-to-sample (DMS). In the RT task, a single stimulus appeared for 1000 ms in one of three locations at the bottom of the display (left, center, or right), and participants pressed one of three buttons to indicate the location where the stimulus appeared. There were 12 trials in each 40 s RT block. In PMT, a sample stimulus appeared centrally in the upper portion of the screen along with three choice stimuli located in the lower part of the screen. The task was to indicate which of the three choice stimuli matched the sample. The ATT task, consisted of a sample stimulus appearing in the center of the upper part of the screen. After an inter stimulus interval (ISI) of 500 ms an arrow pointing either to the right or to the left was presented for 1500 ms in the lower part of the screen. After another 500 ms interval, two stimuli appeared in the right and left locations for 3000 ms. The task was to attend only to the cued location and press one of two buttons to indicate whether or not the cued target stimulus matched the sample. Finally, in the DMS task, a sample stimulus was presented for 1500 ms in the center of the upper portion of the screen followed by a delay of 2500 ms (blank screen). Then, three choice stimuli were presented for 3000 ms in the lower portion of the screen and the participants had to press one of three buttons to indicate which of the three stimuli matched the previously seen sample.

Four runs were acquired for each subject using a block design with eight alternating task-fixation conditions (FIX) per run (20 scans/task-period alternating with 10 scans/fixation-period, TR = 2 s) for four tasks (ATT, DMS, PMT, RT) with two repetitions each. We have used the RT task to demonstrate our gCCA approach as less brain areas are activated during this task. The image data scanned in the inter-trial interval were not included in the analysis. The full experimental design is reported elsewhere (Grady et al., 2010); the 10 subjects used in this paper have ages from 20 to 30 years.

The imaging protocol was as follows. T2* functional images (TE = 30 ms, TR = 2000 ms, flip angle = 70°, FOV = 200 mm, 64 × 64 pixels/slice) were acquired using an EPI pulse sequence. Each functional volume-image consisted of 28 5-mm thick axial slices, positioned to image the whole brain. A T1-weighted anatomical volume using SPGR (TE = 2.6 ms, TR = 2000 ms, FOV = 256 mm, slice thickness = 1 mm) was also acquired for co-registration with the functional images.

Preprocessing

We created an unbiased, non-linear average anatomical image volume (Kovacevic et al., 2005), which we refer to as the template. Functional data was slice-time corrected using AFNI (afni.nimh.nih.gov/afni) and motion corrected using AIR (bishopw.loni.ucla.edu/AIR5/; Woods et al., 1998). For each run, the mean functional volume after motion correction was registered with each subject's structural volume using a rigid body transformation. Appropriate transform concatenations were performed: from the initial volume to the reference volume within each run, from the mean-run volume to the structural volume, and from the structural into the template space. These concatenated transforms were applied to register the data using a direct non-linear transform from each initial fMRI volume into the template space with a voxel size of $(4\ \text{mm})^3$. Additional pre-processing steps applied to each voxel's time series were: removing each voxel's time-series mean; 0–5th order polynomial detrending; dividing each voxel's time-series by its standard deviation; and regressing out an average white matter time-series and removing the means of the fMRI scans. Normalizing each time-series to its standard deviation removes inter-subject in-homogeneities to some extent and increases reproducibility.

For the white matter regression, we used the probabilistic white matter map from the ICBM probabilistic tissue atlas (http://www.loni.ucla.edu/ICBM/Downloads/Downloads_ICBMprobabilistic.shtml) as a starting point. This white matter map was transformed into the template space which has the voxel size of $(4\ \text{mm})^3$, and a binary mask of white matter mask was generated by thresholding the transformed map based on a threshold of 0.8 (≥80% chance of being white matter). This initial mask which has the voxel size of $(4\ \text{mm})^3$ was then eroded using a 3 × 3 square morphological structuring element. We calculated the spatial average of voxels in the resulting conservative white matter mask in each subject and used the resulting averaged time series as a regressor in a GLM for each voxel and subject separately. To test the performance of our gCCA approach RT and fixation blocks were used for further processing. In addition, in order to have equal numbers of scans for RT and fixation only the two fixation blocks (FIX) adjacent to the RT blocks were used from each run of the experiment. We removed the first two transition scans of each RT and FIX blocks. No spatial smoothing was applied to the data sets before data analysis.

Resting-state fMRI data

We used the resting-state fMRI data available online at www.nitrc.org/projects/fcon_1000/. During the acquisition of this experimental data, the subjects were instructed to keep their eyes closed. Ten subjects (five females) were used to show the performance of the proposed approach. The fMRI images acquired using a 3 T MRI whole-body scanner (Siemens 3.0 Tesla Allegra). In the fMRI experiment, T2* functional images (FOV = 192 mm, 3 mm isotropic resolution) were acquired using an EPI pulse sequence (TE = 25 ms, TR = 2000 ms, flip angle = 90°). The functional images were acquired during a rest period of 390 sec. A T1-weighted anatomical volume was also acquired using MPRAGE (TE = 13 ms, TR = 2500 ms, flip angle = 30, FOV = 192 × 256 × 256 mm) for co-registration with the functional images.

In the preprocessing of the resting state data, motion correction was performed on each subject using FSL (http://www.fmrib.ox.ac.uk/fsl/). A single transfer matrix is obtained and used to register fMRI data (EPI volumes) of each subject to the MNI atlas (using the subsequent registration of EPI volume to anatomical volume and the anatomical volume to MNI by FSL). In addition, low frequency fluctuations and high frequency noise were removed using a Butterworth band-pass filter with the pass band from 0.01 Hz to 0.1 Hz. We also removed the mean of each scan. No spatial smoothing was applied to the datasets before data analysis.

Canonical correlation analysis of two data sets

Canonical correlation analysis (CCA) is a technique for finding a pair of basis vectors that maximize the correlation between two projected sets of variables in which each set of variables is projected onto its corresponding basis vector. Let $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ be two known $V \times t_1$ and $V \times t_2$ full-rank data matrices. Canonical correlation analysis (CCA)

can be defined as the problem of finding two canonical vectors: $\boldsymbol{a}_1$ of size $t_1 \times 1$ and $\boldsymbol{a}_2$ of size $t_2 \times 1$, such that canonical variates $\boldsymbol{z}_1 = \boldsymbol{X}_1\boldsymbol{a}_1$ and $\boldsymbol{z}_2 = \boldsymbol{X}_2\boldsymbol{a}_2$ are maximally correlated.

$$(\boldsymbol{a}_1, \boldsymbol{a}_2) = \arg\max\left(\frac{\boldsymbol{z}_1^{\mathrm{T}}\boldsymbol{z}_2}{||\boldsymbol{z}_1||||\boldsymbol{z}_2||} = \frac{\boldsymbol{a}_1^{\mathrm{T}}\boldsymbol{C}_{12}\boldsymbol{a}_2}{\sqrt{\boldsymbol{a}_1^{\mathrm{T}}\boldsymbol{C}_{11}\boldsymbol{a}_1\boldsymbol{a}_2^{\mathrm{T}}\boldsymbol{C}_{22}\boldsymbol{a}_2}}\right) \tag{1}$$

where $\boldsymbol{C}_{11} = \boldsymbol{X}_1^{\mathrm{T}}\boldsymbol{X}_1$, $\boldsymbol{C}_{22} = \boldsymbol{X}_2^{\mathrm{T}}\boldsymbol{X}_2$, and $\boldsymbol{C}_{12} = \boldsymbol{X}_1^{\mathrm{T}}\boldsymbol{X}_2 = \boldsymbol{C}_{21}^{\mathrm{T}}$ are estimates of the within set and between set covariance matrices respectively. Eq. (1) is equivalent to the maximization of $q = \boldsymbol{a}_1^{\mathrm{T}}\boldsymbol{C}_{12}\boldsymbol{a}_2$ subject to constraints $\boldsymbol{a}_1^{\mathrm{T}}\boldsymbol{C}_{11}\boldsymbol{a}_1 = \boldsymbol{a}_2^{\mathrm{T}}\boldsymbol{C}_{22}\boldsymbol{a}_2 = 1$. The solution to this problem is given by the eigenvector corresponding to the largest eigenvalue of the following generalized eigenvalue problem (Mardia et al., 1979)

$$\begin{bmatrix} 0 & \boldsymbol{C}_{12} \\ \boldsymbol{C}_{21} & 0 \end{bmatrix}\boldsymbol{a} = q\begin{bmatrix} \boldsymbol{C}_{11} & 0 \\ 0 & \boldsymbol{C}_{22} \end{bmatrix}\boldsymbol{a} \tag{2}$$

where q is the canonical correlation and $\boldsymbol{a} = [\boldsymbol{a}_1^{\mathrm{T}}, \boldsymbol{a}_2^{\mathrm{T}}]^{\mathrm{T}}$ is the eigenvector. In the next subsection we review the extension of CCA to N data matrices ($\boldsymbol{X}_k$, $1 \le k \le N$).

Generalization of CCA to several sets (gCCA)

Any generalization of the CCA to several sets has to be equivalent to the CCA in the case of 2 sets. Kettenring (1971) studied five versions of generalized CCA in which the following functions of paired-wise correlation coefficients are maximized: (1) sum of correlations (SUMCOR), (2) maximum variance (MAXVAR), (3) sum of squared correlations (SSQCOR), (4) minimum variance (MINVAR), and (5) generalized variance method (GENVAR). Also, in a recent generalization (Tenenhaus and Tenenhaus, 2011), the authors proposed an approach to maximize the sum of absolute values of correlations (SABSCOR).

In this section, the classical sum of correlations (SUMCOR) generalization is summarized and gCCA is used as an abbreviation for SUMCOR maximization approach throughout the paper. The gCCA approach can be formulated as the problem of sequentially maximizing the generalized canonical correlation:

$$q = \frac{1}{N(N-1)}\sum_{\substack{k,l=1 \\ k\neq l}}^{N} q_{kl} \tag{3}$$

where q_{kl} is the correlation coefficient between $\boldsymbol{z}_k = \boldsymbol{X}_k\boldsymbol{a}_k$ and $\boldsymbol{z}_l = \boldsymbol{X}_l\boldsymbol{a}_l$. $\boldsymbol{a}_k$ ($1 \le k \le N$) are unknown weights that have to be estimated for each matrix $\boldsymbol{X}_k$ with size $V \times t_k$ ($1 \le k \le N$). $\boldsymbol{X}_k$ are N known full-rank data matrices, and we want to calculate the canonical correlation between them. The full-rank constraint of data matrices may be relaxed by regularizing the estimated covariance matrices (Ledoit and Wolf, 2004). We may rewrite Eq. (3) as:

$$q = \frac{1}{N(N-1)}\boldsymbol{a}^{\mathrm{T}}(\boldsymbol{C}-\boldsymbol{D})\boldsymbol{a}, \tag{4}$$

where $\boldsymbol{C}$, $\boldsymbol{D}$ and $\boldsymbol{a}$ are

$$\boldsymbol{C} = \begin{bmatrix} \boldsymbol{C}_{11} & \cdots & \boldsymbol{C}_{1N} \\ \vdots & \ddots & \vdots \\ \boldsymbol{C}_{N1} & \cdots & \boldsymbol{C}_{NN} \end{bmatrix}, \boldsymbol{D} = \begin{bmatrix} \boldsymbol{C}_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \boldsymbol{C}_{NN} \end{bmatrix}, \boldsymbol{a} = \left[\boldsymbol{a}_1^{\mathrm{T}}, \ldots, \boldsymbol{a}_N^{\mathrm{T}}\right]^{\mathrm{T}} \tag{5}$$

and $\boldsymbol{C}_{lk} = \boldsymbol{X}_l^{\mathrm{T}}\boldsymbol{X}_k$. We may use the following energy constraint to avoid trivial solutions in maximizing the cost function in Eq. (3) or (4).

$$\frac{1}{N}\sum_{k=1}^{N}\boldsymbol{a}_k^{\mathrm{T}}\boldsymbol{C}_{kk}\boldsymbol{a}_k = \frac{1}{N}\boldsymbol{a}^{\mathrm{T}}\boldsymbol{D}\boldsymbol{a} = 1 \tag{6}$$

The solution of this generalized CCA problem can be obtained using Lagrange multipliers. Using the Lagrange multiplier λ, the cost function J is formed as below, and the unknown vector $\boldsymbol{a}$ is found to maximize it.

$$J = \frac{1}{N(N-1)}\boldsymbol{a}^{\mathrm{T}}(\boldsymbol{C}-\boldsymbol{D})\boldsymbol{a} + \lambda\left(1 - \frac{1}{N}\boldsymbol{a}^{\mathrm{T}}\boldsymbol{D}\boldsymbol{a}\right) \tag{7}$$

Taking the derivative with respect to $\boldsymbol{a}$, one can write

$$\frac{1}{(N-1)}(\boldsymbol{C}-\boldsymbol{D})\boldsymbol{a} = \lambda\boldsymbol{D}\boldsymbol{a}. \tag{8}$$

Eq. (8) is a general eigenvalue decomposition problem and the largest eigenvector maximizes the cost function in Eqs. (4) and (3). Therefore, the solution for the MAXVAR generalization of CCA is the eigenvector corresponding to the largest eigenvalue in the general eigen-decomposition of Eq. (8). In order to obtain λ, Eq. (8) can be left multiplied by $\boldsymbol{a}^{\mathrm{T}}$, which, applying the constraint from Eq. (6), implies $\lambda = q$. We use gCCA as the abbreviation for this generalization of canonical correlation analysis. The general eigenvalue decomposition in Eq. (8) has $m = \min(t_1, \ldots, t_N)$ non-negative eigenvalues with the assumption that all $\boldsymbol{X}_k$ are of full rank. We may calculate $\boldsymbol{z}_k$ by projecting $\boldsymbol{X}_k$ on the calculated $\boldsymbol{a}_k$ to acquire the most reproducible map. The average canonical variate $\boldsymbol{z} = \frac{1}{N}\sum_{k=1}^{N}\boldsymbol{z}_k$ may be used as the common canonical variate of the N datasets and it represents the most reproducible pattern in all N datasets that may be acquired by a linear transform.

Considering the eigenvectors in Eq. (8), we have m estimates for $\boldsymbol{a}$ (and thus for $\boldsymbol{a}_k$'s) acquired using the generalization of CCA and therefore m common canonical variates ($\boldsymbol{z}$) are available. In this paper, the most reproducible results related to the largest eigenvalue are reported by default, and will show examples of the first several reproducible maps.

Estimating a reproducible fMRI map using gCCA

Let us assume that fMRI datasets from M subjects are available and the effect of head motion and the low frequency nuisance effects are removed from the data. The data is also assumed to be registered to a template and $\boldsymbol{Y}_k$ is the resulting $V \times t_k$ matrix of fMRI data for subject k ($1 \le k \le M$). Thus each row of $\boldsymbol{Y}_k$ is the time-series of an intra-cerebral voxel. All subjects contain the same number of intra-cerebral voxels (V) but may have different numbers of scans (t_k). In order to have a robust estimate of reproducibility and to obtain a standard robust map, we use NPAIRS split-half resampling. We partition M subjects into two independent split-half sets (i.e., $N = M/2$ for M even, and $N = M/2 \pm 1/2$ for M odd), and apply the gCCA on the datasets in each split half separately (see Fig. 1 for a block diagram of the method).

The matrix $\boldsymbol{Y}_k$ may not be full rank and the solution of the generalized eigen-decomposition in Eq. (8) may not be stable. Here, we use a dimension reduction approach such as principal component analysis (PCA) to reduce the dimension of data in $\boldsymbol{Y}_k$ to $V \times p_k$, $p_k < t_k$.

The PCA dimension reduction approach provides a set of principal components and their corresponding images in the data space of each subject. These components are usually ranked based on their variance (or eigenvalues). In order to detect the most informative components in each subject, we use a leave-one-out based approach.

A PCA is applied on the data of each subject and another PCA is applied on the data of the remaining subjects in the group. The principal components of the remaining subjects are calculated by concatenating subjects' data matrices in the time direction and applying PCA to the concatenated data matrix. Hence, there are sets of single-subject and group images for each subject. The images from two sets which have the maximum cross-correlation (cc) coefficient are paired. The

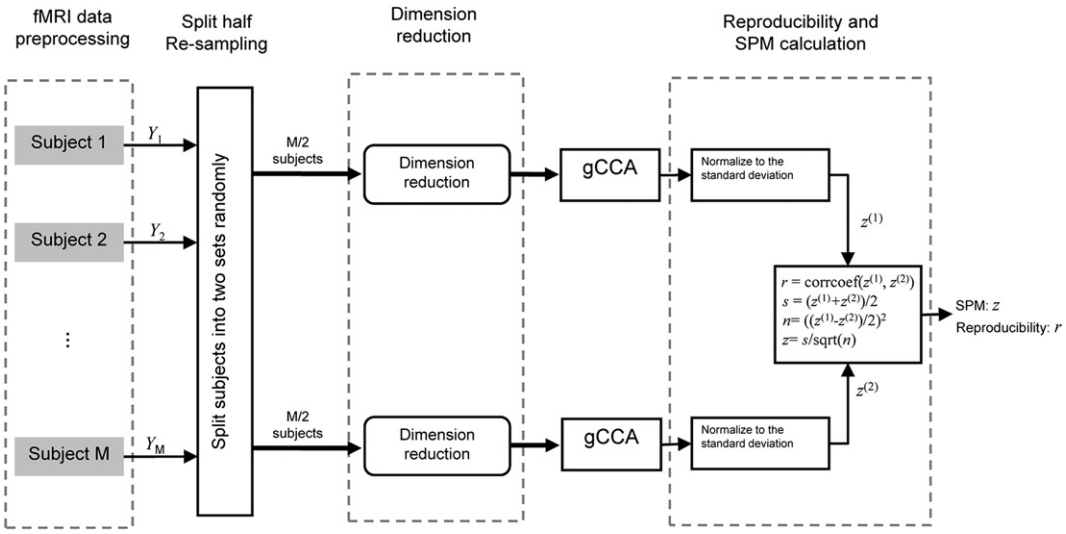


Fig. 1. The block diagram of the gCCA reproducibility enhancement approach in the NPAIRS framework (NPAIRS-gCCA) for one split-half resampling. The datasets ($\boldsymbol{Y}_k$, $k=1,\ldots,N$) are randomly split into two sets. The dimension of the fMRI datasets ($\boldsymbol{Y}_k$, $k=1,\ldots,N$) in each set is reduced using a leave-one out cross validation based PCA ($\boldsymbol{X}_k=\boldsymbol{Y}_k\boldsymbol{W}_k$, $k=1,\ldots,N$). Each set is analyzed using gCCA and two eigenimages are calculated. The eigenimages are mean-removed and normalized to their standard deviation and the correlation coefficient between the resultant images measures their reproducibility. The final z-scored SPM is calculated using uncorrelated noise rescaling as in the NPAIRS framework. The processes after PCA dimension reductions may be repeated several times to obtain a robust estimate of the reproducibility and the SPM (see the text).

single-subject images that have a maximum cc score higher than a significance threshold are assumed to be robust and informative images. The PCA components of these images in the single subject are chosen as dimension reduced data of this subject and the procedure is repeated for all subjects. As the number of samples used in calculating cc is very large and equal to the number of voxels (V), cc follows a Gaussian distribution with zero mean and standard deviation of $1/\sqrt{V-3}$ under the null hypothesis (when no significant correlation exists between the paired images). The threshold (λ) for cc in this step is set such that the false alarm rate is controlled at 0.001. Note that the test performed here assumes no spatial correlation between voxels and may be biased. As this test is used to choose some principal components, the probable bias is not a major problem. We have summarized the PCA dimension reduction in the following pseudo-code:

1. For $k=1$:N (Number of subjects in a split half)
   a. Derive the principal components and their corresponding eigen-images of subject k.
   b. Concatenate the time series data of the remaining subjects and derive the principal components and the corresponding eigen-images of the resulting data matrix.
   c. For $j=1$: Number of principal components of the subject k;
      c.1 Calculate the cross correlation between the jth eigen-image (of subject k) and the eigen-images obtained in part b.
      c.2 Denote the maximum of the cross correlations obtained in c.1 as cc.
      c.3 Keep the jth principal component if $cc>\lambda$.
   d. End
   e. Put the remaining principal components of subject k in $\boldsymbol{W}_k$.
2. End For.

The PCA is performed in each split as a nested cross-validation and reproducibility is optimized with respect to the number of PCA components on each subset in a nested cross-validation loop and in each fold separately.

This resampling based approach for PCA dimension reduction is described in Mei et al. (2008) and used in Varoquaux et al. (2010). In Varoquaux et al. (2010), the authors re-sample each subject fMRI dataset separately and choose the informative eigen-images of the fMRI dataset of the subject by performing a statistical test. The resampling is performed by randomly choosing a user-defined proportion of fMRI scans of the subject (i.e., bootstrapping) and calculating the principal components of the re-sampled data and the original data. Those eigen-images of the original data are chosen that have a high correlation coefficient with one of the eigen-images in the re-sampled data. In this paper, for each subject in a split-half, we use other subjects' fMRI data in that split-half to find informative eigen-images. To this end, we calculate the principal components of a subject. The principal components of the remaining subjects are calculated by initially concatenating subjects' data in the time direction and performing principal component analysis (PCA). Those eigen-images of the subject are chosen that have a high correlation coefficient with one of the eigen-images in the concatenated data of the remaining subjects.

The dimension reduced dataset of subject k is shown by $\boldsymbol{X}_k$ which is a $V\times p_k$ matrix. Dimension reduction using PCA may reduce noise and decrease the computational complexity of solving the generalized eigenvalue decomposition in Eq. (8). Let $\boldsymbol{W}_k$ be the linear transform acquired using the above mentioned PCA dimension reduction and we may write $\boldsymbol{X}_k=\boldsymbol{Y}_k\boldsymbol{W}_k$. Fig. 2 shows the process in which the removal of some of the PCA components from each subject decreases the size of the covariance matrix $\boldsymbol{C}$.

Furthermore, a seed based version of gCCA may be developed by replacing PCA dimension reduction transforms ($\boldsymbol{W}_k$, $k=1,\ldots,N$) with *supervised* dimension reduction transforms. For example, we may replace the column of $\boldsymbol{W}_k$ with the time-series of one or more seed voxels in regions of interest (ROI). In this case, the gCCA based method extracts the most reproducible spatial maps depicting the voxels functionally correlated with these seed voxels.

In addition, we may replace the columns of $\boldsymbol{W}_k$ with some regressors that model hemodynamic response to the experiment stimulation pattern. Hossein-Zadeh et al. (2003) characterized the hemodynamic response over a wide range of parameters via three basis signals derived using PCA. These basis signals may be convolved with the stimulation pattern to generate three regressors. We use these three regressors in the columns of $\boldsymbol{W}_k$ to reduce the data dimensionality and to extract a reproducible spatial map depicting voxels whose hemodynamic responses follow the stimulation pattern, i.e., task mode network (see Results and Fig. 8).

The estimated weight vectors $\boldsymbol{a}_k$ ($1\le k\le N$) in Eq. (8) may be projected to the original time-series by left multiplying $\boldsymbol{W}_k^T$ and for each subject to obtain the time-series ($\boldsymbol{b}_k=\boldsymbol{W}_k\boldsymbol{a}_k$) that are associated with

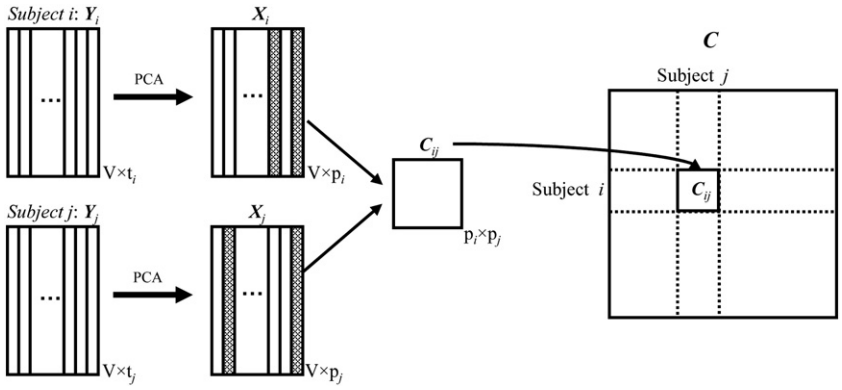


Fig. 2. A schematic of calculating block covariance matrix ($\boldsymbol{C}$): The dimension of fMRI data of the subjects i and j are reduced using PCA in a leave-one-out cross-validation approach. The covariance matrix between dimension-reduced data matrices ($\boldsymbol{X}_i$, $\boldsymbol{X}_j$) is calculated and denoted by $\boldsymbol{C}_{ij}$. The matrix $\boldsymbol{C}_{ij}$ is an element in the block covariance matrix $\boldsymbol{C}$.

the estimated reproducible map. As there are no constraints on the temporal pattern vectors $\boldsymbol{a}_k$ $(1 \le k \le \mathrm{N})$, the vectors $\boldsymbol{b}_k$ $(1 \le k \le \mathrm{N})$ may have totally different temporal patterns. The estimated $\boldsymbol{b}_k$ $(1 \le k \le \mathrm{N})$ show how variable are the calculated time-series in different subjects and it might be used as a measure of *temporal reproducibility* of an analysis technique. We use the estimated $\boldsymbol{b}_k$ to compare the temporal reproducibility of gCCA based analysis technique with that of CVA analysis technique implemented in the NPAIRS framework (see Results).

After applying the gCCA approach on each split half separately, two average canonical images ($\boldsymbol{z}^{(1)}$ and $\boldsymbol{z}^{(2)}$) are available. The correlation coefficient (r) between the canonical images of the two sets shows the reproducibility of the gCCA and provides a robust transformation of the associated reproducible, consensus canonical image to a z-score map. The split-half canonical images are normalized through dividing them by their standard deviation. It is possible to repeat the split-half resampling approach for multiple split-half partitions of the subjects and calculate the median of r as robust estimate of the z-reproducibility of gCCA.

In order to calculate z-values and test the statistical significance of activity, a voxel-wise estimation of the signal amplitude and the noise power is used. In NPAIRS-gCCA, subjects are split in two groups and two maps ($\boldsymbol{z}^{(1)}$, $\boldsymbol{z}^{(2)}$) are calculated from the two groups. These maps assign two values to each voxel in the brain ($z^{(1)}(i)$, $z^{(2)}(i)$, $i = 1,\ldots,\mathrm{V}$). For voxel i, we estimate the signal term as the sum of these values ($s(i) = z^{(1)}(i) + z^{(2)}(i)$) and the noise power as the square of the difference between them ($n(\mathrm{i}) = (z^{(1)}(i) - z^{(2)}(i))^2$). A robust estimate of the signal is calculated as the mean of the signal terms ($s(i)$) across different split half resamplings and similarly a robust estimate of the noise power is made as the mean noise power ($n(i)$) across different split-half resamplings. z-Values for the voxel i ($z(i)$) are calculated by dividing the signal by the root square of the noise power (see Fig. 1 for a block diagram of the method). This gCCA approach in the NPAIRS framework is called NPAIRS-gCCA below.

As the generalization of CCA solves a generalized eigen-value decomposition, it may provide several estimates for ($\boldsymbol{z}^{(1)}$, $\boldsymbol{z}^{(2)}$) related to different eigen-values. Let us assume we have $m^{(1)}$ estimate for $\boldsymbol{z}^{(1)}$ and $m^{(2)}$ estimate for $\boldsymbol{z}^{(2)}$ therefore we will have $\min(m^{(1)}, m^{(2)})$ z-value maps. In order to calculate the z-value map ($\boldsymbol{z}$), we use a component matching procedure. We calculate the absolute values of correlation coefficients (cc) between all components of the first split and the components of the second split. The two components from two splits that have the highest absolute value of cc are matched. We then seek for the next two components with the highest absolute cc and match them until $\min(m^{(1)}, m^{(2)})$ components from the two splits are matched. In this paper, the first z-value maps related to the maps in each split half with the highest absolute cc are reported by default and examples of the later eigenvectors (and estimates of reproducible maps) are shown.

Results

To apply NPAIRS-gCCA to the real fMRI data, we reduced the dimension of each subject's fMRI data set using PCA, by choosing the p first principal components. We performed NPAIRS-CVA for 20 splits of the 10 subjects ($\mathrm{M} = 10$) into two 5-subject sets ($\mathrm{N} = 5$). We then searched over the number of principle components (p) per subject from 2 to 40 by step of 2 to find the most reproducible, consensus canonical variate.

We also applied the NPAIRS-CVA analysis with 2 classes defined by scans in the FIX and RT blocks, implemented in the NPAIRS software package (we use NPAIRS-CVA code available at http://code.google.com/p/plsnpairs/). Before applying the NPAIRS-CVA, the first two transition scans after the FIX and the RT blocks ended were dropped to reduce within-class noise from hemodynamic transition scans. An initial singular value decomposition step of all the data combined was performed only for the NPAIRS-CVA approach and 30% of all components (i.e., 480 out of 1600 components) ordered from the largest to smallest eigenvalues were kept to remove noise and decrease the computational cost for subsequent split-half sets (Evans et al., 2010). As for gCCA we performed NPAIRS-CVA for 20 splits of the 10 subjects into two 5 subject sets and changed the number of eigenimages per split to find the maximum reproducibility (r) as a subspace optimization step (Strother et al., 2010). In addition to our two multivariate NPAIRS techniques, we applied GLM implemented in the NPAIRS framework (NPAIRS-GLM) to acquire the group active areas (see Kjems et al., 2002; Zhang et al., 2008).

In Fig. 3 the reproducibility curve for NPAIRS-CVA is plotted versus the number of principal components per split half data set of 5 subjects. The NPAIRS-gCCA has reproducibility of 0.78 and NPAIRS-CVA has maximum reproducibility of 0.46 at 34 principal components, while NPAIRS-GLM produced a reproducibility of 0.41. As expected based on its target cost function, NPAIRS-gCCA provides the most reproducible map in comparison with NPAIRS-CVA and NPAIRS-GLM for this spatially unsmoothed data, which translates to a large increase in global pattern SNR; $\mathrm{gSNR} = \sqrt{2r/(1-r)}$, with values of 2.7, 1.3 and 1.2 for NPAIRS-gCCA, NPAIRS-CVA and NPAIRS-GLM, respectively (Strother et al., 2010).

To test statistical significance of the reproducibility value acquired using NPAIRS-gCCA, we generate surrogate data using the approach proposed in Breakspear et al. (2004). Their approach is based on the spatiotemporal resampling of the data in the wavelet domain. The resulting surrogate datasets approximately have the same spatial correlation structure of the original/experimental data. We applied the NPAIRS-gCCA approach to a collection of 50 surrogated data (each containing 20 surrogate subjects) and obtained the median reproducibility

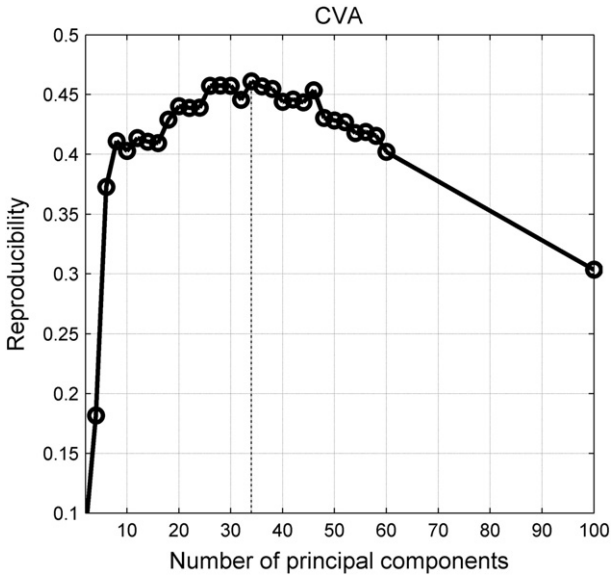


Fig. 3. The reproducibility curve versus the number of principal components per split half for NPAIRS-CVA for the RT task (note vertical axis scale change). The maximum reproducibility is obtained at 34 principal components. The maximum reproducibilities of NPAIRS-CVA and NPAIRS-gCCA are 0.44 and 0.78, respectively, for unsmoothed RT fMRI data sets.

of 0.02. We then used a non-parametric test to test whether the reproducibility (r) acquired for the real fMRI datasets is significantly different from those generated from surrogate datasets and the test rejected the null hypothesis ($p \approx 0$).

In Figs. 4a–c, some of the detected activation areas are shown for SPMs using NPAIRS-GLM, NPAIRS-CVA and some of the reproducible areas outlined using NPAIRS-gCCA, respectively. NPAIRS-GLM detects significant voxels around the central sulcus, particularly near the so-called "left hand knob" of the precentral gyrus (Yousry et al., 1997), in the supplementary motor cortex particularly on the left, and bilaterally in the frontal eye fields, but no voxels from the classic regions of the default mode network (Buckner et al., 2008). NPAIRS-CVA detects more significant voxels around the central sulcus in the same area near the left hand knob, similar results for the supplementary motor cortex and frontal eye fields, and a few significant voxels potentially from the default mode network. NPAIRS-gCCA outlines few significant voxels as reproducible near the left hand knob peaks of NPAIRS-GLM and NPAIRS-CVA, and major decreased activation areas of the default mode network (posterior cingulate cortex, inferior parietal lobe, medial prefrontal cortex, medial cerebellum), and task-positive network (supplementary motor cortex, pre-motor areas, primary motor cortex, insula, putamen and Thalamus). These results suggest that there are some networks in the brain that are more spatially-reproducible than the task-mode network(s). The mean of the paired-wise correlation coefficients is a global measure which is maximized over all voxels in this paper. As the regions of the default mode network have coherent activation and this network is a (spatially) larger network than the task related network for the RT task, the proposed approach extracted this network. This implies that this (spatially large) network has a greater effect on the reproducibility than the task related network that has higher signal intensity in each voxel. Since NPAIRS-gCCA does not use a temporal model for hemodynamic system response, it outlines those networks easily but has difficulty in segmenting primary task-mode activations such as those near the hand knob in Fig. 4.

As mentioned before, NPAIRS-gCCA provides several z-value maps. For the multi-task datasets, other z-value maps do not show any significant activation at the false alarm rate of 0.001. Therefore, we showed some outlined areas by the second, third, and forth maps at the false alarm rate of 0.05 in Fig. 5. The active areas outlined by the second map with the reproducibility of 0.67 are mainly located at the edge of brain ventricles and white matter. This pattern probably caused by uncorrected motion artifacts in the datasets. The third

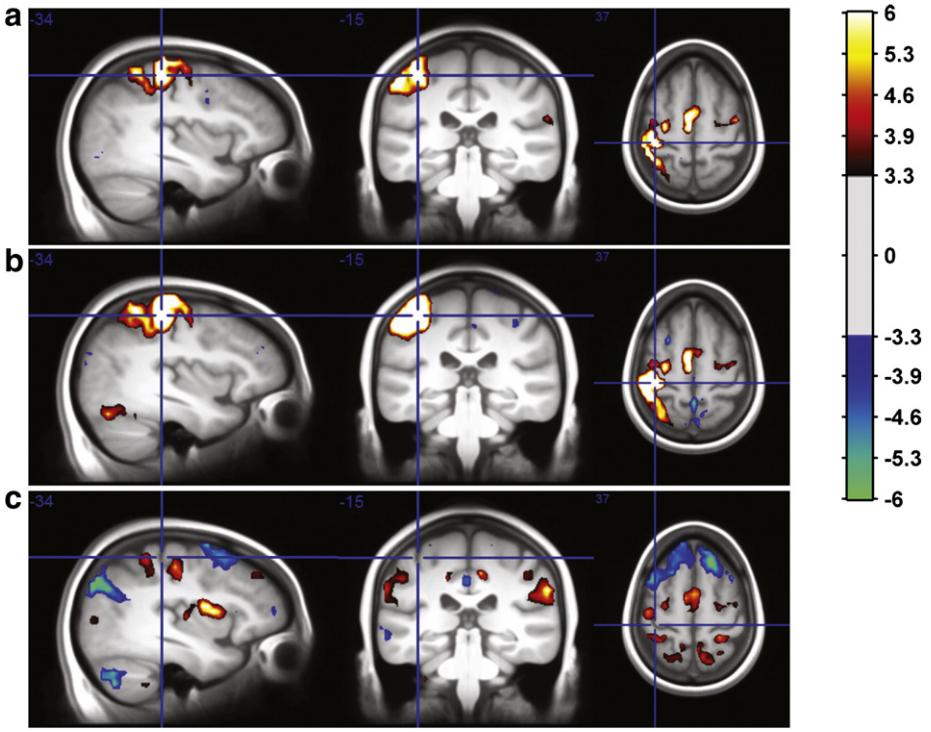


Fig. 4. Detected activation areas for the RT task using (a) NPAIRS-GLM with some voxels detected in the primary motor cortex (the area shown by crosshairs); (b) NPAIRS-CVA with more significant voxels in the primary motor and supplementary motor cortices than with NPAIRS-GLM; (c) NPAIRS-gCCA outlines the most reproducible activation pattern in the brain that reflects primarily default mode regions (posterior cingulate cortex, inferior parietal lobe, medial prefrontal cortex, medial cerebellum and caudate nucleus) and is less obviously task-related than the NPAIRS-GLM and NPAIRS-CVA results.

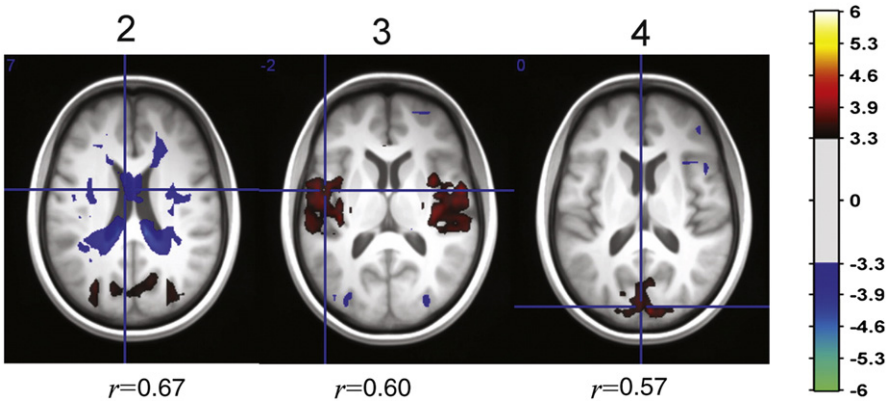


Fig. 5. Some outlined areas by the second (left), third (middle), and fourth (right) reproducible maps acquired using gCCA-NPAIRS (thresholded at the false alarm rate of 0.05).The active areas outlined by the second map with the reproducibility of 0.67 are mainly located at the edges of the ventricles and white matter. The third map outlines some activation in Insula with the reproducibility of 0.60 and the forth map outlines some activation in the visual cortex with the reproducibility of 0.57.

map outlines some activation in insula with the reproducibility of 0.60 and the forth map outlines some activation in the visual cortex with the reproducibility of 0.57. These outlined areas are detected at false alarm rate of 0.05 and may not be statistically significant. This means that other z-value maps calculated by NPAIRS-gCCA are also informative. We will show later that other z-value maps outline significant activations for the resting-state dataset.

In addition, we ranked the following brain regions from most reproducible to less reproducible according to their maximum absolute z-values over the gCCA components for the RT task: medial frontal gyrus ($|z|=9.5$), precentral gyrus ($|z|=8.3$), insula ($|z|=8.2$), inferior parietal lobe ($|z|=7.9$), posterior cingulate cortex ($|z|=7.4$), middle frontal gyrus ($|z|=7.1$), Thalamus ($|z|=6.7$), prefrontal cortex ($|z|=6.1$), superior frontal gyrus ($|z|=6.1$), medial cerebellum ($|z|=6.1$), left inferior occipital gyrus ($|z|=5.4$), parahippocampal gyrus ($|z|=5.6$), visual cortex ($|z|=4.8$), putamen ($|z|=4.3$), pre-central gyrus ($|z|=2.9$), post-central gyrus ($|z|=2.6$), and caudate ($|z|=2.4$).

Finally, a seed based version of NPAIRS-gCCA was applied. We chose a 3×3 patch of voxels as seed voxels in the posterior cingulate cortex with the center voxel coordinates of $(-12, -48, 36)$ in the MNI space for each subject's dataset. Instead of reducing the dimensionality of data using PCA, we projected the time-series of each dataset to the nine seed time-series. We applied the NPAIRS-gCCA based approach to these reduced dimension datasets and calculated the z-score map. The canonical images for 20 splits were calculated and a median reproducibility of 0.76 was obtained. In Fig. 6, the outlined voxels, using a seed-based NPAIRS-gCCA are shown. The post cingulate cortex, preforontal cortex, caudate and hippocampus are

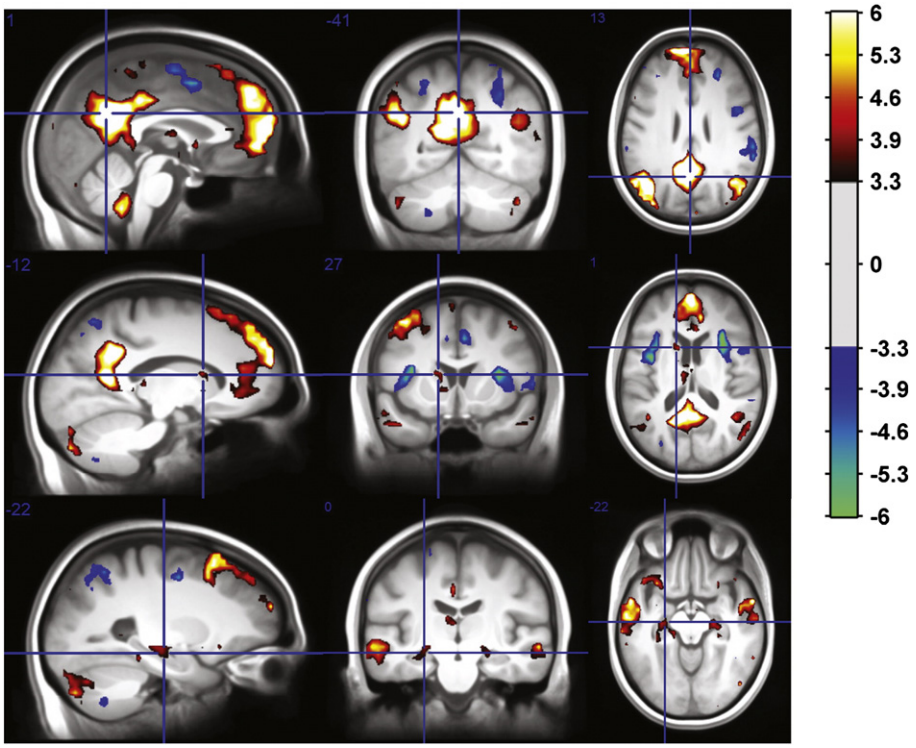


Fig. 6. Some of the outlined area using seed based NPAIRS-gCCA at the false alarm rate of 0.001. Time-series of a 3×3 patch of voxels as seeds are selected from post cingulate cortex for each subject and fMRI data of each subject is projected to those time-series. Using NPAIRS-gCCA the patterns of the subjects are combined to create a map with reproducibility of 0.76. Post cingulate cortex and preforontal cortex (first row), caudate (second row) and hippocampus (third row) are outlined as a part of default mode network during the RT task. Hot colors show the areas positively connected with post cingulate cortex and cold colors show the areas negatively connected with post cingulate cortex.

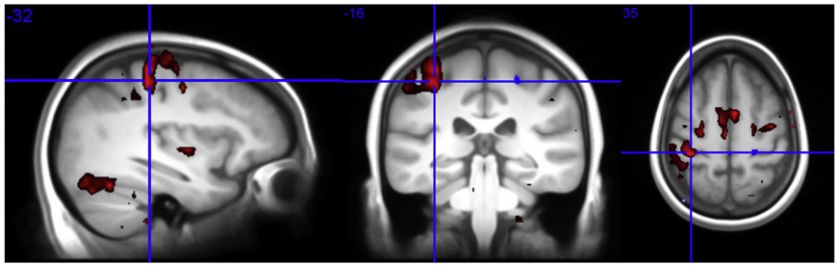


Fig. 7. Some of the outlined areas using model based NPAIRS-gCCA at false alarm rate of 0.001. The fMRI data of each subject are projected to three regressors. The regressors are acquired by convolving the stimulation pattern with three basis functions proposed in Hossein-Zadeh et al. (2003). The area shown by crosshairs is primary motor cortex. The reproducibility of the extracted map is 0.48.

outlined as a part of the regions that are connected to posterior cingulate cortex during the RT task.

In addition, we chose three basis signals proposed in Hossein-Zadeh et al. (2003) to model the hemodynamic response. These regressors are acquired by convolving the three basis signals with the stimulation pattern (two blocks of FIX-RT-FIX in each run). Instead of reducing the dimensionality of data using PCA we projected the time-series of each dataset onto the estimated regressors. The canonical images for 20 splits were calculated and a median reproducibility of 0.48 was obtained. In Fig. 7, the outlined areas using this model-based version of NPAIRS-gCCA at the false alarm rate of 0.001 are shown. The primary motor cortex is detected as a part of the task mode network as also shown with GLM in Fig. 4(top row).

We applied NPAIRS-gCCA on the resting-state datasets. The four maps with maximum reproducibility were selected and thresholded at the false alarm rate of 0.001 to show the most reproducible activity maps among the subjects. Fig. 8 shows some of the areas outlined by these maps (from left to right) with the reproducibility of 0.79, 0.74, 0.68, and 0.61, respectively. Visual cortex, preforontal cortex, and Thalamus (first map), medial frontal gyrus (second map), post cingulated gyrus (third map), and hippocampus (fourth map) are outlined. Other components did not outline significant activations at the false alarm rate of 0.001 and are not reported. The outlined areas are consistent with the results of other studies (see Fax and Raichle, 2007 for a review). We also ranked following areas from most reproducible to less reproducible according to their maximum absolute z-values over the gCCA components for the resting-state dataset: primary visual cortex ($|z|=11.4$), inferior occipital gyrus ($|z|=9.2$), inferior parietal lobe ($|z|=8.3$), fusiform gyrus ($|z|=7.2$), caudate ($|z|=7.8$), temporal anterior cingulate cortex ($|z|=8.1$), pre-frontal cortex ($|z|=5.3$), insula ($|z|=5.2$), hippocampus ($|z|=4.4$), Thalamus ($|z|=4.2$), primary motor cortex ($|z|=3.3$), and medial cerebellum($|z|=3.2$).

For one block of FIX (8 scans)-RT (20 scans)-FIX (8 scans), Fig. 9a shows the signal associated with NPAIRS-gCCA ($\boldsymbol{b}_k$, $k=1,\ldots,M=10$) and Fig. 10b shows the signal associated with the NPAIRS-CVA map with 34 principal components (canonical variates) respectively. Boxes around each time point show the 25th to the 75th percentiles of the distribution of signal variation across subjects and runs. As expected these plots show that the temporal signal outlined using NPAIRS-CVA, which includes task class prediction over time/scans in its cost function, represents a consistent group task response. In contrast the temporal signal outlined using NPAIRS-gCCA doesn't match with the stimulus pattern.

Discussion

We proposed a multivariate method based on a generalized version of canonical correlation (gCCA) to extract a highly reproducible spatial map from fMRI data sets composed of two or more subjects. We used principal component analysis to reduce the dimensions of the data, decrease the computational cost and avoid inverting large matrices. Using gCCA, a different linear combination of the principal components from each subject is estimated such that the resulting spatial map is reproducible across all subjects. We assumed that the spatial map associated with the desired BOLD signal is more reproducible than artifacts such as motion physiological noise. It is very important to remove any spatially reproducible nuisance effects before applying our proposed gCCA-based method. This will guarantee that the most reproducible, functionally connected map associated with the BOLD signal appears in the first component of the generalized eigenvalue decomposition in Eq. (8).

NPAIRS-gCCA extracts a map for each subject such that the mean of the pair-wise correlation coefficients between the maps is maximized. These maps project the fMRI data of each subject onto a vector. The vectors are calculated by the gCCA approach. In this approach, the mean of the pair-wise correlation coefficients between the maps is used as a measure of reproducibility. Using other measures might

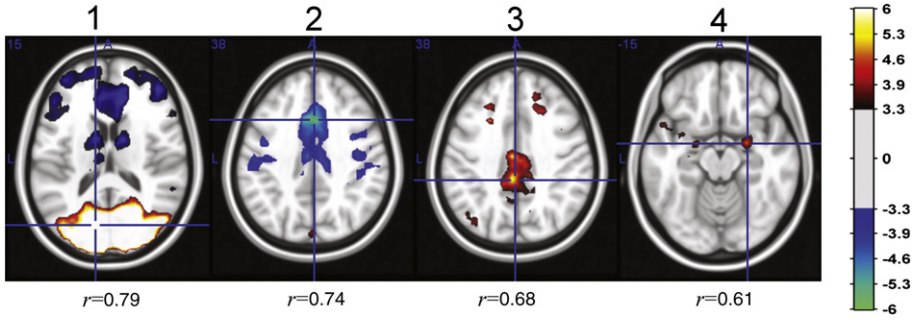


Fig. 8. Each axial figure shows some of the outlined area using 4 components of NPAIRS-gCCA for the resting-state datasets at the false alarm rate of 0.001 (first to forth gCCA components are shown from left to right). Using NPAIRS-gCCA the data of the subjects are combined to create four maps with significant activations and reproducibility of 0.79, 0.74, 0.68 and 0.61 for the first to the fourth components. Visual cortex, preforontal cortex and Thalamus (column 1: first component), medial frontal gyrus (column 2: second component), post cingulated gyrus (column 3: third component) and hippocampus (column 4: fourth component) are outlined. Other components did not outline significant activations and are not reported ($p>0.02$).

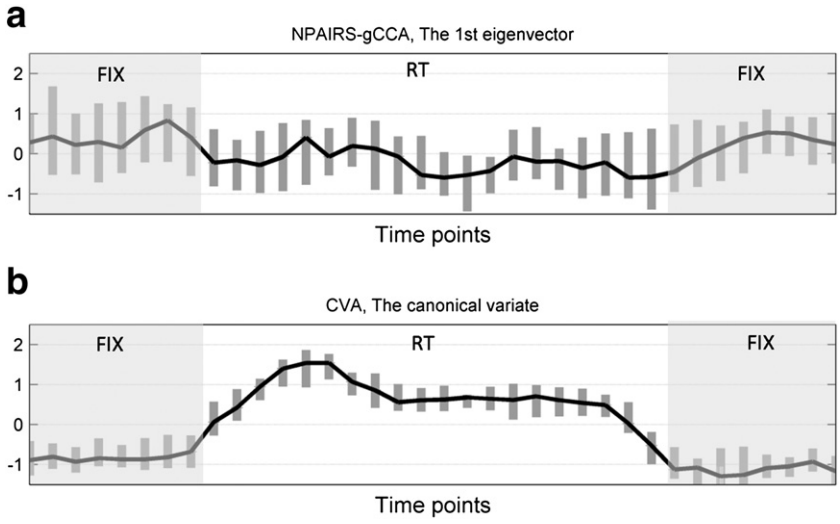


Fig. 9. (a) The BOLD signal associated with the first reproducible map using NPAIRS-gCCA. The box-plots in each time-point show the variation of the BOLD signal across different subjects and runs. (b) The BOLD signal associated with the detected map using NPAIRS-CVA for 34 principal components. The box-plots in each time-point show the variation of the BOLD signal across different subjects and runs. The BOLD signal associated with the outlined network using NPAIRS-gCCA is not reproducible in time but the BOLD signal detected using NPAIRS-CVA is highly reproducible in time.

lead to different results. Also, maximization of other measures might not be as straightforward as pair-wise correlation coefficients.

The proposed multivariate gCCA method maximizes the reproducibility of the derived spatial map but does not put any constraint on the temporal pattern of the BOLD signal, which may differ from subject to subject. However, the CVA implemented in the NPAIRS software package is a multivariate method with a cost function that focuses on temporal prediction. The different constrains in the time- and space-domains used in CVA and gCCA cause somewhat different activation maps to be estimated. CVA extracts a spatial map with an associated temporal BOLD signal that is predictive within and between subjects as a result of being correlated with the common stimulus pattern. Therefore, it may not be applicable when there is high variability in the subjects' BOLD responses like in resting-state experiments. In these experiments, each subject shows their own temporal pattern for the BOLD response as there are no common experimental stimuli for all subjects. This also might be the case in experiments with slow or challenging tasks or loosely defined paradigms where the actual time course of the cognitive event is not perfectly reproducible. On the other hand, gCCA finds such a reproducible map with a corresponding temporal pattern that may be different across subjects. As a result gCCA may not perform well when the goal is to outline the network that is responsible for performing the task. This difference in the extracted maps of gCCA and CVA represents the reproducibility-prediction trade-off described by LaConte et al. (2003) and Strother et al. (2004). These and subsequent results indicate that without selecting voxel subsets or enforcing sparsity an analysis technique may not be both highly predictive in time domain and reproducible in the spatial domain at the same time.

CVA and gCCA may represent two extreme cases of fMRI data processing approaches. The cost function in gCCA maximizes the spatial reproducibility without any constraint on the temporal reproducibility. On the other hand, CVA maximizes the temporal reproducibility without any constrain on the spatial reproducibility. Therefore, it appears desirable to add a penalty function to the CVA and gCCA cost functions to adjust the temporal and spatial reproducibility to desired levels. To some extent this appears to be happening with tuning temporal prediction and spatial reproducibility trade-offs by changing the number of principal components as a hyper-parameter in the NPAIRS-CVA framework (LaConte et al., 2003; Shaw et al., 2003; Strother et al., 2004, 2010; Yourganov et al., 2011; Zhang et al., 2008, 2009). We are currently investigating the relationship between the prediction–reproducibility tradeoffs seen in NPAIRS-CVA and gCCA, which is beyond the scope of this paper.

Also, we examined the possibility of using *supervised* dimensionality reduction approaches to extract a spatial reproducible map. Using a seed-based version of the gCCA approach, we limit the temporal response of each subject to follow a linear combination of their seed time-series. By choosing the seed voxels in the posterior cingulate cortex, we estimated a reproducible map depicting the areas functionally connected to the seed voxels. Also, using a model-based version of gCCA approach, we limit the temporal response of each subject to follow a linear combination of three regressors modeling the hemodynamic response to the stimulation pattern. The model-based dimensionality reduction in gCCA addresses both issues of modeling within and between subjects' hemodynamic response variability (Aguirre et al., 1998; Handwerker et al., 2004; Neumann et al., 2003) and improving the spatial reproducibility at the same time. Each subject in each run may have a different hemodynamic response to the stimulation pattern and therefore different linear combinations might be acquired for different subjects in different runs in a model based gCCA approach.

Furthermore, the model-based version of gCCA outlines the primary motor cortex (p-value = 0.00005) and the extracted map has a reproducibility of 0.48. This reproducibility is lower than the reproducibility of the extracted map ($r = 0.78$) using the NPAIRS-gCCA with PCA dimension reduction and the reproducibility of the seed-based NPAIRS-gCCA ($r = 0.76$). It seems that the default mode network includes more voxels than the task mode network for the RT task, which is a simple task and fewer brain resources/regions are utilized to perform it, especially in young subjects (Grady et al., 2010). Therefore, the most reproducible eigen-image assigns significant statistics to the voxels in the default mode network areas in NPAIRS-gCCA.

Finally, we measure the numbers of the principal components chosen for all subjects in leave-one-out cross-validation PCA dimension reduction. The number of principal components over subjects and splits nearly follows a Gaussian-like distribution with the minimum of 17, maximum of 49, and mean of 32. We note that whether adjusted per subject for maximum reproducibility with no temporal constraints in NPAIRS-gCCA, or for split-half groups of five subjects for maximum reproducibility with temporal task constraints, the

optimal PCA subspace is very similar containing 32 and 34 PCs, respectively. This is suggestive of there being a shared signal subspace of approximately this dimensionality across young, normal subjects performing a simple motor task. These results are directly relevant to the current efforts to represent the basic network structure of the brain using ICA of resting state data or through decomposition of the BrainMap database (Smith et al., 2009). How best to measure this number (e.g., Yourganov et al., 2011), and whether or not it varies significantly with tasks with tasks are important research questions that are accessible using the techniques we have outlined above. Such measurements may represent fundamental subspace size that reflects the dynamic repertoire of the brain (McIntosh et al., 2010) that is relatively conserved whether the brain is at rest or performing a specific task.

Conclusion

In this paper, we propose a method based on generalized CCA to process fMRI datasets. The gCCA based analysis technique is a linear method which produces high reproducible statistical parametric maps. We applied the gCCA based method on real fMRI dataset and compare it with CVA implemented in NPAIRS and NPAIRS-GLM by comparing the reproducibility of their generating maps. We show that the maps obtained by gCCA have higher reproducibility than CVA and NPAIRS-GLM. However the associated BOLD signal of the generate map using gCCA is less reproducible than that of the generated map using CVA. CVA detects the task related network in the brain better than gCCA but gCCA is more successful in segmenting the default mode network.

Acknowledgment

This research was supported in part by the University of Tehran under Grant no. 8101079/01/01. Partial support was also provided by CIHR Operating Grant (MOP 84483), and SCS gratefully acknowledges the support of the Heart and Stroke Foundation of Ontario through the Centre for Stroke Recovery. We thank Dr. Cheryl Grady for permission to use her reaction time fMRI data.

References

Aguirre, G.K., Zarahn, E., D'Esposito, M., 1998. The variability of human, BOLD hemodynamic responses. Neuroimage 8, 360–369.
Ardekani, B.A., Bachmana, A.H., Strother, S.C., Fujibayashic, Y., Yonekura, Y., 2004. Impact of inter-subject image registration on group analysis of fMRI data. Int. Congr. Ser. 1265, 49–59.
Branch, M.N., 1999. Statistical inference in behavior analysis: some things significance testing does and does not do. Behav. Anal. 22, 87–92.
Breakspear, M., Brammer, M.J., Bullmore, E.T., Das, P., Williams, L.M., 2004. Spatiotemporal wavelet resampling for functional neuroimaging data. Hum. Brain Mapp. 23, 1–25.
Buckner, R.L., Andrews-Hanna, J.R., Schacter, D.L., 2008. The brain's default network: anatomy, function, and relevance to disease. Ann. N. Y. Acad. Sci. 1124, 1–38.
Carver, R.P., 1993. The case against statistical significance testing, revisited. J. Exp. Educ. 61, 287–292.
Churchill, N., Oder, A., Herve, A., Fred, T., Wayne, L., Christopher, T., Graham, S.J., Strother, S.C., 2011. Optimizing preprocessing and analysis pipelines for single-subject fMRI: 1. Standard temporal motion and physiological noise correction methods. Hum. Brain Mapp. 33 (3), 609–627.
Evans, J.W., Todd, R.M., Taylor, M.J., Strother, S.C., 2010. Group specific optimisation of fMRI processing steps for child and adult data. Neuroimage 50 (2), 479–490.
Fax, M.D., Raichle, M.E., 2007. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. Nat. Rev. Neurosci. 8, 700–711.
Genovese, C.R., Noll, D.C., Eddy, W.F., 1997. Estimating test–retest reliability in functional MR imaging I: statistical methodology. Magn. Reson. Med. 38 (3), 497–507.
Grady, C.L., Protzner, A.B., Kovacevic, N., Strother, S.C., Afshin-Pour, B., Wojtowicz, M., Anderson, J.A.E., Churchill, N., McIntosh, A.R., 2010. A multivariate analysis of age-related differences in default mode and task-positive networks across multiple cognitive domains. Cereb. Cortex 20, 1432–1447.
Handwerker, D.A., Ollinger, J.M., D'Esposito, M., 2004. Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses. Neuroimage 21, 1639–1651.
Hossein-Zadeh, G.A., Ardekani, B.A., Soltanian-Zadeh, H., 2003. A signal subspace approach for modeling the hemodynamic response function in fMRI. Magn. Reson. Imaging 21, 835–843.
Kettenring, J.R., 1971. Canonical analysis of several sets of variables. Biometrik 58, 433–451.
Kjems, U., Hansen, L.K., Anderson, J., Frutiger, S., Muley, S., Sidtis, J., Rottenberg, D., Strother, S.C., 2002. The quantitative evaluation of functional neuroimaging experiments: mutual information learning curves. Neuroimage 15 (4), 772–786.
Kovacevic, N., Henderson, J.T., Chan, E., Lifshitz, N., Bishop, J., Evans, A.C., Henkelman, R.M., Chen, X.J., 2005. A three-dimensional MRI atlas of the mouse brain with estimates of the average and variability. Cereb. Cortex 15 (5), 639–645.
LaConte, S., Anderson, J., Muley, S., Ashe, J., Frutiger, S., Rehm, K., Hansen, L.K., Yacoub, E., Hu, X., Rottenberg, D., Strother, S.C., 2003. The evaluation of preprocessing choices in single-subject BOLD fMRI using NPAIRS performance metrics. Neuroimage 18 (1), 10–27.
Ledoit, O., Wolf, M., 2004. A well-conditioned estimator for large-dimensional covariance matrices. J. Multivariate Anal. 88 (2), 365–411.
Li, Y.O., Adal, T., Wang, W., Calhoun, V.D., 2009. Joint blind source separation by multiset canonical correlation analysis. IEEE Trans. Signal Process. 57 (10), 3918–3929.
Liou, M., Su, H.R., Lee, J.D., Aston, J.A.D., Tsai, A.C., Cheng, P.E., 2006. A method for generating reproducible evidence in fMRI studies. Neuroimage 29, 383–395.
Liou, M., Su, H.R., Savostyanov, A.N., Lee, J.D., Aston, J.A.D., Chuang, C.H., Cheng, P.E., 2009. Beyond p-values: averaged and reproducible evidence in fMRI experiments. Psychophysiology 46, 367–378.
Mardia, K.V., Kent, J.T., Bibby, J.M., 1979. Multivariate Analysis. Academic press.
McIntosh, A.R., Kovacevic, N., Lippe, S., Garrett, D., Grady, C., Jirsa, V., 2010. The development of a noisy brain. Arch. Ital. Biol. 148 (3), 323–337.
Mei, L., Figl, M., Rueckert, D., Darzi, A., Edwards, P., 2008. Statistical shape modelling: how many modes should be retained? 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp. 1–8.
Neumann, J., Lohmann, G., Zysset, S., von Cramon, D.Y., 2003. Within-subject variability of BOLD response dynamics. Neuroimage 19 (3), 784–796.
Nickerson, R.S., 2000. Null hypothesis significance testing: a review of an old and continuing controversy. Psychol. Methods 5, 241–301.
Salli, E., Korvenoja, A., Visa, A., Katila, T., Aronen, H.J., 2001. Reproducibility of fMRI: effect of the use of contextual information. Neuroimage 13, 459–471.
Shaw, M.E., Strother, S.C., Gavrilescu, M., Podzebenko, K., Waites, A., Watson, J., Anderson, J., Jackson, G., Egan, G., 2003. Evaluating subject specific preprocessing choices in multisubject fMRI data sets using data-driven performance metrics. Neuroimage 19 (3), 988–1001.
Smith, L.D., Best, L.A., Cylke, V.A., Stubbs, D.A., 2000. Psychology without p values. Am. Psychol. 55, 260–263.
Smith, S.M., Fox, P.T., Miller, K.L., Glahn, D.C., Fox, P.M., Mackay, C.E., Filippini, N., Watkins, K.E., Toro, R., Laird, A.R., Beckmann, C.F., 2009. Correspondence of the brain's functional architecture during activation and rest. Proc. Natl. Acad. Sci. U. S. A. 106 (31), 13040–13045 4.
Strother, S.C., Lange, N., Anderson, J.R., Schaper, K.A., Rehm, K., Hansen, L.K., Rottenberg, D.A., 1997. Activation pattern reproducibility: measuring the effects of group size and data analysis models. Hum. Brain Mapp. 5 (4), 312–316.
Strother, S.C., Anderson, J., Hansen, L.K., Kjems, U., Kustra, R., Sidtis, J., Frutiger, S., Muley, S., LaConte, S., Rottenberg, D., 2002. The quantitative evaluation of functional neuroimaging experiments: the NPAIRS data analysis framework. Neuroimage 15 (4), 747–771.
Strother, S.C., La Conte, S., Hansen, L.K., Anderson, J., Zhang, J., Pulapura, S., Rottenberg, D., 2004. Optimizing the fMRI data-processing pipeline using prediction and reproducibility performance metrics: I. A preliminary group analysis. Neuroimage 23 (S1), S196–S207.
Strother, S., Oder, A., Spring, R., Grady, C., 2010. The NPAIRS computational statistics framework for data analysis in neuroimaging. In: Lechevallier, Y., Saporta, G. (Eds.), 19th International Conference on Computational Statistics: Keynote, Invited and Contributed Papers. Physica-Verlag, Springer, Paris, France, pp. 111–120.
Swallow, K.M., Braver, T.S., Snyder, A.Z., Speer, N.K., Zacks, J.M., 2003. Reliability of functional localization using fMRI. Neuroimage 20, 1561–1577.
Tegeler, C., Strother, S.C., Anderson, J.R., Kim, S.G., 1999. Reproducibility of BOLD-based functional MRI obtained at 4 T. Hum. Brain Mapp. 7 (9), 267:283–267:.
Tenenhaus, A., Tenenhaus, M., 2011. Regularized generalized canonical correlation analysis. Psychometrika 76 (2), 257–284.
Thirion, B., Pinel, P., Tucholka, A., Roche, A., Ciuciu, P., Mangin, J.F., Poline, J.B., 2007. Structural analysis of fMRI data revisited: improving the sensitivity and reliability of fMRI group studies. IEEE Trans. Med. Imaging 26 (9), 1256–1269.
Varoquaux, G., Sadaghiani, S., Pinel, P., Kleinschmidt, A., Poline, J.B., Thirion, B., 2010. A group model for stable multi-subject ICA on fMRI datasets. Neuroimage 51 (1), 288–299.
Wang, Z., Peterson, B.S., 2009. Partner-matching for the automated identification of reproducible ICA components from fMRI datasets: algorithm and validation. Hum. Brain Mapp. 29, 875–893.
Woods, R.P., Grafton, S.T., Holmes, C.J., Cherry, S.R., Mazziotta, J.C., 1998. Automated image registration: I. General methods and intrasubject, intramodality validation. J. Comput. Assist. Tomogr. 22, 139–152.
Yang, Z., LaConte, S., Weng, X., Hu, X., 2008. Ranking and averaging independent component analysis by reproducibility (RAICAR). Hum. Brain Mapp. 29, 711–725.
Yourganov, G., Xu, C., Lukic, A., Grady, C., Small, S., Wernick, M., Strother, S.C., 2011. Dimensionality estimation for optimal detection of functional networks in BOLD fMRI Data. Neuroimage 56 (2), 531–543.

Yousry, T.A., Schmid, U.D., Alkadhi, H., Schmidt, D., Peraud, A., Buettner, A., Winkler, P., 1997. Localization of the motor hand area to a knob on the precentral gyrus — a new landmark. Brain 120, 141–157.
Zeng, W., Qiu, A., Chodkowski, B., Pekar, J.J., 2009. Spatial and temporal reproducibility-based ranking of the independent components of BOLD fMRI data. Neuroimage 46, 1041–1054.
Zhang, J., Liang, L., Anderson, J.R., Gatewood, L., Rottenberg, D.A., Strother, S.C., 2008. Evaluation and comparison of GLM- and CVA-based fMRI processing pipelines with Java-based fMRI processing pipeline evaluation system. Neuroimage 41 (4), 1242–1252.
Zhang, J., Anderson, J.R., Liang, L., Pulapura, S.K., Gatewood, L., Rottenberg, D.A., Strother, S.C., 2009. Evaluation and optimization of fMRI single-subject processing pipelines with NPAIRS and second-level CVA. Magn. Reson. Imaging 27 (2), 264–278.