



HAL
open science

Automatic Cognitive Workload Classification Using Biosignals for Distance Learning Applications

Rui Varandas, Hugo Gamboa, Inês Silveira, Patrícia Gamboa, Cláudia Quaresma

► **To cite this version:**

Rui Varandas, Hugo Gamboa, Inês Silveira, Patrícia Gamboa, Cláudia Quaresma. Automatic Cognitive Workload Classification Using Biosignals for Distance Learning Applications. 12th Doctoral Conference on Computing, Electrical and Industrial Systems (DoCEIS), Jul 2021, Costa de Caparica, Portugal. pp.254-261, 10.1007/978-3-030-78288-7_24 . hal-03685934

HAL Id: hal-03685934

<https://inria.hal.science/hal-03685934>

Submitted on 2 Jun 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Automatic Cognitive Workload Classification Using Biosignals for Distance Learning Applications

Rui Varandas^{1,2}, Hugo Gamboa^{1,2}, Inês Silveira², Patrícia Gamboa^{1,2},
Cláudia Quaresma²

¹ PLUX Wireless Biosignals S.A., Lisbon, Portugal

² LIBPhys-UNL, Departamento de Física, Faculdade de Ciências e Tecnologia,
FCT, Universidade Nova de Lisboa, Caparica, Portugal
rvarandas@plux.info

Abstract. Current e-learning platforms provide recommendations by applying Artificial Intelligence algorithms to model users' preferences based on content, by collaborative filtering, or both, thus, do not consider users' states, such as boredom. Biosignals and Human-Computer Interaction will be used in this study to objectively assess the state of the user during a learning task. Preliminary data was obtained from a small sample of young adults using physiological sensors (e.g., electroencephalogram, EEG, and functional near infrared spectroscopy, fNIRS) and computer interfaces (e.g., mouse and keyboard) during cognitive tasks and a Python tutorial. Using Machine Learning (ML), Cognitive Workload was classified considering EEG and fNIRS. The results show that it is possible to automatically distinguish cognitive states with accuracy around 84%. This procedure will be applied to adjust the difficulty level of learning tasks, model user preferences, and ultimately optimize the distance learning process in real-time, in a future e-learning platform.

Keywords: distance-learning, biosignals, artificial intelligence, machine learning, human-computer interaction.

1 Introduction

E-learning has seen an increase in popularity in recent years, however, typical systems fail to monitor students' engagement [1]. On e-learning, the absence of a tutor's in-person supervision can lead to students' inattention. Thus, this study will consider the best variables that maximize engagement: attention, cognitive workload (CW), and emotions, which enables the creation of more comfortable and successful learning atmospheres [1],[2],[3],[4].

Our investigation proposes cognitive workload and user's emotional state assessment by using a set of sensors: functional Near-Infrared Spectroscopy (fNIRS), Electroencephalography (EEG), Accelerometer to measure the movement of the head, Respiratory Inductive Plethysmography (RIP), Electrocardiogram (ECG) and finally, Electrodermal Activity (EDA).

The PhD that inspires this work focuses on answering the question: How can the internal state of a person be included in e-learning to improve and optimize engagement and, thus, learning? Therefore, the proposed hypothesis is that the enumerated biosignals will allow to access the internal state of the learner and help to build a system tailored to each individual in different moments. Artificial intelligence (AI) algorithms, namely machine learning (ML) and deep learning (DL) models will be used for the classification of the state of the learners.

This work will focus on CW assessment using a low number of channels of fNIRS and EEG sensors to validate their performance, which contrasts with most current approaches that use a high number of channels, such as pointed out in [9] and [10]. Thus, the objectives of the current work are (1) to define an appropriate data acquisition scenario that properly distinguishes rest and task periods, (2) to analyse fNIRS and EEG signals during different states, namely, engagement vs. baseline and finally (3) to use ML algorithms to automatically identify the state of the participants during the data acquisition procedure. If the third objective is met, then this simpler setup, can be used in e-learning to adequate the contents to the instantaneous state of the user, allowing to optimize engagement and the learning experience.

This article is structured as follows: (2) description of the relationship between this work and the conference theme; (3) state-of-the-art about cognitive monitoring in learning environments; (4) methods and materials applied to this study; (5) results of our analysis; (6) discussion of the results; (7) final conclusions and future work.

2 Relationship to Applied Artificial Intelligence Systems

Since tutors are not present to guide the learning process, Recommendation Systems may be applied in e-learning platforms, referred as Intelligent Tutoring Systems (ITS), which can model learners' interests and states through AI and adapt course contents to maximize learning [9]. This work innovation lies on the integration of different biosignals from learners to identify cognitive states' changes and, subsequently, model learners' optimal state (for learning) and adapt e-learning content accordingly, through the application of AI models.

In this preliminary study, we will focus on the analysis of two specific biosignals with low number of channels - EEG and fNIRS - allied with one ML model to validate the possibility of applying them for automatic cognitive state assessment.

3 Literature Review

Attention plays a crucial role in learning performance [1],[2]. The work of [1] aimed at assessing the student's attention levels based on features extracted from γ , β , α and θ waves of EEG signals, achieving an average accuracy rate of 89.52%.

Related to attention, CW can be defined as the relationship between mental processing capabilities and the performed task demands, being inversely related to performance [5],[6]. From a neurophysiological point of view, CW is directly

associated with the variation of the concentration of oxygenated haemoglobin (HbO) and inversely with the concentration of deoxygenated haemoglobin (Hb) - effect of neurovascular coupling [7],[10].

To measure CW, fNIRS has been increasingly used [11] and the prefrontal cortex (PFC) is commonly the area assessed as it encompasses executive functions, i.e. high-order cognitive processes [12].

In [13] fNIRS was used to monitor participants during various psychological tasks, with accuracy results between 62-71% for the discrimination of tasks and baseline. In turn, the study [16] obtained an accuracy of 78% using only N-Back tasks. According to the study, CW induced by a memory task, N-Back, and fNIRS-based evaluation of PFC constitute a realistic method for passive Brain Computer Interface - BCI [14]. In [6], EEG was used, and results suggested that α waves were the most appropriate brain waves for measuring CW in learning-based cognitive tasks and for discriminating different learning states. Furthermore, recent literature has highlighted the significance of emotions in e-learning environments [4]. Some studies advise research on emotions in e-learning to focus on negative activating emotions (e.g. frustration) and on positive activating emotions (e.g. enjoyment) [15],[16]. While most studies on this topic rely on surveys, [3] adds that EDA and ECG might be the future means of accessing emotions.

4 Materials and Methods

4.1 Experimental Setup

According to the literature, EEG and fNIRS are the sensors that better allow to measure cognitive processes while ECG and EDA can be used to infer emotional regulation responses. We used these sensors along with a RIP band around the upper abdominal region and an accelerometer attached to the head of the participants to measure changes in posture and head movements.

EEG and fNIRS sensors were employed using two channels each (reducing the number of channels to create a close to natural learning-environment), positioned around the F7 and F8 positions of the 10-20 system [17]. Electronic devices and noise sources such as WiFi or Bluetooth devices were not near the sensors to avoid noise. The ECG sensor was applied to measure the lead I of the Einthoven system. Lastly, the EDA sensor was placed in the palm of the non-dominant hand, reducing movement constraints during computer interaction.

N-Back (described in [14]) and Mental Subtraction, were employed as standard cognitive tasks using PsychoPy [18]. Rest periods of 60 s were implemented before, between and after the two main tasks and of 20 s between tasks' explanations and procedure to avoid contamination from reading the instructions. Finally, a rest period of 10 s between difficulty levels of N-Back and between subtraction periods was introduced. Regarding the number of trials, the N-Back task consisted of 4 levels with 60 trials each and mental subtraction had 20 subtraction periods of 10 s, in which participants had to continuously subtract a given number from the result of the previous subtraction while a visual cue was shown.

Finally, participants performed a close-to-real learning task that involved the execution of a Python tutorial with theoretical and practical examples. During this part of the acquisition procedure, besides the physiological sensors, human-computer interaction (HCI) was monitored using Latent [19], which monitors mouse movement, keyboard presses, screenshots, audio, and snapshots using the webcam of the computer. A questionnaire was filled in at the end of the tutorial to get a rough individual self-assessment on thoughts regarding the tutorial, which served as ground-truth.

4.2 Data Recording

First, due to the current pandemic situation, all the equipment and room were disinfected at each acquisition, following government guidelines set by a public health entity (*Direção-Geral de Saúde*). Also, all ethical procedures were guaranteed. Data was collected using two biosignalsplux acquisition devices (from PLUX Wireless Biosignals) at 1000 Hz and 16-bit resolution. Regarding the fNIRS sensor, specifically, the two optodes emitted at 660 nm and 860 nm and the distance between emitters and detector was 2 cm. In this preliminary study, data acquisition was performed on a sample of 8 volunteering subjects (4 females), aged between 20 and 27 years-old ($M=22.9$, $SD=2.1$). All participants were right-handed, none reported to suffer from psychological or neurological disorders or taking medication other than contraceptive pills. All provided written informed consent and none objected wearing the sensors.

4.3 Signals Preprocessing and Features Extraction

To remove physiological artifacts (e.g., heart rate) and long period shifts, a 4th order Butterworth filter with cut-off frequencies of 0.01 and 0.2 Hz was first applied to the raw fNIRS signals. Thereafter, to reduce spikes noise of HbO and experimental errors due to optodes' movement and motion artifact, a 3rd order Savitzky-Golay filter, using a window size of 11 s was applied. The raw optical density signals were then converted to the concentration variations of HbO and Hb via modified Beer-Lambert law. Concerning EEG signals, data was filtered using a baseline shift and a bandpass filter with cut-off frequencies of 3 and 30 Hz. Finally, both fNIRS and EEG signals were normalized and segmented into windows of 2 s, thus originating 2583 samples.

Initially, 176 features from the temporal, statistical and spectral domain were extracted from the HbO, Hb and total haemoglobin (HbT) segments, but only the 15 most relevant features were included, by the feature selection method Select K Best, provided by the scikit-learn python library [20]. This method scores the features based on the distance between the means of different classes and variance of each single class [20].

5 Results

The differences in the fNIRS signals in different conditions will be explored (EEG waves will be explored in Section 6). For this purpose, the mean wave of the fNIRS signal in the F7 and F8 positions are shown in Fig. 1 in different procedural states, namely during the mental subtraction task and during rest states. These signals (from one participant) are a good example of the neurovascular coupling effect.

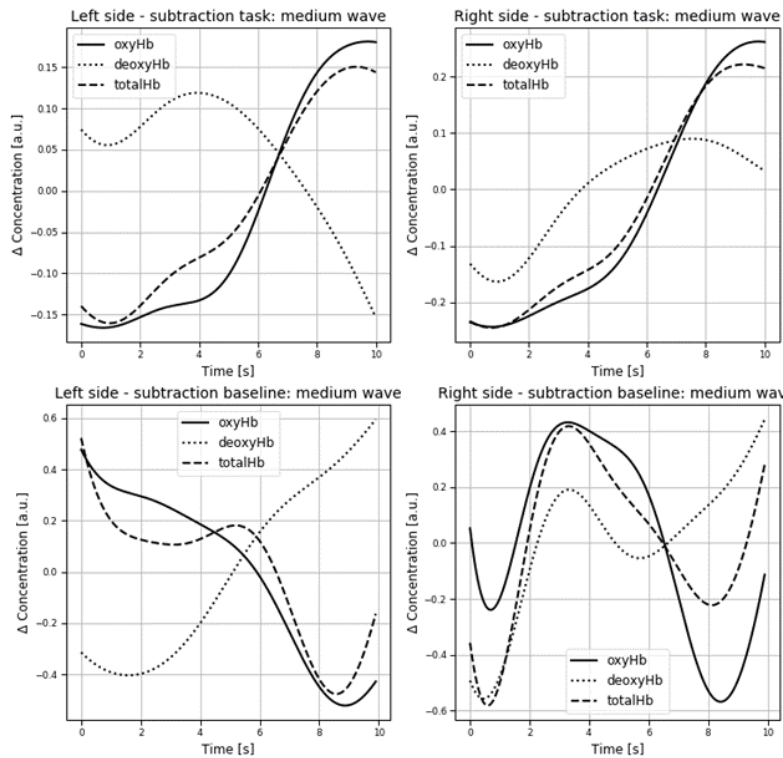


Fig. 1. Mean Waves of fNIRS signals from a single participant in the left side (F7) and right side (F8) of the PFC and during subtraction task and rest.

The classification performance using Random Forest method will be discussed taking into account the accuracy score, f-score, precision, recall and AUC-ROC. Random Forest consists of a combination of decision trees and uses averaging to improve accuracy and avoid over-fitting [21]. The chosen criterion was entropy and the number of estimators 100. Table 1 shows the results for the classification of the signals of all participants using only the fNIRS signal, only the EEG signals and using both signals. The results displayed correspond to the mean and standard deviation percentage of the 10-Fold cross-validation.

Table 1. Classification results using Random Forest, considering fNIRS, EEG and both signals after preprocessing and features extraction and selection.

Signal	Accuracy	F-Score	Precision	Recall	AUC-ROC
fNIRS	80.05 ± 1.59	88.33 ± 0.91	80.96 ± 0.98	97.17 ± 0.83	65.70 ± 5.04
EEG	83.55 ± 1.33	90.21 ± 0.81	83.84 ± 1.18	97.66 ± 1.68	82.50 ± 2.16
fNIRS + EEG	83.55 ± 1.33	90.21 ± 0.81	83.84 ± 1.18	97.66 ± 1.68	82.50 ± 2.16

Moreover, to assess if differences in the tasks may influence the classification results, the classification was performed in separate for the N-Back task and the Mental Subtraction task using both sensors and always comparing to the rest state (Table 2).

Table 2. Classification results using Random Forest, considering fNIRS and EEG signals after preprocessing and features extraction and selection considering the tasks in separate.

Task	Accuracy	F-Score	Precision	Recall	AUC-ROC
N-Back	76.16 ± 1.99	84.10 ± 1.50	76.92 ± 0.98	92.80 ± 2.76	76.30 ± 3.98
Subtraction	77.24 ± 3.43	82.21 ± 2.48	77.28 ± 3.60	88.00 ± 3.48	83.60 ± 4.82

6 Discussion

Our findings indicate that it is possible to differentiate the cognitive state of the monitored participants while performing a subtraction task and at rest state using a low number of fNIRS channels. The effects of the neurovascular coupling are clearly visible in these signals, expressed by an increase of the variation of the concentration of HbO and a decrease of Hb during task and the inverse for the baseline case. Moreover, PSD analysis of the EEG signals revealed differences during task and at rest states, enabling a distinction between both states. In fact, alpha (α) band power during mental subtraction is $2.92 \pm 1.62/3.78 \pm 3.12$, while at rest this value rises to $3.92 \pm 2.48/3.21 \pm 1.21$, in F7 and F8 positions, respectively. Since α activity is related to a relaxation state and to a decreased amount of attention resources allocated to the task, one would have expected a greater resting α power [1],[2],[22]. Additionally, an increased θ activity means higher mental fatigue and CW [1],[2],[23]. In effect, θ power during mental subtraction was $9.03 \pm 8.16/11.41 \pm 10.06$ and $8.90 \pm 7.65/7.78 \pm 6.88$ during rest, in F7 and F8 positions, respectively. Thus, the task load index, given by the θ/α ratio was 3.09/3.02 during the task and 2.27/2.42 at the resting state, in F7 and F8 positions, respectively [23].

Regarding the third identified objective of this study, ML algorithms could effectively contribute to detect CW changes and, more specifically, differentiate task and rest states. According to Table 1, a hybrid fNIRS + EEG approach resulted in a predictive model as good as the one using only EEG, reaching an accuracy of $83.55 \pm 1.33\%$, while using only fNIRS an accuracy score of $80.05 \pm 1.59\%$ was achieved. This happened because the Select K Best algorithm selected the same features in the hybrid dataset as in the dataset consisting only of EEG features. In addition, in line with Table 2, N-Back and subtraction task provided both a successful CW states differentiator, in which accuracies of $76.16 \pm 1.99\%$ and $77.24 \pm 3.43\%$ were accomplished. The higher

accuracy score of the whole dataset is due to the larger number of samples when compared to each task in separate.

7 Conclusions and Future Work

This work proposed a data acquisition scenario that mimics learning in order to develop passive BCI's, that can be used to design better user-adjustable e-learning platforms, which contrast with current solutions. A set of physiological sensors that help to infer attention, CW and emotional states, important in learning processes, was identified. In this preliminary study, only two sensors related to attention and CW, fNIRS and EEG, were analysed and allowed to conclude that brain functioning activity related to learning could in fact be inferred by their signals. Finally, it is possible to conclude that ML algorithms could automatically detect different states, specifically using both fNIRS and EEG signals. Furthermore, N-Back was more adequate to induce mental state changes than the mental subtraction case, based on the results of the ML algorithm. Notwithstanding the positive results, more participants will be included in future work, allowing to have more data to train the ML classifiers and to increase the statistical relevance of the work, currently limited by the low number of participants.

Next steps of investigation will include the use of other sensors to understand if they can, in fact, expand further the current results. Furthermore, the Python tutorial data, including HCI data should be analysed to understand if the results of the cognitive tasks can be extrapolated to real e-learning scenarios. Finally, the resulting classifiers will be used in an e-learning context, where users will be monitored and the classifiers will infer about learners' cognitive states, allowing for the continuous update of the e-learning content in real-time to optimize engagement and, thus, learning.

Acknowledgements. This work was partly supported by Fundação para a Ciência e Tecnologia, under PhD grants PD/BDE/150304/2019 and PD/BDE/150672/2020 and by PLUX Wireless Biosignals S.A.

References

1. Chen, M.C, Wang, Y.J. and Yu, M.C.: Assessing the attention levels on students by using a novel attention aware system based on brainwave signals. *British Journal of Educational Technology* 48.2, 348-369 (2015)
2. Liu, H.N., Chiang, Y.C. and Chu, C.H.: Recognizing the degree of human attention using EEG signals from mobile sensors. *Sensors* 13.2, 10273--10386 (2013)
3. Mayer, E.R.: Searching for the role of emotions in e-learning. *Learning and Instruction* 70, 101--213 (2020)
4. D'Errico, F., Paciello, M. and Cerniglia, L.: When emotions enhance students' engagement in e-learning processes. *Journal of e-Learning and Knowledge Society* 12.4 (2016)
5. Maior, H., Pike, M., Wilson. L. M. and Sharples, S.: Continuous detection of workload overload. *Contemporary Ergonomics and Human Factors*. 450-457 (2014)

6. Mazher, M., Aziz, A.A., Malik, S.A. and Amin, U.H.: An EEG-Based Cognitive Load Assessment in Multimedia Learning Using Feature Extraction and Partial Directed Coherence. *IEEE Access* 5, 14819-14829 (2017)
7. Herold, F., Wiegel, P., Scholkmann, F. and Müller, N.G.: Applications of Functional Near-Infrared Spectroscopy (fNIRS) Neuroimaging in Exercise–Cognition Science: A Systematic, Methodology-Focused Review. *J. Clin. Med.* (2018)
8. C. Q. Lai, H. Ibrahim, M. Z. Abdullah, J. M. Abdullah, S. A. Suandi, and A. Azman, "Literature survey: Recording set up for electroencephalography (EEG) acquisition," *ISCAIE 2018 - 2018 IEEE Symp. Comput. Appl. Ind. Electron.*, pp. 333–338, 2018.
9. S. Kulkarni and S. F. Rodd, "Context Aware Recommendation Systems: A review of the state of the art techniques," *Comput. Sci. Rev.*, vol. 37, p. 100255, 2020.
10. Bracken, B., Festa, E., Sun, M.H., Leather, C. and Strangman, G.: Validation of the fNIRS Pioneer™, a Portable, Durable, Rugged functional Near-Infrared Spectroscopy (fNIRS) Device. In: *Special Session on Real-world Assessment of Individuals During Everyday Routines*, 521--531 (2019)
11. Pinti, P., Tachtsidis, I., Hamilton, A., Hirsch, J., Aichelburg, C., Gilbert and Burgess, W. P.: The present and future use of functional near-infrared spectroscopy (fNIRS) for cognitive neuroscience. *Ann. N.Y. Acad. Sci.* (2018)
12. Moriguchi, Y. and Hiraki, K.: Prefrontal cortex and executive function in young children: a review of NIRS studies. *Front. Hum. Neurosci.* (2013)
13. Herff, C., Dominic, H., Felix, P., Johannes, H., Fortmann, O. and Schultz, T.: Classification of mental tasks in the prefrontal. In: *35th Annual International Conference of the IEEE EMBS*, pp. 3--7, Japan (2013)
14. Herff, C., Heger, D., Fortmann, O., Hennrich, J., Putze, F. and Schultz, T.: Mental workload during n-back task—quantified in the prefrontal cortex using fNIRS. *Front. Hum. Neurosci.* (2014)
15. Duffy, L., PekM. C., Pekrun, S. P., Pekrun, R., and Lachapelle, K.: Emotions in medical education: Examining the validity of the Medical Emotion Scale (MES) across authentic medical learning environments. *Learn. Instr.* 70 (2020)
16. Harley, J. M., Lajoie, S. P., Tressel, T. and Jarrell, A.: Fostering positive emotion and history knowledge with location-based augmented reality and tour-guide prompts. *Learn. Instr.* (2018)
17. Acharya, N. J., Hani, A., Cheek, J., Thirumala, P. and Tsuchida N. T.: American Clinical Neurophysiology Society Guideline 2: Guidelines for Standard Electrode Position Nomenclature. *J. Clin. Neurophysiol.* 33.4, pp. 308--311 (2016)
18. Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., Lindeløv, J. (2019). *PsychoPy2: experiments in behavior made easy.* *Behavior Research Methods.* 10.3758/s13428-018-01193.
19. Cepeda, C.: *Latent: A Flexible Data Collection Tool to Research Human Behavior in the Context of Web Navigation.* *IEEE Access* 7, pp. 77659--77673 (2019)
20. Pedregosa, F., Varoquaux, G., Gramfort, A., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. and Duchesnay, E. *Scikit-learn: Machine Learning in Python.* *J. Mach. Learn. Resear.* 12, 2825-2830 (2011)
21. Breiman, L.. *Random Forests.* *Machine Learning* 45, pp. 5--32 (2001)
22. Benedek, M., Schickel, J. R., Jauk, E., Fink, A. and Neubauer, C.A.: Alpha power increases in right parietal cortex reflects focused internal attention. *Neurop.* 56.100, 393--400 (2014)
23. Fernandez, R., Rojas, E., Debie, E., Fidock, J., Barlow, M., Kasmarik, K., Anavatti, S., Garratt, M. and Abbass, H.. *Electroencephalographic Workload Indicators During Teleoperation of an Unmanned Aerial Vehicle Shepherding a Swarm of Unmanned Ground Vehicles in Contested Environments.* *Front. Neurosci.* 14 (2020)