Towards Edge-Computing Assessment of Cognitive Workload Using fNIRS Data

Tanya S. Paul
Thales Research and Technology Canada
Québec, Canada
tanya.paul@thalesgroup.com

Laura Salvan
Thales Research and Technology Canada
Québec, Canada
laura.salvan@thalesgroup.com

Maëlle Kopf
Thales Research and Technology Canada
Québec, Canada
maelle.kopf@thalesgroup.com

Danielle Benesch
Thales Research and Technology Canada
Québec, Canada
danielle.benesch@thalesgroup.com

Alexandre Marois
Thales Research and Technology Canada
Québec, Canada
alexandre.marois@thalesgroup.com

***Abstract*— Situations of high cognitive workload can induce errors. This is widely prevalent in safety-critical domains where operators must make high-stake decisions while facing complex, dynamic situations. Therefore, it seems critical to detect instances of high cognitive workload to provide timely assistance to operators when they face such error-prone situations. This, however, requires capacities for real-time evaluation, which can be performed using brain-computer interface systems driven by measures of the central nervous system. Functional near-infrared spectroscopy (fNIRS) measures have been shown to be associated with variations in cognitive workload, but these are typically processed and analyzed offline in a post-hoc fashion. Besides, to be used in operational situations, workload evaluation should be performed with mobile devices, supported by edge-computing devices. The goal of this paper is to present the first steps toward a real time, mobile assessment of cognitive workload using fNIRS data. We present two models developed using the Open Neural Network Exchange format that allows edge inference of cognitive workload with fNIRS data. We present prediction performance of the models and show the possibility to integrate the model into the Sensor Hub platform, a real-time sensor-agnostic data integration, synchronization, and processing nexus that allows sampling data from multiple sensors and users simultaneously for operational use cases. More specifically, we show that the model can produce workload inference from the fNIRS signal within the one-second latency requirement of the Sensor Hub platform.**

Keywords—Cognitive workload, fNIRS, Edge computing, Real time

I. Introduction

Human cognitive resources have been widely shown to be limited. Cognitive workload represents the association between the amount of mental processing resources and the amount required by a task [1]. When demands exceed the amount of resources available to process information, errors can occur. Cognitive workload can be a key component for safety-critical operational domains (e.g., in aviation, defence, driving, or command and control) where high-stake decisions must be made in highly complex situations. Functional near-infrared spectroscopy (fNIRS), a low cost, portable system allowing for measures of brain blood oxygenation can be used to measure brain activity according to variations in cognitive workload [2]. It uses near-infrared light to determine changes in oxyhemoglobin (HbO_2) and deoxyhemoglobin (HbR). Proportion of hemoglobin saturation in tissue (StO_2) can also be calculated. This study presents the first step towards real-time, edge-computing assessment of cognitive workload with fNIRS data. Our study optimizes the implementation of our model with the Open Neural Network Exchange (ONNX) format, which is integrated into a Java-based platform allowing for which allows for a complete pipeline from signal acquisition, signal preprocessing, feature extraction and selection, to modelling and model deployment, onto wearable devices. To reach this goal, we report two proof-of-concept machine-learning models to show our capability to predict cognitive workload from measures collected from an fNIRS mobile device while participants performed a task involving two levels of cognitive workload during an *n*-back task.

II. Method

Eight participants performed dual *n*-back, i.e. a working memory task that requires responding when a given visual and/or audio stimulus presented is the same as the one presented *n* positions before. On each trial, squares were presented sequentially for 500 ms, with an inter-stimulus interval of 2,500 ms, within a 3 × 3 matrix on the computer screen. A letter was concurrently spoken in the speakers. In the 1-back condition (low load), participants were asked to respond when a square was presented at the same position as the trial before and when the letter spoken in the speakers was the same as the one presented the trial before. In the 2-back condition (high load), the same was asked but with two trials before. Participants pressed the “A” key on the keyboard for the visual targets and the “L” key for the audio targets.

Participants’ brain activity was measured with the Artinis OctaMon fNIRS. As shown in Fig. 1, the device is composed of eight transmitters (Tx), transmitting two wavelengths (±760 nm and ±850 nm), and two receivers (Rx). The OxySoft software provided by Artinis collected the raw data from the fNIRS and was used to inspect the data prior to signal processing. Then, the data was preprocessed. The OxySoft software applies an algorithm based on the modified Beer-Lambert law in order to transform the raw signal into concentration change in hemoglobin. The fNIRS signal was also passed through a third-order digital Butterworth filter with a low cut-off 0.008 Hz and high cut-off 0.2 Hz [3].

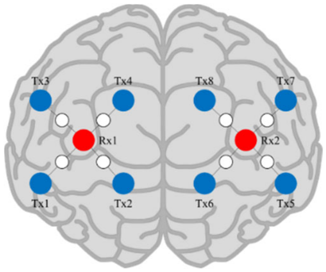

Fig. 1. Optode placement for the transmitter (Tx) and receiver (Rx).

The following features were extracted: i) HbO_2 values for each optode; ii) HbR values for each optode; iii) StO_2 values for each optode; and iv) mean StO_2 by sub-regions with mean saturation for the right PFC (Subregion A with Tx1, Tx2, Tx3 and Tx4), median PFC (Subregion B with Tx1, Tx2, Tx4, Tx5, Tx6 and Tx8), and left PFC (Subregion C with Tx5, Tx6, Tx7 and Tx8). Forward sequential feature selection followed by random-search based hyperparameter tuning was applied to identify the best set of features and hyperparameters using the leave-one-out method (K-fold = 8) for comparing accuracy of the random forest models generated. A second random forest model was also developed using moving time windows of 20 sec (with a hop size of 1 sec) for extracting and transforming the fNIRS signal into mean, kurtosis and skewness values. To validate significant connection between the data and class labels, we ran permutation tests where the final model selected was trained and evaluated with 100 permutations of the class labels using scikit-learn [4]. The model characterized by the best performance was then extracted to the ONNX format. The ONNX model pipeline (pre-processing and feature extraction) was implemented in Java to be fully operational within the Sensor Hub [5], a real-time sensor-agnostic data integration, synchronization, and processing nexus that allows sampling data from multiple sensors and users. This allowed for a proof-of-concept for providing real-time workload assessments using data collected by the fNIRS headband.

III. Results and Discussion

Results on the dual *n*-back task showed that the two load tasks induced the anticipated effects. Indeed, the mean performance for the low load condition (1-back), 94.81% (SD = 6.19), was significantly inferior to that of the high load (2-back) condition, 69.94% (SD = 15.49), p = .012. Such a difference supports that participants faced higher workload in the 1-back condition, compared with the 2-back condition.

The best model produced with the leave-one-out approach was comprised of only the HbR values of the eight optodes. The model reached a prediction accuracy of 62.5% following forward sequential selection procedure of the three best features (Tx8 [$\beta = 0.41$], Tx7 [$\beta = 0.36$], and Tx3 [$\beta = 0.23$]). A permutation test ran on 100 permutations of the class labels suggested a significant connection between the data and labels (p = .009). The mean, kurtosis, and skewness of the three features selected by the previous model were calculated within the 20-sec epoch. Of the resulting nine features, the following were empirically selected, leading to a prediction accuracy of 66.9%: Tx8_Mean ($\beta = 0.43$), Tx8_Kurtosis ($\beta = 0.29$), and Tx3_Kurtosis ($\beta = 0.28$). Permutation tests supported the model's ability to provide significant classifications (p = .01). This outlines the possibility to predict two cognitive workload levels from fNIRS measures. The fact that these models were reached with very few participants also provide support for the validity of this approach. Besides, the importance ascribed to the Tx8 and Tx7 optodes is consistent with previous literature outlining important activity in the left prefrontal cortex in periods of high cognitive workload [6].

To test our ability to perform real-time assessments on an edge-computing device, we integrated the first model within the Sensor Hub by exporting the model in an ONNX format and implementing data preprocessing through the solution, in Java format. Simulations of the fNIRS device driver were conducted in order to confirm that the preprocessing and prediction could run with continuously streaming data. Tests showed that the model could provide cognitive workload inference from the fNIRS signal within the one-sec latency requirement of the platform. This supports the operational relevance of favouring an edge-computing approach such as the one developed within this study.

Overall, the study supports the possibility for real-time collection, extraction, preprocessing, and analysis of fNIRS data to infer cognitive workload. The opportunity to run an entire pipeline within an edge-computing device offers potential for adaptive teaming or for workload levels depiction on context-tailored dashboards to give actionable information on end users. This could provide many benefits for uses cases such as fitness for duty evaluation, neuropsychological status monitoring, personnel selection, and self-regulation biofeedback [7]. Future work will aim to integrate baseline capacities as well as models with epochs in the Sensor Hub. Model improvement will also be reached by collecting further data on real end users to train a more generalizable model from a dataset closer to relevant use cases.

IV. References

[1] S. G. Hart, and L. E. Staveland, "Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research," Adv. Psychol., vol. 52, pp. 139–183, 1988.

[2] S. Karmakar, S. Kamilya, P. Dey, P. K. Guhathakurta, M. Dalui, T. K. Bera, S. Halder, C. Koley, T. Pal, and A. Basu, "Real time detection of cognitive load using fNIRS: A deep learning approach," Biomed. Signal Process., vol. 80, 104227, Feb. 2023.

[3] S. Montero-Hernandez, F. Orihuela-Espina, L. E. Sucar, P. Pinti, A. Hamilton, P. Burgess, and I. Tachtsidis, "Estimating functional connectivity symmetry between oxy- and deoxy-haemoglobin: implications for fNIRS connectivity analysis," Algorithms, vol. 11, 70, May 2018.

[4] M. Ojala, and G. C. Garriga, "Permutation tests for studying classifier performance," J. Mach. Learn. Res., vol. 11, pp. 1833–1863, June 2010.

[5] J.-F. Gagnon, D. Lafond, M. Rivest, F. Couderc, and S. Tremblay "Sensor-Hub: A real-time data integration and processing nexus for adaptive C2 systems," ADAPTIVE 2014, pp. 63–67, 2014.

[6] M. J. Saikia, W. G. Besio, and K. Mankodiya, "The validation of a portable functional NIRS system for assessing mental workload," Sensors, vol. 21, 3810, May 2021.

[7] L. Salvan, A. Marois, M. Kopf and J.-F. Gagnon, "Sensors-Enabled Human State Monitoring System for Tactical Settings," 2022 IEEE CogSIMA, Salerno, Italy, 2022, pp. 55–61.