PhD thesis: Multi-modal explainable machine learning for exploring consciousness recovery of comatose patients

Context:

While consciousness is currently associated with brain processes, the subjective experience we have also appears embodied and situated (Bayne et al., 2020). Enactivism is the naturalistic theory that addresses the “hard problem of consciousness” by studying the dynamic interaction between the brain, the body and the environment in relation to subjective experience (Thompson et Varela, 2001).

One possible scientific way of evaluating such a theory is to study the recovery of interactions when consciousness returns, as following a coma. Coma is a state of unconsciousness from which patients can wake up, sometimes passing through different disorders of consciousness (DoC).

We hypothesize that the recovery of consciousness of comatose patients can be predicted by studying the coupling between the brain, the rest of the body (particularly behaviour and cardiac activity) and the environment.

The first aim of the project is to apply, in healthy participants and patients, new behavioral tools we have developed thanks to different computer vision methods (body and face analysis from video) (Michelot et al. 2025), and to characterize the behavioral responses associated with naturalistic conditions, thanks to machine-learning methods. Then, the first models should be extended to more modalities to characterize body-brain-environment interactions, i.e. how brain (from high-density EEG connectivity), body (from video and electrocardiogram) responses are coupled/decoupled according to the environment (emotional and social vs. neutral stimuli).

The second aim is to design a new multi-modal machine learning model (e.g. based on Transformers) and train it on even more heterogeneous data coming from different sources (video, EEG, ECG, but also IRM, clinical data, hormonal dosages, actimetry, etc.) and from different cohorts and recorded at different time scales (continuously or at specific instants).

Overall, this work will allow a better care of comatose and DoC patients and more generally a better understanding of the multimodal bases of consciousness. It will further advance the state of the art in multi-modal Machine Learning on scarce and heterogeneous data and explainable models allowing to extract interpretable features of these modalities as well as correlations and inter-relations and potentially to acquire new medical and psychological insights.

Objectives:

We will build upon our previous works, where state-of-the-art computer vision algorithms (e.g. OpenPose, OpenFace) for human pose estimation and face and gaze analysis have been integrated and adapted to the specific context and acquisition conditions.
The first experiments based on recurrent (LSTM) neural networks, CNN and Transformers have given promising results for classifying different situations (in terms of interaction with the environment and overall behavioral phenomena) (Michelot et al. 2025) which we will further improve and make explicit, i.e. interpretable by medical specialists and neuroscientists.
The first objective of the PhD student will be to study and improve multi-modal Machine Learning models, for the fusion of video (body/face point time series), EEG and ECG data, to predict the situations of our healthy control group and try to adapt the models to patients. Based on our preliminary work on multi-modal LSTM and Transformer models, the aim would be to find characteristic patterns and correlations in the data that represent the different emotional or interactive situations, using eXplainable AI (XAI) techniques such as Integrated Gradient or SHAP.
To train a model for patients (either patient-specific models or a global one), two approaches will be investigated: first, adapt the control group models to the more complex data distribution using transfer learning and domain adaptation techniques. Second, unsupervised pre-training (e.g. self-supervised learning) of the different modalities on larger unannotated datasets, individually and potentially combined, on healthy subjects and fine-tuning on patient data for diagnosis or prognosis of coma recovery. The principal challenge is to identify relevant features or representation spaces and/or pre-text tasks from the heterogeneous and noisy data that can generalize to the scarce patient data having a very high inter-subject variability.
The second objective is to extend these models to even more heterogeneous data from more modalities such as IRM or other physiological measures that may be more sparse (clinical data, hormonal dosages, actimetry, etc.). The challenge is to fuse these highly variable data recorded at different time scales and to leverage their complementarity.
To render this tractable, we will develop new specific multi-modal Transformer architectures and will investigate different unsupervised and self-supervised training strategies to effectively exploit data from different modalities and experiment with transfer learning approaches to adapt pre-trained models from healthy subjects to patient data or from one group of patients to another one.
Such approaches based on unsupervised multi-modal representation learning and XAI may give rise to new methodologies both in AI/ML and in Neuroscience.

Environnement:

This PhD thesis is part of an multidisciplinary research project called **AgoraComa** (funded by ShapeMed@Lyon), and will be supervised by Stefan Duffner from the LIRIS lab (IMAGINE team), Fabien Perrin CRNL lab (CAP team), Carole Lartizien from the CREATIS lab (MYRIAD team) and Florent Gobert, hospital practitioner at the HCL, Neurology hospital.
The PhD student will closely interact with other PhD students in Neuroscience at the CRNL and in Computer Science at LIRIS and CREATIS.
The main work will be done at LIRIS (INSA Lyon, Campus de la Doua), at Intensive Care Unit of the Neurological hospital (HCL, Bron) and at CRNL (Bron).

Prerequisites:

- A strong background in Computer Science or Applied Mathematics and Machine Learning, both from the theoretical and from the practical side (python programming)
- Scientific curiosity and autonomy
- Interest in the Medicine and Neuroscience fields and willing to participate in the interdisciplinary approach of the AgoraComa project

Contact:

stefan.duffner@liris.cnrs.fr

fabien.perrin@univ-lyon1.fr

carole.lartizien@creatis.insa-lyon.fr

florent.gobert01@chu-lyon.fr