Neuroscience Informatics 2 (2022) 100042

Contents lists available at ScienceDirect

Neuroscience Informatics

www.elsevier.com/locate/neuri

CINTDA

Statistical valuation of cognitive load level hemodynamics from functional near-infrared spectroscopy signals

Farzana Khanam [a,*], A.B.M. Aowlad Hossain [b], Mohiuddin Ahmad [c]

[a] Department of Biomedical Engineering, Khulna University of Engineering & Technology (KUET), Khulna-9203, Bangladesh
[b] Department of Electronics and Communication Engineering, Khulna University of Engineering & Technology (KUET), Khulna-9203, Bangladesh
[c] Department of Electrical and Electronic Engineering, Khulna University of Engineering & Technology (KUET), Khulna-9203, Bangladesh

ARTICLE INFO

Article history:
Received 10 December 2021
Received in revised form 12 January 2022
Accepted 13 January 2022

Dataset link: https://www.nature.com/articles/sdata20183

Keywords:
Functional near-infrared spectroscopy (fNIRS)
Hemodynamics
Cognitive load
n-back test
ANOVA test
Grand average

ABSTRACT

Human cognitive load level assessment is a challenging issue in the field of functional brain imaging. This work aims to study different cognitive load levels statistically from brain hemodynamics. Since the functional brain activities can be evaluated by functional near-infrared spectroscopy (fNIRS), a renowned fNIRS dataset is considered for this work. The dataset contains fNIRS data of three types of *n*-back tasks (0-back, 2-back, and 3-back) of twenty-six healthy volunteers. The fNIRS signals were pre-processed and separated according to the tasks and trials. The mean changes of oxygenated hemoglobin (HbO_2) and deoxygenated hemoglobin (dHb) are calculated from each trial corresponding to the tasks and tested for significant inference among three levels utilizing analysis of variance (ANOVA). From the outcomes of the ANOVA ($p < 0.005$), two significant channels (AF7 (frontal) and C3h (motor)) were figured out. The significance of these two channels was further justified using the property consistency test by three different time intervals of hemodynamics inside the total task period. The latter result also explored the functional pattern of the hemodynamics of AF7 and C3h positions. Moreover, two-level cognitive load (due to easy i.e., 0-back test and hard i.e., 2-back and 3-back task) is classified using support vector machine and found classification accuracy in average 73.40%±0.076 for HbO_2 data and 71.48%±0.061 for dHb data. The study signposts the collective role played by both fNIRS signals and statistical valuation of functioning cognitive load efficacy to use fNIRS as a cognitive load assessment biomarker.

1. Introduction

Human brain health is a significant issue in the modern era. The basic and important part of the human brain signifies cognitive health which features the ability to think, learn, and remember clearly. But the use of modern technology in regular working and social life builds a lot of pressure on cognitive health and demands more cognitive activity. As a consequence, the over-cognitive load may compromise human performance and welfare through increasing error rates and fatigue, a decline in motivation including higher reaction times [1–3]. Since human cognitive characteristics and limits are in the acute phase which yields progress in the design and operation of brain-machine interfaces (BMI), are reactive to the changes of human cognitive load [3–6].

The cognitive activity comes from the term 'Cognition', which means the psychological action or process of knowledge acquisition, understanding the utilized thoughts and senses; also respond perceiving all actions [7–11]. According to cognitive psychology, the used amount in working memory denotes cognitive load [12]. Basically, the elderly, students, and children experience different and sometimes higher amounts of cognitive load. At this time humans are more exposed to experience high cognitive load through excessive smartphone use, multi-tasking, unhealthy lifestyle with increased distractions, which can converse all technological achievements of mankind [13]. That's why many researchers advanced several ways to determine perceived mental effort signifies cognitive load [14]-[15].

Cognitive load roots basic brain functions which are related to the variation of oxygen saturation of the blood called hemodynamics. Concentration variation of oxygenated hemoglobin (HbO_2) and deoxygenated hemoglobin (dHb) provide information about brain functioning. Functional brain imaging performs well based on hemodynamic properties. Functional neuroimaging provides the functional activities of the brain, not the structure of the brain tissue. Considering different attributes with other techniques, functional near-infrared spectroscopy (fNIRS) is a hemodynamics-based functional brain imaging modality owing to its' good spatial resolution (1∼1.5 cm). The fNIR modality also involves modest temporal

* Corresponding author.
E-mail address: farzanabme@just.edu.bd (F. Khanam).

https://doi.org/10.1016/j.neuri.2022.100042

resolution (up to 100 Hz), the worthy value of the signal to noise ratio (SNR), fewer motion artifacts including low cost and portable facility [16,17]. Several scientists have paid attention to fNIRS as a medium to detect human cognitive load detection. In [18], the researchers have investigated different cognitive levels using *n*-back tasks through electroencephalography (EEG) and fNIRS signals with limited participants. The oxygenation activity level of fNIRS over different *n*-back tasks has been determined in [19] using only eight subjects. In [20], it has been studied that the prefrontal cortex area has been used as a location to detect mental workload with ten participants, and similarly *n*-back tasks were used as earlier researches. In the work described in [21], the results of analyzing the cognitive load with a set of characteristic tasks with smartphone activity via six subjects.

From the aforementioned literature review, it is quite obvious to realize about the limited participants' study with lack of females have been mined about cognitive load analysis. To overcome this concern, we have studied a trustworthy fNIRS dataset (available in [22]) of twenty-six healthy participants including both male and female. The dataset is comprised of three types of *n*-back tasks (0-back, 2-back, and 3-back test) including nine trials for each subject for the cognitive workload detection. At first, the dataset was organized according to subjects and trials. For preprocessing, the raw signals were filtered as well as corrected the baselines. The sensitive areas of the brain have been located according to the different levels of brain stimulation. Statistical procedures were involved to signify hemodynamic responses from the context of oxygenation and de-oxygenation level including detection of the signal pattern. To meet the requirement of statistical inference, a one-way analysis of variance (ANOVA) test was performed to signify the effective channels responsible for different cognitive loads. Based on the significant channels, the pattern of the relative changes in the concentration of HbO_2 and dHb is valuated. We also found that the frontal and motor area of the brain surface directly functions conferring to the hemodynamics activity due to different cognitive loads. To recognize the input data conferring three categories of *n*-back tasks, an SVM classifier is applied to all participants for binary classification. The promising classification accuracies validated fNIRS signals based brain functional load analysis. The outcomes of this study guide to utilizing the fNIRS technology as a biomarker for different cognitive load level investigations.

The remaining paper is organized as follows: Section 2 describes the materials and methodology. In section 3, experimental results and respective discussions are provided. Finally, the work is concluded in section 4.

2. Materials and methods

2.1. Materials

In our experimental study, visual *n*-back task, a classic memory game, was selected for the cognitive load assessment. Among different cognitive load tasks, *n*-back is one of the most common brain stimulators which was first introduced by Kirchner in 1958 [18,24]. In this task, a sequence of single letters is visualized back-to-back with a certain time interval. After each letter, the participants have to decide by pressing the key whether it is identical to the letter that appeared *n* items back in the sequence. Literature-wise, the 0-back task is used as a control state of the brain. With the ascending value of *n* e.g., 1-back, 2-back, 3-back, etc. tasks become tougher for the human brain to remember gradually.

For experimental protocol, a benchmark fNIRS-EEG signal dataset of different cognitive load tasks is considered, whereas only fNIRS data are extracted for an *n*-back task using cognitive load

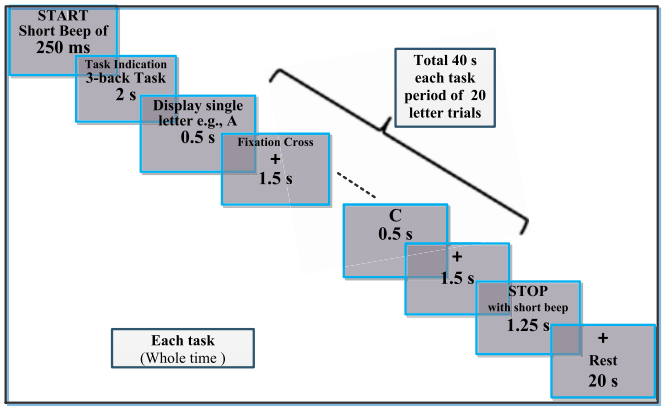

Fig. 1. The time sequence of each *n*-back task.

evaluation. The open-access dataset is collected from nature scientific data journal website [22] and it can be freely downloaded on top of open access repository [23]. The dataset consists of three sessions, each containing three series of 0-, 2-, and 3-back tasks in the following order of $\{(3 \rightarrow 2 \rightarrow 0) \rightarrow (2 \rightarrow 0 \rightarrow 3) \rightarrow (0 \rightarrow 3 \rightarrow 2)\} \rightarrow \{(2 \rightarrow 0 \rightarrow 3) \rightarrow (0 \rightarrow 3 \rightarrow 2) \rightarrow (3 \rightarrow 2 \rightarrow 0)\} \rightarrow \{(0 \rightarrow 3 \rightarrow 2) \rightarrow (3 \rightarrow 2 \rightarrow 0) \rightarrow (2 \rightarrow 0 \rightarrow 3)\}$. Here, the first bracket of *n*-back tasks indicates one series and the second bracket implies one session. Therefore, each subject performed (3 sessions $\times$ 3 series) = a total of 9 series for each *n*-back task (i.e., 0, 2, and 3) which means for a total of three *n*-back tasks, over-all 27 trials were accomplished. According to each task, every trial schedule contains three parts. The first part starts with 2 s instruction orders indicating the type of the task (0-, 2- or 3-back). Then, the second part is scheduled for a 40 s task period, and the remaining part is a 20 s rest period.

Elaborately, the task begins with a short beep-beep of 250 ms indicating the beginning and for the end of the task period beep was also provided to sign finish. Then, after 2 s instruction order, for each *n*-back task period, the monitor displayed a random one-digit number in every 2 s. For each number, display time was 0.5 s and the remaining time 1.5 s is given to the subject for target fixation. The word 'STOP' was shown on the monitor for 1 s at the termination of each task period. The aforesaid timing arrangement of each *n*-back task of one trial is shown in Fig. 1. Throughout the rest period, the monitor exhibited fixation cross sign '+'. After each session of 80 s gap was executed for rest. The targets appeared with a 30% chance (70% non-targets) and twenty trials were repeated for each task. For target fixation in each task, participants pressed number 7 as the 'target' button with their right index finger. Otherwise, for the 'non-target' case, participants pushed the number 8 button with the right middle finger of the keyboard. Fig. 2 depicts the whole timing sequence of the entire experimental protocol.

2.2. Participants

In this research, we involved twenty-six (26) right-handed healthy participants. Among them, 9 males and 17 females, and their average age were 26.1$\pm$3.5 years (mean$\pm$standard deviation). The age ranges of the participants are 17 to 33 years. All the participants were informed earlier and gave written consent prior to the experiment. The experimental protocol was conducted to the Helsinki declaration and was approved by the Ethics Committee of the Institute of Psychology and Ergonomics, Berlin Institute of Technology (approval number: SH_01_20150330). Any diseases (psychological, nerve, or other brain-related diseases) that may confuse the result, were not reported by the participants.

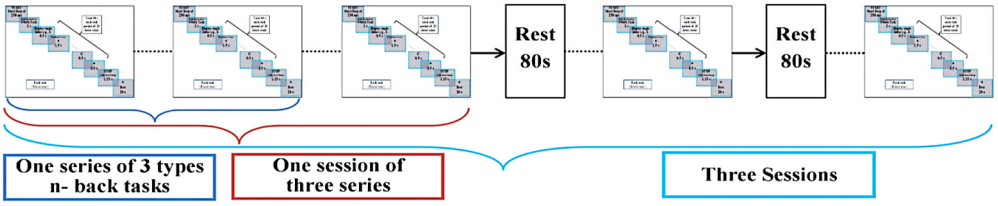

Fig. 2. Total timing sequence of the entire experiment. Detail time sequence and clear view of *n*-back task block is shown in Fig. 1.

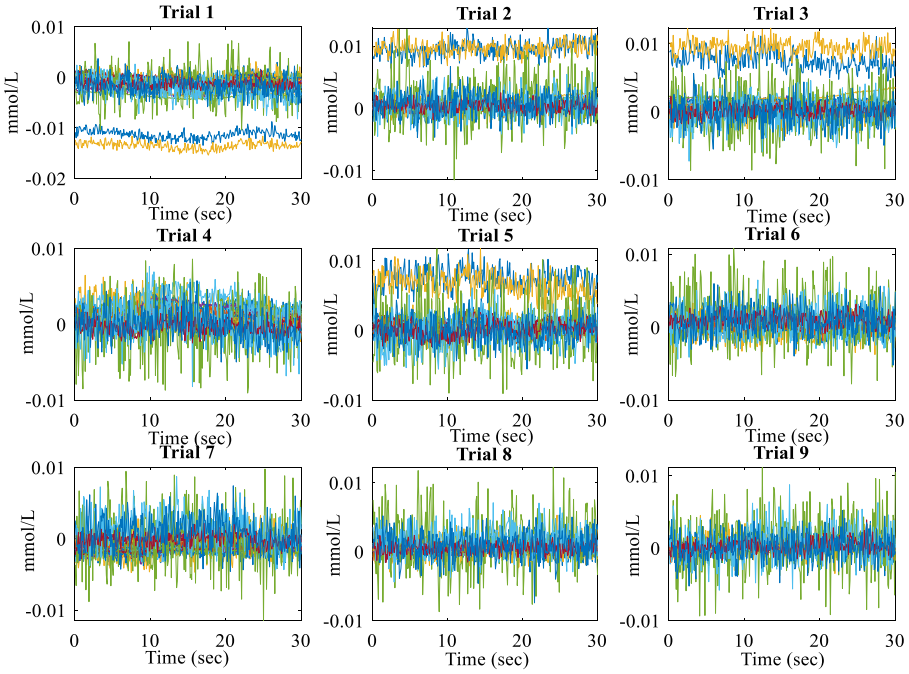

Fig. 3. The deoxy-hemoglobin concentration changes of raw data through nine trials of subject 1 in the context of the 0-back task.

2.3. Experimental protocol

During the experiment procedure, each participant sat down on an armchair in a relaxed way forward-facing an LCD monitor. There was approximately a 1.2 m distance between the eyes and the monitor. The participants were trained to preserve their eyes on the monitor and abstain from moving body as much as possible throughout the data recording. The right arm's index and middle finger positions were on numeric keyboard buttons number 7 and 8, respectively.

2.4. fNIRS data acquisition and pre-processing

Maintaining a sampling rate of 10 Hz, the fNIRS dataset was recorded with a NIRScout (NIRx Medizintechnik GmbH, Berlin, Germany) stating that one adjacent source-detector pair configures one fNIRS channel. At the frontal area, sixteen source-detector pairs were positioned (16 channels around AFz, AF3, AF4, AF7, and AF8). Motor part (4 channels each around C3 and C4 = total 8 channels), parietal area (4 channels each around P3 and P4 = total 8 channels), and occipital area (4 channels around POz) were placed, which yields a total of 36 channels. For all the 36 channels, the distance of the source-detector pair was fixed to 30 mm [25–27]. Utilizing the provided markers, the fNIRS data have separated the trials accordingly. Therefore, the data dimensions are structured as time-duration$\times$36$\times$9$\times$26 for each type of *n*-back task where time-duration or samples denote the cognitive task duration through samples, 36 is the total channel number, 9 is the number of trials performed by each participant and the number of participants is 26. The organized signals are filtered and baselines are corrected for all signals, oxy and deoxyhemoglobin concentration changes (HbO_2 and dHb) were figured out using the modified Beer-Lambert law.

2.5. Methods

fNIRS follow modified Beer-Lambert law measures the light attenuation between the source and detector can be formulated as [28,29] as,

$$I_{out} = I_{in} 10^{-OD_\lambda} \tag{1}$$

Here, OD_λ is the optical density at the wavelength, λ. From modified Beer-Lambert law, a transformation from the changes of light output leads to the changes in blood chromophore concentration (HbO_2 and dHb). As HbO_2 and dHb signals show a negative correlation, so high concentration of HbO_2 or low concentration of dHb represents the activity of the region due to cognitive load. From the aforementioned study of fNIRS signals, it is quite practical that the extracted raw data contains many artifacts such as biological, technical, etc. Some cardiovascular effects like fast heart-beat, respiration, slow waves such as Mayer waves may affect the acquired data. Subject's movement artifacts may alter the contact position of the optodes which creates pointless spikes in almost all fNIRS datasets. For example, the representation of raw data of subject 1 while participating in the 0-back task of nine trials is given in Fig. 3. In this figure, we can see huge pointless spikes that have overshadowed the original data. As a result, the research becomes

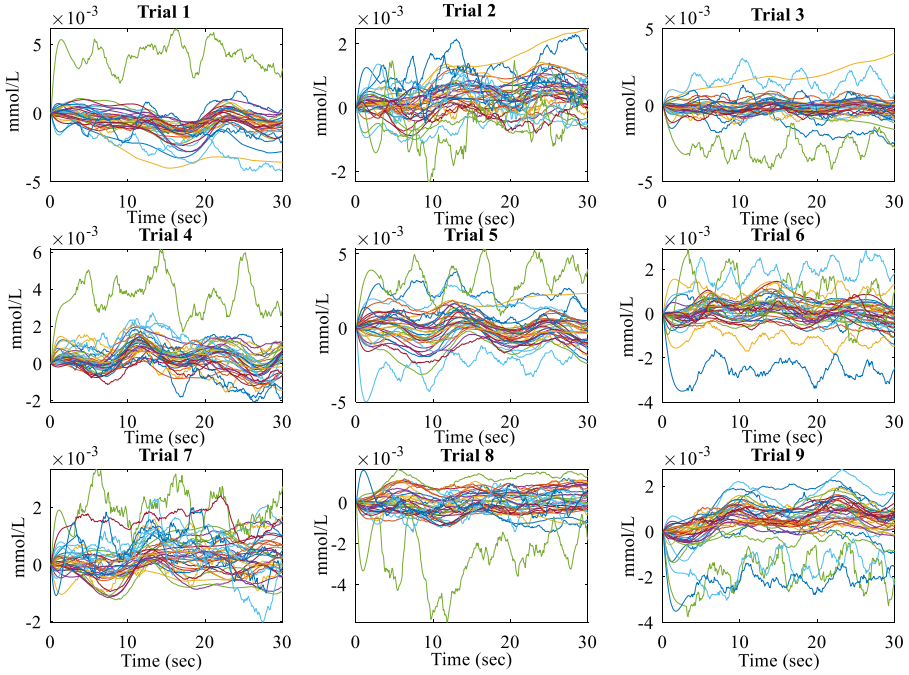

Fig. 4. The deoxy-hemoglobin concentration changes of baseline-corrected data through nine trials of subject 1 in the context of the 0-back task.

challenging to extract the appropriate information. To get rid of the artifacts, we have filtered the whole data with Savitzky-Golay (SG) filter to remove all the signal components greater than 0.1 Hz. In the SG filtering procedures, the least error estimation is considered to fit the data. Consider, a set of $2m+1$ input samples including an approximate interval with a fixed set of weighting coefficients. The input samples can be computed for specified polynomial order and frame length of n and $2m+1$, respectively [30,31]. The output samples can be expressed in terms of discrete convolution form given in equation (2).

$$Y[n] = \sum_{m=-M}^{M} h[m]X[n-m] \tag{2}$$

To accomplish the whole filtering purposes, SG filter stated with 3rd order and 121 frame length was used. After that, an important pre-processing technique called baseline correction has been accomplished to separate true spectroscopic signals from interference effects. In this technique, subtracting the baseline from the raw signal in such a form of hemodynamic activation remains at the baseline value of zero [16]. Fig. 4 represents the baseline-corrected data which was displayed in Fig. 3. From Fig. 4, it is very much easy to understand the fNIR signals of 36 channels including the whole nine trials separately.

Then, the signals were considered for statistical feature analysis. The data were separated according to 0-back, 2-back, and 3-back tasks. Six significant features of fNIRS signals are extracted that are mean, minimum, maximum, standard deviation (SD), slope, and skewness. After extracting the filtered dataset as $\mathrm{Y} = \{y_1, y_2, \ldots\ldots, y_n\}$, the arithmetic mean, $\bar{\mathrm{Y}}$ [32] is stated in equation (3),

$$\bar{\mathrm{Y}} = \frac{1}{N}\sum_{i=1}^{N} y_i \tag{3}$$

While in the case of a discrete random variable from the dataset Y with each value having the same probability, the SD, σ [33] is expressed in equation (4),

$$\sigma = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(y_i - \bar{\mathrm{Y}})^2} \tag{4}$$

During SD computing, the minimum and maximum values were also determined. To fit the data as a linear function in x and y coordinates, slope [34] is represented as in equation (5),

$$\mathrm{Y} = \mu\mathrm{X} + \gamma, \text{where slope is } \mu = \frac{\Delta\mathrm{Y}}{\Delta\mathrm{X}} \tag{5}$$

The last feature skewness measures the deviation of a random variable's given distribution from the normal distribution [35]. Considering Y a univariate dataset, simple formula for skewness, $\tilde{\mu}_3$ of the third standardized moment [36] is stated in (6),

$$\tilde{\mu}_3 = \frac{\sum_{i=1}^{N}(y_i - \bar{\mathrm{Y}})^4/N}{\sigma^4} \tag{6}$$

Where, N is the number of data points, $\bar{\mathrm{Y}}$ is mean, and σ is the SD. To check the significant statistical inference among the n-back tasks or classes, one way ANOVA test was performed according to features. After the statistical significance test, we have checked two findings: a) most significant channels, b) how many channels are responsive based on features. That means the feature indicating maximum channels supported through significant results and most common significant channels according to all features will be our concluding point of the entire investigation. We also extracted the functional pattern of each n-back task changing the size of the window and shifting the window, respectively. While the grand average was performed to detect the whole patterns' mean for the generalized perseverance.

To realize the efficacy of fNIRS data based on three types of tasks for identifying cognitive load, the measurement of classification accuracy is obligatory [37]. Several supervised machine learning algorithms for classification are developed. In our study, SVM (Support vector machine) standard machine learning method is preferred due to its high prediction accuracy through higher-dimensional space mapping and easy implementation [38]. SVM classifier establishes a procedure that maximizes the distance between the separating hyper-planes and the nearest training points.

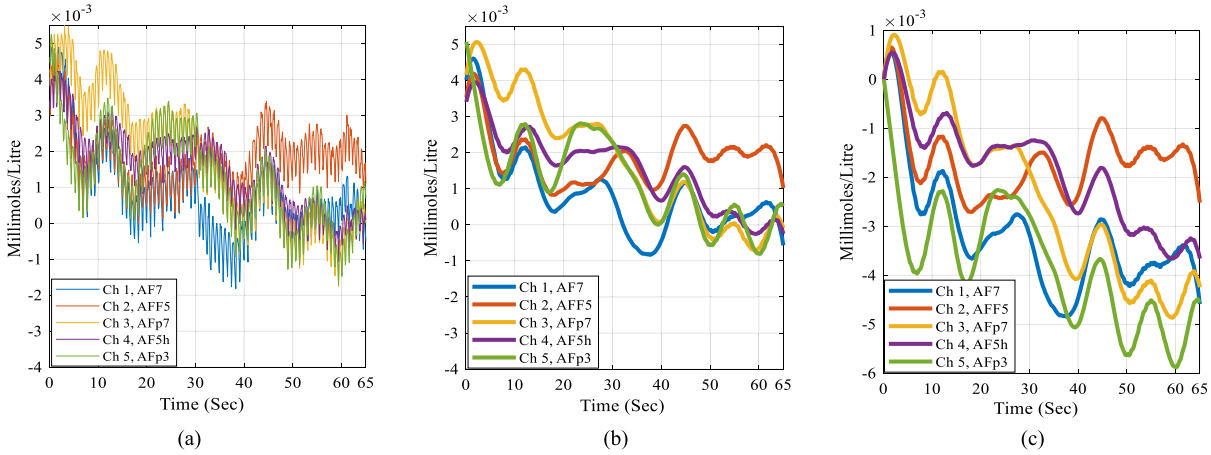

Fig. 5. HbO_2 fNIRS signal of frontal five channels (a) 0_back raw (b) 0_back filtered (c) 0_back baseline corrected.

Table 1
Channel Significance for dHb Concentration.

Name of the features	Statistically Inferring Channels	Total Significant Channels
Mean	**1**,2,3,4,11,**13**,21,22,23,24,33,34	12
Minimum	**1**,2,3,4,6,11,**13**,21,22,34	10
Maximum	**1**,7,10,**13**,24,25	6
Slope	**1**,2,8,12,**13**,15,16,17,19,23,31	11
Standard Deviation	**1**,2,4,6,7,10,11,**13**,20,21,25	11
Skewness	**1**,2,3,4,12,**13**,14,32,34,35	10

The nearest training points are called support vectors [39]-[40]. In the 2D feature space, the above-mentioned dividing hyper-planes can be formulated as,

$$\psi(x) = m.x + c \tag{7}$$

According to (7), m, x belongs to 2D real coordinate space, c belongs to 1D real coordinate space. The optimization of m^* yields to maximize the distance between separating hyper-planes and the nearest training points. Observing the cost function in (8), the maximization method is acquired.

$$\psi(m, \zeta) = \frac{1}{2}\|m\|^2 + H.\sum_{n=1}^{N}\zeta_n \tag{8}$$

Here in (8), H is the optimistic regularization parameter defined by the users, which is based on the penalty factors of classification errors.

Above mentioned cost function should obey the following conditions expressed in (9).

$$\left.\begin{array}{c}(x_n.m + c) \geq 1 - \zeta_n \text{ for } y_n = +1\\ (x_n.m + c) \geq \zeta_n - 1 \text{ for } y_n = -1\\ \zeta_n = 0 \forall_n\end{array}\right\} \tag{9}$$

Here $m^T m = \|m\|^2$ and ζ_n denotes the error of training period. N denotes the misclassified samples, and y_n is the class label for the n^{th} sample in the case of binary classification either $+1$ or -1. We used kernel trick to create non-linear classifiers for optimum-margin hyper-planes [41].

To establish a nonlinear classification rule for the trained data points, $\varphi(q_n)$ where kernel function κ should satisfy $\kappa(q_n, q_m) = \varphi(q_n).\varphi(q_m)$. The classification vector, u is formulated as,

$$u = \sum_{n-1}^{N} x_n y_n \varphi(q_n) \tag{10}$$

Here, x_n are achieved by solving the optimization problem given (7)–(9).

In this work, we used the Matlab toolbox in place of the one versus all approach of SVM. In every case of HbO_2 and dHb, the classification accuracy was calculated with a 5-fold cross-validation technique for all twenty-six participants. The average classification accuracy inclusive standard deviation was also calculated for both cases justifying the effectiveness of fNIRS data.

3. Results and discussion

To confirm the filtering assessment, we have depicted five frontal area channel signals: AF7, AFF5, AFp7, AF5h, and AFp3 and compared the filtered and baseline corrected signals with raw data. Fig. 5(a) represents oxygenated data of five frontal channels participating in the 0-back task by subject 1 where Fig. 5(b) shows the SG filtered data contains associated extractive information and Fig. 5(c) shows the baseline corrective value of the same channels comprising relative difference more than filtered value.

Six sets of features were extracted from each trial for all 36 channels of 26 subjects. The features are as follows: mean value, minimum, maximum, slopes, standard deviation, and skewness of HbO_2 and dHb concentration. To justify the proposed methodology, a one-way ANOVA test was performed upon 36 channels to all subjects based on the features. Table 1 symbolizes the statistically inferring channels conferred to six features due to changing concentration of dHb. Whereas, Table 2 denotes the statistically inferring channels for HbO_2 concentration of the same features. These tables illustrate that the feature, mean encompasses maximum significant channels such as 18 channels for HbO_2 and 12 channels for dHb concentration compared to other features. Conversely, Channel 1 and 13 *i.e.*, AF7 and C3h are the most common significant channels conferred to all features and hemoglobin concentration. From each set of extracted features, we can apprehend that the mean value is significantly different from the context of 0-back, 2-back, and 3-back tasks. The second finding was AF7 and C3h channels represented significantly different ANOVA results

Table 2
Channel Significance for HbO_2 Concentration.

Name of the features	Statistically Inferring Channels	Total Significant Channels
Mean	**1**,4,6,7,8,9,10,11,**13**,15,20,21,23,24,25,27,33,34	18
Minimum	**1**,4,5,6,7,8,9,**13**,15,20,21,34	12
Maximum	**1**,9,10,**13**,24,25	6
Slope	**1**,2,8,12,**13**,15, 23,31, 32,34,35	11
Standard Deviation	**1**,2,12,**13**,25,28	6
Skewness	**1**,4,**13**,21,24	5

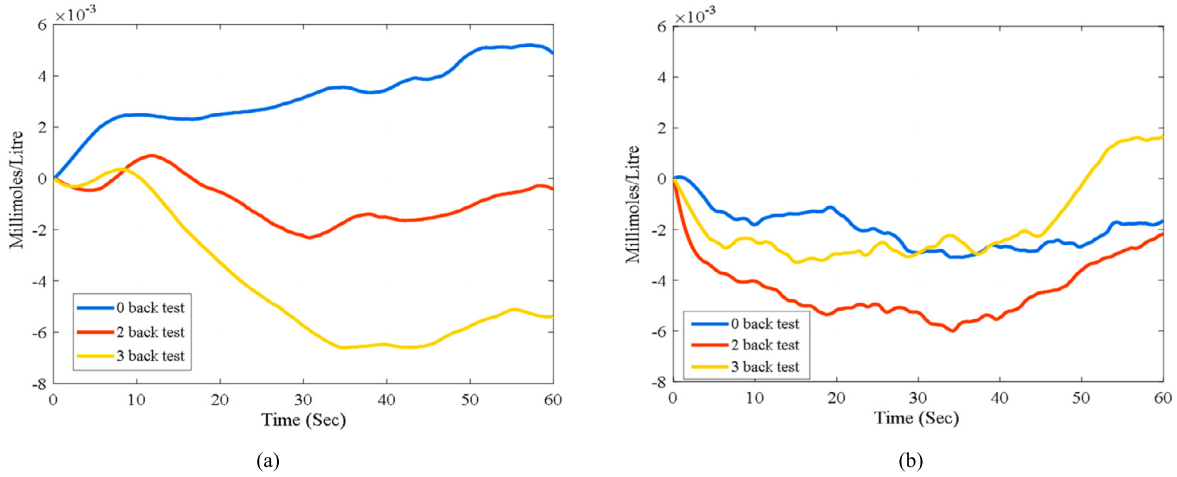

Fig. 6. Grand average performance of normalized dHb concentration during three types of *n*-back task over all trials and subjects for (a) AF7 channel and (b) C3h channel.

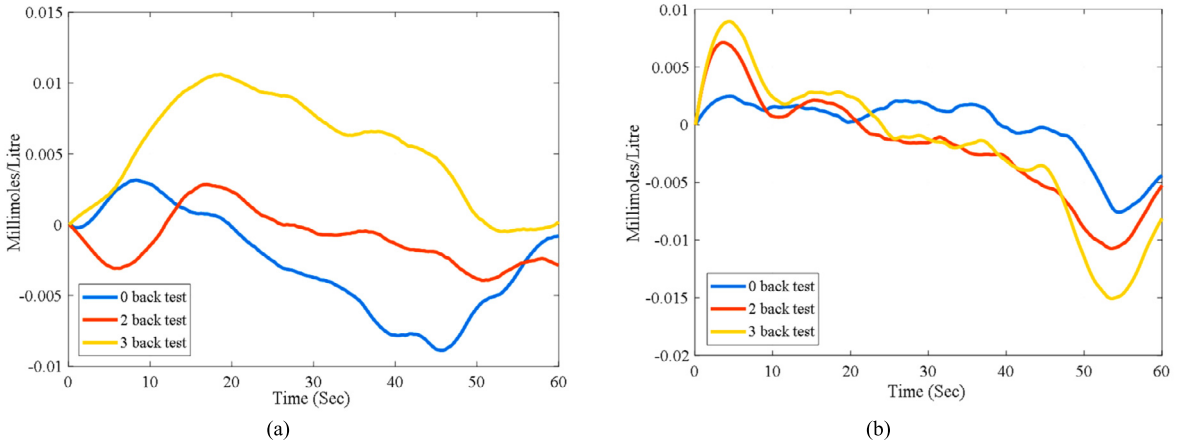

Fig. 7. Grand average performance of normalized HbO_2 concentration during three types of *n*-back task over all trials and subjects for (a) AF7 channel and (b) C3h channel.

among the 3 types of *n*-back test with a p-value < 0.005. The two locations denote the prefrontal and occipital area. The above outcomes are also mentioned in the dataset [22] which justifies the findings. Subsequently, the Grand average was performed considering subsamples have the same number of data points. Fig. 6 and Fig. 7 epitomize the grand average results based on the change in dHB and HbO_2 concentration of AF7 and C3h channels. From these figures, we found that changes in AF7 are more significant than C3h.

To establish fNIRS technology as a valid biomarker for analyzing the mental load assessment, consistent result providing is a key demand from the signals conveyed by fNIRS. Therefore, the signals of three different stress levels of twenty-six subjects are considered for analyses through the statistical point of view and found two significant channels among 36 channels. These two channels became the point of interest to represent the fNIRS as a valid biomarker with a consistent result. On this contrary, a one-minute signal is divided into the three-time interval and the mean value of the concentration of HbO_2 and dHb (change in the concentration) are considered for consistency check. In the case of channel AF7 shown in Fig. 8, it is found that the change in the concentrations of dHb is decreasing (3-back<2-back<0-back) consistently in three different time intervals. Inversely, the change in the concentrations of HbO_2 is increasing (3-back>2-back>0-back) in the 2nd (21-40 sec) and 3rd (41-60 sec) time interval, although a slightly different result is found in the 1st (1-20 sec) time interval. It may be due to the lack of initiating the change in the concentration for the assigned mental work. Different increasing and decreasing patterns of HbO_2 and dHb are observed in the case of significant channel C3h presented in Fig. 9. In time interval 1st (1-20 sec) and 2nd (21-40 sec), the pattern of decreasing of Δ[dHb] and increasing of $\Delta[HbO_2]$ are found consistent but the unusual fact found here is the changing nature of 2-back is greater than 3-back (especially for Δ[dHb]) which is supposed to be 3-back>2-back. Since

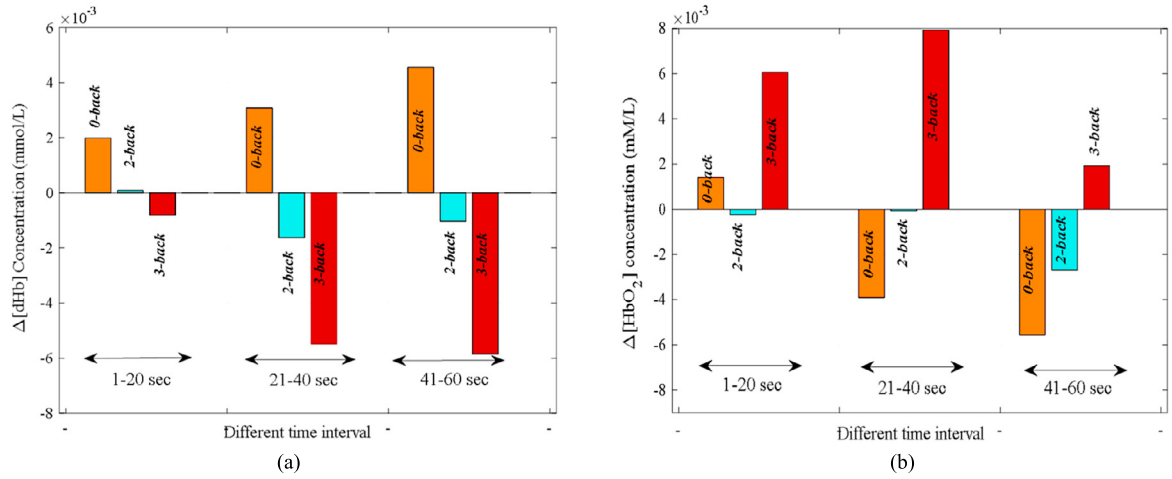

Fig. 8. AF7 channel performance during three types of *n*-back task over different time slots context of (a) Mean deoxygenated concentration and (b) Mean oxygenated concentration.

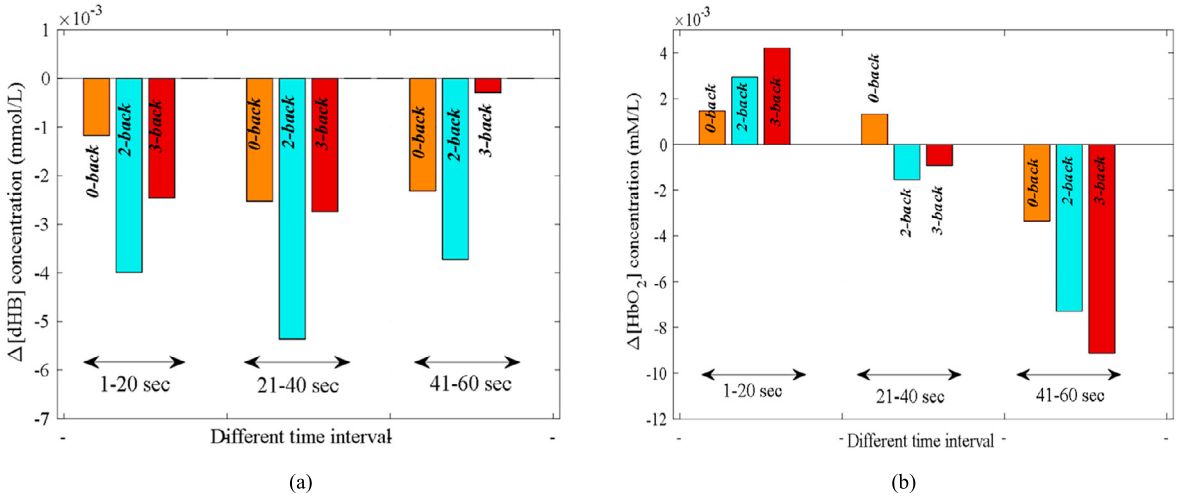

Fig. 9. C3h channel performance during three types of *n*-back task over different time slots context of (a) Mean deoxygenated concentration and (b) Mean oxygenated concentration.

the error probability for 3-back is more than 2-back, this event was found in channel C3h. In the case of the 3-back test, the participant often couldn't play properly, and then he/she presses the button randomly to complete the game that means he/she was not taking the mental load properly. Perhaps this nature is not flourished in channel AF7 but C3h. From these analyses, we found the activity locations of the human brain stimulated by different cognitive levels. Because channel AF7 indicates the frontal area whereas, C3h locates the motor area according to the dataset acquisition information [22]. The active location scenario of the brain surface is depicted in Fig. 10. This figure establishes a positive relationship between growing cognitive load hemodynamics and the oxygenation activity is observed in the frontal and motor cortex.

To achieve another goal, *i.e.*, cognitive load level identification through a soft machine learning approach, SVM is used as the cognitive load level predictor based on supervised learning of subject-based fNIRS data. From the primary investigation, it is found that the functional brain activity between easy task (0-back test) and hard task (2-back and 3-back test) is significantly separable. On the other hand, finding the inference between two hard tasks i.e., 2-back and 3-back tests, is quite complex and subtle. On the contrary, the classification problem is considered as a binary problem, and the 0-back is considered as a class and both 2-back and 3-back are considered as another class. This assumption proceeds to

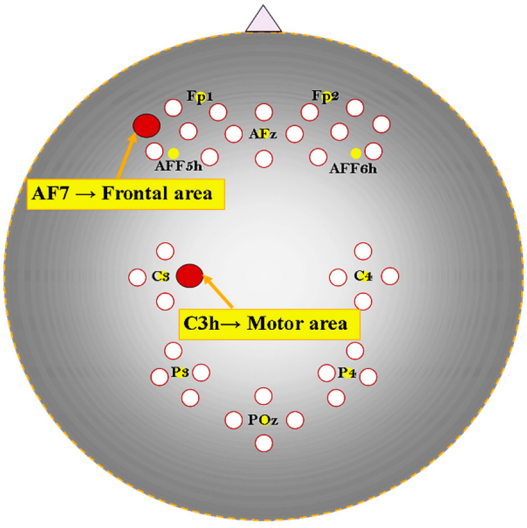

Fig. 10. Location of fNIRS signal activity due to different cognitive level hemodynamics.

predict two-level cognitive loads (easy and hard) from the fNIRS signal. In this work, SVM is modeled for a binary classifier with the

Table 3
Subject wise classification results of cognitive load level from HbO_2 and dHb data.

HbO_2		dHb	
Subject ID	Classification Accuracy (%)	Subject ID	Classification Accuracy (%)
Sub 1	91.67%	Sub 1	83.33%
Sub 2	75%	Sub 2	75%
Sub 3	75%	Sub 3	75%
Sub 4	66.67%	Sub 4	75%
Sub 5	66.67%	Sub 5	75%
Sub 6	75%	Sub 6	66.67%
Sub 7	66.67%	Sub 7	66.67%
Sub 8	83.33%	Sub 8	75%
Sub 9	83.33%	Sub 9	75%
Sub 10	91.67%	Sub 10	83.33%
Sub 11	66.67%	Sub 11	58.33%
Sub 12	66.67%	Sub 12	66.67%
Sub 13	75%	Sub 13	66.67%
Sub 14	75%	Sub 14	66.67%
Sub 15	75%	Sub 15	66.67%
Sub 16	75%	Sub 16	66.67%
Sub 17	66.67%	Sub 17	75%
Sub 18	75%	Sub 18	66.67%
Sub 19	75%	Sub 19	66.67%
Sub 20	66.67%	Sub 20	66.67%
Sub 21	66.67%	Sub 21	66.67%
Sub 22	66.67%	Sub 22	75%
Sub 23	75%	Sub 23	75%
Sub 24	75%	Sub 24	83.33%
Sub 25	75%	Sub 25	75%
Sub 26	58.33%	Sub 26	66.67%
	Average = 73.40%±0.076		**Average = 71.48%±0.061**

3^{rd} order polynomial kernel function. The classification accuracy is estimated by a 5-fold cross-validation technique. The resulting classification accuracies are given in Table 3.

In this table, the functional changes in the concentration of both HbO_2 and dHb are considered for classification accuracy estimation and presented separately. From the average classification accuracy, it is found that HbO_2 data shows slightly higher accuracy (73%) than that of dHb (71%). Associating the results of classifiers, the proposed study has been found considerable for statistical valuation of three types of *n*-back tasks through fNIRS data as cognitive load assessment.

4. Conclusions

This work thoroughly investigated an fNIRS dataset of cognitive load which was on 3 different cognitive loads (0-back, 2-back, and 3-back test). Some faulty data are found in almost every trial among the 36 channels, randomly. For the final processing of the dataset in this work, the faulty data were avoided, and based on the rest of the channels; the relative changes of dHb and HbO_2 were taken into consideration for statistical analysis. From the statistical (ANOVA test) analysis, only two significant channels were found that showed the statistical inference among three different cognitive loads. However, it is found from the grand average of the relative change in the concentration of dHb and HbO_2 that the patterns of inference of the claimed two channels are different. As a consequence, we can claim that the frontal and motor area of the human brain cerebral cortex conveys an optimistic relationship regarding cognitive load hemodynamics. We examined and evaluated the whole outcomes according to statistical methods. The study established the findings through a greater number of volunteers' participation which justifies the outcomes rather than other related studies.

Although from the outcomes of this dataset, we got important information about the cognitive load behaviors of the fNIRS signal, some queries are still confusing. To get the proper brain functionality assessment through the fNIRS signals, some other datasets of different cognitive loads are also necessary to investigate and compare. In our future work, such a comparative study of different cognitive loads and their corresponding functional patterns will be presented. Besides, other functional brain imaging modalities like EEG can also be added in the future work to differentiate the cognitive load level more precisely.

CRediT authorship contribution statement

F. Khanam conceptualized, designed, and analyzed the work and wrote the manuscript. A.B.M. Aowlad Hossain and M. Ahmad supervised the work and reviewed the manuscript. All authors read the final manuscript and gave consent for publication.

Declaration of competing interest

The authors declare they have no potential conflicts of interest with respect to the research, authorship, and publication of this article.

Data availability and ethical approval

The authors of this study used a renowned dataset (available in: https://www.nature.com/articles/sdata20183) that was conducted obeying the Helsinki declaration and also approved by the Ethics Committee of the Institute of Psychology and Ergonomics, Berlin Institute of Technology (approval number: SH_01_20150330).

Acknowledgements

The authors would like to thank Md. Asadur Rahman, Ph.D., Assistant Professor, Dept of BME, MIST, Dhaka, Bangladesh for his guidelines to complete this work.

References

[1] H. Aghajani, M. Garbey, A. Omurtag, Measuring mental workload with EEG+fNIRS, Front. Human Neurosci. 11 (359) (2017), https://doi.org/10.3389/fnhum.2017.00359.

[2] S. Midha, H.A. Maior, M.L. Wilson, S. Sharples, Measuring mental workload variations in office work tasks using fNIRS, Int. J. Hum.-Comput. Stud. 147 (2021) 1–98, https://doi.org/10.1016/j.ijhcs.2020.102580.

[3] M.S. Young, N.A. Stanton, Malleable attentional resources theory: a new explanation for the effects of mental underload on performance, Hum. Factors 44 (3) (2002) 365–375, https://doi.org/10.1518/0018720024497709.

[4] C.Y. Wong, G. Seet, Workload, awareness and automation in multiple-robot supervision, Int. J. Adv. Robot. Syst. 14 (3) (2017) 1–16, https://doi.org/10.1177/1729881417710463.

[5] R. Parasuraman, G.F. Wilson, Putting the brain to work: neuroergonomics past, present, and future, Hum. Factors 50 (3) (2008) 468–474, https://doi.org/10.1518/001872008X288349.

[6] J.F. Gagnon, G. Durantin, F. Vachon, M. Causse, S. Tremblay, F. Dehais, Anticipating human error before it happens: towards a psychophysiological model for online prediction of mental workload, in: Human Factors: a View from an Integrative Perspective: Proceedings HFES Europe Chapter Conference, 2012, pp. 57–66.

[7] M.W. Eysenck, M. Brysbaert, Fundamentals of Cognition, 3rd ed., Routledge, 2018.

[8] G. Strube, Cognitive Science: Overview Chapter, International Encyclopedia of the Social & Behavioral Sciences, 2001, pp. 2158–2166.

[9] R. Revlin, Cognition: Theory and Practice, 1st ed., Worth Publishers, 2012.

[10] H.G. Liddell, R. Scott, A Greek-English Lexicon, H.S. Jones, R. McKenzie (Eds.), Clarendon Press – via Perseus Project, Oxford, 1940.

[11] S. Franchi, F. Bianchini, On the historical dynamics of cognitive science: a view from the periphery, in: The Search for a Theory of Cognition: Early Mechanisms and New Ideas, in: Cognitive Science Series, vol. 238, Brill, 2011.

[12] T.V. Gog, F. Paas, J. Sweller, Cognitive load theory: advances in research on worked examples, animations, and cognitive load measurement, Educ. Psychol. Rev. 22 (2010) 375–378, https://doi.org/10.1007/s10648-010-9145-4.

[13] S.T. Frein, L.S. Jones, J.E. Gerow, When it comes to Facebook there may be more to bad memory than just multitasking, Comput. Hum. Behav. 29 (6) (2013) 2179–2182, https://doi.org/10.1016/j.chb.2013.04.031.

[14] F.G.W.C. Paas, J.J.G.V. Merriënboer, The efficiency of instructional conditions: an approach to combine mental effort and performance measures, Hum. Factors 35 (4) (1993) 737–743, https://doi.org/10.1177/001872089303500412.

[15] A. Skulmowski, G.D. Rey, Measuring cognitive load in embodied learning settings, Front. Psychol. 8 (2017) 1191, https://doi.org/10.3389/fpsyg.2017.01191.

[16] M.A. Rahman, M.S. Uddin, M. Ahmad, Modeling and classification of voluntary and imagery movements for brain-computer interface from fNIR and EEG signals through convolutional neural network, Health Inf. Sci. Syst. 7 (22) (2019), https://doi.org/10.1007/s13755-019-0081-5.

[17] M.A. Rahman, M. Ahmad, Towards a portable near-infrared spectroscopy system for bedside monitoring of a patient brain, in: Proceedings of 4th Annual Conference of Bangladesh Medical Physics Society, 2015.

[18] H. Aghajani, A. Omurtag, Assessment of mental workload by EEG+FNIRS, in: 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2016, pp. 3773–3776.

[19] H. Ayaz, M. Izzetoglu, S. Bunce, T. Heiman-Patterson, B. Onaral, Detecting cognitive activity related hemodynamic signal for brain computer interface using functional near infrared spectroscopy, in: 3rd International IEEE/EMBS Conference on Neural Engineering, 2007, pp. 342–345.

[20] C. Herff, D. Heger, O. Fortmann, J. Hennrich, F. Putze, T. Schultz, Mental workload during n-back task—quantified in the prefrontal cortex using fNIRS, Front. Human Neurosci. 7 (2014) 935, https://doi.org/10.3389/fnhum.2013.00935.

[21] L. Cabañero, R. Hervás, I. González, J. Fontecha, T. Mondéjar, J. Bravo, Analysis of cognitive load using EEG when interacting with mobile devices, in: Proc. 13th Int. Conf. on Ubiquitous Computing and Ambient Intelligence (UCAmI), Vol. 31, No. 1, 2019, p. 70.

[22] J. Shin, A. von Lühmann, D.W. Kim, J. Mehnert, H.J. Hwang, K. Müller, Simultaneous acquisition of EEG and NIRS during cognitive tasks for an open access dataset, Sci. Data 5 (2018), https://doi.org/10.1038/sdata.2018.3.

[23] J. Shin, et al., Technische Universität Berlin, 2016.

[24] W.K. Kirchner, Age differences in short-term retention of rapidly changing information, J. Exp. Psychol. 55 (4) (1958) 352–358, https://doi.org/10.1037/h0043688.

[25] S.D. Power, A. Kushki, T. Chau, Towards a system-paced near-infrared spectroscopy brain-computer interface: differentiating prefrontal activity due to mental arithmetic and mental singing from the no-control state, J. Neural Eng. 8 (6) (2011), https://doi.org/10.1088/1741-2560/8/6/066004.

[26] N. Naseer, N.K. Qureshi, F.M. Noori, K.S. Hong, Analysis of different classification techniques for two-class functional near-infrared spectroscopy-based brain-computer interface, Comput. Intell. Neurosci. 2016 (2016) 11, https://doi.org/10.1155/2016/5480760.

[27] H.J. Hwang, J.H. Lim, D.W. Kim, C.H. Im, Evaluation of various mental task combinations for near-infrared spectroscopy based brain-computer interfaces, J. Biomed. Opt. 19 (7) (2014), https://doi.org/10.1117/1.JBO.19.7.077005.

[28] M. Izzetoglu, S.C. Bunce, K. Izzetoglu, B. Onaral, K. Pourrezaei, Functional brain imaging using near-infrared technology, IEEE Eng. Med. Biol. Mag. 26 (4) (2007) 38–46, https://doi.org/10.1109/MEMB.2007.384094.

[29] A. Bozkurt, A. Rosen, H. Rosen, B. Onaral, A portable near infrared spectroscopy system for bedside monitoring of newborn brain, Biomed. Eng. Online 4 (29) (2005) 1–11.

[30] A. Savitzky, M.J.E. Golay, Smoothing and differentiation of data by simplified least-squares procedures, Anal. Chem. 36 (8) (1964) 1627–1639, https://doi.org/10.1021/ac60214a047.

[31] M.A. Rahman, M.A. Rashid, M. Ahmad, Selecting the optimal conditions of Savitzky–Golay filter for fNIRS signal, Biocybern. Biomed. Eng. 39 (3) (2019) 624–637, https://doi.org/10.1016/j.bbe.2019.06.004.

[32] V. Kulkarni, Basics of Statistics - Arithmetic Mean, Kindle ed., 2020, p. 16.

[33] E. Rukmangadachari, E.K. Reddy, Probability and Statistics, O'Reilly online learning, 2021.

[34] C.S. Chicago, Treatise on Plane Co-Ordinate Geometry as Applied to the Straight Line and the Conic Sections: With Numerous Examples, 1st ed., Forgotten Books, 2012.

[35] S. Brown, Measures of Shape: Skewness and Kurtosis, BrownMath.com, 2020.

[36] M.A. Rahman, M. Ahmad, Movement related events classification from functional near infrared spectroscopic signal, in: Int. Conf. on Computer and Information Technology (ICCIT), Dhaka, Bangladesh, 18-20 December 2016, Dhaka, Bangladesh.

[37] A. Sarkar, A. Singh, R. Chakraborty, A deep learning-based comparative study to track mental depression from EEG data, Neurosci. Inform. 2 (4) (2022) 100039, https://doi.org/10.1016/j.neuri.2022.100039.

[38] M.A. Rahman, M.A. Rashid, M. Ahmad, A. Kuwana, H. Kobayashi, Activation modeling and classification of voluntary and imagery movements from the prefrontal fNIRS signals, IEEE Access 8 (2020) 218215–218233, https://doi.org/10.1109/ACCESS.2020.3042249.

[39] M.A. Rahman, M.F. Hossain, M. Hossain, R. Ahmmed, Employing PCA and t-statistical approach for feature extraction and classification of emotion from multichannel EEG signal, Egypt. Inform. J. 21 (1) (March 2020) 23–35, https://doi.org/10.1016/j.eij.2019.10.002.

[40] Y. Zhang, Q. Deng, W. Liang, X. Zou, An efficient feature selection strategy based on multiple support vector machine technology with gene expression data, BioMed Res. Int. 2018 (2018) 11, https://doi.org/10.1155/2018/7538204.

[41] A. Sharma, S. Kaur, N. Memon, A.J. Fathima, S. Ray, M.W. Bhatt, Alzheimer's patients detection using support vector machine (SVM) with quantitative analysis, Neurosci. Inform. 1 (3) (2021), https://doi.org/10.1016/j.neuri.2021.100012.