



## Transport on Riemannian Manifold for Functional Connectivity-based Classification

Bernard Ng, Martin Dressler, Gaël Varoquaux, Jean-Baptiste Poline, Michael Greicius, Bertrand Thirion

### ► To cite this version:

Bernard Ng, Martin Dressler, Gaël Varoquaux, Jean-Baptiste Poline, Michael Greicius, et al.. Transport on Riemannian Manifold for Functional Connectivity-based Classification. MIC-CAI - 17th International Conference on Medical Image Computing and Computer Assisted Intervention, Sep 2014, Boston, United States. Springer, 2014. <hal-01058521>

**HAL Id: hal-01058521**

**<https://hal.inria.fr/hal-01058521>**

Submitted on 27 Aug 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Transport on Riemannian Manifold for Functional Connectivity-based Classification

Bernard Ng<sup>1,2</sup>, Martin Dressler<sup>3</sup>, Gael Varoquaux<sup>1</sup>, Jean Baptiste Poline<sup>1</sup>,  
Michael Greicius<sup>2</sup>, and Bertrand Thirion<sup>1</sup>

<sup>1</sup>Parietal team, Neurospin, INRIA Saclay, France

<sup>2</sup>FIND Lab, Stanford University, United States

<sup>3</sup>Donders Institute for Brain, Cognition and Behaviour, Nijmegen, Netherlands  
bernardyng@gmail.com

**Abstract.** We present a Riemannian approach for classifying fMRI connectivity patterns before and after intervention in longitudinal studies. A fundamental difficulty with using connectivity as features is that covariance matrices live on the positive semi-definite cone, which renders their elements inter-related. The implicit independent feature assumption in most classifier learning algorithms is thus violated. In this paper, we propose a matrix whitening transport for projecting the covariance estimates onto a common tangent space to reduce the statistical dependencies between their elements. We show on real data that our approach provides significantly higher classification accuracy than directly using Pearson's correlation. We further propose a non-parametric scheme for identifying significantly discriminative connections from classifier weights. Using this scheme, a number of neuroanatomically meaningful connections are found, whereas no significant connections are detected with pure permutation testing.

**Keywords:** Classification, Connectivity, fMRI, Riemannian manifold

## 1 Introduction

Recent years have witnessed significant growth in the application of resting state functional magnetic resonance imaging (RS-fMRI) for neuroscience research [1]. Besides addressing the fundamental question of how the brain is wired, substantial effort has been placed on translating RS-fMRI for classification tasks, such as separating diseased and healthy subjects [1]. Recently, there is a growing interest in using RS-fMRI to examine the effects of intervention on intrinsic connectivity [1]. For example, comparing the functional connectome before and after motor learning can provide insights into neuroplasticity [1]. In this work, we focus on the task of classifying connectivity patterns before and after intervention in longitudinal studies.

Directly using conventional Pearson's correlation, i.e. normalized covariance, as features for classifier learning has a fundamental limitation. Since covariance matrices,  $\Sigma$ , live on the space of positive semidefinite cone, i.e.  $\mathbf{v}^T \Sigma \mathbf{v} \geq 0$  for all vectors  $\mathbf{v}$ , elements of  $\Sigma$  are inherently inter-related [2, 3], which violates the independent feature assumption implicit in most classifier learning algorithms. In particular,

many classifier learning algorithms become unstable in the face of correlated features, i.e. small perturbations on the training data can alter the relative weighting of the features [4]. Worsening the situation is the lack of training samples. Connectivity-based classifier learning is usually performed on data from multiple subjects to keep scan time reasonable for each subject, i.e. reliable covariance estimation requires at least several minutes of RS-fMRI data. For typical studies, the number of subjects is no more than thirty or so. Hence, the sample size is rather limited. As a result, the estimated classifier weights would have high variance. Combined with the effects of correlated features, the generalizability of the classifiers is limited, and identification of significantly discriminative connections from these classifiers is nontrivial.

In this paper, we present a Riemannian approach for classifying connectivity patterns before and after intervention in longitudinal studies. The heart of our approach is to treat the positive definite cone as a Riemannian manifold and project the covariance estimates onto a common tangent space of this manifold. On the tangent space, elements of the covariance matrices are no longer linked by the positive definite constraint [3]. Thus, the impact of correlated connectivity features on classifier learning is alleviated. To bring the covariance estimates of all subjects to a common tangent space, we propose a matrix whitening transport. The underlying idea is to find a subject-specific base covariance estimate that is close to all covariance estimates of each subject and use it for matrix whitening. The resulting covariance estimates of all subjects would be close to the identity matrix. Thus, we can use the tangent space at the identity matrix as the common space for projection. For comparison, we examine the concept of parallel transport from differential geometry [5]. To perform parallel transport, we use the Schild’s ladder algorithm [5, 6]. For validation, we apply our approach to RS scans of fifty one subjects before and after a memory task, and compare it against directly using Pearson’s correlation and its regularized variants as features. Furthermore, we propose a non-parametric scheme that combines bootstrapping and permutation testing for identifying significantly discriminative connections from the classifier weights. Detection of neuroanatomically relevant connections is shown.

## 2 Methods

Treating the positive definite cone as a Riemannian manifold and using the associated manifold operations (Section 2.1), we propose a matrix whitening transport for covariance projection (Section 2.2), and compare it against parallel transport with Schild’s ladder (Section 2.3). Various methods for estimating base covariance matrices are examined (Section 2.4). A non-parametric scheme is devised for identifying significantly discriminative brain connections from classifier weights (Section 2.5).

### 2.1 Manifold Operations for the Space of Positive Definite Matrices

Let  $\mathbf{A} \in S^{++}$  be a  $d \times d$  matrix that lives on the positive definite cone,  $S^{++}$ . Due to the positive definite constraint:  $\mathbf{v}^T \mathbf{A} \mathbf{v} > 0$  for all vectors  $\mathbf{v}$ , elements of  $\mathbf{A}$  are inter-related.

One way to decouple the elements of  $\mathbf{A}$  is to treat  $S^{++}$  as a Riemannian manifold and project  $\mathbf{A}$  onto the tangent space at a  $d \times d$  base point,  $\mathbf{B} \in S^{++}$ , using the *Log* map [3]:

$$\text{Log}_{\mathbf{B}}(\mathbf{A}) = \mathbf{B}^{1/2} \text{logm}(\mathbf{B}^{-1/2} \mathbf{A} \mathbf{B}^{-1/2}) \mathbf{B}^{1/2}, \quad (1)$$

where  $\text{logm}(\cdot)$  denotes matrix logarithm and  $\text{Log}_{\mathbf{B}}(\mathbf{A})$  is the tangent vector at  $\mathbf{B}$  “pointing towards”  $\mathbf{A}$  with  $\mathbf{B}$  assumed to be close to  $\mathbf{A}$  [5]. The key property of  $\text{Log}_{\mathbf{B}}(\mathbf{A})$  for our purposes is that its elements are not linked by the positive definite constraint [3], which helps alleviate the impact of correlated features on classifier learning. More generally, a major difficulty with working in  $S^{++}$  is that the resultant from even standard operations, such as subtraction, may not reside in  $S^{++}$ . An elegant solution to this problem is to operate in the tangent space and project the resultant back onto  $S^{++}$  using the inverse mapping, i.e. the *Exp* map (2), which guarantees positive definiteness [3].

$$\text{Exp}_{\mathbf{B}}(\mathbf{A}) = \mathbf{B}^{1/2} \text{expm}(\mathbf{B}^{-1/2} \mathbf{A} \mathbf{B}^{-1/2}) \mathbf{B}^{1/2}, \quad (2)$$

where  $\text{expm}(\cdot)$  denotes matrix exponential. By combining (1) and (2), one can compute the geodesic, i.e. local shortest path, between two positive definite matrices [3]:

$$\gamma(t) = \text{Exp}_{\mathbf{B}}(t \cdot \text{Log}_{\mathbf{B}}(\mathbf{A})), \quad t \in [0,1]. \quad (3)$$

This concept of geodesic will be important for parallel transport (Section 2.3).

## 2.2 Proposed Matrix Whitening Transport

Let  $\mathbf{C}(c, s) \in S^{++}$  be the  $d \times d$  covariance matrix of subject  $s$  associated with scan  $c$ , where  $d$  is the number of brain regions. Also, let  $\overline{\mathbf{C}}(s) \in S^{++}$  be a  $d \times d$  base covariance matrix of each subject  $s$  that is close to  $\mathbf{C}(c, s)$  for all  $c$ . We note here again that our goal is to classify connectivity patterns derived from scans with interventions interleaved in between. Now, if we simply apply (1) to project  $\mathbf{C}(c, s)$  onto the tangent space at  $\overline{\mathbf{C}}(s)$ , all subjects’ projected covariance matrices will lie in different tangent spaces (i.e.  $\overline{\mathbf{C}}(s)$  are different across subjects), hence not comparable with each other. Instead, under the assumption that  $\overline{\mathbf{C}}(s)$  is close to all  $\mathbf{C}(c, s)$  of subject  $s$ , we can use  $\overline{\mathbf{C}}(s)$  to approximately whiten  $\mathbf{C}(c, s)$  for each subject, so that the resulting covariance matrices,  $\overline{\mathbf{C}}(s)^{-1/2} \mathbf{C}(c, s) \overline{\mathbf{C}}(s)^{-1/2}$ , would be close to the identity matrix,  $\mathbf{I}_{d \times d}$ , for all subjects. Thus, the tangent space at  $\mathbf{I}_{d \times d}$  can serve as the common space across subjects for projection, which reduces to taking the matrix logarithm:

$$d\mathbf{C}(c, s) = \text{logm}(\overline{\mathbf{C}}(s)^{-1/2} \mathbf{C}(c, s) \overline{\mathbf{C}}(s)^{-1/2}), \quad (4)$$

since  $\text{Log}_{\mathbf{I}_{d \times d}}(\mathbf{A}) = \mathbf{I}_{d \times d}^{1/2} \text{logm}(\mathbf{I}_{d \times d}^{-1/2} \mathbf{A} \mathbf{I}_{d \times d}^{-1/2}) \mathbf{I}_{d \times d}^{1/2} = \text{logm}(\mathbf{A})$ . The proposed matrix whitening transport is summarized in Fig. 1(a). In effect, we are isolating the distinctive connectivity attributes by removing commonalities between scans in a nonlinear

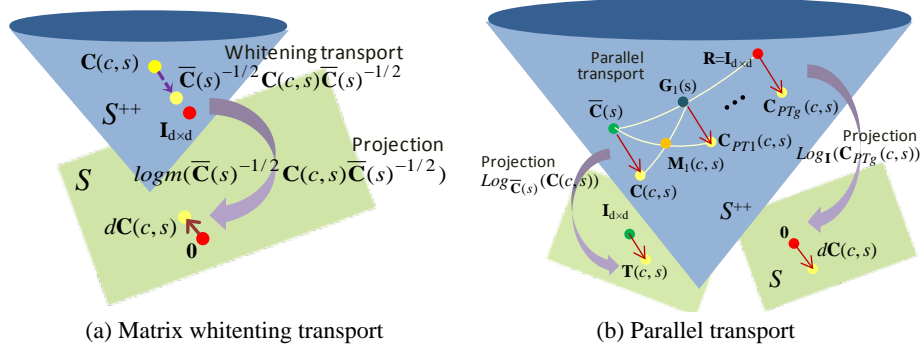


Fig. 1. Proposed matrix whitening transport versus parallel transport.

fashion, and using the residual for classification. It turns out that (4) is analogous to the manifold operation deployed in [2] for building one-class classifier to discriminate subject types, but we are proposing a new perspective on the entailed matrix multiplication as a transport to the neighborhood of  $\mathbf{I}_{d \times d}$ , which justifies why projection onto the tangent space at  $\mathbf{I}_{d \times d}$  is legitimate. To generate well-conditioned, positive definite covariance estimates, we employ oracle approximating shrinkage (OAS), which has a closed-form solution for determining the optimal level of  $l_2$  regularization [7].

To examine the necessity of matrix whitening, we compare (4) with the case of setting  $d\mathbf{C}(c, s)$  to  $\text{logm}(\mathbf{C}(c, s))$ , i.e. no transport before projection, which is exactly the Log-Euclidean approach employed by the diffusion MRI community [8]. Another simplification to (4), which we refer to as the Euclidean approximation, is to remove the commonalities between the  $\mathbf{C}(c, s)$ 's in a linear fashion and apply a first order approximation to the matrix logarithm, i.e.  $\text{logm}(\mathbf{A}) \approx \mathbf{A} - \mathbf{I}_{d \times d}$ , resulting in  $d\mathbf{C}(c, s) = \mathbf{C}(c, s) - \bar{\mathbf{C}}(s)$  for all off-diagonal elements. Since linear subtraction can result in non-positive definite matrices, the first order  $\text{logm}$  approximation is necessary. Also, for connectivity-based classification, only the off-diagonal elements are of interest.

### 2.3 Parallel Transport with Schild's Ladder

Another way of transporting the covariance matrices of all subjects to a common tangent space is to employ parallel transport, which provides the least deforming way of moving geometric objects along a curve on a manifold [5]. One way of performing parallel transport is to use the Schild's ladder algorithm [5, 6], as summarized in Fig. 1(b). Let  $\mathbf{T}(c, s) = \text{Log}_{\bar{\mathbf{C}}(s)}(\mathbf{C}(c, s))$  be a  $d \times d$  projected covariance matrix, i.e. a tangent vector, at  $\bar{\mathbf{C}}(s)$ . To parallel transport  $\mathbf{T}(c, s)$  from the tangent space at  $\bar{\mathbf{C}}(s)$  to the tangent space at  $\mathbf{R} \in S^{++}$  using Schild's ladder, we first determine using (3) the point on  $S^{++}$  that is unit distance from  $\bar{\mathbf{C}}(s)$  along the geodesic uniquely defined by  $\mathbf{T}(c, s)$  [3], which is exactly  $\mathbf{C}(c, s)$  since  $\mathbf{C}(c, s) = \text{Exp}_{\bar{\mathbf{C}}(s)}(1 \cdot \text{Log}_{\bar{\mathbf{C}}(s)}(\mathbf{C}(c, s)))$ .

Next, we generate a discretized geodesic from  $\bar{\mathbf{C}}(s)$  to  $\mathbf{R}$ :  $\mathbf{G}_i(s)$ ,  $i = \{1, \dots, g\}$ , where  $g$  is the number of points along the geodesic, and find the midpoint of the geodesic joining  $\mathbf{G}_1(s)$  and  $\bar{\mathbf{C}}(c, s)$ :  $\mathbf{M}_1(c, s) = \text{Exp}_{\bar{\mathbf{C}}(c, s)}(0.5 \cdot \text{Log}_{\bar{\mathbf{C}}(c, s)}(\mathbf{G}_1(s)))$ . We then construct a geodesic from  $\bar{\mathbf{C}}(s)$  to  $\mathbf{M}_1(c, s)$  and move twice the distance to find the one-step parallel transported covariance matrix:  $\mathbf{C}_{\text{PT}_1}(c, s) = \text{Exp}_{\bar{\mathbf{C}}(s)}(2 \cdot \text{Log}_{\bar{\mathbf{C}}(s)}(\mathbf{M}_1(c, s)))$ .

This procedure is repeated for all  $\mathbf{G}_i(s)$ ,  $i = \{1, \dots, g\}$  until we reach  $\mathbf{R}$ , where we perform projection to find the parallel transported  $\mathbf{T}(c, s)$ :  $d\mathbf{C}(c, s) = \text{Log}_{\mathbf{R}}(\mathbf{C}_{\text{PT}_g}(c, s))$ . In effect, the Schild’s ladder algorithm is parallel transporting  $\mathbf{T}(c, s)$  by forming parallelograms on  $S^{++}$  but with geodesics in place of straight lines. To enable direct comparisons with the proposed transport, we set  $\mathbf{R}$  to  $\mathbf{I}_{d \times d}$ . As for  $g$ , we have tried values from 1 to 10 with close to exactly the same classification results obtained.

## 2.4 Base Covariance Estimation

Let  $\mathbf{X}(c, s)$  be a  $t \times d$  matrix containing  $d$  brain region time courses of subject  $s$  associated with scan  $c = 1$  to  $N$ . For estimating a base covariance matrix that is close to all covariance estimates of each subject  $s$ , we examine three mean covariance estimation methods: Euclidean mean, Log-Euclidean mean, and time course concatenation. We exclude the Frechet mean [3] due to the observed numerical instability, e.g. using Euclidean mean versus Log-Euclidean mean of the covariance estimates as initialization result in different Frechet mean estimates. The Euclidean mean is simply given by:  $\Sigma_c \mathbf{C}(c, s)/N$ , but does not preserve the spectral properties of the covariance estimates [8]. One way to mitigate this problem is to first apply matrix logarithm, take the mean, and apply matrix exponential to bring the mean back to  $S^{++}$ , i.e.  $\text{expm}(\Sigma_c \text{logm}(\mathbf{C}(c, s)))/N$  [8]. Yet, another way for estimating a mean covariance matrix is to concatenate  $\mathbf{X}(c, s)$  across scans into a  $ct \times d$  matrix and apply OAS [7].

## 2.5 Discriminative Brain Connection Identification

Critical to neuroscience studies is result interpretability. For identifying significantly discriminative connections from classifier weights, we propose here a non-parametric scheme that combines bootstrapping with permutation testing. Bootstrapping enables identification of the more stable discriminative features, while permutation testing facilitates generation of a null distribution. Importantly, the chance of assigning large weights to the same brain connections for different bootstrap samples would presumably be lower with labels permuted. This intuition is exploited in our proposed scheme, which proceeds as follows. Let  $\mathbf{w}_{pq}$  be the classifier weights for scan  $p$  versus scan  $q$  learned from all subjects’  $d\mathbf{C}(c, s)$  for  $c = p$  and  $q$ . We first randomly permute the scan labels  $p$  and  $q$  10,000 times. For each permutation, we perform classifier learning on each of  $B = 500$  bootstrap samples (with replacement). Denoting the classifier weights for each bootstrap sample  $b$  as  $\mathbf{w}_{pq}^b$ , we compute the normalized mean over bootstrap samples:  $\bar{\mathbf{w}}_{pq} = 1/B \cdot \Sigma_b \mathbf{w}_{pq}^b / \text{std}(\mathbf{w}_{pq}^b)$ , and store the maximum element of  $\bar{\mathbf{w}}_{pq}$  for each permutation. We then compute  $\bar{\mathbf{w}}_{pq}$  without label permutation,

and declare its elements as statistically significant if they are greater than the 95th percentile of the maximum normalized mean distribution, corresponding to a p-value threshold of 0.05. We highlight that using maximum statistics implicitly accounts for multiple comparisons [9]. The same procedure is applied for finding significantly negative elements of  $\mathbf{w}_{pq}$  but with maximum replaced by minimum.

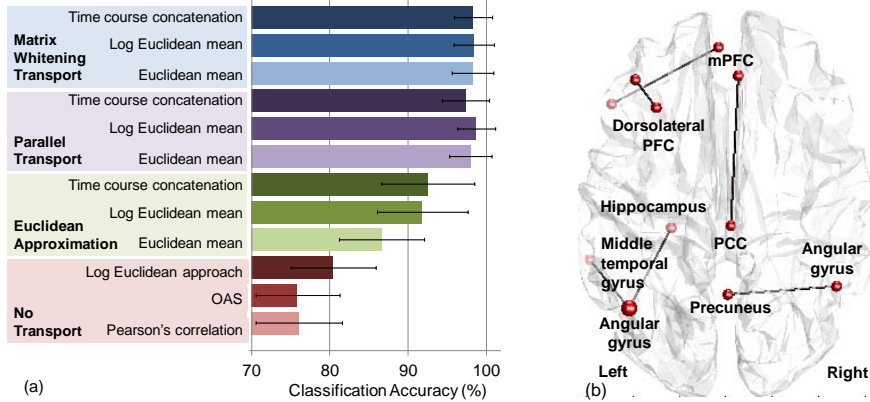
### 3 Materials

RS-fMRI data were collected from 51 healthy subjects. All subjects underwent two scans of 8 min each. Between the scans, the subjects were asked to memorize 72 words in 12 min and were tested immediately after encoding in a recognition task of 12 min. Acquisition was performed using a 3.0T GE scanner with TR = 2.5 s, TE = 30 ms, and flip angle = 90°. The RS-fMRI data were motion corrected using FSL and spatially normalized using SPM8. Motion artifacts, white matter and cerebrospinal fluid confounds, and principal components from high variance voxels extracted using CompCor [10] were regressed out from the voxel time courses. A bandpass filter with cutoff frequencies at 0.01 and 0.1 Hz was subsequently applied. For defining brain regions, we employed the atlas in [11], which comprises 90 functional regions that span 14 large-scale networks. We restricted our analysis to 78 of these regions due to incomplete cerebellum coverage. Gray matter voxel time courses within each region were averaged to generate brain region time courses. These time courses were normalized by subtracting the mean and dividing by the standard deviation.

### 4 Results and Discussion

To evaluate the gain of decoupling connectivity features, we compared Pearson’s correlation and its OAS regularized variant against using projected covariance estimates for classifying the connectivity patterns before and after the described memory task. We used  $l_2$ -regularized linear SVM for classifier learning with the soft margin parameter left at its default value of 1. We opted to use SVM since  $l_1$ -regularized classifiers tend to be unstable with limited samples [4]. For estimating classification accuracy, we performed repeated subsampling over 10,000 random splits: 34 subjects used for training and 17 subjects used for testing. Subsampling on subjects prevents  $d\mathbf{C}(c, s)$  of different  $c$  from the same subject to be used for training and testing, which avoids introducing correlations between the training and test samples.

Both Pearson’s correlation and its OAS regularized variant resulted in an accuracy of 76%. Projection using the Log Euclidean approach [8] without transport obtained an accuracy of 80%. Using the Euclidean approximation to the matrix whitening transport obtained accuracy of 87%–92%. Accounting for the manifold structure of  $S^{++}$  using the matrix whitening transport and parallel transport achieved accuracy of 98%, which is a 22% increase in accuracy compared to the standard approach of directly using Pearson’s correlation as features. Also, the substantial improvement compared to the Log Euclidean approach suggests the necessity for transporting prior to projection. Further, lower variability in accuracy was observed compared to



**Fig. 2.** Classification accuracy and significantly discriminative connections.

the Euclidean approximation. Overall, our results show that treating  $S^{++}$  as a Riemannian manifold and using the associated operations to decouple connectivity features is highly beneficial for connectivity-based classification. Since the matrix whitening transport and parallel transport provided similar performance, from a computational standpoint, the proposed transport is preferred considering the increasing temporal resolution offered by recent acquisition protocols that enable larger connectivity matrices to be reliably estimated. We note that using sparse Gaussian graphical model for covariance estimation resulted in close to exactly the same classification accuracy.

To identify significantly discriminative brain connections, we applied our proposed non-parametric scheme (Section 2.5) on classifier weights learned with all subjects'  $dC(s, c)$  as features. We present here only results for  $dC(s, c)$  estimated using the matrix whitening transport with mean covariance computed from concatenated time courses. Statistical significance is declared at a p-value threshold of 0.05 with multiple comparisons implicitly corrected for using maximum statistics [9]. We note that no significant connections were found with pure permutation testing on classifier weights derived from the same  $dC(s, c)$ , hence illustrating superior detection sensitivity with our proposed scheme for discriminative connection identification.

Using our non-parametric scheme, a number of significant connections between memory-relevant regions were found (Fig. 2(b)). The posterior cingulate cortex (PCC), precuneus, angular gyrus, and medial prefrontal cortex (mPFC) are key regions of the default-mode network (DMN) [12], which are closely associated with memory processes [13] and targeted early in Alzheimer's disease, i.e. the quintessential disorder of memory impairment. The dorsolateral PFC is implicated in episodic memory encoding and retrieval [14], and the hippocampus, which is often found to be coupled with the DMN and nearby temporal regions, is critical for encoding and retrieval of recent information [15]. The detected connections thus conform to what one would expect for the memory task employed. An interesting implication from our results is that even within a short duration of 30 minutes, detectable functional rewiring appears to be present, which provides further evidence for the shorter temporal scale of functional plasticity compared to its anatomical counterpart.



## 5 Conclusions

We presented a Riemannian approach for classifying functional connectivity patterns before and after interventions. By using manifold operations to decouple connectivity features, significantly higher classification accuracy was obtained compared to directly using connectivity estimates as features, which are inherently inter-related. Also, higher detection sensitivity was shown with our proposed discriminative connection identification scheme compared to conventional permutation testing. The overall classification framework thus provides both accuracy and interpretability.

## References

1. Castellanos, F.X., Di Martino, A., Craddock, R.C., Mehta, A.D., Milham, M.P.: Clinical Applications of the Functional Connectome. *NeuroImage* 28, 527–540 (2013)
2. Varoquaux, G., Baronnet, F., Kleinschmidt, A., Fillard, P., Thirion, B.: Detection of Brain Functional-connectivity Difference in Post-stroke Patients Using Group-level Covariance Modeling. In: Jiang, T., Colchester, A., Duncan, J. (eds.) *MICCAI 2010. LNCS*, vol. 6361, pp. 200–208, Springer, Heidelberg (2010)
3. Pennec, X., Fillard, P., Ayache, N.: A Riemannian Framework for Tensor Computing. *Int. J. Comp. Vision* 66, 41–66 (2006)
4. Tolosi, L., Lengauer, T.: Classification with Correlated Features: Unreliability of Feature Ranking and Solutions. *Bioinformatics* 27, 1986–1994 (2011)
5. Hauberg, S., Lauze, F., Pedersen, K.S.: Unscented Kalman Filtering on Riemannian Manifolds. *J. Math. Imaging & Vision* 46, 103–120 (2013)
6. Lorenzi, M., Ayache, N., Pennec, X.: Schild’s Ladder for the Parallel Transport of Deformations in Time Series of Images. In: Szekely, G., Hahn, H.K. (eds.) *IPMI 2011. LNCS*, vol. 6801, pp. 463–474, Springer, Heidelberg (2011)
7. Chen, Y., Wiesel, A., Eldar, Y.C., Hero, A.O.: Shrinkage Algorithms for MMSE Covariance Estimation. *IEEE Trans. Sig. Proc.* 58, 5016–5029 (2010)
8. Arsigny, V., Fillard, P., Pennec, X., Ayache, N.: Fast and Simple Calculus on Tensors in the Log-Euclidean Framework. In: Duncan, J., Gerig, G. (eds.) *MICCAI 2005. LNCS*, vol. 8, pp. 115–122, Springer, Heidelberg (2005)
9. Nichols, T., Hayasaka, S.: Controlling the Familywise Error Rate in Functional Neuroimaging: a Comparative Review. *Stat. Methods Med. Research* 12, 419–446 (2003)
10. Behzadi, Y., Restom, K., Liau, J., Liu, T.T.: A Component based Noise Correction Method (CompCor) for BOLD and Perfusion based fMRI. *NeuroImage* 37:90–101(2007)
11. Shirer, W.R., Ryali, S., Rykhlevskaia, E., Menon, V., Greicius, M.D.: Decoding Subject-driven Cognitive States with Whole-brain Connectivity Patterns. *Cereb. Cortex* 22, 158–165 (2012)
12. Greicius, M.D., Krasnow, B., Reiss, A.L., Menon, V.: Functional Connectivity in the Resting Brain: A Network Analysis of the Default Mode Hypothesis. *Proc. Natl. Acad. Sci.* 100, 253–258 (2003)
13. Kim, H. Dissociating the roles of the default-mode, dorsal, and ventral networks in episodic memory retrieval. *NeuroImage* 50, 1648–1657 (2010)
14. Sandrini, M., Cappa, S.F., Rossi, S., Rossini, P.M., Miniussi, C.: The Role of Prefrontal Cortex in Verbal Episodic Memory: rTMS Evidence. *J. Cog. Neurosci.* 15, 855–861 (2003)
15. Squire, L.R.: *The neuropsychology of memory*. Springer (1984)