NeuroImage: Clinical 4 (2014) 718–729

Contents lists available at ScienceDirect

NeuroImage: Clinical

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

Longitudinal deformation models, spatial regularizations and learning strategies to quantify Alzheimer's disease progression

Jean-Baptiste Fiot [a,b,*,1], Hugo Raguet [b], Laurent Risser [d], Laurent D. Cohen [b], Jurgen Fripp [c], François-Xavier Vialard [b], for the Alzheimer's Disease Neuroimaging Initiative [2]

[a] IBM Research, Smarter Cities Technology Centre, Damastown, Dublin 15, Ireland
[b] CEREMADE, UMR 7534 CNRS, Université Paris Dauphine, PSL★, France
[c] CSIRO Preventative Health National Research Flagship ICTC, The Australian e-Health Research Centre — BioMedIA, Royal Brisbane and Women's Hospital, Herston, QLD, Australia
[d] CNRS, Institut de Mathématiques de Toulouse, UMR 5219, France

ARTICLE INFO

Article history:
Received 19 September 2013
Received in revised form 22 January 2014
Accepted 14 February 2014
Available online 1 April 2014

Keywords:
Alzheimer's disease
Brain imaging
Deformation model
LDDMM
Disease progression
Karcher mean
Transport
Logistic regression
Spatial regularization
Coefficient map

ABSTRACT

In the context of Alzheimer's disease, two challenging issues are (1) the characterization of local hippocampal shape changes specific to disease progression and (2) the identification of mild-cognitive impairment patients likely to convert. In the literature, (1) is usually solved first to detect areas potentially related to the disease. These areas are then considered as an input to solve (2). As an alternative to this sequential strategy, we investigate the use of a classification model using logistic regression to address both issues (1) and (2) simultaneously. The classification of the patients therefore does not require any a priori definition of the most representative hippocampal areas potentially related to the disease, as they are automatically detected. We first quantify deformations of patients' hippocampi between two time points using the *large deformations by diffeomorphisms* framework and transport these deformations to a common template. Since the deformations are expected to be spatially structured, we perform classification combining logistic loss and *spatial regularization* techniques, which have not been explored so far in this context, as far as we know. The main contribution of this paper is the comparison of regularization techniques enforcing the coefficient maps to be spatially smooth (Sobolev), piecewise constant (total variation) or sparse (fused LASSO) with standard regularization techniques which do not take into account the spatial structure (LASSO, ridge and ElasticNet). On a dataset of 103 patients out of ADNI, the techniques using spatial regularizations lead to the best classification rates. They also find coherent areas related to the disease progression.

1. Introduction

Large scale population studies aim to improve the understanding of the causes of diseases, define biomarkers for early diagnosis, and develop preventive treatments. An important challenge for medical imaging is to analyze the variability in MRI acquisitions of normal control (NC), mild cognitive impairment (MCI), and Alzheimer's disease (AD) patients. For Alzheimer's disease, several classification strategies have been proposed to separate patients according to their diagnosis. These methods can be split into three categories: voxel-based (Fan et al., 2007, 2008a,b; Klöppel et al., 2008; Lao et al., 2004; Magnin et al., 2009; Vemuri et al., 2008), cortical-thickness-based (Desikan et al., 2009; Klöppel et al., 2008; Querbes et al., 2009) and hippocampus-based (Chupin et al., 2007, 2009; Gerardin et al., 2009) methods. While decent classification rates can be achieved to separate AD from NC or NC from p-MCI (progressive MCI patients, i.e. converting to AD), all methods perform poorly at separating s-MCI (stable MCI patients, i.e. non-converting to AD) and p-MCI. A recent review comparing these methods can be found in Cuingnet et al. (2011).

In the case of longitudinal analysis, it is not anymore the shapes that are compared but their evolutions in time. To extract information between two successive time-points, we use a one-to-one deformation which maps the first image onto the second one. Different registration algorithms are available to compute plausible deformations in this context. However, only one, the *large deformations via diffeomorphisms (LDDMM)* (Beg et al., 2005), provides a Riemannian setting that enables to represent the deformations using tangent vectors: initial velocity

* Corresponding author at: IBM Research, Smarter Cities Technology Centre, Damastown, Dublin 15, Ireland.
E-mail address: jean-baptiste.fiot@centraliens.net (J.-B. Fiot).

[1] Now affiliated to IBM Research. The work and preparation of this article were mostly done while being affiliated to Université Paris Dauphine.

[2] Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (http://adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

http://dx.doi.org/10.1016/j.nicl.2014.02.002

fields or equivalently initial momenta. This can be used in practice to retrieve local information and to perform statistics on it as presented in Vaillant et al. (2004) and Wang et al. (2007). In this direction, it is worth mentioning the study of Singh et al. (2010) which shows the correlation between principal modes of deformation and diagnosis. In Qiu et al. (2008) the authors estimate the typical deformation of several clinical groups from the deformations between baseline and follow-up hippocampus surfaces. In order to compare this information across the population, we need to define a common coordinate system. This implies (1) the definition of a template and (2) a methodology for the transport of the tangent vector information. Note finally that, as far as the authors know, no paper explores binary classification using logistic regression in this context.

Quality of shape descriptors with regard to the disease is often evaluated through statistical significance tests or classification performance. In this paper, we evaluate descriptors on a binary classification task using logistic regression.

In addition to its simplicity, it has the advantage of providing a map of coefficients weighting the relevance of each voxel. Such map can be used to localize the hippocampus deformations that are related to AD. However, the dimensionality of the problem (i.e. number of voxels p) being much higher than the number of observations (i.e. number of patients n, $p \sim 10^6 \gg n \sim 10^2$), the problem requires proper regularization.

Now standard regularization methods such as *ridge* (Hoerl and Kennard, 1970), *LASSO* (Tibshirani, 1994) and *Elastic Net* (Zou and Hastie, 2005) do not take into account any spatial structure of the coefficients.

In contrast, spatial models for regularizing supervised learning methods have been proposed in the literature (Grosenick et al., 2013; Jenatton et al., 2012; Ng and Abugharbieh, 2011). *Total variation* was used to regularize a logistic regression on functional MRI (fMRI) data (Michel et al., 2011). This method promotes coefficient maps with spatially homogeneous clusters. *Fused LASSO* was also used on fMRI data (Baldassarre et al., 2012; Gramfort et al., 2013). Similar ideas can be found in Cuingnet et al. (2012) where the authors defined the notion of spatial proximity to regularize a linear SVM classifier.

In Durrleman et al. (2013), the authors introduce sparse parametrization of the diffeomorphisms in the LDDMM framework. Our goal is different: we want spatial properties (smoothness, sparsity, etc.) to be found *across* the population (i.e. on the common template) and we want this coherence to be driven by the disease progression.

In this paper, we investigate the use of total variation, Sobolev and fused LASSO regularizations in 3D volumes. Compared to total variation, Sobolev enforces smoothness of the coefficient map, whereas fused LASSO adds a sparsity constraint.

The deformation model used to assess longitudinal evolutions in the population is presented in Section 2. Machine learning strategies are discussed and the model of classification with logistic loss and spatial regularization is described in Section 3. The dataset used and numerical results are presented in Section 4. We illustrate that initial momenta capture information related to AD progression, and that spatial regularizations significantly increase classification performance. Section 5 concludes the paper.

2. Longitudinal deformation model for population analysis

2.1. Global pipeline

Let us assume that we have a population of patients and the binary segmentation of their hippocampus at two different time points, called *screening* and *follow-up*. Let us also assume that all patients have the same diagnosis at the screening time point, and only a part of them have converted to another diagnosis at the follow-up time point. Our goal is to compare patient evolutions, and classify them with regard to disease progression, i.e. stable diagnosis versus progressive diagnosis. From a machine learning point of view, we need to build features encoding the evolutions of the patients.

We use the pipeline summarized in Fig. 1. First, the evolution descriptors are computed locally for each patient (independently). To be able to compare these descriptors, one needs to transport them into a common space. To do so, a population template is computed, towards which all the local descriptors are transported. Finally, classification is performed to separate progressive from stable patients.

2.2. Diffeomorphic registration via geodesic shooting

As mentioned in Sections 1 and 2.1, local deformation descriptors are computed to model the evolutions of the patients. In this section, we describe how we use diffeomorphic registration via geodesic shooting Vialard et al. (2012a) to compute these local deformation descriptors.

2.2.1. Definitions

To register a source image $I : \Omega \subset \mathbb{R}^3 \to \mathbb{R}$ towards a target image $J : \Omega \subset \mathbb{R}^3 \to \mathbb{R}$, the LDDMM framework (Beg et al., 2005) introduces the following minimization problem

$$\underset{v \in L^2([0,1],\mathcal{H}_K)}{\operatorname{argmin}} \frac{1}{2}\|I \circ \phi_{0,1}^{-1} - J\|_{L^2}^2 + \lambda \int_0^1 \|v_t\|_K^2 dt, \tag{1}$$

where $v : (t,\omega) \in [0,1] \times \Omega \subset \mathbb{R}^3 \to \Omega$ is a time dependent velocity field that belongs to a reproducing kernel Hilbert space $\mathcal{H}_K$ of smooth enough vector fields defined on Ω, and of associated kernel K and norm $\|\ \|_K$, and $\lambda \geq 0$ is a regularization coefficient. For $(t,\omega) \in [0,1] \times \Omega$, we note $v_t(\omega) = v(t,\omega)$. The deformation $\phi : [0,1]^2 \times \Omega \subset \mathbb{R}^3 \to \Omega$ is given by the flow of v_t

$$\forall (t,\omega) \in [0,1] \times \Omega, \quad \begin{cases} \dfrac{\partial \phi_{0,t}}{\partial t}(\omega) = v_t \circ \phi_{0,t}(\omega) \\ \phi_{t,t}(\omega) = \omega\ , \end{cases} \tag{2}$$

where $\phi_{t1,\,t2}$ is the deformation from $t = t_1$ to $t = t_2$. Such approach induces a Riemannian metric on the orbit of I, i.e. the set of all deformed images by the registration algorithm (Miller et al., 2006). The first term in formula (1) is a similarity term controlling the matching quality whereas the second one is a smoothing term controlling the deformation regularity. Now noting $I_t \overset{\text{def.}}{=} I \circ \phi_{0,t}^{-1}$ and $J_t \overset{\text{def.}}{=} J \circ \phi_{t,1}$,

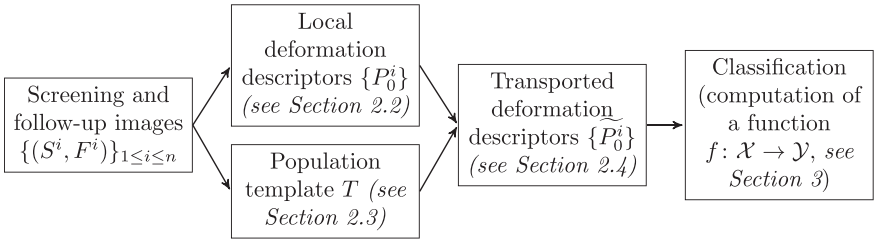

Fig. 1. Four steps are needed to classify patient evolutions using local descriptors of shape deformations: (1) the local descriptors are computed for each patient independently, (2) a population template is computed, (3) all local shape deformation descriptors are transported towards this template, and (4) classification is performed.