NeuroImage 139 (2016) 470–479

Contents lists available at ScienceDirect

NeuroImage

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

Domain adaptation for Alzheimer's disease diagnostics

Christian Wachinger [a,b,c,*], Martin Reuter [b,c], for the Alzheimer's Disease Neuroimaging Initiative [1], the Australian Imaging Biomarkers and Lifestyle flagship study of ageing [2]

[a] Department of Child and Adolescent Psychiatry, Psychosomatic and Psychotherapy, Ludwig-Maximilian-University, Munich, Germany
[b] Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA
[c] Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA, USA

ARTICLE INFO

Article history:
Received 11 January 2016
Revised 11 May 2016
Accepted 22 May 2016
Available online 2 June 2016

Keywords:
Computer-aided diagnosis
Alzheimer's disease
Classification
Domain adaptation

ABSTRACT

With the increasing prevalence of Alzheimer's disease, research focuses on the early computer-aided diagnosis of dementia with the goal to understand the disease process, determine risk and preserving factors, and explore preventive therapies. By now, large amounts of data from multi-site studies have been made available for developing, training, and evaluating automated classifiers. Yet, their translation to the clinic remains challenging, in part due to their limited generalizability across different datasets. In this work, we describe a compact classification approach that mitigates overfitting by regularizing the multinomial regression with the mixed ℓ_1/ℓ_2 norm. We combine volume, thickness, and anatomical shape features from MRI scans to characterize neuroanatomy for the three-class classification of Alzheimer's disease, mild cognitive impairment and healthy controls. We demonstrate high classification accuracy via independent evaluation within the scope of the CADDementia challenge. We, furthermore, demonstrate that variations between source and target datasets can substantially influence classification accuracy. The main contribution of this work addresses this problem by proposing an approach for supervised domain adaptation based on instance weighting. Integration of this method into our classifier allows us to assess different strategies for domain adaptation. Our results demonstrate (i) that training on only the target training set yields better results than the naïve combination (union) of source and target training sets, and (ii) that domain adaptation with instance weighting yields the best classification results, especially if only a small training component of the target dataset is available. These insights imply that successful deployment of systems for computer-aided diagnostics to the clinic depends not only on accurate classifiers that avoid overfitting, but also on a dedicated domain adaptation strategy.

Introduction

Alzheimer's disease (AD) is the most common form of dementia with incidence rates further increasing in the future due to increasing life expectancy. Early and accurate diagnosis of AD is a key objective as it can help patients to access supportive therapies earlier allowing them to maintain independence for longer (Paquerault, 2012). When treatment options that directly interfere with disease pathways finally become available, intervention will likely be most effective in early preclinical or presymptomatic disease stages. Furthermore, early identification of high-risk individuals can already support selection into promising drug trials, inform patient stratification, as well as aid the identification of risk and preserving factors. Magnetic resonance imaging (MRI) is an important tool for AD diagnosis because the atrophy measured in MRI correlates with neuron loss and can indicate the onset of the impairment in close temporal proximity (Jack et al., 2013). Computer-aided diagnosis of dementia based on MRI is an active research field as indicated by 50 articles reviewed on this topic by Falahati et al. (2014). The deployment of automated system for diagnosis of AD in the clinic promises several advantages: (i) the improvement of diagnosis in places with limited neuroradiological know-how, (ii) a faster diagnosis without compromising accuracy by avoiding lengthy specialist investigations, and (iii) a more objective diagnostic assessment based increasingly on quantitative information in contrast to traditionally more subjective diagnostic impression (Klöppel et al., 2012). Computational diagnostics promise to

* Corresponding author at: Waltherstr. 23, 80337 München, Germany.
E-mail address: christian.wachinger@med.uni-muenchen.de (C. Wachinger).

[1] Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (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

[2] Data used in the preparation of this article was obtained from the Australian Imaging Biomarkers and Lifestyle flagship study of aging (AIBL) funded by the Commonwealth Scientific and Industrial Research Organisation (CSIRO) which was made available at the ADNI database (www.loni.usc.edu/ADNI). The AIBL researchers contributed data but did not participate in analysis or writing of this report. AIBL researchers are listed at www.aibl.csiro.au.

http://dx.doi.org/10.1016/j.neuroimage.2016.05.053
1053-8119/

be particularly useful for screening purposes to identify individuals with preclinical disease.

Large, multi-center datasets are available for studying Alzheimer's disease and for supporting the training of complex classification models. A challenge for such models is generalizability, i.e., the ability to transfer a model that is trained on one dataset to another dataset while retaining high prediction accuracy. In an attempt to provide an objective assessment of state-of-the-art methods for AD classification, the CADDementia challenge has been organized recently (Bron et al., 2015). The task was to differentiate between patients with Alzheimer's disease, mild cognitive impairment (MCI), and healthy controls (CN) based on T1-weighted MRI data. Classification accuracy of a variety of submissions was evaluated on an independent test dataset with hidden diagnosis. Intriguingly, the study showed that all participating groups overestimated the accuracy of their method. One of the main reasons for the overestimation may be overfitting to the training data. Neuroimaging applications are susceptible to overfitting due to a potentially large number of features extracted from images and a restricted number of samples available for training. Overfitting is further aggravated by complex classification models with many degrees-of-freedom that easily fine-tune to a specific population but overestimate the performance on the general population (Adaszewski et al., 2013; Mwangi et al., 2014). In our classifier we employ methods that mitigate overfitting by (i) using sparsity constraints to estimate a compact model and by (ii) choosing a linear classification model based on multinomial regression to further limit the number of free parameters. Yet, in spite of these efforts, the bias towards overestimating performance on the training set still prevails, indicating that overfitting may not solely be responsible. Here, we identify another cause for reduced classification accuracy on the final test set: the differences in the distribution between training and test data.

The main contributions of this work are twofold: We introduce a compact classifier for Alzheimer's disease that incorporates shape information and evaluate its performance on an independent test setting. We further demonstrate that variations in source and target datasets have a large impact on classification accuracy and present a novel algorithm for domain adaptation that re-weights samples from the source dataset.

Computer-aided diagnosis of dementia

Predicting or classifying dementia based on structural MRI is an active field of research. Cuingnet et al. (2011) compare several approaches for the discrimination of AD and MCI patients using the cortical thickness, the hippocampus and voxel-based methods. Falahati et al. (2014) review the literature for the classification of individuals with dementia. The extensive list of articles discussed in the review illustrates the wide interest in the research field. In this work, we introduce an algorithm for AD classification that is based on *BrainPrint* (Wachinger et al., 2015) for quantifying brain morphology, which naturally extends the region of interest (ROI)-based volume and thickness analysis with shape information (Reuter et al., 2006). Anatomical shape features contribute valuable information to the characterization of brain structures, which are only coarsely represented by their volume. Finding representative and descriptive features is crucial for automatic classification as it is well known in pattern recognition that the prediction accuracy is primarily driven by the representation (Dickinson, 2009).

Both of the review articles mentioned above refer to a total of only three publications that employ shape information, indicating that shape is not commonly used. Most previous work that includes shape analysis, typically focus on a single structure, predominantly the hippocampus. More precisely, Gerardin et al. (2009) approximate the hippocampal shape by a series of spherical harmonics. Ferrarini et al. (2009) use permutation tests to extract surface locations that are significantly different among patients with AD and controls. Costafreda et al. (2011) incorporate shape information by deriving thickness measurements of the hippocampus from a medical representation. Shen et al. (2012) use statistical shape models to detect hippocampal shape changes. Bates et al. (2011) investigated spectral signatures for AD classification, with a focus on right hippocampus, right thalamus and right putamen. Other structures of interest for shape analysis were the cortex and ventricles: Kim et al. (2014) use multi-resolution shape features with non-Euclidean wavelets for the analysis of cortical thickness, King et al. (2010) analyze the fractal dimension of the cortical ribbon, and Gutman et al. (2013) model surface changes of the ventricles in a longitudinal setup with a medial representation. In contrast to all these studies, we incorporate an ensemble of both cortical and subcortical structures. This extensive characterization of brain anatomy is promising in diagnosing Alzheimer's disease, which is associated with wide-spread atrophy across the entire brain.

Domain adaptation

As described above, differences between source and target datasets can significantly reduce classification accuracy. In traditional cross-validation, where a single dataset is split into subsets, such variations are negligible, as the subsets tend to represent the data well. However, when an independent dataset is used for testing, differences in the distributions can have a dramatic impact on the classification accuracy. Such problems are studied in domain adaptation (Pan and Yang, 2010), where the model is learned on a source dataset and then transferred to a target dataset with different properties. In fact, we believe that domain adaptation is crucial for the translation of computer-aided diagnostic methods to the clinic, where the source dataset usually consists of large, possibly multi-center, data and the target dataset is the (limited) data acquired at the specific hospital, where the system is deployed. There are clearly several factors that can contribute to variations between source and target datasets arising from location and selection biases.

Here, we assume a supervised domain adaptation scenario, where a subset of the target dataset is available for training, replicating the situation that a small, local dataset from the clinic is available to support training. Based on this small target training set we weight samples from the source dataset to match distributional properties of the target dataset. The proposed instance weighting presents a general framework, where naïve strategies for combining source and target training data (e.g. the union or selecting one vs. the other) can be derived by setting the weights to appropriate constants. We measure a variation in classification accuracy of more than 20% across strategies, highlighting the importance of domain adaptation. Domain adaptation with instance weighting has previously been described in the machine learning literature (Bickel et al., 2007; Jiang and Zhai, 2007). An unsupervised domain adaptation strategy for AD classification was used by Moradi et al. (2014). This strategy applies discriminative clustering on the source and target domain, where a feature weighting is learned by optimizing the mutual information (Shi and Sha, 2012). In contrast, we use a supervised domain adaptation strategy and do not weight features but instances. Further related are approaches that assume a semi-supervised classification setting (Zhao et al., 2014; Adeli-Mosabbeb et al., 2015), yet they operate on the same domain.

Domain adaptation has previously been successfully used in medical image analysis. van Opbroek et al. (2015) proposed transfer learning for image segmentation across scanners and image protocols with support vector machines and AdaBoost. Heimann et al. (2013) used domain adaptation for the localization of ultrasound transducers in X-ray images with probabilistic boosting trees. Schlegl et al. (2014) applied domain adaptation for lung tissue classification with convolutional neural networks.

Methods

In this section, we introduce our approach to AD classification with domain adaptation. The classification task is to predict the diagnostic