NeuroImage xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

NeuroImage

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

Individualized Gaussian process-based prediction and detection of local and global gray matter abnormalities in elderly subjects

G. Ziegler [a,b,*], G. Ridgway [a], R. Dahnke [b], C. Gaser [b,c], for The Alzheimer's Disease Neuroimaging Initiative [1]

[a] Wellcome Trust Center for Neuroimaging, Institute of Neurology, London, UK
[b] Department of Psychiatry, Jena University Hospital, Jena, Germany
[c] Department of Neurology, Jena University Hospital, Jena, Germany

ARTICLE INFO

Article history:
Accepted 4 April 2014
Available online xxxx

Keywords:
Brain morphology
Lifespan brain aging
Gaussian processes
Single case analysis
Bayesian inference

ABSTRACT

Structural imaging based on MRI is an integral component of the clinical assessment of patients with potential dementia. We here propose an individualized Gaussian process-based inference scheme for clinical decision support in healthy and pathological aging elderly subjects using MRI. The approach aims at quantitative and transparent support for clinicians who aim to detect structural abnormalities in patients at risk of Alzheimer's disease or other types of dementia. Firstly, we introduce a generative model incorporating our knowledge about normative decline of local and global gray matter volume across the brain in elderly. By supposing smooth structural trajectories the models account for the general course of age-related structural decline as well as late-life accelerated loss. Considering healthy subjects' demography and global brain parameters as informative about normal brain aging variability affords individualized predictions in single cases. Using Gaussian process models as a normative reference, we predict new subjects' brain scans and quantify the local gray matter abnormalities in terms of Normative Probability Maps (NPM) and global z-scores. By integrating the observed expectation error and the predictive uncertainty, the local maps and global scores exploit the advantages of Bayesian inference for clinical decisions and provide a valuable extension of diagnostic information about pathological aging. We validate the approach in simulated data and real MRI data. We train the GP framework using 1238 healthy subjects with ages 18–94 years, and predict in 415 independent test subjects diagnosed as healthy controls, Mild Cognitive Impairment and Alzheimer's disease.

Introduction

Magnetic Resonance Imaging (MRI) and computational morphometry have become invaluable tools for in-vivo exploration of the underlying changes in healthy and pathological brain aging (Fjell and Walhovd, 2010; Frisoni et al., 2010). Consistent findings show that regional gray matter volume, as well as cortical thickness, exhibit substantial decline as a process of healthy aging (Fjell and Walhovd, 2010; Raz and Rodrigue, 2006). Importantly, studies observed considerable variability of age-related structural trajectories across brain regions and healthy elderly individuals (Raz et al., 2005, 2010; Walhovd et al., 2011). An open question in clinical practice still is, how to efficiently identify local pathological brain aging in individuals at risk of developing Alzheimer's disease (AD) or other types of dementia. Due to the large individual differences of normative age-related decline, the visual assessment of healthy vs. pathological local atrophy is a challenging task even for experienced radiologists. While single case studies are long-standing practice in neuropsychology (for overview of methods see e.g. Crawford and Garthwaite, 2012; McIntosh and Brooks, 2011), there is also an increasing number of neuroimaging studies using Voxel-based Morphometry (VBM) (Ashburner and Friston, 2000; Mechelli et al., 2005) that focus on single cases in comparison to a reasonably sized group of control subjects. These studies explored voxelwise macroanatomy in patients with neurological disorders like aphasia, Huntington disease, lesions, focal cortical dysplasia, epilepsy, cortical atrophy, and dementia (Colliot et al., 2006; Maguire et al., 2010; Mehta et al., 2003; Migliaccio et al., 2012; Mühlau et al., 2009; Mummery et al., 2000; Salmond et al., 2003; Scarpazza et al., 2013; Seghier et al., 2008; Sehm et al., 2011; Woermann et al., 1999).

In order to provide statistical measures of suspicious local brain volumes (or cognitive test scores) in single case studies, several parametric techniques have been proposed. A simple approach is to calculate z-scores using the control sample mean and standard deviation. If the observed z-score is found to be less than a certain percentile of the standard normal distribution, e.g. $z < -1.645$ (corresponding to a

* Corresponding author at: Wellcome Trust Center for Neuroimaging, 12 Queen Square, WC1N 3BG London, UK.
E-mail address: g.ziegler@ucl.ac.uk (G. Ziegler).

[1] Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (www.loni.ucla.edu/ADNI). 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. Complete listing of ADNI investigators available at http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

http://dx.doi.org/10.1016/j.neuroimage.2014.04.018
1053-8119/

one-tailed 95% percentile), the deviation might be considered statistically significant. Unfortunately, the z-score approach lacks the ability to account for the uncertainty of the control sample statistics, which might inflate type I errors especially in small samples (Crawford and Howell, 1998). Thus, the more conventional parametric approach to single case inference is the two sample t-test using a 'pooled' estimate of the variance (for details see e.g. Mühlau et al., 2009). The t-test statistic in the special case with n controls and one single patient reduces to $t = \left(\mu_c - \mu_p\right) / \left(\sigma_c \sqrt{1/n+1}\right)$ with control sample standard deviation σ_c. Previous studies explored methodological issues using this type of unbalanced parametric design. In particular, small samples have been found to reduce sensitivity for detection of structural differences in single subjects (Mühlau et al., 2009). Unfortunately, for unbalanced designs the above difference score might be particularly affected by non-normality, rendering the t-test invalid (Salmond et al., 2002; Viviani et al., 2007). Robustness of the tests was found to be increased (type I errors reduced) by using larger smoothing kernels or appropriate transformations of the data. However, for inference in elderly subjects, the approaches often do not address the underlying developmental process, e.g. age-related effects in the control sample (see also Dukart et al., 2011), as well as variations due to other relevant covariates, e.g. global volume differences (Peelle et al., 2012).

Gaussian process (GP) models have emerged as a flexible and elegant approach for prediction of continuous, i.e. $y \in \mathbb{R}$, or binary, i.e. $y \in [0, 1]$ variables (Kim and Ghahramani, 2006; Rasmussen, 1996; Rasmussen and Williams, 2006). Recently,GPs were successfully introduced to the neuroimaging community. The potential applications range from spatial priors (Groves et al., 2009), cortical maps (Macke et al., 2011), image denoising (Zhu et al., 2012), parameter estimation (Wang et al., 2012), white matter fiber clustering (Wassermann et al., 2010) and meta-analysis (Salimi-Khorshidi et al., 2011). GP models were shown to be particularly powerful for clinical applications, providing probabilistic predictions of symptom severity, pain states, recovery, cognitive and disease states using regression (Doyle et al., 2013a; Hope et al., 2013; Marquand et al., 2010) and classification (Hahn et al., 2011; Marquand et al., 2010; Mourao-Miranda et al., 2012; Pyka et al., 2012; Young et al., 2013) using functional and structural MR images as inputs. In addition to the common application as decoding or recognition models, i.e. making inference about causes of functional and structural brain states based on images (Friston et al., 2008), GPs might be particularly useful for generative modeling of individual differences of brain morphometry (see also Ashburner and Klöppel, 2011; Friston and Ashburner, 2004).

Here we propose a new approach to support individualized clinical decisions about an elderly patient's brain structure by providing quantitative, unbiased and highly transparent maps of local gray matter abnormalities and global volume z-scores for gray matter, white matter and cerebrospinal fluid. That means, the maps and z-scores aim at information support rather than providing fixed patient-level predictions about disease states derived from 'black-box' classifiers. GPs are used to implement a normative generative model of elderly subjects' local and global volumes in terms of a non-parametric function of subjects' covariates. The model captures normative age-related trajectories and effects of covariates typically observed in control samples. This implicitly assumes smooth structural trajectories without imposing strong constraints on the developmental model and thus allows more flexibility than low degree polynomial expansions (for discussion of quadratic fits see e.g. Fjell et al., 2010). At the same time it accounts for region specific late life accelerated gray matter shrinkage, which is shown to be part of healthy brain aging (Fjell et al., 2012, 2013; Walhovd et al., 2011). The substantial individual differences of local and global volumes in elderly brains (i.e. even at the same age and fixed covariates) and the measurement noise are modeled in terms of Gaussian distributions and accounted for in individualized predictions. After model optimization in a large control sample, the local GP priors are conditioned on scans of new single subjects at risk of developing AD or other types of dementia. Training with a large pooled MRI database of 1238 healthy subjects with ages 18–94 years, and testing with an independent sample from the Alzheimer's Disease Neuroimaging Initiative dataset including subjects with MCI and AD, we show that the obtained normative probability maps (NPM) and global z-scores provide a powerful clinical application by quantitatively characterizing the single patient's abnormalities as compared to age-matched neurologically normal controls. This implements a Bayesian single case inference about structural abnormalities that flexibly accounts for predictive uncertainty in practical situations of different control data sample sizes, different data noise levels, and individual patient covariates, i.e. age, brain sizes, etc.

Methods

A Gaussian process model of cross-sectional gray matter observations in healthy elderly

Ideally, a generative model of the normative structural aging process accurately predicts the local gray matter volume y of an elderly study participant based on the age and a set of informative covariates $\mathbf{x} = [age, sex, \ldots]$,i.e. forming a low dimensional covariate space $\mathcal{D} \in \mathbb{R}^m$. The predictions require availability of most covariates for all cases in the training and test samples. Thereby, we here restrict our local generative model to six covariates summarized in $\mathbf{x}_i = [\text{age, sex, tgmv, twmv, tcsf, fstr}]$ for subject i, including demography and global parameters, i.e. total gray matter volume (tgmv), total white matter volume (twmv), and total cerebrospinal fluid (tcsf) obtained from MRI preprocessing. Furthermore, for inference about global atrophy an additional generative model for global brain parameters tgmv, twmv, and tcsf was applied using four covariates $\mathbf{x}_i = [\text{age, sex, ticv, fstr}]$ with ticv = tgmv + twmv + tcsf. Note, the proposed framework also naturally extends to physiological and behavioral factors, as well as subject independent but scan specific variables, e.g. the signal to noise ratio of the scan. In order to afford pooling across samples from 1.5 and 3 Tesla MRI scanners, we also included a field strength variable (fstr). The whole training sample covariate data is further denoted by $\mathbf{X}$, which was obtained from concatenation of rows $\mathbf{x}_i$ for all n training subjects. The rows of brain data matrix $\mathbf{Y}$ (with entries y_{ij}) refer to the GMV images of all n training subjects, and $\mathbf{y}_j$ is used to denote its j-th column, i.e. the data of voxel j from all subjects. Then the lifespan generative model of gray matter in voxel j follows

$$y_{ij} = g\left(\mathbf{x}_i, \boldsymbol{\theta}_j\right) + \epsilon_{ij}, \quad \epsilon_{ij} \sim \mathcal{N}\left(0, \sigma_j^2\right) \tag{1}$$

with subject index i and hyperparameter $\boldsymbol{\theta}_j$, an additive independent identically distributed Gaussian noise (also called the likelihood model) with variance σ_j^2. The latent (or noise free) variables $g(\mathbf{x}, \boldsymbol{\theta})$ incorporate our knowledge about aging and variability in different locations $\mathbf{x}$ of the covariate space $\mathcal{D}$. We now exploit the function space perspective and define a GP prior, which implements our assumption about smoothness of the latent trajectories $g(\mathbf{x}, \boldsymbol{\theta})$. Technically, a GP is a distribution of functions, which is fully specified by its mean and its covariance function (for a technical introduction see Rasmussen and Williams, 2006)

$$g \sim \mathcal{GP}(m, cov). \tag{2}$$

The following specification of the prior mean m and covariance function cov implies a distribution over latent structural trajectories and their individual differences in voxel j

$$m\left(g\left(\mathbf{x}_p, \boldsymbol{\theta}_j\right)\right) = 0 \tag{3}$$

Please cite this article as: Ziegler, G., et al., Individualized Gaussian process-based prediction and detection of local and global gray matter abnormalities in elderly subjects, NeuroImage (2014), http://dx.doi.org/10.1016/j.neuroimage.2014.04.018