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

R&D MASTER INTERNSHIP

Contacts: Gabriel Jimenez – gabriel.jimenez@icm-institute.org
Anuradha Kar - anuradha.kar@icm-institute.org
Daniel Racoceanu - daniel.racoceanu@sorbonne-universite.fr

Application deadline: **December 10th, 2021**
Beginning of the internship: February/March, 2022 (according to the university regulation)
Duration: **6 months** (longer duration possible condition to an agreement with the University/School)
Financial support: **regular internship gratification / StratifIAD project (BBT3-ICM)**
Locations of the internship: **Paris Brain Institute** (**Institut du Cerveau-ICM), 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:

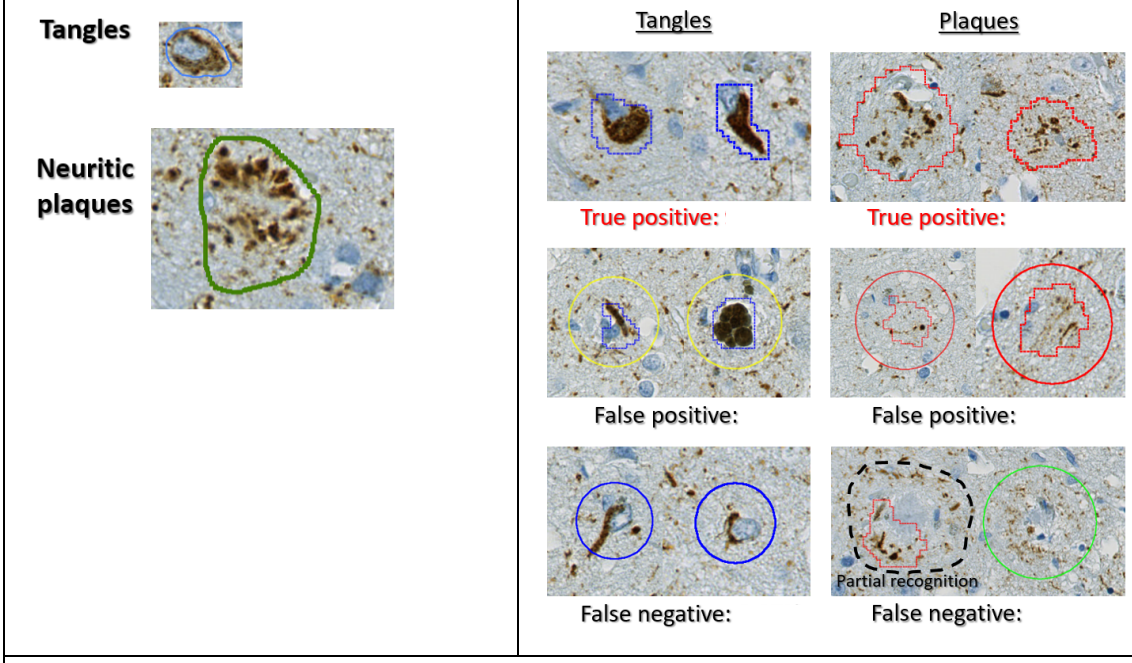
Alzheimer's disease (AD), the most frequent neurodegenerative disease, is defined by the misfolding and accumulation of Aß peptides and of tau proteins in the brain (see Figure). Clinically, sporadic Alzheimer's disease (AD) most commonly presents in later life as an amnestic syndrome. However, the clinical 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 two teams with complementary expertise:

- 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 in particular develops image analysis and pattern recognition methods with application in computational pathology.

- The team Alzheimer's and prion diseases (Dr. Benoît Delatour and Dr. Lev Stimmer), 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 analyzed within the ICM.



Automatic segmentation of Tau aggregates using a commercial software (based on a deep learning approach with a U-Net architecture).

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. Still some objects were incorrectly classified.

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 work with a large dataset of whole slide 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 ≈30,000 annotated aggregates. All WSI were annotated with neuritic plaques.

The tasks planned for the internship include the following:

- Benchmarking three deep learning architectures for semantic/instance segmentation. It will incorporate reinforcement learning, unsupervised and supervised deep learning (there is an existing supervised baseline pipeline for detection and segmentation using a Unet architecture).
- The outputs of the tested architectures should be compared using a common set of metrics (F1 score, semantic segmentation metrics – Dice score, Jaccard index). Also, the outputs should indicate label uncertainty of the segmentations.
- Explaining the decisions of one or more of the benchmarked architectures. One first idea to test will be the attention maps (especially for the Unet).
- Estimate morphological features from segmented objects in the WSI and study differences in its statistics across different slides. The estimation will be done after selecting the best model from the benchmarking task.
- Study the correlation between the patient's characteristics and the morphological statistics from the WSI, to deeply understand the evolution of the patients' disease as the patients' stratification.

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.
- Programming language: Python. Libraries: PyTorch, Open CV, CUDA;
- 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;

Concerning ICM:

The project will be carried out at the ICM (Paris Brain Institute). The research in this center is devoted to the study and treatment of neurological disorders, with a strategy, pluridisciplinary by nature, integrating cellular biology, neurophysiology, neuropathology, behavioral analysis, neuroimaging, mathematical modelling, and molecular-genetic approaches. This center 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's group @ Aramis lab).