

M2 Internship + PhD opportunity - Deep Learning / Neuroimaging

From brain injury to cognitive outcome after stroke

Lesion-disconnectome modelling for individualized prediction

H. Maaref, V. Vigneron

IBISC EA 4526, Université d'Évry, Université Paris-Saclay

Partners: IBISC, Centre Hospitalier Sud-Francilien (CHSF)

Duration: 5-6 months - Indicative period: February-August 2027 - Gratification: legal rate in force

Subject

Develop an interpretable machine-learning framework linking post-stroke brain injury to cognitive function. The project will combine focal lesion information from MRI, lesion-induced structural disconnection (the *disconnectome*), and available clinical variables to predict global and domain-specific cognitive outcomes at follow-up.

Position of the problem

Cognitive sequelae after stroke cannot be explained by lesion volume alone. The functional impact of an infarct depends on its location, the cortical and subcortical regions it damages, and the distributed networks it disconnects. Network-level abnormalities have been associated with deficits in memory, attention, language and executive functions, and structural disconnectome maps provide a principled way to quantify the remote consequences of a focal lesion [1, 2, 3]. Recent lesion-mapping studies also show that distinct cognitive domains are associated with partially distinct patterns of focal damage and network disconnection [4].

The central hypothesis of the internship is therefore that a patient-specific representation

$$\hat{\mathbf{y}} = f_{\theta}(\mathbf{z}_{\text{lesion}}, \mathbf{z}_{\text{disc}}, \mathbf{c})$$

where $\mathbf{z}_{\text{lesion}}$ describes the infarct, $\mathbf{z}_{\text{disc}}$ the induced disconnection and $\mathbf{c}$ the clinical covariates, will predict cognitive outcome more accurately and more informatively than models based only on lesion volume or conventional clinical scores.

Internship objectives

- Structural disconnectome estimation. Project each lesion onto a normative structural connectome or tract atlas to estimate the probability of interruption of major white-matter pathways and network edges. Compare voxel-wise disconnectome maps, tract-level lesion load and graph-based descriptors.
- Prediction of cognitive outcome. The prediction target will be the cognitive status measured at follow-up, using the neuropsychological variables available in the cohort. The primary formulation will be regression of continuous scores: a global cognitive score and, when available, domain-specific scores for memory, executive function, attention, language and visuospatial abilities. With several endpoints, the model will predict a vector

$$\hat{\mathbf{y}} = (\hat{y}_{\text{global}}, \hat{y}_{\text{memory}}, \hat{y}_{\text{executive}}, \hat{y}_{\text{attention}}, \dots),$$

using either separate models or multi-task learning with shared representations. Inputs will come from acute/subacute lesion descriptors, disconnectome features and clinical covariates, so that the task is genuinely prognostic: predict later cognitive function from information available earlier after stroke. Interpretable baselines (regularized regression, PLS, SVR and tree ensembles) will be compared with multimodal neural models. Performance will be quantified with MAE/RMSE, R^2 and score correlation, using patient-wise nested cross-validation and bootstrap confidence intervals. If validated clinical cut-offs are available, a secondary analysis will classify patients as cognitively impaired versus non-impaired and assess discrimination and calibration.

- Explainability and uncertainty. Identify which lesion locations, tracts and network components drive each prediction. Estimate prediction confidence using bootstrap, ensembles or calibrated predictive distributions, with particular attention to patients outside the training distribution.
- Controlled comparison. Quantify the incremental value of each information level: (i) clinical variables, (ii) lesion volume, (iii) lesion topography, (iv) structural disconnectome, and (v) multimodal fusion. Evaluation will use patient-level nested cross-validation, confidence intervals and clinically interpretable error measures.

Application and expected impact

The work addresses a clinically relevant question: why do patients with apparently comparable focal injuries show very different cognitive trajectories? A model that separates the effects of local tissue loss from network disconnection could support individualized prognosis, identify patients at risk of persistent cognitive impairment, and help orient follow-up and rehabilitation. The project is also intended to provide mechanistic insight rather than only a black-box prediction.

Data, methodology and expected deliverables

- Data. Retrospective multimodal MRI, lesion annotations and clinical/cognitive information from stroke cohorts available to IBISC and its clinical partners. Cognitive endpoints will be defined from the available neuropsychological assessments (global score and, when possible, memory, executive, attentional, language and visuospatial domains).
- Pipeline and lesion representation. As a methodological prerequisite, acute/subacute MRI and lesion masks will be registered to a common anatomical space. Patient-specific lesion descriptors will then be derived, including volume, laterality, anatomical load and cortical/subcortical involvement; DWI/ADC/FLAIR and other available sequences will be used according to the cohort. When available, markers of pre-existing brain vulnerability such as white-matter hyperintensities or atrophy may also be incorporated. The computational pipeline will combine Python/PyTorch, neuroimaging registration and parcellation, atlas-based tract/disconnectome analysis, statistical learning and deep learning, from MRI and lesion masks to cognitive prediction.
- Validation strategy. Avoid information leakage through strict patient-wise splitting. Compare models using repeated or nested cross-validation and bootstrap confidence intervals. If several acquisition sites are available, study robustness to site/domain shift.
- Deliverables. (i) a documented preprocessing and disconnectome pipeline; (ii) a benchmark of lesion-only versus disconnection-aware models; (iii) interpretable patient-level maps linking injury to cognitive outcome; (iv) a short research report and, depending on results, preparation of a manuscript for a neuroimaging / medical-AI journal or conference.

Indicative work plan

- Months 1-2: literature review, cohort description, quality control, registration and lesion descriptors.
- Months 2-3: construction of tract-level and voxel-wise disconnectome representations; exploratory lesion-cognition analyses.
- Months 3-5: predictive modelling, multimodal fusion, ablation studies and uncertainty/interpretability analysis.
- Months 5-6: final validation, figures, reproducible code and preparation of the research report / manuscript.

Candidate profile

We are looking for a strongly motivated M2 student in applied mathematics, statistics, computer science, signal/image processing, physics or engineering. A solid background in machine learning, probability/statistics and linear algebra is expected, together with good scientific programming skills, preferably in Python. Experience with deep learning, medical imaging, graph methods or neuroimaging is appreciated but not mandatory. The candidate should be interested in combining methodological work with a concrete clinical question.

What we offer

- A research problem at the interface of machine learning, neuroimaging, network neuroscience and precision medicine.
- Hands-on work with real stroke imaging and cognitive outcome data.
- Close scientific supervision within the IBISC laboratory and interaction with clinical partners.
- Possibility of co-authoring a scientific publication and continuing toward PhD studies depending on results and funding opportunities.

Practical information and application

The intern will be mainly hosted at the UFR Sciences et Technologies, Université d'Évry Paris-Saclay, 40 rue du Pelvoux, 91020 Evry-Courcouronnes. Short periods at a partner hospital may be planned according to the project needs. Internship gratification will follow the legal rate in force.

To apply, send a CV, a short motivation letter and University transcripts since the first year of the BSc/Licence to:

`vincent.vigneron@univ-evry.fr` and `hichem.maaref@iut.univ-evry.fr`.

Selected references

- [1] J. S. Siegel et al., "Disruptions of network connectivity predict impairment in multiple behavioral domains after stroke," *Proc. Natl. Acad. Sci. USA*, 113(30):E4367-E4376, 2016. doi:10.1073/pnas.1521083113.
- [2] J.-S. Lim, J.-J. Lee, C.-W. Woo, "Post-Stroke Cognitive Impairment: Pathophysiological Insights into Brain Disconnectome from Advanced Neuroimaging Analysis Techniques," *Journal of Stroke*, 23(3):297-311, 2021. doi:10.5853/jos.2021.02376.
- [3] K. K. Kolskär et al., "Structural disconnectome mapping of cognitive function in poststroke patients," *Brain and Behavior*, 12(8):e2707, 2022. doi:10.1002/brb3.2707.
- [4] M. J. Moore, J. B. Mattingley, N. Demeyere, "Multivariate and network lesion mapping reveals distinct architectures of domain-specific post-stroke cognitive impairments," *Neuropsychologia*, 204:109007, 2024. doi:10.1016/j.neuropsychologia.2024.109007.
- [5] O. Godefroy et al., "Are we ready to cure post-stroke cognitive impairment? Many key prerequisites can be achieved quickly and easily," *European Stroke Journal*, 10(1):22-35, 2025. doi:10.1177/23969873241271651.