Proposition de recherche doctorale

Unsupervised learning from neuroimaging data to identify disease subtypes in Alzheimer’s disease and related disorders

Mots clés : Array

- **Directeur de thèse** : Didier DORMONT
- **Co-encadrant(s)** :
- **Unité de recherche** : Centre de recherche de l'Institut du cerveau et de la moelle épinière
- **Ecole doctorale** : École Doctorale Informatique, Télécommunications, Électronique de Paris
- **Domaine scientifique principal**: Sciences et technologies de l'information et de la communication

Résumé du projet de recherche (Langue 1)

Neurodegenerative diseases are a major public health concern for our societies. In particular, neurodegenerative dementias, such as Alzheimer’s disease (AD), affect over 20 million people world-wide and this number is expected to reach 80 million by 2040. The development of new treatments is hampered by the heterogeneity of these diseases. For example, an Alzheimer’s pathology may correspond to a typical profile of prominent memory loss and medial temporal lobe alterations, but also to more atypical posterior, language or visual presentations. While the best characterized and most typical phenotypes have been known by clinicians for a long time, the heterogeneity of diseases remains inadequately characterized. In particular, it is unclear if additional clusters exist beyond those that can be identified clinically and to which extend the putative clusters overlap. In the past decade, large datasets of patients explored with multiple modalities (clinical data, MRI, PET, genetics…) have been gathered. This opens the possibility to study the heterogeneity of neurodegenerative diseases in a data-driven manner, by learning the main components and clusters. To that purpose, it is necessary to design new unsupervised learning approaches.

The objective of this PhD thesis is to develop and evaluate clinically-relevant approaches for unsupervised learning to characterize disease heterogeneity in AD and related dementias. Specific objectives include: 1) To adequately account for normal variability. For instance, in a clustering approach, the aim would be to cluster the deviations from normal variability, rather than the raw characteristics of the patients. 2) To design approaches that can handle the structure and high-dimensionality of data of neuroimaging data. 3) To define clinically-relevant measures to assess the results of the unsupervised learning. At the beginning of the PhD, we propose to use predefined input features composed of regional measures of gray matter atrophy and white matter lesions. This will have the advantage to keep the dimensionality to a tractable level. However, regional measures are not optimal because their boundaries may not match those of the alterations. Therefore, in a second part, we will aim to use image data as input. To that purpose, we will investigate deep learning approaches that have been particularly successful for learning features from image data. We will aim at designing deep learning approaches that are adapted to the specificities of brain imaging data.

To design and evaluate the approaches, we will make use of the ADNI (Alzheimer’s Disease Neuroimaging Initiative), a large, publicly-available, database with over 1500 participants, as well as of smaller datasets of patients with either atypical AD phenotypes or other types of dementia from the ICM and the Pitié-Salpêtrière hospital. The team has extensive experience with using these datasets and preprocessing and feature extraction pipelines have already been implemented.

Informations complémentaires (Langue 1)

Environment of the PhD thesis
The PhD thesis will be conducted within the ARAMIS team at the Brain and Spine Institute. ARAMIS is a joint team between CNRS, INRIA, Inserm and University Pierre et Marie Curie. The ICM is a recently created neuroscience research center within Pitié-Salpêtrière hospital in Paris. It gathers over 700 researchers covering the full spectrum of neuroscience. ARAMIS is the methodological research team of the ICM. It is a multidisciplinary research team gathering computer scientists and medical doctors. The team develops cutting-edge machine learning and image analysis approaches for multimodal medical data (neuroimaging, clinical, genetic data), in order to create new tools for diagnosis, prognosis and monitoring of brain disorders. The team has close collaborations with several clinical teams of the ICM and the Pitié Salpêtrière hospital to apply these methods to the study of neurodegenerative diseases including Alzheimer's disease, fronto-temporal dementia and Parkinson's disease. The team has a very strong network of international collaborations and in particular currently participates or coordinates two large-scale European projects funded under Horizon 2020 and two US-French grants co-funded by NIH (USA), NSF (USA) and ANR (France).

Profile of the candidate
The candidate should have a strong background in engineering (image and signal processing, machine learning, biomedical engineering). He/she should have a strong interest for multidisciplinary collaborations in the medical field. Previous experience with brain imaging data would be a plus but is not mandatory.