

PhD thesis position

Artificial intelligence tools for clinical data warehouses in neuroimaging

Keywords: Medical imaging, deep learning, image analysis, quality control, clinical data warehouse, anomaly detection

The project

In recent years, very large clinical data warehouses (CDW) have been created containing the medical data of millions of patients. The AP-HP (Assistance Publique-Hôpitaux de Paris) CDW brings together data from multiple hospitals in the Paris region, including clinical data, diagnoses, medical reports and medical imaging data. CDWs provide fantastic opportunities to revolutionize digital healthcare. However, harnessing CDWs for research raises major challenges, among which controlling for data quality, biases and dealing with the full range of possible disorders and medical conditions.

Our team is a pioneer on the topic of neuroimaging in CDWs. We have built the first automatic quality control system for T1-weighted brain MRI of CDWs [1,2], that we subsequently extended to FLAIR MRI [3]. We have demonstrated that AI models trained on research data failed to generalize to clinical routine data and that, when data quality is not adequately taken into account, this leads to the catastrophic phenomenon of “short-cut” learning where the AI model learns to recognize image quality in place of the radiological features of the disorders [4].

The general objective of this PhD thesis project is to develop AI-based tools to harness the full potential of neuroimaging data in CDWs and to demonstrate that they can be used to develop trustworthy and unbiased AI-assisted reading systems for neuroradiology. Specific objectives are:

- **Objective 1: build AI-based quality control (QC) tools for different types of MRI sequences**, including T2-weighted, T2*-weighted, susceptibility weighted imaging and diffusion weighted imaging. To that purpose, we will develop new approaches based on transfer learning and synthetic data generation.
- **Objective 2: build a proof-of-concept unbiased AI-assisted reading system for neuroradiology.** To that purpose, we will develop unsupervised anomaly detection to assist the detection of lesional areas, leveraging previous work from the team [5,6] while exploring new approaches such as diffusion models.

The work will also include data management and preparation tasks, installation of code and dependencies in specific environments, and performance benchmarking. The methodological developments will be integrated into ClinicaDL (<https://clinicadl.readthedocs.io>), an open-source software platform designed to enable reproducible neuroimaging processing with deep learning.

This project is expected to have a major impact on several aspects. It will allow researchers to fully exploit the very rich but complex neuroimaging data in CDWs, potentially leading to major new discoveries in various fields such as a better understanding of the factors influencing healthy and pathological brain aging. The project will result in next generation AI-based diagnostic tools that are expected to generalize well and thus have a high potential for translation to the clinic.

- [1] Simona Bottani, Ninon Burgos, Aurélien Maire, Adam Wild, Sébastien Ströer, Didier Dormont, Olivier Colliot. Automatic quality control of brain T1-weighted magnetic resonance images for a clinical data warehouse. Medical Image Analysis, 2021. <https://hal.inria.fr/hal-03154792v3>
- [2] Sophie Loizillon, Simona Bottani, Aurélien Maire, Sebastian Ströer, Didier Dormont, Olivier Colliot, Ninon Burgos, APPRIMAGE Study Group. "Automatic motion artefact detection in brain T1-weighted magnetic resonance images from a clinical data warehouse using synthetic data". Medical Image Analysis, 2024. <https://hal.inria.fr/hal-03910451v2>
- [3] Sophie Loizillon, Olivier Colliot, Lydia Chougar, Sebastian Stroer, Yannick Jacob, Aurélien Maire, Didier Dormont and Ninon Burgos. Semi-supervised Domain Adaptation for Automatic Quality Control of FLAIR MRIs in a Clinical Data Warehouse. In DART 2023 - 5th MICCAI Workshop on Domain Adaptation and Representation Transfer, 2023 <https://inria.hal.science/hal-04273997>
- [4] Simona Bottani, Ninon Burgos, Aurélien Maire, Dario Saracino, Sebastian Stroer, Didier Dormont, Olivier Colliot. Evaluation of MRI-based machine learning approaches for computer-aided diagnosis of dementia in a clinical data warehouse. Medical Image Analysis, 2023. <https://inria.hal.science/hal-03656136v2>
- [5] Sophie Loizillon, Yannick Jacob, Aurélien Maire, Didier Dormont, Olivier Colliot, Burgos Ninon "Detecting Brain Anomalies in Clinical Routine with the beta-VAE: Feasibility Study on Age-Related White Matter Hyperintensities." 2024. <https://openreview.net/forum?id=YFfOvLf2T1>
- [6] Ravi Hassanaly, Maëlys Solal, Olivier Colliot, and Ninon Burgos. Pseudo-healthy image reconstruction with variational autoencoders for anomaly detection: A benchmark on 3D brain FDG PET, 2024. <https://inria.hal.science/hal-04445378>

A vibrant scientific, technological and clinical environment

You will work within the ARAMIS Lab (www.aramislab.fr) at the Paris Brain Institute (<https://institutducleveau-icm.org/en>), one of the world top research institutes for neurosciences. The institute is ideally located at the heart of the Pitié-Salpêtrière hospital, downtown Paris. The ARAMIS Lab, which is also part of Inria (the French National Institute for Research in Digital Science and Technology), is dedicated to the development of new computational approaches for the analysis of large neuroimaging and clinical data sets. With about 40 people, the lab has a multidisciplinary composition, bringing together researchers in machine learning and statistics and medical doctors (neurologists, neuroradiologists). You will interact locally with the PhD students and engineers of the lab, as well as our medical collaborators at the Pitié-Salpêtrière hospital.

The PhD thesis will be co-directed by Ninon Burgos (Research Scientist, HDR) and Olivier Colliot (Research Director). The position is funded through the GALAN project, a large-scale national grant in collaboration between the ARAMIS Lab, the Lille Neurosciences and Cognition Research Team, the departments of neuroradiology of the Pitié-Salpêtrière hospital and of the CHU of Lille, and the teams in charge of the CDWs of AP-HP and CHU of Lille. You will be involved in these collaborations and interact with the different partners.

Your profile

- Master's degree or engineering degree with computer science, signal/image analysis and/or applied mathematics profile
- Strong interest for medical applications
- Knowledge of deep learning
- Knowledge in digital image processing and medical imaging
- Good programming skills in Python
- Good writing skills
- Good relational and communication skills to interact with professionals from various backgrounds
- Proficient in French language

Ready to take up the challenge?

Send your CV to Ninon Burgos (ninon.burgos@cnrs.fr), Olivier Colliot (olivier.colliot@cnrs.fr) and Sophie Loizillon (sophie.loizillon@icm-institute.org), explaining your motivations.