

Classification of propofol-induced sedation states using brain connectivity analysis

Dheeraj Rathee*, Hubert Cecotti *Senior Member, IEEE*, Girijesh Prasad *Senior Member, IEEE*

Abstract—Brain connectivity measurements can provide key information about ongoing brain processes. In this paper, we propose to investigate the performance of the binary classification of Propofol-induced sedation states using partial granger causality analysis. Based on the brain connectivity measurements obtained from EEG signals in a database that contains four sedation states: baseline, mild, moderate, and recovery, we consider eight sensors and evaluate the area under the ROC curve with five classifiers: the k-nearest neighbor (density method), support vector machine, linear discriminant analysis, Bayesian discriminant analysis, and a model based on extreme learning machine. The results support the conclusion that the different Propofol-induced sedation states can be identified with an AUC of around 0.75, by considering signal segments of only 4 second. These results highlight the discriminant power that can be obtained from scalp level connectivity measures for online brain monitoring.

I. INTRODUCTION

Sedation is a common practice during majority of non-invasive or minimally invasive medical interventions performed in intensive care units (ICUs). However, current available methods for the evaluation of sedation level in ICU environment are majorly of subjective nature i.e., using sedation scales such as Richmond agitation-sedation scale (RASS) and Sedation-Agitation Scale (SAS) [1]. Furthermore, these methods are highly dependent on the clinicians' expertise and thus, may lead to unreliable assessments. A strong association between excessive and/or inadequate sedation and adverse healthcare outcomes have been confirmed by several recent scientific studies and clinical surveys [2]. Moreover, accurate evaluation of patients' sedation level during their ICU stay may lead to substantial reduction in mortality, mechanical ventilation, duration of stay, and improved healthcare resource employment which may collectively result in substantial decrease in the overall healthcare cost [3].

The process of altering the consciousness of patients for clinical purpose using anaesthetic drugs like Propofol is commonplace in medicinal science. The association of Propofol-induced sedation and functional brain networks has been proved by several previous studies [4], [5], [6], [7]. However,

D.R. was supported by Ulster University Vice Chancellor's Research scholarship (VCRS). G.P. and H.C. were supported by the NIFBM Facility project (1303/101154803). G.P. is also supported by the UKIERI DST Thematic Partnership project (DST-UKIERI-2016-17-0128).

D.R. and G.P. are with the Intelligent Systems Research Centre, School of Computing & Intelligent Systems, Ulster University, Derry~Londonderry, N. Ireland, UK. HC is with the Department of Computer Science, College of Science and Mathematics, Fresno State, Fresno, CA, USA. Corresponding author: Dheeraj Rathee: rathee-d@ulster.ac.uk

the main focus of these studies was to establish a relationship between varying levels (or states) of consciousness and reorganisation of brain networks, thus, the majority of them utilized functional magnetic resonance imaging (fMRI) signals rather than EEG signals as the former provides better spatial resolution. However, its application for assessment of sedation in clinical settings possess resilient confounds such as claustrophobic working environment, high susceptibility to head movement related artifacts, and incompatibility for real-time implementation in ICU settings. Nevertheless, fewer attempts have been reported in the literature for the classification of different sedation levels based on brain functional networks.

Estimating connectivity between different brain regions using electroencephalography (EEG) is a popular alternative to study the causal communication mechanisms between distinct neuronal systems at scalp level. However, the process is a non-trivial task, with several potential problems caused by inherent characteristics of the EEG signals and the limitations of the connectivity methods employed. In this study, in line with our previous work [7], we implemented time-domain partial granger causality (PGC) analysis to estimate the directed EEG connectivity networks during four different levels of sedation. These levels were determined based on the collective assessment of Propofol blood concentration and the behavioral responses of healthy participants. Further, several state-of-the-art classifiers have been implemented for distinguishing the different sedation levels using connectivity features from frontal, parietal, and occipital scalp regions of the brain.

The remainder of the paper is organized as follows: the EEG database, the formal description of PGC, and the methodology are presented in section II. Then, the results are detailed in section III and finally discussed in section IV.

II. MATERIAL AND METHODS

A. Database

In this study related to the classification of Propofol-induced sedation states, we considered an EEG database acquired with 20 healthy adult participants who got injected a target-controlled infusion of Propofol. The database and the experimental protocol are fully described in [8]. Based on the behavioral responses, the complete database is divided into two groups: responsive group (13 participants) and drowsy group (7 participants). For further analysis, we considered only the drowsy group so as to have confirmed ground truth of different states of sedation i.e., baseline (Ba), mild sedation (Mi), moderate sedation (Mo), and recovery (Re). The

database consists of approximately 7 minutes of 128-channel high-density EEG data acquired in eyes-closed resting state condition, sampled at 250 Hz, and referenced to the vertex.

B. Time-domain Partial Granger Causality Analysis

Time-domain PGC is a variation of Granger causality wherein causal relationships between the different signals can be analyzed through multivariate autoregressive (MVAR) analysis. PGC gives better estimation of the true interactions between signals by reducing the effect of confounding variables contrary to bivariate GC and conditional GC [9]. We consider three time series signals: $S_1(t)$, $S_2(t)$ and $S_3(t)$. To analyze the effective connectivity between $S_1(t)$ and $S_2(t)$, conditioned on $S_3(t)$, based on PGC rules, the reduced model is defined by:

$$S_1(t) = \sum_{p=1}^k (a_{(1,p)}X(t-p)) + \sum_{p=1}^k (c_{(1,p)}Z(t-p)) + (1) \\ \epsilon_1(t) + \epsilon_1^E(t) + \beta_1(L)\epsilon_1^L(t)$$

$$S_2(t) = \sum_{p=1}^k (b_{(1,p)}Y(t-p)) + \sum_{p=1}^k (d_{(1,p)}Z(t-p)) + (2) \\ \epsilon_2(t) + \epsilon_2^E(t) + \beta_2(L)\epsilon_2^L(t)$$

where p is the model order, $\epsilon_i(t)$ is the prediction error, $\epsilon_i^E(t)$ and $\beta(L)\epsilon_i^L(t)$ are the residual errors. $\epsilon_i^E(t)$ and $\beta(L)\epsilon_i^L(t)$ correspond to the exogenous (E) and the latent (L) inputs, respectively. The full model that includes past values of the sink variable, and past values of source variable conditioned on rest of variables, is expressed by:

$$S_1(t) = \sum_{p=1}^k (a_{(2,p)}S_1(t-p)) + \sum_{p=1}^k (b_{2,p}S_2(t-p)) + (3) \\ \sum_{p=1}^k (c_{(2,p)}S_3(t-p)) + \epsilon_3(t) + \epsilon_3^E(t) + \beta_3(L)\epsilon_3^L(t)$$

$$S_2(t) = \sum_{p=1}^k (d_{(2,p)}S_1(t-p)) + \sum_{p=1}^k (e_{2,p}S_2(t-p)) + (4) \\ \sum_{p=1}^k (f_{(2,p)}S_3(t-p)) + \epsilon_4(t) + \epsilon_4^E(t) + \beta_4(L)\epsilon_4^L(t)$$

The overall prediction errors are expressed as:

$$\mu_i = \epsilon_i(t) + \epsilon_i^E(t) + \beta_i(L)\epsilon_i^L(t) \quad (5)$$

with $1 \leq i \leq 4$. The prediction error covariance matrix for the reduced model can be obtained by:

$$R = \begin{bmatrix} \text{var}(\mu_1(t)) & \text{cov}(\mu_1(t), \mu_2(t)) \\ \text{cov}(\mu_2(t), \mu_1(t)) & \text{var}(\mu_2(t)) \end{bmatrix} \quad (6)$$

and the covariance matrix for the full model can be generated can be:

$$L = \begin{bmatrix} \text{var}(\mu_3(t)) & \text{cov}(\mu_3(t), \mu_4(t)) \\ \text{cov}(\mu_4(t), \mu_3(t)) & \text{var}(\mu_4(t)) \end{bmatrix} \quad (7)$$

Finally, the PGC indices are estimated by taking the log ratio of the partial variance of prediction error from the reduced model and the partial variance of prediction error from the

full model. The two following expressions give the PGC indices for $S_2(t)$ causing $S_1(t)$, and $S_1(t)$ causing $S_2(t)$:

$$\mathcal{G}_{S_2 \rightarrow S_1 | S_3} = \ln\left(\frac{R_{1,1} - R_{1,2}R_{2,2}^{-1}R_{2,1}}{L_{1,1} - L_{1,2}L_{2,2}^{-1}L_{2,1}}\right) \quad (8)$$

$$\mathcal{G}_{S_1 \rightarrow S_2 | S_3} = \ln\left(\frac{R_{2,2} - R_{2,1}R_{1,1}^{-1}R_{1,2}}{L_{2,2} - L_{2,1}L_{1,1}^{-1}L_{1,2}}\right) \quad (9)$$

C. Signal processing pipeline

To remove the movement related noise, EEG channels placed on neck and face were excluded and the remaining data was visually inspected and cleaned for noisy segments. EEG data from remaining channels were converted to their scalp current source density representations by using spherical spline interpolation [10], [11]. Our recent work showed significant improvement in classification of motor-imagery related brain responses with CSD estimation [12], [13]. The data was bandpassed in the frequency range of 8-12 Hz using a 4th order, zero-phase forward and backward bandpass Butterworth filter. Previous studies reported better discrimination of sedation states within alpha (8-12 Hz) frequency band [8].

Furthermore, the data were decomposed into independent components using Infomax ICA algorithm and components related to eye-blink and ECG were rejected using an in-house pre-processing method. The clean components were back-projected to the sensor space and the artefact-cleaned data were segmented having a sliding window of 4 s with 1 s overlap. The process generated approximately 414, 410, 410, and 419 overlapping segments (averaged over participants) for Ba, Mi, Mo, and Re, respectively. For further analysis, we selected a subset of 8 EEG channels related to the frontal (FP1, FP2, F3, F4), parietal (P3, P4), and occipital (O1, O2) scalp brain regions. Next, for each segment, time-domain PGC method was implemented to estimate the causal interactions between the eight scalp channel data. The coefficients of the MVAR model were estimated using the LWR algorithm [14]. Schwarz bayesian information criterion (SBIC) [15] was used for the estimation of the optimal value of the model order p . The complete analysis has been previously described in [16].

As the number of selected sensors is $N_s = 8$, the total number of obtained connectivity features for each segment (N_f) is $56 = N_s * N_s - N_s$ as there is no self-connectivities. Before the classification, the input features obtained from PGC are normalized to obtain their z-value, i.e., by removing the mean and dividing by the standard deviation across the 56 features. A classifier was trained for each participant by considering a 5-fold cross validation procedure. As the four experimental conditions happen over different time, we addressed the problem of the detection of the segments for 6 binary classification tasks, from one condition to one of the next conditions. We evaluated the performance of six possible pairwise comparisons: baseline vs mild (Ba↔Mi), baseline vs moderate (Ba↔Mo), baseline vs recovery (Ba↔Re), mild vs moderate (Mi↔Mo), mild vs recovery (Mi↔Re), and moderate vs recovery (Mo↔Re).

The performance of the binary classifiers is assessed with the area under the receiver-operator characteristic (ROC) curve (AUC) [17]. The computational analyses were implemented with custom MATLAB (V8.6) scripts using functions from GCCA (V2.9) [18] and EEGLAB (V14.0.0b) toolboxes [19] on an Intel Core i7-4790 processor with 16 GB of memory.

In this study, we compare the performance of the k -nearest neighbor (k NN) classifier (with $k=10$), support vector machine (SVM) (linear kernel), extreme learning machine (ELM), linear discriminant analysis (LDA), Bayesian linear discriminant analysis (BLDA), and stepwise LDA. The k -nearest neighbors algorithm is a non-parametric density estimation method which considers the k closest training examples in the feature space. The examples are classified by a majority vote of their neighbors: the test example is assigned to the class most common among its k nearest neighbors. SVM is a supervised technique that performs the classification by constructing a multidimensional hyperplane, which optimally discriminates between two classes by maximizing the margin between two data clusters [20]. LDA and BLDA classifiers are widely used linear classifiers [21]. Stepwise LDA adds and removes terms from a multi-linear model based on their statistical significance in a regression. ELM networks are a special type of artificial feedforward neural networks that embeds a single hidden layer [22]. Moreover, ELM can be applied for both classification and regression [23]. It is based on a linear combination of non-linear representations of the input data by using non-linear activation functions such as the sigmoid function. A key advantage of ELM models is the speed for training.

III. RESULTS

The AUC values for the 6 binary classification problems and the 6 classifiers are presented in Tables I to VI. Each cell in the table represents the mean of the performance obtained for a participant. The LDA based techniques provide the same type of results (about 0.776), followed by k NN, ELM, and SVM. Figure 1 depicts the simulation of the performance (AUC and accuracy) that can be obtained by the combination of several decisions, i.e. through the sum of the decisions of several examples. The simulated results were obtained by considering two standard normal distributions separated by a d-prime equivalent to an AUC=0.750, i.e., d-prime being defined by $\sqrt{2} * icdf(pd, AUC)$ where $icdf$ represents the inverse cumulative distribution function and pd represents a standard normal distribution (mean=0, standard deviation=1). The results indicate that an accuracy of 90% could be reached after 30 second of continuous examples from the same state, i.e., after accumulated decisions of 7 examples.

IV. DISCUSSION AND CONCLUSION

In this paper we have compared the performance of binary classification between multiple Propofol-induced sedation states using brain connectivity analysis with PGC method. Whereas the typical analysis is achieved across subjects to find the relative effect of Propofol, we have analyzed the extent to which it is possible to obtain a reliable decision

TABLE I
CLASSIFICATION PERFORMANCE FOR k NN.

	Ba↔Mi	Ba↔Mo	Ba↔Re	Mi↔Mo	Mi↔Re	Mo↔Re
S01	0.816	0.879	0.721	0.717	0.798	0.792
S02	0.785	0.818	0.783	0.861	0.700	0.879
S03	0.755	0.786	0.744	0.710	0.799	0.752
S04	0.670	0.635	0.738	0.663	0.759	0.802
S05	0.785	0.737	0.849	0.691	0.732	0.821
S06	0.736	0.807	0.735	0.747	0.679	0.763
S07	0.755	0.836	0.802	0.658	0.679	0.615
Mean	0.757	0.785	0.767	0.721	0.735	0.775

TABLE II
CLASSIFICATION PERFORMANCE FOR SVM.

	Ba↔Mi	Ba↔Mo	Ba↔Re	Mi↔Mo	Mi↔Re	Mo↔Re
S01	0.670	0.853	0.589	0.736	0.754	0.744
S02	0.652	0.737	0.480	0.856	0.606	0.834
S03	0.795	0.702	0.643	0.543	0.677	0.593
S04	0.545	0.553	0.664	0.584	0.655	0.667
S05	0.646	0.531	0.603	0.569	0.675	0.755
S06	0.663	0.701	0.661	0.639	0.515	0.683
S07	0.577	0.823	0.558	0.703	0.597	0.505
Mean	0.650	0.700	0.600	0.661	0.640	0.683

TABLE III
CLASSIFICATION PERFORMANCE FOR ELM.

	Ba↔Mi	Ba↔Mo	Ba↔Re	Mi↔Mo	Mi↔Re	Mo↔Re
S01	0.763	0.790	0.678	0.711	0.748	0.728
S02	0.715	0.737	0.641	0.783	0.651	0.783
S03	0.741	0.739	0.699	0.675	0.729	0.682
S04	0.637	0.601	0.704	0.632	0.714	0.742
S05	0.703	0.656	0.733	0.692	0.668	0.786
S06	0.720	0.766	0.741	0.703	0.665	0.731
S07	0.684	0.777	0.702	0.703	0.701	0.624
Mean	0.709	0.724	0.700	0.700	0.697	0.725

TABLE IV
CLASSIFICATION PERFORMANCE FOR LDA.

	Ba↔Mi	Ba↔Mo	Ba↔Re	Mi↔Mo	Mi↔Re	Mo↔Re
S01	0.820	0.862	0.749	0.797	0.814	0.817
S02	0.779	0.807	0.695	0.861	0.730	0.836
S03	0.805	0.813	0.751	0.751	0.806	0.752
S04	0.692	0.692	0.785	0.689	0.799	0.804
S05	0.776	0.740	0.800	0.763	0.792	0.840
S06	0.787	0.808	0.771	0.780	0.741	0.799
S07	0.737	0.832	0.780	0.755	0.753	0.672
Mean	0.771	0.793	0.762	0.771	0.776	0.789

TABLE V
CLASSIFICATION PERFORMANCE FOR BLDA.

	Ba↔Mi	Ba↔Mo	Ba↔Re	Mi↔Mo	Mi↔Re	Mo↔Re
S01	0.826	0.867	0.756	0.800	0.814	0.810
S02	0.778	0.804	0.689	0.860	0.732	0.835
S03	0.807	0.817	0.749	0.739	0.804	0.748
S04	0.694	0.684	0.785	0.680	0.799	0.814
S05	0.782	0.739	0.804	0.762	0.791	0.835
S06	0.786	0.808	0.772	0.779	0.741	0.798
S07	0.731	0.833	0.781	0.766	0.758	0.672
Mean	0.772	0.793	0.762	0.769	0.777	0.787

(AUC>0.75) regarding the current state of the participant based on only 4 s of EEG signals. The results suggest that the detection of a Propofol-induced sedation state is possible by considering only a short segment of EEG data from frontal,

TABLE VI
CLASSIFICATION PERFORMANCE FOR STEPWISE LDA.

	Ba↔Mi	Ba↔Mo	Ba↔Re	Mi↔Mo	Mi↔Re	Mo↔Re
S01	0.802	0.850	0.715	0.789	0.786	0.811
S02	0.775	0.798	0.698	0.849	0.725	0.831
S03	0.794	0.807	0.742	0.713	0.786	0.737
S04	0.707	0.644	0.774	0.702	0.795	0.790
S05	0.786	0.744	0.778	0.741	0.777	0.837
S06	0.787	0.793	0.760	0.790	0.716	0.782
S07	0.721	0.820	0.766	0.748	0.740	0.647
Mean	0.767	0.779	0.748	0.762	0.761	0.776

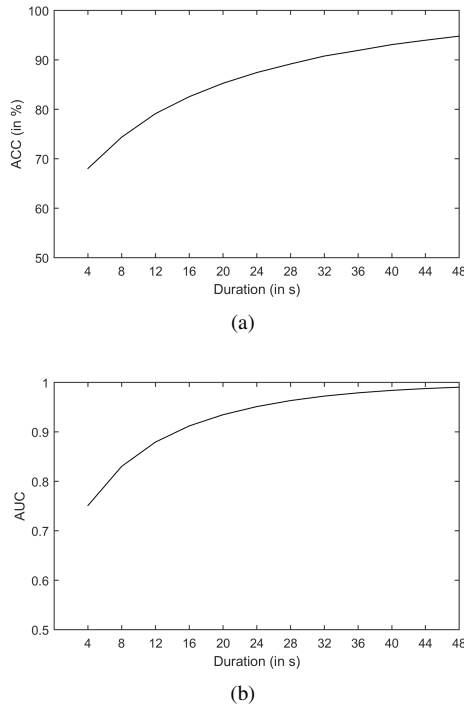


Fig. 1. Simulation results of the performance analysis with accumulated decision of multiple data segments. (a) Accuracy and (b) AUC

parietal and occipital scalp regions. Moreover, it is shown that PGC provides discriminant features for the classification. The results highlight that state-of-the-art methods such as k NN, LDA with its variants, provide the best performance in relation to other more advanced classifiers.

A typical classification problem, the training phase requires data and corresponding labels for all the available classes to train the machine learning model, however, for monitoring sedation states in ICU settings, the training data may not be possibly available for the classifier. Although the possible solutions to this problem involves using the data recorded in previous sedation sessions of the patient itself or from other patients. Nevertheless, these results confirm that brain connectivity analysis provides features that can be exploited at a relatively short time scale for assessment of sedation level. Further work will include the evaluation of other brain connectivity techniques that can be more computationally efficient for online analysis. Finally, a key challenge for exploiting these results in clinical settings is to efficiently separate the Propofol-induced sedation states

from their effect on the level of consciousness.

REFERENCES

- [1] C. Gurudatt, "Sedation in intensive care unit patients: Assessment and awareness," *Indian journal of anaesthesia*, vol. 55, no. 6, p. 553, 2011.
- [2] Y. Shehabi *et al.*, "Sedation depth and long-term mortality in mechanically ventilated critically ill adults: a prospective longitudinal multicentre cohort study," *Intensive care medicine*, vol. 39, no. 5, pp. 910–918, 2013.
- [3] L. M. Aitken and A. P. Marshall, "Monitoring and optimising outcomes of survivors of critical illness," *Intensive and Critical Care Nursing*, vol. 31, no. 1, pp. 1–9, 2015.
- [4] G. Gürkan *et al.*, "Analysis of brain connectivity changes after propofol injection by generalized partial directed coherence," *Digital Signal Processing*, vol. 25, pp. 156–163, 2014.
- [5] P. Guldenmund *et al.*, "Propofol-induced frontal cortex disconnection: A study of resting-state networks, total brain connectivity, and mean bold signal oscillation frequencies," *Brain connectivity*, vol. 6, no. 3, pp. 225–237, 2016.
- [6] X. Liu *et al.*, "Variation of the default mode network with altered alertness levels induced by propofol," *Neuropsychiatric disease and treatment*, vol. 11, p. 2573, 2015.
- [7] D. Rathee *et al.*, "Propofol-induced sedation diminishes the strength of frontal-parietal-occipital eeg network," in *Engineering in Medicine and Biology Society (EMBC), 2017 39th Annual International Conference of the IEEE*. IEEE, 2017, pp. 4463–4466.
- [8] S. Chennu *et al.*, "Brain connectivity dissociates responsiveness from drug exposure during propofol-induced transitions of consciousness," *PLoS Comput Biol*, vol. 12, no. 1, p. e1004669, 2016.
- [9] S. Guo *et al.*, "Partial granger causality-eliminating exogenous inputs and latent variables," *J. of Neuroscience Methods*, vol. 172, no. 1, pp. 79–93, 2008.
- [10] F. Perrin *et al.*, "Spherical splines for scalp potential and current density mapping," *Electroencephalogr. Clin. Neurophysiol.*, vol. 72, no. 2, pp. 184–187, 1989.
- [11] J. Kayser, "Current source density (CSD) interpolation using spherical splines–CSD toolbox (version 1.1)," *New York State Psychiatric Institute: Division of Cognitive Neuroscience*, 2009.
- [12] D. Rathee *et al.*, "Current source density estimation enhances the performance of motor-imagery-related brain–computer interface," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 12, pp. 2461–2471, 2017.
- [13] —, "Estimation of effective fronto-parietal connectivity during motor imagery using partial granger causality analysis," in *Neural Networks (IJCNN), 2016 International Joint Conference on*. IEEE, 2016, pp. 2055–2062.
- [14] M. Morf *et al.*, "Recursive multichannel maximum entropy spectral estimation," *IEEE Trans. Geosci. Electron*, vol. 16, no. 2, pp. 85–94, 1978.
- [15] G. Schwarz, "Estimating the dimension of a model," *The annals of statistics*, vol. 6, no. 2, pp. 461–464, 1978.
- [16] D. Rathee *et al.*, "Single-trial effective brain connectivity patterns enhance discriminability of mental imagery tasks," *Journal of neural engineering*, vol. 14, no. 5, p. 056005, 2017.
- [17] T. Fawcett, "An introduction to roc analysis," *Pattern recognition letters*, vol. 27, no. 8, pp. 861–874, 2006.
- [18] A. K. Seth, "A MATLAB toolbox for granger causal connectivity analysis," *Journal of neuroscience methods*, vol. 186, no. 2, pp. 262–273, 2010.
- [19] A. Delorme and S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis," *Journal of neuroscience methods*, vol. 134, no. 1, pp. 9–21, 2004.
- [20] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, pp. 273–297, 1995.
- [21] D. J. C. MacKay, "Bayesian interpolation," *Neural Comput.*, vol. 4, no. 3, pp. 415–447, 1992.
- [22] G. B. Huang *et al.*, "Extreme learning machine for regression and multiclass classification," *IEEE Trans. on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 42, no. 2, pp. 513–529, 2012.
- [23] M. Alhamdoosh and D. Wang, "Fast decorrelated neural network ensembles with random weights," *Information Sciences*, vol. 264, pp. 104–117, 2014.