Multi-modal explainable machine learning for exploring consciousness recovery of coma patients

Context:

While consciousness is currently seen as the result of processes in the brain, ordinary human experience is in fact embedded in a web of causal relations that link the brain to the body and the environment (Bayne et al., 2020). Embodied cognition is a naturalistic theory in which consciousness is associated with a dynamic interaction between brain, body and environment (BBE) (Thompson et Varela, 2001). Indeed, from an evolutionary point of view, the nervous system appears to be dedicated to perceptual and motor processes that allow interaction with the environment (Thompson et Varela, 2001). According to this model, the world of which we are aware is enacted by our past interactions with the environment. In other words, consciousness would be better explained in terms of being in the world at a specific time and place. From a biological point of view, this would be associated with both information shared across cortical regions and non-reflexive behaviors (Kanai et al., 2019).

One way to better understand consciousness is to study its disorders and recovery. Indeed, when such patients recover, they go through different clinical states that are characterized by the recovery of arousal and/or awareness and by the recovery of BBE interactions. Coma is a state of unconsciousness in which patients cannot be awakened. Those who recover could transit through a disorder of consciousness (DOC).

We hypothesize that the analysis of synchronized brain-behavior recordings to naturalistic stimulations will better predict consciousness recovery, than the current analyses of one of the two modalities in neutral contexts. Specifically, we hypothesize that DOC patients who will recover consciousness are those whose cerebral connectivity abilities are associated with non-reflexive behaviors to emotional and/or social stimuli (listening to preferred music, visit from relatives).

The first aim of the project is to apply, in healthy participants and DOC patients, new behavioral tools we have developed thanks to different computer vision methods (body and face analysis from video), and to characterize the behavioral responses associated with naturalistic conditions, thanks to machine-learning methods. The second aim is to characterize, in the same populations, the BBE interactions, i.e. how brain (from high-density EEG connectivity), body (from video and ECG -electrocardiogram) responses are coupled/decoupled according to the environment (emotional and social vs. neutral stimuli). This work will allow a better care of DOC patients and more generally a better understanding of the cerebral-behavioral bases of consciousness.

Objectives:

We will build upon existing preliminary 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.

Subjects are successively put into 3 different situations:

- 🎬 listening to sound,
- 🎬 rest,
- 🎬 interactions with the instructor,

and for each situation an emotional and a non-emotional version is performed (e.g. for sound: listening to favorite music). The first experiments based on recurrent (LSTM) neural networks have already given promising results for classifying different situations (in terms of interaction with the environment and overall behavioral phenomena). which we will further improve and make explicit, i.e. interpretable by medical specialists and neuroscientists. For example, at this point, we do not know yet, what parts of the body or face are likely to produce more relevant indicators of consciousness state and which type of behaviour, motion or gestures are informative. This will be one concrete expected result of the internship.

We have video, ECG and hd-EEG data from 20 healthy subjects and 60 DOC patients, which will allow the development of more precise and robust machine learning models.

Nevertheless, this represents relatively few data given the large variability between patients. Thus strong priors and regularisation are needed to avoid overfitting. Also specific pre- and post-processing methods will help to reduce the noise and decrease the dimensionality (e.g. by learning compact feature embeddings, or by features selection algorithms). For example, we developed a specific filtering algorithm on the output of OpenPose that effectively removes undesired oscillations (due to pixel quantisations) and produces body pose and motion estimates that are much more realistic and that contain fewer artifacts.

Furthermore, different learning strategies and models will be developed to deal with the large amount of noise in general and the imbalance between the amount of relevant data compared to irrelevant data, for example by "rebalancing" the data using other methods or by specific learning mechanisms such as the Multiple Instance learning framework or some type of self-supervised learning.

The combination of these different modalities using new deep learning models as well as the adaptation of our existing models for unsupervised learning multi-variate time series (Berlemont et al. 2017) will allow us to further analyse complex correlations and co-occurences of characteristics and, by focusing on explainable methods and results (explainable AI), give insights into BBE interactions and further give rise to new neuroscientific hypotheses. The findings of such correlation patterns will be another major result that we expect from this project.

Thus, the nature of this research is clearly exploratory, and the expected results will concern both methodological contributions in AI and original methods leading to new knowledge in Neuroscience.

Environnement:

This internship will be supervised by Stefan Duffner from the LIRIS lab (IMAGINE team) and the Fabien Perrin CRNL lab (CAP team). The intern will also interact with the recently hired PhD student in Neuroscience at the CRNL. The main work will be done at LIRIS (Lyon, Campus de la Doua) but the candidate is free to move between the two laboratories.

Contact:

stefan.duffner@liris.cnrs.fr

fabien.perrin@univ-lyon1.fr