STRATIFICATION OF ALZHEIMER DISEASE’S PATIENTS BY AUTOMATED DETECTION OF PEPTIDE ACCUMULATION IN WHOLE SLIDE IMAGES USING DEEP LEARNING

R&D MASTER / ENGINEERING INTERNSHIP
PHD OPPORTUNITY IN CASE OF A SUCCESSFUL INTERNSHIP

Contacts: Lev Stimmer – lev.stimmer@icm-institute.org
Benoit Delatour - benoit.delatour@upmc.fr
Stanley Durrleman - stanley.durrleman@icm-institute.org
Daniel Racoceanu - daniel.racoceanu@sorbonne-universite.fr

Application deadline: January 15th, 2019
Beginning of the internship: February, 2020 (according to the university regulation)
Duration: 6 months (longer duration possible condition to an agreement with the University/School)
Financial support: regular internship gratification
Locations of the internship: Institut du Cerveau et de la Moelle Épinière, ARAMIS team, Hôpital Pitié Salpêtrière, 47 Bd. Hôpital, 75013 Paris

Keywords: Alzheimer Disease, TAU and Aß positive markers, Computational Pathology, Whole Slide Images, BioMedical Image Analysis, Neuropathology, Deep Learning (DL), U-net, Convolutional Neural Networks (CNN), Generative Adversarial Networks (GAN), Python, PyTorch, TensorFlow.

Context of the internship:
Alzheimer’s disease (AD), the most frequent neurodegenerative disease, is defined by the misfolding and accumulation of Aß peptides and of tau proteins in the brain (see Figure). Clinically, sporadic Alzheimer’s disease (AD) most commonly presents in later life as an amnestic syndrome. However, the clinical presentation of the patients is more heterogeneous and different subtypes or clusters of brain lesions have been described. In particular, the rapidly progressive subtype of AD (rpAD) is frequently misdiagnosed as Creutzfeldt-Jakob disease. The team “Alzheimer’s and prions diseases” at the ICM has contributed to describe specific traits associated with rpAD not observed in standard AD cases with slower progression.

The overarching aim of this project is to understand to which extent the topography and morphology of the different peptide aggregates present in the brain can predict the diversity of symptoms observed in the patients. To this purpose, our pathologists have acquired and extensively annotated a unique set of histological images of postmortem brains from the very rare form of rapidly progressive AD patients (rpAD) and most common forms of AD. The counting and description of the aggregates in the images is done manually or semi-automatically using black-box proprietary software. There is therefore a urgent need to develop precise and robust tools to locate, annotate and characterize the different types of aggregates in histological images.

To address this need, the internship will take place at the interface of three teams with complementary expertise:
- The team Alzheimer’s and prion diseases (Dr. Benoît Delatour and Dr. Lev Stimmer), who is a world-leading team for the study of the molecular and cellular mechanisms responsible for prion and prion-like diseases like AD. Together with the cellular imaging core facility of the ICM,
they have developed a unique expertise in the acquisition and analysis of histologic images. Images are acquired and analysed within the ICM.

- The Brain Development team (Prof. Daniel Racoceanu) who has a long-run expertise in digital/computational pathology, being at the origin of the first digital pathology challenges (MITOS @ ICPR 2012) and involved in the European Society for Digital and Integrative Pathology (ESDIP : https://digitalpathologysociety.org/) supporting from 15 years the European Congress on Digital Pathology. Prof. Racoceanu (http://daniraco.free.fr) in particular develops image analysis and pattern recognition methods with application in computational pathology.

- The ARAMIS lab (Dr. Stanley Durrleman), a joint team between ICM and Inria, develops statistical and computational methods for the analysis of neuroimaging data and the construction of digital models of brain diseases.

A. The 2 main tau lesions in AD brains (tangles and neuritic plaques)
B. Quality control of tangles and neuritic plaques detection. The majority of lesions were correctly detected (80-90% of true positives). Still some objects were incorrectly classified (false positive and false negative detections).

Project – Operational automatic quantification system for TAU and Aß objects in Whole Slide Images for Alzheimer Disease’s Patients’ stratification

The project aims at developing a fully automated method for the location and characterization of the tau and Aß aggregates in histological images of the brain. The topography and morphology of the aggregates is heterogeneous with Aß accumulation taking the form of focal deposits or diffuse plaques; tau lesions forming the so-called neurofibrillary tangles but also with different morphologies in dendrites or axons (see Figure).
The candidate will design a deep learning method to achieve similar performances than the ones obtained with proprietary software before human corrections. The candidate will use as training data a set of 6 very-high resolution histological images (average size: 1,000,000x60,000 pixels, ndpi format). The images have been fully annotated by neuropathologist with a degree of confidence, yielding a set of ≈15,000 annotated aggregates. An additional set of 15 images may be used to test unsupervised learning approaches. If need be, additional annotation might be performed.

To go further, the candidate will improve the method to have higher performance and ease its usability in biological research. There are several directions that the candidate could follow, such as (i) provide a level of uncertainty in the label that matches the degree of uncertainty of a histologic expert, (ii) account for the neighboring structures in the classification which may include competitive segmentation approaches, (iii) provide shape and texture characteristics as output, (iv) explore semi- or non-supervised approaches for the clustering of the different types of aggregates, which may include auto-encoding approaches possibly in combination with adversarial approaches.

These research directions may be investigated in more depth in a follow-up PhD thesis.

The candidate will work in close interaction with the histologists and pathologists to understand their needs and the precise type of information that needs to be provided. The method developed will be tested and evaluated by our experts in the cellular imaging core facility. Depending on the results, the software may be deployed routinely in the core facility and proposed as a solution to similar research organization worldwide.

**Competencies (selection) required to reinforce our R&D projects:**
- Image Classification, Pattern recognition, Machine and Deep Learning (CNN, U-net, GAN, MIL)
- Rapid prototyping of ergonomic, modern software interfaces;
- Biomedical image analysis, Statistical analysis and validation of the results.

**Applicant profile:**
- University Master or Engineering School student (last year of study) with computer science, image analysis and/or applied mathematics profile;
- Interest for multidisciplinary projects, curiosity, learning capability and creativity are qualities.
- Interest for neuroscience research and brain diseases
  We do appreciate;
- Positive spirit, communication skills and ability to work in a team, if necessary;
- Autonomy, dynamism and motivation to advance his/her own part of the project;
- Excellent methodological and hands-on computer programming skills.
- Facility of understanding and manipulating mathematical models in a biological context.

**Expected deliverables:**
- Development of software components (data collection, expertise formalization/modelling, study of the state of the art and technology intelligence, design, test, validation);
- Proof of concept in interaction with the partners, according to the progress of the project;
- Possible publications and patents, with the prior consent of ICM and partners;
- Internship report (including methods used, results and perspectives) according to university guidelines;
- Consistent and effective user manual of the software/code developed.

Remarks:
- A careful assessment of general, methodological and programming skills will be carried out by e-conference or face-to-face (depending on availability);
- Regular (weekly) meetings will be organized, with synthetic presentations of the last advances, problems encountered, potential and proposed solution(s) as the necessary support;
- Periodic (monthly) meetings will be organized by involving a larger group of partners, according to the projects’ advances and perspectives;
- In case of satisfactory results, possibility of a Sorbonne University PhD recruitment within ICM. This type of pathway will be favoured for future recruitment of our permanent scientific collaborators.

Concerning ICM:
The project will be carried out at the ICM (Brain and Spine Institute, Paris). The research in this centre is devoted to the study and treatment of neurological disorders, with a strategy, pluridisciplinary by nature, integrating cellular biology, neurophysiology, neuropathology, behavioural analysis, neuroimaging, mathematical modelling, and molecular-genetic approaches. This centre is equipped with cutting edge technology and scientific expertise required for the completion of the project.
The Master internship will involve different ICM research teams with an expertise in the field of Alzheimer’s disease neuropathology (Lev Stimmer, Benoît Delatour) and IA-based image analysis (Daniel Racoceanu, Stanley Durrleman).