



HAL
open science

Disentangling spatio-temporal patterns of brain changes in large-scale brain imaging databases through Independent Gaussian Process Analysis

Clement Abi Nader, Nicholas Ayache, Valeria Manera, Philippe Robert,
Marco Lorenzi

► To cite this version:

Clement Abi Nader, Nicholas Ayache, Valeria Manera, Philippe Robert, Marco Lorenzi. Disentangling spatio-temporal patterns of brain changes in large-scale brain imaging databases through Independent Gaussian Process Analysis. 12ème Conférence Francophone d'Epidémiologie Clinique (EPICLIN) et 25èmes Journées des statisticiens des Centre de Lutte Contre le Cancer (CLCC), May 2018, Nice, France. *Revue d'Épidémiologie et de Santé Publique* (66), pp.S159, 2018, 10.1016/j.respe.2018.03.108 . hal-01826517

HAL Id: hal-01826517

<https://hal.science/hal-01826517>

Submitted on 29 Jun 2018

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.

Disentangling spatio-temporal patterns of brain changes in large-scale brain imaging databases through Independent Gaussian Process Analysis

Clement Abi Nader¹, Nicholas Ayache¹, Valeria Manera^{2,3}, Philippe Robert^{3,4}, Marco Lorenzi¹

¹ UCA, Inria Sophia Antipolis, Epione Research Project

² UCA, Inria Sophia Antipolis, Stars Research Project

³ CoBTeK, University of Nice Sophia Antipolis
Centre Memoire, CHU de Nice

Abstract. The morphological changes affecting the brain over time are related to several biological processes, governed either by healthy aging or by pathological factors. Since these processes are to date largely unknown, we need statistical approaches to automatically identify these latent morphological evolutions through the analysis of structural brain Magnetic Resonance Images (MRIs). Such approaches require to scale to high-dimensional volumetric observations (around 10e6 voxels) with application to the analysis of large scale biomedical databases such as UK BIOBANK (thousands of individuals across time). We present a novel spatio-temporal analysis method, aiming at automatically estimating latent spatio-temporal patterns of brain changes from collections of brain MRIs over time. This approach extends standard methods (such as ICA) to encode priors on spatial and temporal properties of the signal measured in brain images. The application to currently available large-scale biomedical datasets is addressed by focusing on scalable and distributed learning methods. Our method considers the observed data as a matrix factorization of both temporal and spatial sources. The temporal sources are treated as independent Gaussian Processes to promote smoothness in time and model a plausible aging evolution. The spatial sources are modeled as Gaussian random fields to encode the spatial continuity of the brain sub-structures. This particular structure allows to factorize the spatial covariance matrix as a Kronecker product over the three spatial dimensions, which greatly simplifies computations and reduces the dimensions of the matrices we work on. The overall model is efficiently optimized through stochastic variational inference. We tested our model on synthetic and real data. We generated statistically independent temporal sources and spatial sources as smooth heatmaps. Then we trained our model so that it disentangles the observed data in two matrices that best fit the generated mixed observations. The figure below shows on plot A) the original temporal sources (in red) and their approximation by the independent Gaussian Processes (in blue). Likewise, plot B) shows on top the heatmaps manually generated and under them the maps generated by the algorithm. We can observe that the method is able to capture these raw sources from the noisy observations given as target.

Eventually this method may provide an ideal exploratory tool for analyzing large-scale medical imaging datasets such as the UK Biobank. Indeed it allows to efficiently scale to both high-dimensionnal data and large sample sizes, and also identifies hidden spatio-temporal processes in a completely unsupervised manner.