Analysis of Cognitive Load - Importance of EEG Channel Selection for Low Resolution Commercial EEG Devices

Aniruddha Sinha, Debatri Chatterjee, Diptesh Das, Arijit Sinharay
Innovation Labs,
Tata Consultancy Services Ltd.
Kolkata, India
{aniruddha.s, debatri.chatterjee, diptesh.das, arijit.sinharay}@tcs.com

***Abstract*— Measurement of cognitive load using brain signals is an important area of research in human behavior and psychology. Recently, there have been attempts to use low cost, commercially available Electroencephalogram (EEG) devices for the analysis of the cognitive load. Due to the reduced number of leads, these low resolution devices pose major challenges in signal processing as well as in feature extraction. In this paper, we investigate the significant leads or channels that are useful for the analysis of the cognitive load. We use a standard matching test and n-back memory test imparting low and high cognitive loads respectively. The investigation is based on the analysis of variance (ANOVA) of Alpha and Theta frequency band signals for various combinations of leads. Comparisons have been done between the previously reported leads and those obtained using a few feature selection algorithms. Results indicate that for a given stimulus, though the significant leads are very much dependent on the subjects, the leads corresponding to the left frontal lobe and right parieto-occipital lobe are in general most significant across majority of subjects for analysis of the cognitive load.**

Keywords—electroencephalogarphy; channel selection; cognitive load; commercial EEG; feature selection

I. INTRODUCTION

Cognitive load is a measure of the processing done using working memory [1] of the brain. The effectiveness of an activity is dependent on the amount of cognitive load experienced by a subject. Various fields of applications ranging from personalized learning to air-traffic monitoring, require monitoring the cognitive load in real time. The mental state changes due to the level of imparted cognitive load and the performance of a subject may drastically degrade if the load exceeds beyond a critical point [2]. Analysis of various physiological signals to detect and analyze human mental states is an emerging and widely recognized technique. Physiological measurements give more unbiased, reliable and accurate metrics than performance based indices. Basic cardiovascular measures like heart-rate is found to significantly increase with increase in attention and mental work load [3]. Similarly changes in skin conductance levels are also found to be highly co-related with the one's present mental condition [4].

Though the physiological measurements have good correlation with the changes in mental state and cognitive load, it still remains the effect caused due to the brain activities, hence, serves as an indirect way of measurement. Obviously, if permitted, it is always desirable to analyze the brain signals directly as it is the root cause for the changes in physiological/ cognitive signals. Several methods exists to analyze the brain signals namely, Electroencephalogram (EEG) [5], functional Magnetic Resonance Imaging (fMRI) [6], functional Near Infrared Spectroscopy (fNIRS) [7] etc. The neurophysiological changes in the brain, to a given stimuli, can be successfully used to differentiate between the human thinking processes triggered for different levels of effortful cognitive tasks. Among the various methods, EEG signal has the advantage of higher time-resolution and portability enabling a real-time analysis of the cognitive load in practical scenarios. Moreover, low cost wireless EEG headsets productized by various companies made EEG modality particularly attractive. Since the low cost EEG devices come with lower number of EEG leads/ channels (i.e. low spatial resolution) many of the standard preprocessing steps like Independent Component Analysis, Common Spatial Pattern Filtering etc. cannot be performed efficiently to enhance the signal quality. Hence, finding out the most sensitive EEG leads positions plays a crucial part in addressing subject variability and artifact removals to enhance the measurement quality. This motivates us to carry out our research showing the importance of lead/channel selections to work with low resolution EEG devices.

The organization of this paper is as follows. Section II presents the literature review on cognitive load analysis using various channels along with few feature selection algorithms. Section III presents experimental setup and details of the tool used for EEG data capture. The detailed methodology for the analysis of the channels and features are given in Section IV. Results are presented and discussed in Section V followed by conclusion in section VI.

978-1-4799-7502-0/14 $31.00 © 2014 IEEE
DOI 10.1109/BIBE.2014.28

II. LITERATURE REVIEW

The cortex or cerebrum region is the largest part of human brain. This region is associated with different brain functions like critical thinking, perception, decision making etc. Different lobes of cerebral cortex are responsible for different brain functions e.g. occipital lobe is associated with visual perception [8], temporal lobe is associated with perception and recognition of auditory stimuli [9] etc. A list of literatures suggest that cognitive load variations for tasks having different difficulty levels are most clearly visible if we consider frontal and parietal lobes [10], more precisely the theta wave variations at frontal lobe and alpha wave variation in parietal lobe [11, 12]. In general it has been widely reported that alpha power decreases with increased cognitive load. This effect is most prominent at central parietal (Pz) location, whereas theta power increases with increased cognitive load and is most prominent at central frontal (Fz) location. Attempts have been done to use low-resolution EEG headsets (where Pz and Fz locations are not available) to detect cognitive loads [13, 14] as well. For example, the work of Anderson et al. [15], which combined both theta and alpha power in an intelligent way to come up with a novel feature, also looks very promising. However, no systematic study to select proper EEG lead positions has been found. It is a well-known fact that EEG signals are person specific to some extent and to get highest accuracy specific EEG leads need to be identified. An alternative solution would be to use data driven spatial filters to address this subject variability [16]. This method might not be optimal for low resolution system like Emotiv as the leads are separated far apart which eventually eliminates existence of true neighboring leads to help the spatial filters to work efficiently. The only option left is to directly find out the best lead positions among the given sets.

Channel selection for cognitive load analysis is studied by Tian et al. [17] using two types of load with three subjects. They have proposed mutual information based channel selection. Findings indicate that the selected channels vary from subject to subject, however a global selection of 10 leads provide satisfactory classification accuracy. However, the findings are quite different from the one proposed by previous psychological studies by Russel et al. [18]. There have been few works on channel selection [19] in the area of Brain Computer Interface (BCI). Moreover, the feature and lead selection is highly dependent on the task being performed as they directly relate with the activated lobes of the brain [20]. In addition to this, most of the previous works on lead selection is based on high resolution EEG devices with 32 or more leads.

Feature selection is a widely researched topic in various applications. In the area of people identification, connectionist system is used for feature selection algorithm along with Adaptive Neural Network (ANN) classifier [21]. Similarly, in the area of clinical document classification, MIC based feature selection algorithm is used by Chen et al. [22]. In the current paper, we have used the concepts of these feature selection algorithms and proposed two methods of channel selection algorithm for EEG. Moreover, we have compared and analyzed few of the existing lead selection methods using Alpha and Theta sub bands. The primary focus of this paper is on lead selection while using low resolution 14-lead EEG device. The lead selection and channel selection are interchangeably used and they mean the same.

III. EXPERIMENTAL SETUP

We have used a 14-channel low cost EEG device named Emotiv[1]. The data capture has been done using a Python based in-house EEG capture tool which presents the cognitive tasks to the subject as well as collects EEG data and trial video data synchronously. In this paper, the raw EEG data are analyzed using multiple sub-sets of the 14 sensor channels for specific frequency bands. These subsets include the following:

- All 14 Channels (shown in Fig 1.a)
- Four channels on left frontal lobe (AF3, F7, F3, FC5)
- Four channels on right frontal lobe (AF4, F8, F4, FC6)
- Frontal and Parietal lobe as reported in [3, 23, 11, 12]
- Channels suggested by psychological literature [18]
- Channels and bands derived using a modified Adaptive Neural Network (ANN) feature selection (FS) algorithm [24]
- Channels and bands derived using Maximal Information Coefficient (MIC) algorithm [25]

Further to that, this section details the stimulus used for the experiments, profiles of the subjects who participated in the experiments and the tool used for the data collection. The placement of the leads for 14-channel Emotiv EEG device is shown in Fig. 1(a). The device follows standard 10-20 electrode system for channel locations. The CMS and DRL, placed in the location of P3 and P4, do not generate any signal and used for reference only. Remaining 14 leads shown in the figure generate the EEG signals at 128 Hz sampling rate. The device has an in-built notch filter at power line frequency.

A. Stimulus Design

Two elementary cognitive tasks are used for low and high mental workload on the subjects. A “Finding number” task is used for low load condition and “n-back memory” task is used for high low condition [26]. The cognitive index as presented in [15] is used to measure the work load imparted on the users while doing a particular task. For example, two sets of experiments as shown in Fig. 1(b), were designed pertaining low and high cognitive work load on participants. Each experiment consisted of 10 trials. Each trial, having 10 slides each, containing a number between 0-9. Each slide was presented for a duration of 1.6 seconds.

[1] www.emotiv.com

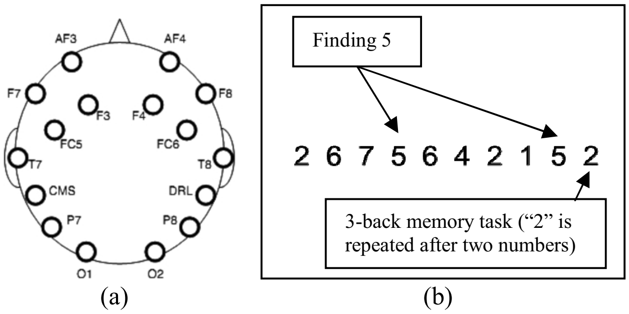


Fig. 1. (a) Lead locations for 14-channel Emotiv EEG device, (b) Stimulus used during the EEG experiments

The trials were separated by an inter-trial interval of 5 seconds. The participants were asked to relax during that period. This period is treated as the baseline period.

For low load trials, termed as "Finding number", users were presented a set of numbers one after another. They were asked to respond by clicking the left mouse button if a pre-defined number (say 5) appeared on the screen as depicted in Fig 1(b).

For high load trials, termed as "3-back memory", users needed to remember the numbers presented in each slide. They were instructed to respond if a number matched with the number presented 3 slides back (as shown in Fig. 1(b)).

Two sets of low and high load trial sessions were conducted for each user.

B. Participants

A group of 10 participants were selected in the age group of 25-30 years. All of them were right-handed male engineers working in a research lab. These ensures minimum variance in level of expertise and brain lateralization across all the participants.

Each participant completed both the sessions of low and high cognitive load trials with a short break of 2 minutes between each trials. Users were asked to relax during that period. For 5 subjects high load tasks were given first and then the low load task. For remaining 5 participants the order of tasks given were reversed. This was done to minimize the learning effect.

For the EEG data capture, we have taken consent of the participants along with the approval from the ethics committee of our company.

C. Data Collection

We have used an in-house python-based data capture tool for this purpose. The tool enables us to present the stimulus as well as collect the raw EEG data, synchronized with trials.

It also introduces some markers in the raw EEG data like EEG start and end times, Stimulus start and end time, baseline start time, User response time and eye-blink events using Emotiv Software Development Kit (SDK). Later, the artifacts due to the eye blinks and muscle motion are removed using the filtering method as stated in [26].The block diagram of the capture tool is shown in Fig. 2 where synchronous EEG and Galvanic Skin Response (GSR) are recoded with event markers (i.e. eye blink event etc.). Please note that the GSR data analysis is not presented as it is out of scope.

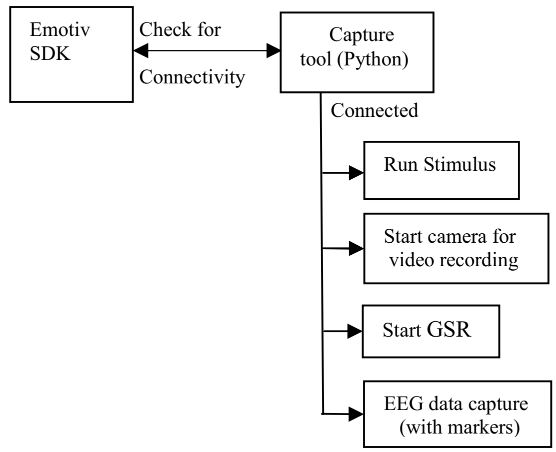


Fig. 2. Block Diagram of the Setup for EEG Capture

IV. Methodology

The detailed preprocessing of raw EEG signal and different methodologies adopted for channel selection are presented in this section.

A. Signal Analysis

The flow of the signal processing steps is depicted in Fig. 3. Here 14 time-series EEG signals (one from each sensor) were first segmented into individual trials. Next these trials were subdivided into inter-trial baseline and trial simulation as explained in [15]. This division was done based on the predefined marker data introduced by the data capture tool while acquiring the data. The segment corresponding to a trial is extracted as a fixed size window of 2.5 seconds around the user response time. This ensures an equal length trial epoch of 5 seconds for all users. The baseline epoch is derived from the 5 second inter-trail interval during which the subjects relax.

Apart from the above, an analysis of continuous 5 seconds trial windows with 50% overlapping is also performed. In this case the baseline is taken from the last relax time before the trial window.

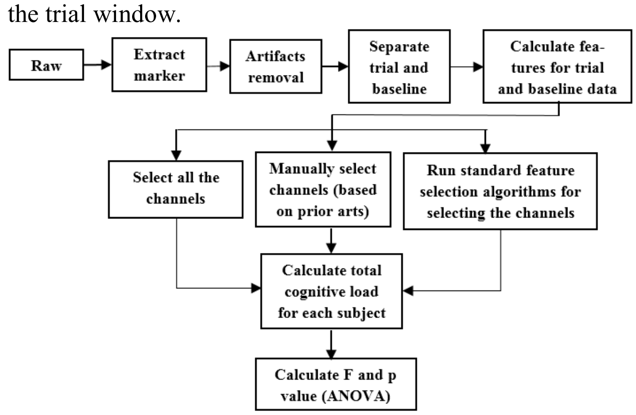


Fig. 3. Signal processing flowchart

Raw EEG data is highly contaminated by eye-blinks. We have followed the process explained by Berka et al. [26] to remove the contaminations due to eye-blink. For detection of eye-blinks, we have relied on the SDK supplied by Emotiv. The trial epochs and the baseline data were then transformed using S-transform. S-transform decomposes a non-stationary signal in time-frequency domain for better precision. Next the alpha and theta band mean frequencies and powers at mean frequencies were calculated for all the selected leads as mentioned in Sec III. These values were then used to derive the total cognitive load as explained in [15]. Finally, we averaged the cognitive loads from all the selected channels to get the single measurement of cognitive load. The mean frequency of alpha, theta band and the total cognitive load were calculated using the formula given by Anderson [15].

The mean frequency is computed by (1).

$$f(\omega)=\frac{\sum_{i=0}^{n-1} I_{\omega(i)} f_{\omega(i)}}{\sum_{i=0}^{n-1} I_{\omega(i)}} \tag{1}$$

where ω is the frequency band in question, n is the number of frequency bins in ω, f_i is the frequency at bin i and I_i is the energy density of ω at frequency bin i.

The mean frequencies for both trial and baseline epochs were calculated. Next the frequency shift between them were calculated. The total cognitive load $L(t)$ was calculated using a combination of power and frequency changes for both alpha (α) and theta (θ) band considering all the selected channels for trail t as given in (2) [9]:

$$L(t)=\Delta|f_t(\alpha)|f_t(\alpha)-\Delta|f_t(\theta)|f_t(\theta) \tag{2}$$

Now, we compute L using three different ways for channel selection as depicted in Fig 3. First, we considered all the 14 channels. Second, we manually selected leads based on the literature studies. This suggests that for specific brain functions like problem solving, visual attention tasks etc, theta waves from frontal lobe (Fz location) and alpha waves from parietal lobe (Pz location) act as the most distinguishing feature. We have tried to find out if these are most discriminative features for Emotiv as well. Since Emotiv does not have any sensor at Fz or Pz locations, we selected P7 and P8 from parietal lobe and F3 and F4 from frontal lobe as they seem to be the closest representative of Pz and Fz. We calculated total cognitive load using (2) considering alpha and theta waves from these channels only. Finally averaging the values obtained from these channels gives a single measure of cognitive load. Similarly, left four channels of frontal lobe are manually selected to measure the cognitive load in the above manner.

Next we use couple of feature selection algorithms to derive the significant channels for both alpha and theta waves. The total cognitive load was again calculated using (2) but only for the leads selected by the feature selection algorithms.

B. Channel Selection

Apart from channels found in the prior literature, we have performed channel selection using two methods – connectionist framework based on ANN and MIC using support vector machine (SVM). The S-transform [15] of each epoch of EEG signal is a time-frequency data representing the frequency response at each instance of time. The energy values (***E***) and the mean frequency (***f***) for each band are computed for each lead at every time instance of the epoch. The maximum, minimum and average values of ***E*** in a given epoch are termed as $E_{\max}^{l,i}, E_{\min}^{l,i}, E_{avg}^{l,i}$ respectively where, $i \in \{\alpha,\theta\}$ frequency bands and $1 \le l \le 14$ denotes the 14 leads of the Emotiv EEG device. Similarly, the maximum, minimum and average values of mean frequencies in a given epoch are termed as $f_{\max}^{l,i}, f_{\min}^{l,i}, f_{avg}^{l,i}$ respectively. The composite feature vector is derived from the energy values and the mean frequency of the α and θ bands as given in (3).

$$F=\{E_{\max}^{l,i}, E_{\min}^{l,i}, E_{avg}^{l,i}, f_{\max}^{l,i}, f_{\min}^{l,i}, f_{avg}^{l,i}\} \tag{3}$$

The feature vector $F \in \Re^{128}$ consists of 6x2x14 = 128 dimension features. These features are then used for the subsequent lead selection whose objective is to find the set of leads for α and θ bands that maximizes the separation of the computed cognitive load L as in (2) for the previously mentioned two types of tasks.

1) Connectionist Framework based channel selection

The features in (3) are used to train a connectionist framework based ANN [24] having 128 input nodes and two output nodes and a hidden layer. The number of nodes for the hidden layer is experimentally chosen to be 25 using the method suggested by Hagiwara [27]. Once the ANN is trained then the selection layer holds a set of weights (W_k) those are proportional to the importance of the features in discriminating the two types of tasks. Though the selection layer is an integral part of the ANN structure, it just holds a linear scale factor for the input features. This motivated the authors to derive the lead selection from the weights (W_k) of the features.

The initial lead assignments for each of the six types of features for all the C channels/ leads are done using (4) to (9). If the weights for the corresponding feature values are greater than a predefined threshold η, then the corresponding lead for the feature is set to 1 else it is set to 0.

$$l_{E_{\max}^{i,l}}=\{0,1\,|\,l_{E_{\max}^{i,l}}=1,\ if\ W_{E_{\max}^{i,l}}\ge\eta\}\ \forall 1\le l\le C,\ i\in\{\alpha,\theta\} \tag{4}$$

$$l_{E_{\min}^{i,l}}=\{0,1\,|\,l_{E_{\min}^{i,l}}=1,\ if\ W_{E_{\min}^{i,l}}\ge\eta\}\ \forall 1\le l\le C,\ i\in\{\alpha,\theta\} \tag{5}$$

$$l_{E_{avg}^{i,l}}=\{0,1\,|\,l_{E_{avg}^{i,l}}=1,\ if\ W_{E_{avg}^{i,l}}\ge\eta\}\ \forall 1\le l\le C,\ i\in\{\alpha,\theta\} \tag{6}$$

$$l_{f_{\max}^{i,l}}=\{0,1\,|\,l_{f_{\max}^{i,l}}=1,\ if\ W_{f_{\max}^{i,l}}\ge\eta\}\ \forall 1\le l\le C,\ i\in\{\alpha,\theta\} \tag{7}$$

$$l_{f_{\min}^{i,l}}=\{0,1\,|\,l_{f_{\min}^{i,l}}=1,\ if\ W_{f_{\min}^{i,l}}\ge\eta\}\ \forall 1\le l\le C,\ i\in\{\alpha,\theta\} \tag{8}$$

$$l_{f_{avg}^{i,l}}=\{0,1\,|\,l_{f_{avg}^{i,l}}=1,\ if\ W_{f_{avg}^{i,l}}\ge\eta\}\ \forall 1\le l\le C,\ i\in\{\alpha,\theta\} \tag{9}$$

In case of Emotiv EEG device, there are 14 leads, hence C is 14. For the present experiments, the threshold η is selected as 0.3. After the initial assignments of the leads for each feature we need to combine them to arrive at the selected leads for the α and θ bands. The method for the combination is given in (10) to (12). Initially, the intersection is taken for the leads corresponding to $E_{\max}^{l,i}$, $E_{avg}^{l,i}$ and $E_{\min}^{l,i}$, $E_{avg}^{l,i}$. Later the union is done for the above. Similar processing is done for the mean frequency (*f*). Finally, the union of the leads is done for the leads identified by energy and mean frequency in (12).

$$l_E^i = U(I(l_{E_{\max}^{l,i}}, l_{E_{avg}^{l,i}}), I(l_{E_{\min}^{l,i}}, l_{E_{avg}^{l,i}})) \tag{10}$$

$$l_f^i = U(I(l_{f_{\max}^{l,i}}, l_{f_{avg}^{l,i}}), I(l_{f_{\min}^{l,i}}, l_{f_{avg}^{l,i}})) \tag{11}$$

$$l^i = U(l_E^i, l_f^i) \qquad \forall \quad \mathrm{i} \in \{\alpha, \theta\} \tag{12}$$

where, *U* denotes the union of sets and *I* denotes the intersection of sets.

The computation of the cognitive load is done using (2) and averaged over the selected leads or channels.

2) MIC based channel selection

The feature vector of 128 dimension derived in (3) is fed into an MIC (Maximal Information Coefficient) based feature selection algorithm. MIC is a statistical tool [25] for measuring the inter-relationship between a pair of dataset based on the concept of binning and grid formation. For every data pair (p, q), if *M* represents the mutual information for a grid *G,* then MIC of a dataset X of sample size n and grid size (p, q) less than b_n is given by (13).

$$MIC(X) = \max_{p,q<b_n} \{M(X)_{p,q}\} \quad \text{where } b_n = n^{0.6} \tag{13}$$

For different distributions of G, $M(X)$ is given by (14).

$$M(X)_{p,q} = \frac{\max\{I(X \mid G)\}}{\log \min(p,q)} \tag{14}$$

Using this MIC value, a gain factor of the k^{th} feature is defined by (15) [28].

$$W_k = \frac{1}{1+e^{-S.(m_k-0.5)}} \tag{15}$$

where S represents the stiffness factor of the sigmoid function which controls the suppression and elevation of the importance of a feature. Therefore a judicious optimization of S is essential for proper evaluation of the gain factor of individual feature. This is realized by using SVM as an optimization function.

To run the SVM, the induced feature vector (V_S), given by (16), is randomized to generate a number of dataset for manifold validation of the SVM classifier and thus choosing the optimum one from a series of gain factors. The randomization is performed using the standard 'randperm' function of Matlab and the typical number (g) of randomized dataset chosen for the experiment is 4.

$$V_S = W_S * F \tag{16}$$

Where W_S represents the gain factor vector corresponding to stiffness value S and V_S is the final induced feature vector. From the accuracy of SVM classifier a Decision Parameter $DP|_S$ is generated as (17).

$$DP|_S = \frac{\max\{R_S(I)\}}{std(R_S(I))} \quad \forall I, \ I_i = \underset{1\le i\le g}{randperm}(r) \tag{17}$$

Here $R_S(I)$ represents the SVM accuracy for a particular S and I.

The final stiffness factor S and thus the corresponding gain factor vector W can be determined as the gain factor vector associated with the maximum Decision Parameter value as given in (18).

$$W = W_S(S \equiv \max(DP|_S)) \tag{18}$$

Once the gain factor vector W is determined, the channel selection can be done in a similar fashion as explained in the previous section. Here the elements of this gain factor vector W are similar as explained in the Connectionist framework based channel selection section.

C. Statistical Analysis

We have used one-way ANOVA to test the significant differences between cognitive loads for low and high load tasks. One-way ANOVA tests the null hypothesis that both the groups are derived from same population and hence there are no significant differences between the class means. If the F value obtained is greater than 1, then it indicates that there is significant difference between the class means. Next the results are tested for statistical significance or p value. Smaller the p-value, lesser is the chance that the test classes belong to the same group. By setting p = 0.01, the critical F value can be determined from a standard lookup table[2]. F value greater than F-critical denotes rejection of null hypothesis. Since, cognitive load directly depends on the frequency bands of the various EEG channels, the ANOVA analysis varies depending on the subset of channels are chosen. Thus there is a need to understand a systematic approach to do the selection. Moreover, there is a need to investigate whether the selection is subject and/or session dependent and repeatable. These are captured and explained in the next section.

V. RESULTS AND DISCUSSIONS

In this section we present a comparative study of various channel selection methods followed by an analysis of the channels responsible for the cognitive load information.

A. Comparison of Various Channel Selection Methods

We present the results obtained by the one-way ANOVA analysis using various approaches of lead selection for 10 subjects undertaking two types of tasks as described in Section III. Table I shows F-value and p-value obtained using ANOVA analysis for the following lead combinations:

[2] http://www.socr.ucla.edu/applets.dir/f_table.html

- All channels: Taking all 14 channels into account.
- Left 4 channels (Left 4): Frontal lobe of left hemisphere i.e., leads AF3, F7, F3 and FC5. We have also experimented with the frontal lobe of right hemisphere (AF4, F8, F4 and FC6) however, they didn't produce mentionable results and hence not reported in this paper. The reason behind this may be that the left hemisphere is responsible for problem solving, language processing, logical thinking, planning etc. [29].
- Frontal and Parietal lobe (FP Lobe): Based on the finding in [11, 12], the cognitive load is calculated taking Theta from leads of frontal lobe (F3, F4) and alpha from leads of parietal lobe (P7, P8).
- Leads from the psychological literature (Phy 7): As reported in [18], there are 7 leads (Cz, P3, P4, Pz, O2, PO4, F7) which are most important for cognitive load. These are used in the experiments by Tian et al. [17] with a 32 lead EEG device. The 14-lead Emotiv EEG device does not have all the 7 leads. In Emotiv, the Cz, Pz are not available; P7 is used instead of P3; P8 is used instead of P4 and PO4. Hence we have taken P7, P8, O2 and F7 for our experiment.
- Connectionist Framework based channel selection (ANN CS): The leads are selected using the weights derived from the connectionist framework of ANN as given in section IV.B.1.
- MIC based channel selection (MIC CS): The leads are selected using the weights derived from the MIC based approach as given in section IV.B.2.
- Mutual Information based channel selection ([17]): As reported by Tian et al. [17] there are global 7 leads which gave good results for 3 subjects using a 32 lead EEG device. Among these, we have used the 6 leads (O1, F8, F7, FC5, FC6, AF3) which are available in Emotiv.

In Table I, the green colored entries correspond to the best F, p and the orange colors correspond the second best for each subject. It can be seen that the "MIC CS" gives best results for 4 subjects, "Left 4" gives best results for 4 subjects and "Phy 7" gives best results for 2 subjects.

The output of ANOVA analysis for continuous 5 seconds window with 50% overlap is shown in Table II. The green colored entries correspond to the best F, p and the orange colors correspond the second best for each subject. It can be seen that for few subjects there are multiple green entries as the F values are very close to each other. The values of F have noticeably increased compared to the one in Table I. This indicates that a subject is continuously experiencing more cognitive load for the "3-back memory" task compared to the "Finding number" task. Hence the continuous window analysis is more suitable to find the cognitive load experienced by a subject for an unknown task. It can also be seen that the "MIC CS" gives best results for 7 subjects. Hence this channel selection algorithm performs the best amongst the compared ones.

TABLE I. RESULTS OF ANNOVA FOR DIFFERENT CHANNEL SELECTION APPROACHES – 5 SEC WINDOW AROUND USER RESPONSE (GREEN COLORED ENTRIES CORRESPOND TO THE BEST F, P AND THE ORANGE COLORS CORRESPOND THE SECOND BEST FOR EACH SUBJECT)

Subjects	All channel	Left 4	FP Lobe	Phy 7	ANN CS	MIC CS	[22]
1	F=0.62 P=0.44	F=9.13 P=0.01	F=7.68 P=0.02	F= 0.01 P= 0.01	F=13.7 P=0.002	F=18.8 P=0.001	F=1.38 P=0.25
2	F=3.70 P=0.06	F=35.3 P = 0	F=11.1 P=0.002	F=25.3 P= 0	F=4.65 P=0.04	F=21.54 P = 0	F=19.5 P = 0
3	F=1.67 P=0.21	F=7.37 P=0.01	F=0.30 P=0.59	F=12.13 P=0.002	F=18.2 P = 0	F=24.7 P = 0	F=1.70 P=0.20
4	F= 0 P=0.95	F=5.04 P=0.03	F=4.56 P=0.04	F=3.93 P=0.06	F=1.85 P=0.19	F=4.61 P=0.04	F=1.49 P=0.23
5	F=1.45 P=0.24	F=2.07 P=0.16	F=3.49 P=0.07	F=11.57 P=0.002	F=0.86 P=0.36	F=4.61 P=0.04	F=4.81 P=0.04
6	F=3.84 P=0.06	F=0.62 P=0.43	F=2.58 P=0.12	F=3.96 P=0.06	F=2.03 P=0.17	F=7.72 P=0.01	F=1.05 P=0.32
7	F=1.74 P=0.20	F=19.6 P= 0	F=7.16 P=0.01	F=13.97 P=0.001	F=0.15 P=0.70	F = 3.21 P = 0.04	F=5.22 P=0.03
8	F=3.94 P=0.06	F=11.9 P=0.002	F=3.55 P=0.07	F=19.33 P= 0	F=12.9 P=0.001	F=21.9 P = 0	F=19.3 P= 0
9	F=1.21 P=0.28	F=23.6 P = 0	F=11.9 P=0.001	F=7.13 P=0.012	F=15.9 P = 0	F=21.13 P = 0	F=9.95 P=0.004
10	F= 0 P=0.99	F=6.0 P=0.02	F=2.06 P= 0.17	F=20.35 P=0	F=6.39 P=0.02	F=15.64 P = 0	F=0.10 P=0.76

TABLE II. RESULTS OF ANNOVA FOR DIFFERENT CHANNEL SELECTION APPROACHES – CONTINUOUS 5 SEC WINDOW WITH 50% OVERLAP(GREEN COLORED ENTRIES CORRESPOND TO THE BEST F, P AND THE ORANGE COLORS CORRESPOND THE SECOND BEST FOR EACH SUBJECT)

Subjects	All channel	Left 4	FP Lobe	Phy 7	ANN CS	MIC CS	[22]
1	F=5.84 P=0.017	F=40.36 P=0	F=40.91 P=0	F=45.96 P=0	F=42.53 P=0	F=56.2 P=0	F=8.84 P=0.003
2	F=15.71 P=0	F=58.14 P=0	F=33.62 P=0	F=58.25 P=0	F=17.13 P=0	F=56.3 P=0	F=37.5 P=0
3	F=0.67 P=0.41	F=32.74 P=0	F=7.53 P=0.006	F=40.94 P=0	F=42.3 P=0	F=47.5 P=0	F=0.14 P=0.71
4	F=2.29 P=0	F =14.73 P=0	F=13.41 P=0	F=46.8 P=0	F=15.9 P=0	F=14.77 P=0	F=10.82 P=0.001
5	F=20.56 P=0.132	F=16.05 P=0	F=22.53 P=0	F=19.07 P=0	F=18.42 P=0	F=0.69 P=0	F=6.12 P=0.014
6	F=7.85 P=0.005	F=2.33 P=0.128	F=9.46 P=0.002	F=33.81 P=0	F=0.05 P=0.814	F=53.76 P=0	F=0.07 P=0.783
7	F=0.12 P=0.732	F=51.71 P=0	F=18.87 P=0	F=26.34 P=0	F=9.36 P=0.003	F=51.73 P=0	F=12.68 P=0
8	F=1.77 P=0.185	F=32.44 P=0	F=4.58 P=0	F=20.54 P=0	F=31.5 P=0	F=34.6 P=0	F=33.67 P=0
9	F =3.52 P=0.01	F =45.23 P=0	F =23.61 P=0	F =18.72 P=0	F =31.78 P=0	F =48.36 P=0	F =20.85 P=0.001
10	F=5.85 P=0.017	F=20.51 P=0	F=1.69 P=0.195	F=13.22 P=0	F=0.974 P=0.326	F=35.86 P=0	F=3.42 P=0.066

In order to understand the repeatability of the outcome, another round of experiments are performed on 4 subjects out of the 7 subjects for which the F,p are marked as green in "MIC CS" column in Table II. For this the previously selected channels for the corresponding subjects are used. Results indicate that the trend is similar and the deviation of F is within 5% with $p < 0.005$.

Next we consider the lead combinations for each subject corresponding to the highest F values and perform the analysis in the remaining section. Fig. 4 gives a pictorial representation of total cognitive load as per equation (2) for all 10 subjects (S1 to S10) considering the leads having maximum F-value with $p < 0.005$, in the ANOVA analysis. The *red* dots represent the cognitive load corresponding to low load task (Finding number) and *blue* dots correspond to cognitive load for high trials (3-back memory test). The figure clearly indicates a difference in average levels for two tasks for almost all subjects.

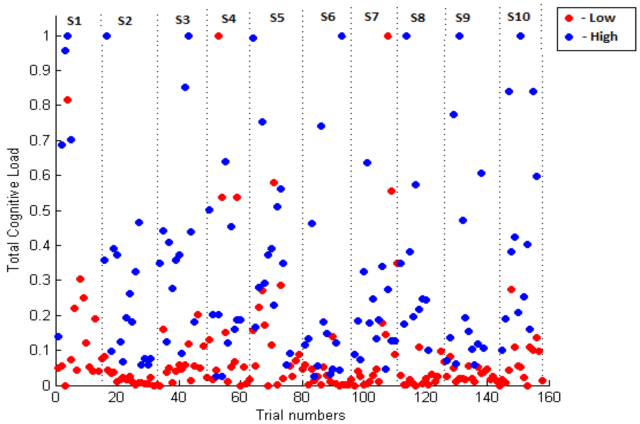


Fig. 4. Total cognitive load (L) for all subjects as per equation (2)

B. *Analysis of Discrimination Power for the channels*

The selected channels obtained from one-way ANOVA analysis can further be analyzed using the color map to derive insights on their discrimination power. The binary representation of selected channels using Alpha activation corresponding to the maximum F value of ANOVA analysis in Table I is shown in Table III, where "S" indicates the subjects from 1 to 10 for 10 rows and columns indicate the leads (1 means selected). Similar set of selected channels are derived for Theta activation.

If we consider the entries in the Table III as entries of a matrix a_{ij}, then we define the channel activation index as given by (19).

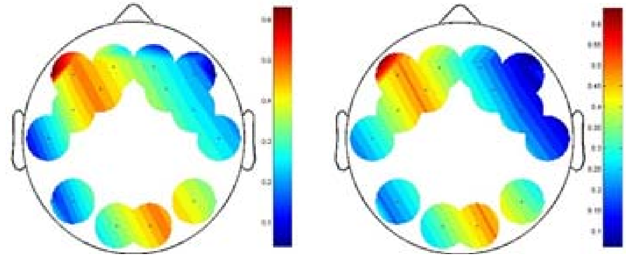
$$C_j = \frac{\sum_{i=1}^{N} a_{ij}}{N}, \quad 1 \le j \le 14 \tag{19}$$

where, N is the number of subjects (in our experiment, N=10) and j is the channel index.

TABLE III. CHANNEL SELECTION FOR ALPHA ACTIVATION CORRESPONDING TO MAXIMUM F-VALUE

S	AF3	F7	F3	FC5	T7	P7	O1	O2	P8	T8	FC6	F4	F8	AF4
1	0	1	0	1	1	0	1	1	0	1	0	0	0	0
2	1	1	1	1	0	0	0	0	0	0	0	0	0	0
3	1	0	1	0	0	1	0	1	0	1	0	1	0	0
4	1	1	1	1	0	0	0	0	0	0	0	0	0	0
5	0	1	0	0	0	1	0	1	1	0	0	0	0	0
6	0	0	0	1	0	0	1	0	1	0	0	1	0	0
7	1	1	1	1	0	0	0	0	0	0	0	0	0	0
8	0	0	0	1	0	1	0	1	0	0	0	0	1	0
9	1	1	1	1	0	0	0	0	0	0	0	0	0	0
10	0	1	0	0	0	1	0	1	1	0	0	0	0	0

The C_j is used to plot the channel intensity color map for Alpha activation in Fig. 5(a). *Red* indicates highest and *blue* indicates lowest discrimination power respectively. Similarly, the activation index for Theta band is depicted in Fig. 5(b). Both the figures depict a strong discrimination power for left frontal lobe and partially for parieto-occipital lobe of right hemisphere.



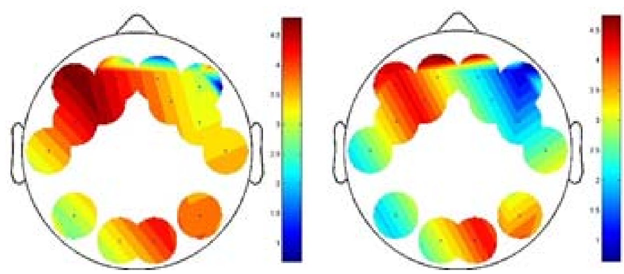
(a) (b)

Fig. 5. Channel Discrimination Map for (a) Alpha and (b) Theta Activation.

Next we perform similar analysis by deriving weighted channel activation index C_j^F using F-value as the weighing factor, shown in (20). The $F_i^{\max}$ is the maximum F value for the subject i in Table II. It is to be noted that for each subject, the selected leads in the rows of the channel selection matrix in Table III, correspond to maximum F values of Table II.

$$C_j^F = \frac{\sum_{i=1}^{N} a_{ij} * F_i^{\max}}{N}, \quad 1 \le j \le 14 \tag{20}$$

The C_j^F is used to plot the channel intensity color map for Alpha and Theta activation in Fig. 6(a) and Fig. 6(b) respectively.



(a) (b)

Fig. 6. Channel Discrimination Map using ANOVA F-value for (a) Alpha and (b) Theta Activation.

In this case also, a strong discrimination map is observed for left frontal lobe and partially for parieto-occipital lobe of right hemisphere. The strength of the map provides the information about the importance of the channels in the cognitive load measurement using a low resolution 14-lead Emotiv EEG device.

We conclude from the results that for few subjects the channels of the left hemisphere perform the best whereas for other subjects the channels and bands selected by the feature selection algorithm work well. It is seen that the prior arts findings (majority of those are reported on high resolution systems) do not comply with the current scenario especially for a low resolution EEG device. It is quite a unique finding in this paper. Hence this approach for channel or lead selection can be successfully used for different scenarios of human behavior understanding using portable low cost head gear devices like the one provided by Emotiv. Moreover, the findings of the lead selection information for a specific subject can be further used to position a single lead EEG device namely Neurosky[3] for the improved performance. However, the analysis on the Neurosky data is out of the scope of this paper.

VI. Conclusions

From the detailed analysis in the paper, it can be inferred that the choice of channel plays a pivotal role in determining the cognitive load experienced by subjects while executing a particular task. Clear separation between different levels of cognitive loads can be achieved from the optimum selection of channel combinations. Though the present analysis focused on two level classifications, it can further be extended to multilevel segregation as well. It is also evident from the above discussion that lead activation is very specific to the type of the task executed and it is very much subject dependent. Therefore, it is advisable to select the subject specific lead combination for achieving a better accuracy. The continuous window analysis with 50% overlapping provides much better result compared to the analysis around subjects' responses. The MIC based channel selection algorithm performs best for 70% of the subjects. The left frontal and right parieto-occipital channels demonstrate the most discriminative power, for the selected two types of tasks, as depicted in the color map representations. As a future scope we plan to investigate the most significant leads indicative of various cognitive functions, which would enable automatic channel selection for a given task.

Acknowledgment

We would like to thank all the subjects participated in the in the EEG data capture.

References

[1] G. A. Miller, "The magical number seven, plus or minus two: Some limits on our capacity for processing information," Psychological Review, 63(2), pp. 81–97, Mar. 1956.

[2] P. Chandler, J. Sweller "Cognitive load theory and the format of instruction. Cognition and Instruction 8 (1991), pp. 293–332.

[3] Kilseop, R. and Rohae, M. "Evaluation of mental workload with a combined measure based on physiological indices during a dual task of tracking and mental arithmetic". International Journal of Industrial Ergonomics 35(11), (2005), 991-1009.

[4] Y. Shi, N. Ruiz , R. Taib, E. Choi, F. Chen: "Galvanic skin response (gsr) as an index of cognitive load." In CHI '07 extended abstracts (2007), ACM, pp. 2651–2656

[5] W. Klimesch, "EEG alpha and Theta oscillations reflect cognitive and memory performance: a review and analysis" in Brain Research Reviews 29, 1999.

[6] F. Lotte, J. Van Dille Dirk, L. Koole, "Tuning down the emotional brain: An fMRI study of the effects of cognitive load on the processing of affective images" in NeuroImage 2009.

[7] Solovey, E., Schermerhorn, P., Scheutz, M., Sassaroli, A., Fantini, S., & Jacob, R. (2012). "Brainput: enhancing interactive systems with streaming fnirs brain input". In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 2193-2202). ACM.

[8] Barbur, J. L., Watson, J. D., Frackowiak, R. S., & Zeki, S. (1993). Conscious visual perception without VI. Brain, 116(6), 1293-1302.

[3] http://neurosky.com/

[9] Kimura, D. (1961). Some effects of temporal-lobe damage on auditory perception.Canadian Journal of Psychology/Revue canadienne de psychologie, 15(3), 156.

[10] Brody, B. A., & Pribram, K. H. (1978). The role of frontal and parietal cortex in cognitive processing: tests of spatial and sequence functions. Brain: a journal of neurology, 101(4), 607-633.

[11] Jensen, O., Gelfand, J., Kounios, J., & Lisman, J. E. "Oscillations in the alpha band (9–12 Hz) increase with memory load during retention in a short-term memory task". Cerebral Cortex, 12(8), pp. 877-882, 2002.

[12] Sauseng, P., Klimesch, W., Schabus, M., & Doppelmayr, M. "Fronto-parietal EEG coherence in theta and upper alpha reflect central executive functions of working memory". International Journal of Psychophysiology, 57(2), pp. 97-103, 2005

[13] Chatterjee, D., Sinharay, A., & Konar, A. "EEG-based fuzzy cognitive load classification during logical analysis of program segments". In IEEE International Conference on Fuzzy Systems (FUZZ), 2013 (pp. 1-6).

[14] A. Sinharay, D. Chatterjee, A. Sinha, "Evaluation of different on-screen keyboard layouts using EEG signals", published in SMC 2013, October 13-16, 2013.

[15] Anderson, Erik W., et al. "A user study of visualization effectiveness using EEG and cognitive load." Computer Graphics Forum. Vol. 30. No. 3. Blackwell Publishing Ltd, 2011.

[16] Blankertz, B., Kawanabe, M., Tomioka, R., Hohlefeld, F., Müller, K. R., & Nikulin, V. V. "Invariant common spatial patterns: Alleviating nonstationarities in brain-computer interfacing". In Advances in neural information processing systems (pp. 113-120) 2007.

[17] Lan, Tian, et al. "Channel selection and feature projection for cognitive load estimation using ambulatory EEG", Computational intelligence and neuroscience 2007, 8-8.

[18] C. A. Russell and S. G. Gustafson, "Selecting salient features of psychophysiological measures," Tech. Rep. AFRL-HE-WP-TR-2001-0136, Air Force Research Laboratory, Rome, NY, USA, 2001.

[19] Arvaneh, M., Guan, C., Ang, K. K., and Quek, C. "Optimizing the channel selection and classification accuracy in EEG-based BCI" IEEE Transactions on Biomedical Engineering, 58.6 pp. 1865-1873, 2011.

[20] Roland, Per E. Brain activation. Wiley-Liss, 1993.

[21] Sinha, Aniruddha, Kingshuk Chakravarty, and Brojeshwar Bhowmick. "Person identification using skeleton information from kinect." ACHI 2013, The Sixth International Conference on Advances in Computer-Human Interactions. 2013.

[22] Lin, Chen, et al. "Maximal information coefficient for feature selection for clinical document classification." ICML Workshop on Machine Learning for Clinical Data. Edingburgh, UK. 2012.

[23] H. Sijuan, "Feature extraction and classification of EEG for imagery movement based on mu/beta rhythms", In Biomedical Engineering and Informatics (BMEI), 2010 3rd International Conference on 16-18 Oct, 2010.

[24] N. R. Pal and K. K. Chintalapudi, "A connectionist system for feature selection", Neural Parallel Sci. Comput. 1997, vol. 5, no. 3, pp. 359–381.

[25] D. N. Reshef, Y. A. Reshef, H. K. Finucane, S. R. Grossman, G. McVean, P. J. Turnbaugh, E. S. Lander, M. Mitzenmacher, and P. C. Sabeti, "Detecting novel associations in large data sets," Science, vol. 334, no. 6062, pp. 1518–1524, 2011.

[26] Berka C., Levendowski D. J., Cvetinovic M. M., Petrovic M. M., Davis G., Lumicao M. N., Zivkovic V. T., Popovic M. V., Olmstead R.: "Real-time analysis of eeg indices of alertness, cognition and memory with a wireless eeg headset". Int'l J. of HCI 17 (2004), 151–170.

[27] M. Hagiwara, "A simple and effective method for removal of hidden units and weights", Neurocomputing, vol. 6, pp. 207-218, 1994.

[28] Banerjee, R., Sinha, A., Choudhury, A. D., & Visvanathan, A. "PhotoECG: Photoplethysmographyto estimate ECG parameters". In IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2014 (pp. 4404-4408).

[29] J. Siegmund, C. K¨astner, C.Parnin, S. Apel, A.Bethmann, T. Leich, G. Saake, A. Brechmann, "Understanding Understanding Source Code with Functional Magnetic Resonance Imaging", International Conference on Software Engineering, 2014.