Do sleep and anesthesia share common multifractal EEG dynamics? Insights from adversarial domain adaptation

Louis Leconte (louis.leconte@ens-paris-saclay.fr)
Ecole Normale Supérieure Paris-Saclay, Paris, France
Computational and Cognitive Neuroscience Lab (CoCo Lab), Dept of Psychology, Université de Montréal, Montreal, QC, Canada
Tarek Lajnef (lajnef.tarek@gmail.com)
Computational and Cognitive Neuroscience Lab (CoCo Lab), Dept of Psychology, Université de Montréal, Montreal, QC, Canada
Thomas Thierry (tbthiery91@gmail.com)
Computational and Cognitive Neuroscience Lab (CoCo Lab), Dept of Psychology, Université de Montréal, Montreal, QC, Canada
George Mashour (gmashour@med.umich.edu)
Center for Consciousness Science, Department of Anesthesiology, University of Michigan, USA
Stefanie Blain-Moraes (stefanie.blain-moraes@mcgill.ca)
School of Physical and Occupational Therapy, McGill University, Montreal, QC, Canada
Perrine Ruby (perrine.ruby@inserm.fr)
Lyon Neuroscience Research Center, Universit Claude Bernard Lyon 1, Universit de Lyon, Lyon, France
Karim Jerbi (karim.jerbi.udem@gmail.com)
Computational and Cognitive Neuroscience Lab (CoCo Lab), Dept of Psychology, Université de Montréal, Montreal, QC, Canada

Abstract:

The brain displays scale-free and multifractal dynamics that change across different states of consciousness. Whether the multifractal properties of EEG data change in a similar way when shifting from a conscious to an unconscious state compared to shifting from wakefulness to sleep is still largely unknown. To address this we ask a slightly different question: How well can we use a classifier trained on sleep EEG multifractality data to correctly discriminate conscious and unconscious states. To this end, we used a Domain Adversarial Neural Network (DANN) framework geared towards discriminating neural signals recorded during conscious vs unconscious states (target domain), based on classification of brain signals recorded during wakefulness vs sleep (source domain). We compare results obtained with naïve transfer learning (no domain adaptation), with supervised and unsupervised domain adaptation. The input data consisted of multifractal parameters computed from EEG recordings. This paper reports two important findings: First, our analyses provide evidence for the feasibility of creating a DANN architecture that can learn to discriminate consciousness from anesthetic-induced unconsciousness via adversarial adaptation of sleep/wakefulness discrimination. Second, by exploring the topographies of the successful classification rates across the EEG array, we were able to identify functional similarities of EEG multifractality patterns across sleep and anesthesia.

Keywords: EEG; multifractality; consciousness; sleep; domain adaptation; neural networks; adversarial learning

Introduction

Background

Sleep and generalized anesthesia are not the same phenomenon, but they share some similarities in terms of diminished arousal for instance. Some studies indicate that the sleep and anesthesia can be associated with similar patterns of oscillatory bbrain dynamics (e.g. Brown, Lydic, & Schiff, 2010). EEG under Propofol anaesthesia shows large slow waves similar in appearance to the slow waves seen during deep sleep (Murphy et al., 2011). Studies that assess large-scale brain dynamics during sleep and anesthesia provide complementary insights into the neural bases of consciousness. Anesthetics are a useful tool for the study of human being unconsciousness as they can be titrated to pharmacologically induce the loss of consciousness. Likewise, the various neural properties properties of distinct sleep stages provide an interesting window on altered states of consciousness. Among the numerous EEG features that are modulated during sleep and under anesthesia, scale-invariance is a promising metric that is receiving increasing attention. In a recent study, we have for instance reported that Sevoflurane-induced light sedation is associated with drops in long-range temporal correlation in beta oscillations over central brain areas, nicely captured with detrended fluctuation analyses (Thiery et al., 2018). A closely related formalism that can be used to assess scaling properties of brain signals is multifractality. Whether the properties it measures (e.g.

self-similarity and extent of multifractality) change in a similar way during sleep and anesthesia is still an open question.

Aim of the study

The goal of the present study is to probe the discrepancies and similarities between changes in EEG multifractility that occur during sleep to those observed in generalized anesthesia. We chose to tackle this question through a transfer learning framework where we assess how well a neural network trained on sleep EEG data can actually discriminate consciousness from anesthetic-induced unconsciousness states. The underlying premise is that EEG features that allow for successfully transfer reflect functional similarities between the source (sleep) and target (anesthesia) domains. In particular, we hypothesize that while the neural underpinnings of sleep and anesthesia are not identical, some similarities in specific scaling properties may be revealed through a transfer learning framework.

Materials and Methods

Sleep dataset

Whole night polysomnography recordings were obtained from 36 subjects collected at the Lyon Neuroscience Research Center (Lyon, France). Each record contains EOG, EMG and 19 scalp-EEG channels and sampled at frequency of 1kHz. The data was scored (Awake, S1, S2, Slow-Wave Sleep (SWS) and REM) both visually and using our automatic staging tools (i.e. decision-tree multi-class SVM). For further details on the data set and pre-processing methods, see Lajnef et al., (2015).

Anesthesia dataset

Sevoflurane EEG data (light sedation) was collected from 10 participants at the Department of Anesthesiology of the University of Michigan, Ann Arbor, MI, USA. EEG signals were collected using a 64-channel biosignal amplifier at a 500 Hz sampling rate. Sevoflurane concentration was gradually increased until loss of consciousness was reached. After at least 10 min of unconsciousness, the reverse protocol was employed until the participant regained consciousness (cf. Thiery et al. 2018 for details). After artefact identification and rejection, data from 7 participants (4 males, 20-23 yrs old) were kept for further analyses.

Feature extraction: multifractal parameters

We used the p-leaders multifractal formalism (Jaffard et al. 2015) to compute 12 multifractal features (log-cumulnants $c(p)\,1$,$c(p)\,2$,$c(p)\,3$ for $p = 1 \rightarrow 4$) for each participant, and each epoch of the EEG data, in both the sleep and anesthesia data sets. These parameters characterize the singularities present in the signal. The p-leader formalism (Jaffard et al, 2015) is thought to outperform multifractal detrended fluctuation analysis (MFDFA) and wavelet leader formalisms (Wendt, 2008). Gathering results for different values of p, enables us to be aware of the disparate types of singularities that are present in EEG signals For the transfer learning task considered here we chose to set p=4 and to limit the multifractal spectrum to its truncated 2nd order polynomial expansion with only two scaling parameters, i.e. c1 (self-similarity, comparable to Hurst exponent) and c2 which corresponds to local fluctuations of scaling behaviors ($c2 < 0$ indicates multifractality, while $c2 = 0$ indicate monofractality).

Transfer learning method

To explore transfer learning between the source domain (wakefulness/sleep) and the target domain (consciousness/unconsciousness), we explored 3 options: (i) naïve transfer (no domain adaptation), (ii) supervised domain adaptation and (iii) unsupervised domain adaptation. As we have 5 different sleep stages, we can train a classifier on all 10 possible binary sleep stage classification problems (e.g. awake vs S2 sleep), and then applied or adapted it to the 2-class (conscious vs unconscious) anesthesia dataset. But, first of all, we determined a geometrical mapping between the 63 electrodes of the anesthesia dataset, and the 19 electrodes of the sleep dataset using (closest electrodes in terms of spatial geometry).

Naive Transfer Our simplest attempt at training a classifier on sleep stage discrimination to then apply it to consciousness vs unconsciousness classification was performed without any domain adaptation. For this approach which we refer to as naïve transfer (or "without domain adaptation") we tested SVM (Support Vector Machine) with RBF kernel, Random Forest, and single hidden layer Neural Network.

Domain Adaptation Learning a discriminative classifier or other predictor in the presence of a shift between training and test distributions is known as Domain Adaptation. Here we implemented an approach proposed by Ganin et al., (2017), known as Domain-Adversarial Neural Network (DANN). The DANN framework elegantly addresses the domain adaptation problem by promoting the emergence of features that are discriminative for the main learning task on the source domain while being indiscriminate with respect to the shift between the domains. Such adaptation behaviour can be achieved using augmented neural network architectures that include a gradient reversal layer (Ganin et al. 2016). In principle, an adversarial DA trains two neural networks, a discriminator that attempts to separate the target domain from the transformed source domain, and a generator that aims to fool the discriminator to make the source and target domains look like one another as much as possible, similar to the philosophy of a conditional GAN (Goodfellow et al.,

2014).

The goal is both to optimize a ”label predictor” that learns to predict class label, and to optimize a ”domain classifier” that learn to discriminate the source domain from the target domain. Hence each part of the neural network (”label predictor” and ”domain classifier”) is competing against each other in the optimization of the min and max equation. To tackle this problem, a stochastic gradient procedure, in which updates are made in the opposite direction of the gradient for the minimizing parameters, and in the direction of the gradient for the maximizing parameters, is implemented by selecting a random subset of data at each step.

Supervised DANN In the supervised DANN approach we optimized the hyper-parameters (coefficient of adaptation and size of the hidden layer). This was achieved using nested cross-validation approach where a ”Leave P Group(s) Out” (repeated 10 times, with 1 participant out at each iteration) was followed by two ”Leave P Group(s) Out” with 2 participants out for the anesthesia data and 7 participants out for the sleep data. For each electrode, the selected classifier was the one that gave the best decoding accuracy on the anesthesia data (the 2 participants left out) after domain adaptation.

Unsupervised DANN Although the nested cross-validation described above was used to rule out bias, we do introduce a priori information by selecting the set of hyperparameters that give the best decoding accuracy on the transfer. We therefore also implemented a fully unsupervised DANN where all hyperparameters were set to their default value and (for 10 times) 2 anesthesia participants and 7 sleep participants were left out to implement both a classifier (fit on sleep pair) without (naïve transfer) and with domain adaptation on the anesthesia dataset.

Results

Semi-supervised vs Naive transfer

The SVM and Random Forest algorithms led to decoding accuracies of 86,89% and 88,59% respectively when classifying pairs of sleep stages. When we subsequently applied these trained models to the anesthesia data (i.e. naïve transfer from sleep classification to consciousness versus unconsciousness), the decoding accuracies obtained were close to the 50% chance-level. These results are expected since the training (sleep data) and test (anesthesia data) sets come from similar but different distributions. However, as shown in Figure 1, domain adaptation using DANN achieves a compromise where the accuracy on the source domain (sleep stages) is reduced while the decoding accuracy on the target domain (anesthesia data) is enhanced.

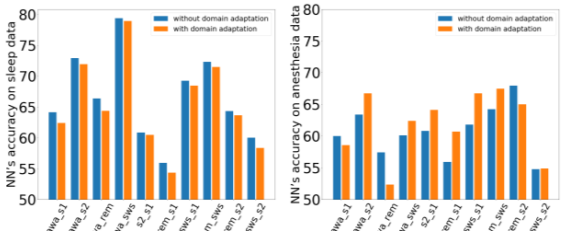

Figure 1: Left panel – Neural Network classification accuracy on source domain (sleep) with (orange) and without (blue) domain adaptation when using various sleep stage pairs for training. Right panel – same but for results of NN on target domain (conscious vs unconscious).

Our results indicate that the classification were now significant on almost all datasets (source and target). It is noteworthy, however, that the decoding accuracy on the source domain stays above 70% only for ”Awake/SWS”, ”Rem/SWS” and ”Awake/S2” datasets. We can tentatively state that this may point towards a similarity between changes of EEG fractality compring shifts between sleep states and shifts between consciousness and unconsciousness, particularly for ”Awake and SWS” stages, ”REM and SWS”, and ”Awake and S2” stages.

Obviously an algorithm that over-adapts to the target classification problem (consciousness vs unconsciousness) to the extent that it fails to reach relevant levels of decoding on the source domain (sleep) is of little value for our research question. To highlight which sleep stage pairs led to a significant decoding accuracy on the anesthesia dataset, while maintaining a good decoding accuracy on the sleep set, we computed for each of the 10 pairs the product of the two (in source and target domain) mean decoding accuracies. The results (akin to a joint distribution) are shown in Figure 2.

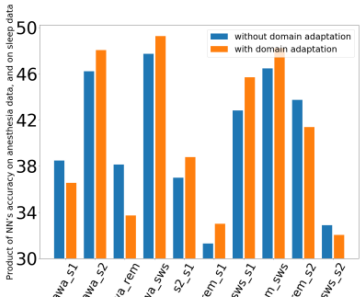

Figure 2: Product of decoding accuracies for source and target domains using neural networks, with and without domain adaptation.

Unsupervised vs Naive transfer

As expected in Figure 3, the decoding accuracy on the source dataset (sleep data) is not different from the one obtained with a semi-supervised method (Figure 1).

Nevertheless, the results on the target dataset (anesthesia data) have substantially decrease for each sleep pair possible.

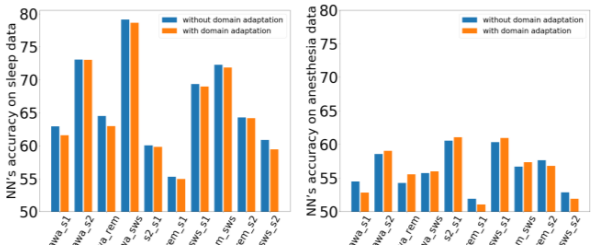

Figure 3 Left panel – Neural Network classification accuracy on source domain (sleep) with and without domain adaptation. Right panel – same but for results of NN on target domain (conscious vs unconscious). By contrast to Fig 1, here domain adaptation was without hyper-parameter optimization (i.e. unsupervised DANN)

Conclusion and future work

Our last findings put to the fore the existence of a link (in terms of multifractality) between the conscious-unconscious shift and the shift between distinct sleep stages. Interestingly, the best domain adaptation results were made possible by adapting an “awake vs deep sleep (SWS)” classifier (Figures 1 and 2). The other prominent source domain models that showed high adaptability to anesthesia classification were ”Awake vs S2” and ”REM vs SWS”. These findings are consistent with the fact that ”SWS” represents a deeper sleep state (Susmakova, 2004) and that ”REM” sleep stage has several brain dynamic properties that ar similar to the Awake state.

More generally, the combination of representation learning and domain adaptation has a wide range of promising applications in many fields especially where the availability of labeled data and sample sizes are limiting factors. Here we illustrate how assessing the feasibility of domain adaptation (i.e. performance and what features of the data are particularly useful for it) can be useful to address questions in neuroscience that seek to uncover relationships between the neural substrates of various cognitive states. Ongoing work involves identifying the brain regions that showing the strongest discrepancies and strongest similarities between sleep and anesthesia. Moreover the validity and generalization of the findings reported here would benefit from incorporating anesthesia data sets with other anesthetics (e.g. Proofol) and with other anesthesia protocols (e.g. deep sedation).

References

Ben-David, S., Blitzer, J., Crammer, K., & Pereira, F. (2006). Analysis of representations for domain adaptation. InNIPS, 137–144.

Brown, E. N., Lydic, R., & Schiff, N. D. (2010). General anesthesia, sleep, and coma. The New England journal of medicine, 363, 26382650.

Ganin, Y., Ustinova, E., Ajakan, H., Germain, P., Larochelle, H., Laviolette, F., ... Lempitsky, V. (2017). Domainadversarial training of neural networks. Domain Adaptation in Computer Vision Applications.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., WardeFarley, D., Ozair, S., ... Bengio, Y. (2014). Generative adversarial nets.

Jaffard, S., Melot, C., Leonarduzzi, R., Wendt, H., Abry, P., Roux, S., & Torres, M. (2015). p-exponent and p-leaders, part ii: Multifractal analysis. relations to detrended fluctuation analysis.

Lajnef, T., Chaibi, S., Ruby, P., Aguera, P.-E., Eichenlaub, J.E., Samet, M., ... Jerbi, K. (2015). Learning machines and sleeping brains: Automatic sleep stage classification using decision-tree multi-class support vector machines.

Murphy, M., Bruno, M., Riedner, B., Boveroux, P., Noirhomme, Q., Landsness, E., ... Boly, M. (2011). Propofol anesthesia and sleep: a high-density eeg study. Sleep, 34.

Susmakova, K. (2004). Human sleep and sleep eeg. MEASUREMENT SCIENCE REVIEW, 4.

Thiery, T., Lajnef, T., Combrisson, E., Dehgan, A., Rainville, P., Mashoure, G., ... Jerbi, K. (2018). Long-rangetemporal correlations in the brain distinguish conscious wakefulness from induced unconsciousness.

Wendt, H. (2008). Contributions of wavelet leaders and bootstrap to multifractal analysis: Images, estimation performance, dependence structure and vanishing moments. condence intervals and hypothesis tests. Thesis manuscript.