Preprint Version. Manuscript submitted to and accepted by IEEE Transactions on Neural Systems and Rehabilitation Engineering.

STEW: Simultaneous Task EEG Workload Dataset

W. L. Lim, O. Sourina, *Member, IEEE*, and L. P. Wang, *Member, IEEE*

***Abstract*—This paper describes an open access electroencephalography (EEG) dataset for multitasking mental workload activity induced by a single-session simultaneous capacity (SIMKAP) experiment with 48 subjects. To validate the database, EEG spectral activity was evaluated with EEGLAB and the significant channels and activities for the experiment are highlighted. Classification performance was evaluated by training a support vector regression model on selected features from neighborhood component analysis based on a 9-point workload rating scale. With a reduced feature dimension, 69% classification accuracy was obtained for 3 identified workload levels from the rating scale with a Cohen's kappa of 0.46. Accurate discrimination of mental workload is a desirable outcome in the field of operator performance analysis and BCI development, thus we hope that our provided database and analyses can contribute to future investigations in this research field.**

***Index Terms*—Electroencephalography (EEG), Mental Workload, Open Access Dataset**

I. Introduction

THE goal of BCI research aims to provide an alternate pathway for users to communicate with devices. In particular, for an EEG based BCI, this is achieved through receiving EEG signals from the user's brain, which should elicit a particular response from the device. To obtain the desired response, the processing algorithm has to be able to correctly identify and classify the user's incoming brain signal such as the detection of the P300 in a BCI speller application [1]. Over the years, many experiments have been performed to develop state of the art processing algorithms that address this requirement of BCI [2-5]. While these studies provide well documented and advanced methods to process BCI data, most of these researches will often choose to validate their methods with their own in-house experimental dataset, usually without releasing the data online. This is undesirable due to two main reasons.

Firstly, it is difficult for other research groups to compare methods and reproduce the stated result if the original database is not provided. In order to provide a workaround, studies usually replicate the methods used in a previous study on their own dataset to serve as a point of comparison [6-8]. To ensure a fair comparison, the proposed method should also be applied on the original referenced datasets, if the two classification contexts are similar.

Secondly, it is resource intensive to conduct a large scale experiment with a sizable number of subjects. Also some research groups might not have the required manpower or resources to establish their own dataset. In order to validate their proposed algorithms, these studies often select a dataset from the EEG databases available for open access [9-11]. However, the current number of databases available is still small and should be expanded upon.

Furthermore, although there are well established open access EEG datasets, each dataset might consider an explicit research area or different modalities and thus might not be applicable depending on a researcher's area of study. For example the DEAP database is a dataset that considers the research area of emotional state [12] while the dataset provided in [13] considers multimodal BCI for a mental workload task and motor imagery. Therefore, it is important that the research community have access to a variety of databases to study. For our dataset, we aim to provide single session EEG data of forty-eight subjects performing multitasking mental workload activity.

We have identified a growing need for the provision of a sizable, open access mental workload EEG dataset for BCI research. Thus, we would like to contribute our dataset toward this goal, with this paper serving as its documentation, providing information on the experimental setup, EEG baseline frequency analysis and classification performance.

II. Related Work

A. Mental Workload

Mental workload (MWL) is defined as the amount of mental or cognitive resources required to meet the current task demands [14]. A high MWL would mean that most or all cognitive resources have been utilized to perform the given task.

The assessment of MWL is an important consideration in

This research is supported by the National Research Foundation, Singapore under its International Research Centres in Singapore Funding Initiative. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not reflect the views of National Research Foundation, Singapore.

W. L. Lim is with the School of Electrical and Electronics Engineering, Nanyang Technological University, 639798, Singapore. (e-mail: WLIM031@e.ntu.edu.sg).

O. Sourina is with the Fraunhofer Institute Singapore, Nanyang Technological University, 639798, Singapore. (e-mail: EOSourina@ntu.edu.sg).

L. P. Wang is with the School of Electrical and Electronics Engineering, Nanyang Technological University, 639798, Singapore. (e-mail: ELPWang@ntu.edu.sg).

DOI:10.1109/TNSRE.2018.2872924

the area of operator performance in order to avoid task errors due to the high workload or "overload" condition [15]. Being able to correctly recognize the MWL of an operator can enhance safety with practical BCI applications. For example, in the area of Air Traffic Management, a passive BCI solution can be implemented to automatically adjust task settings based on the workload of the operator [16].

MWL is traditionally assessed with questionnaires such as the NASA Task load index (NASA-TLX) [17] or Subjective Workload Assessment Technique (SWAT) [18]. As these methods only provide subjective assessment of an operator's workload, the current trend is to complement these ratings with physiological measurements using devices that measure bio-signals such as the EEG or fMRI [14].

In order to properly assess MWL with such devices, there is a need to be able to recognize the workload level of the incoming signal, and this can be achieved with the use of various machine learning techniques [2, 3].

B. Experiments involving Mental Workload

Experiments that assess MWL usually include one of two popular formats to induce workload. The first is that of a task battery, where subjects are to attend to several tasks appearing in two or more separate task windows. This format, which aims to increase MWL by means of multitasking, was first popularized by NASA's Multi-Attribute Task Battery (MATB) [19, 20] with studies involving MWL using the MATB or similarly inspired task [21-26]. The second format is by using mental arithmetic to induce workload, with more complex arithmetic problems for a higher workload level [27-32].

While there are many studies that conduct experiments to induce MWL via EEG, there are few who release their datasets online for further study and validation by other groups. Although there are available datasets such as the BCI competition database and compilation websites [9, 10], the numbers of datasets related to MWL are still limited. A recent open access dataset that provides multi-modal EEG and near-infrared spectroscopy (NIRS) recording of mental arithmetic and motor imagery data is also available for study [13].

Some limitations of the available datasets introduced above include one or a combination of the following. First, the datasets have relatively few subjects, usually less than 10, thus making it difficult to validate generalized MWL activity. Second, the selected subjects are non-uniform, i.e. the subjects are of different gender, age groups or education levels. These variables might adversely affect the uniformity of MWL EEG data collected. For example, as females have lighter skull structures, EEG collected would have higher potential compared to men. If age-groups and education levels are not consistent, subjects performing MWL tasks might display varying results based on individual competency; subjects with a higher education level might find it easier to perform complex arithmetic problems. These datasets are therefore more suited for subject-specific studies, or studies comparing individual differences.

The proposed dataset aims to account for the discussed limitations by selecting male participants from a specific group. This allows for a uniform dataset where studies on general MWL EEG mechanisms across many subjects can be performed.

C. Frequency Bands as Measure of Mental Workload

There are unique characteristics specific to MWL activity found in previous studies, such as the sensitivity to alpha and theta EEG power spectral density (PSD). These are also popular features in EEG signal classification applications [24, 27, 31].

Furthermore, given the prevalence of the frequency power bands in general EEG studies, we shall base the analysis of our dataset on them, as they provide a standard baseline measure in studying the underlying neural mechanisms of the EEG.

III. Methods and Materials

A. Subjects

Fifty male subjects from the university's graduate population participated in this study. Recruitment was performed via open email and all subjects recruited declared to not have any neurological, psychiatric or brain related diseases. They also declared not to have taken part in any prior EEG experiment. Participants were informed of the experimental procedure and written consent was obtained. After the experiment, participants were provided monetary compensation for their time. This study was conducted according to the declaration of Helsinki and was approved by the Institutional Review Board of the Nanyang Technological University (approval number: IRB-2014-04-026).

B. Description of the SIMKAP Experiment

Subjects are asked to perform the Simultaneous Capacity (SIMKAP) test module of the Vienna Test System [33]. SIMKAP is a commercial psychological test created by Schuhfried GmbH for the purpose of assessing an individual's multitasking and stress tolerance. While the test is designed as an assessment tool to screen personnel for their multitasking ability in multitasking heavy occupations such as air traffic management, the test has also been applied in a variety of research scenarios involving multitasking [34-36].

The SIMKAP multitasking test requires subjects to cross out identical items by comparing two separate panes, whilst responding to auditory questions which can be arithmetic, comparison or data lookup in nature. Some cases of auditory questions require subjects to respond at a later time, thus requiring them to monitor a clock on the upper right corner. This multitasking component lasts 18 minutes. The order of questions and tasks in this activity are fixed for all subjects, as designed by the developers of the Vienna Test System. A screenshot of the interface of SIMKAP can be viewed in Fig. 1.

As the test utilizes the task battery format and involves some form of arithmetic problems in addition to other auditory questions, the test follows formats established in previous studies [19-26] and hence is a viable stimulus to induce MWL.

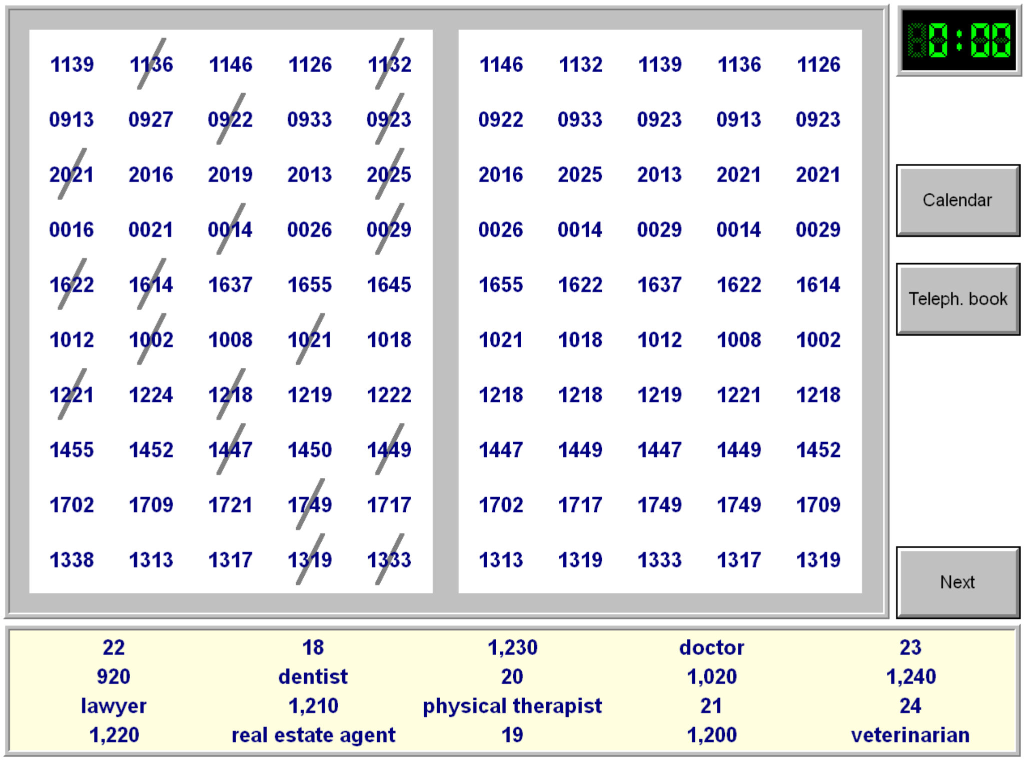

Fig. 1. Screenshot of the SIMKAP multitask test. Subjects are to mark items in the right panel by matching those already crossed out on the left panel. Responses to auditory questions are completed by selecting the correct answer from the bottom panel. Auditory questions include arithmetic problems, comparison problems, and information lookup with calendar or telephone book.

Defining **cognitive workload as the amount of mental effort**, on the scale(1-9) below, rate the cognitive challenge involved in the task of this segment with 1 being the lowest and 9 being the highest.

1	2	3	4	5	6	7	8	9

Fig. 2. Questionnaire on a 1-9 scale for rating of mental workload, which subjects were required to fill after completion of each segment of the experiment.

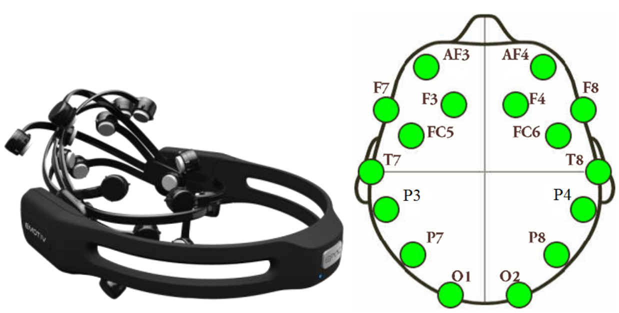

Fig. 3. The Emotiv EEG Device used in this study and electrode positions based on the 10-20 international system.

C. Experimental Procedure

Subjects were seated comfortably; approximately 60cm in front of a 24 inch LED display and were told not to make any unnecessary movements apart from responding to the stimuli during the experiment.

There are two parts to the experiment. First, subjects were asked to maintain a comfortable position with eyes open and not perform any task for 3 minutes. Their EEG was recorded and these 3 minutes of recording is then used as the resting condition. Next subjects were asked to perform the SIMKAP test with EEG being recorded and the final 3 minutes of the recording is used as the workload condition. The first and last 15 seconds of data from each recording was excluded to reduce effects from any between task activity, resulting in recordings of 2.5 minutes. Subjects were asked to rate their perceived MWL after each segment of the experiment on a rating scale of 1 to 9. This was performed as a form of subjective validation that the subject indeed experienced an increase in workload while performing the test as compared to the resting condition. One can perceive a rating of 1-3 as low (lo) workload, 4-6 as moderate (mi) workload and 7-9 as high (hi) workload. The 9-point rating scale [37] used is analogous to the NASA-TLX's 1 to 21 scale and is the most frequently used measure in cognitive load studies according to review in [38]. A screenshot of the questionnaire used can be viewed in Fig. 2.

D. Data Acquisition

EEG data was collected using Emotiv EPOC EEG headset with sampling frequency of 128Hz and 16 bit A/D resolution. The device comprises of fourteen electrodes located at AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4, shown in Fig. 4 according to the 10-20 international system [39]. Data is transmitted to a paired PC desktop via wireless Bluetooth and raw data is recorded with the Emotiv

'TestBench' software.

The Emotiv device was used as it can be easily mounted and provides comparable signal quality to a BioSemi or G-TEC device [40, 41]. A picture of the Emotiv headset and the corresponding electrode positions used for recording in the experiment is shown in Fig. 3.

E. Data Processing

Only 48 of the 50 subjects data was used to form the database as the data of 2 subjects were found to be incomplete. All data processing was done using MATLAB R2018a with EEGLAB, a popular and well documented tool for processing of EEG signals [42].

1) Pre-processing of raw EEG data

It is important to first pre-process raw EEG data to remove artifacts resulting from muscle movement and to clean the noise from data before proceeding with any analysis. Here, we follow the recommended pre-processing pipeline suggested by a developer of EEGLAB [43]. The general steps are:

1. High-pass filter the raw data at 1Hz
2. Remove line noise
3. Perform Artifact Subspace Reconstruction (ASR)
4. Re-reference data to average

The key preprocessing step is the ASR which is a non-stationary method to remove large-amplitude artifacts [44]. Fig. 4 shows sample data before and after pre-processing steps. We observe that the ASR algorithm removed the large amplitude artifact in channel F3 and reconstructed the channel data successfully.

2) Analysis of EEG data with STUDY

We use EEGLAB's STUDY functionality to load the pre-processed datasets to explore the EEG mechanisms across subjects for the different task conditions. We are interested in studying the following "between" conditions:

1. No task vs. SIMKAP task
2. Rating Based Lo vs. Mi vs. Hi MWL

While exploring "between" conditions, we also include the spectral analysis of "within" conditions whilst performing the two above studies. For study 1, we use all 48 subjects' data, but for study 2, we ignore data from S05, S24 and S42 as rating data was not available for these subjects. Study 2 is particularly interesting to see if we can verify subjective ratings with objective EEG spectral data.

For each study, entire length of data from each channel is used to study the significant spectral mechanisms pertaining to each condition and between conditions that contribute to the overall neural activity.

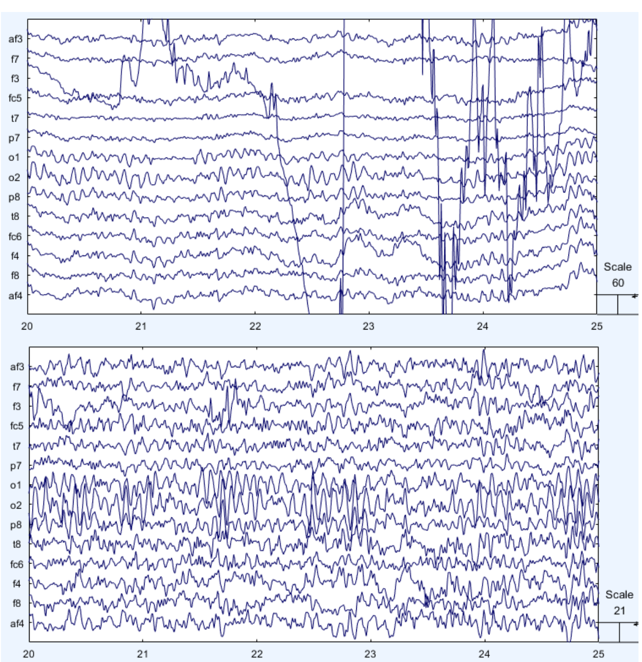

Fig. 4. Sample continuous time EEG channel data before (top) and after completion of preprocessing steps (bottom).

F. Classification Method

We also provide classification performance analysis for the proposed dataset, based on the ratings provided by 45 subjects using PSD features via FFT of the delta, theta, alpha and beta bands. These features are chosen for simplicity and extensive usage in previous studies, hence they serve as a good baseline for analysis. A sliding window of size 512 and shift 128 was used and as all 14 channels are considered, the studied feature set has input dimension of 4 x 14.

A regression problem is considered with the aim to predict the rating of unseen EEG data. 80% of the data (36 subjects) was used to conduct feature selection and training while 20% of the data (9 subjects) was kept as unseen test data.

Feature selection was first performed using Neighborhood Component Analysis (NCA) to select features for regression [45], using 5-fold cross validation. The best features accounting for 75% of the total feature weights across all folds are selected for use to train a Support Vector Regression (SVR) model. The predicted ratings are then converted to labels according to the rating scale: 1-3 as low, 4-6 as moderate and 7-9 as high and classification performance is assessed by comparing with the true labels of the unseen data.

IV. Results

We shall first present the findings from EEG spectral analysis of the two studies with topographical scalp maps, spectral power graphs and regions of significance between conditions. Then, we present results of the feature selection and the resulting classification performance.

A. STUDY results

1) No task vs. SIMKAP task

For the "No task" condition, from the topographical scalp maps, we observe that delta activity is concentrated around the AF3, AF4, F4 and F8 positions, with some activity around the O1 position. Theta activity is present in AF3, AF4, F3, F4, F8 and T8, as well as being present in occipital O1, O2 and parietal P7 and P8 positions. Alpha activity is observed in the AF4, F8, T8, O1, O2, P7 and P8 positions while beta activity is seen in AF4, F8, FC6, T8, O1 and O2 positions.

For the "SIMKAP" condition, activity is present in FC5, AF4, F8 and FC6 for both delta and theta bands. For alpha and beta, activity is present in the same areas as delta and theta, with additional activity in O1 and O2 positions.

Comparing both conditions, we observe higher overall PSD values for the "SIMKAP" condition across all frequencies. Significant frequency regions for each channel are shaded in grey, with the most significant channels being FC5, FC6 and F8. The compiled results of study 1 is displayed in Fig. 5 and those of study 2 is shown in Fig. 6.

2) Rating Based Low vs. Moderate vs. High MWL

For study 2, the frequency band activity is similar to that described in study 1, with the "Low" condition being similar to "No task", with "Moderate" and "High" conditions being similar to the "SIMKAP" condition. This is confirmed by viewing the spectrum graph and observing that the graph for "Moderate" and "High" conditions are almost equal, and similar to the shape of the "SIMKAP" condition. Likewise, the "Low" and "No task" condition graphs are similar. The regions of significance are mostly concentrated in F8 and FC5.

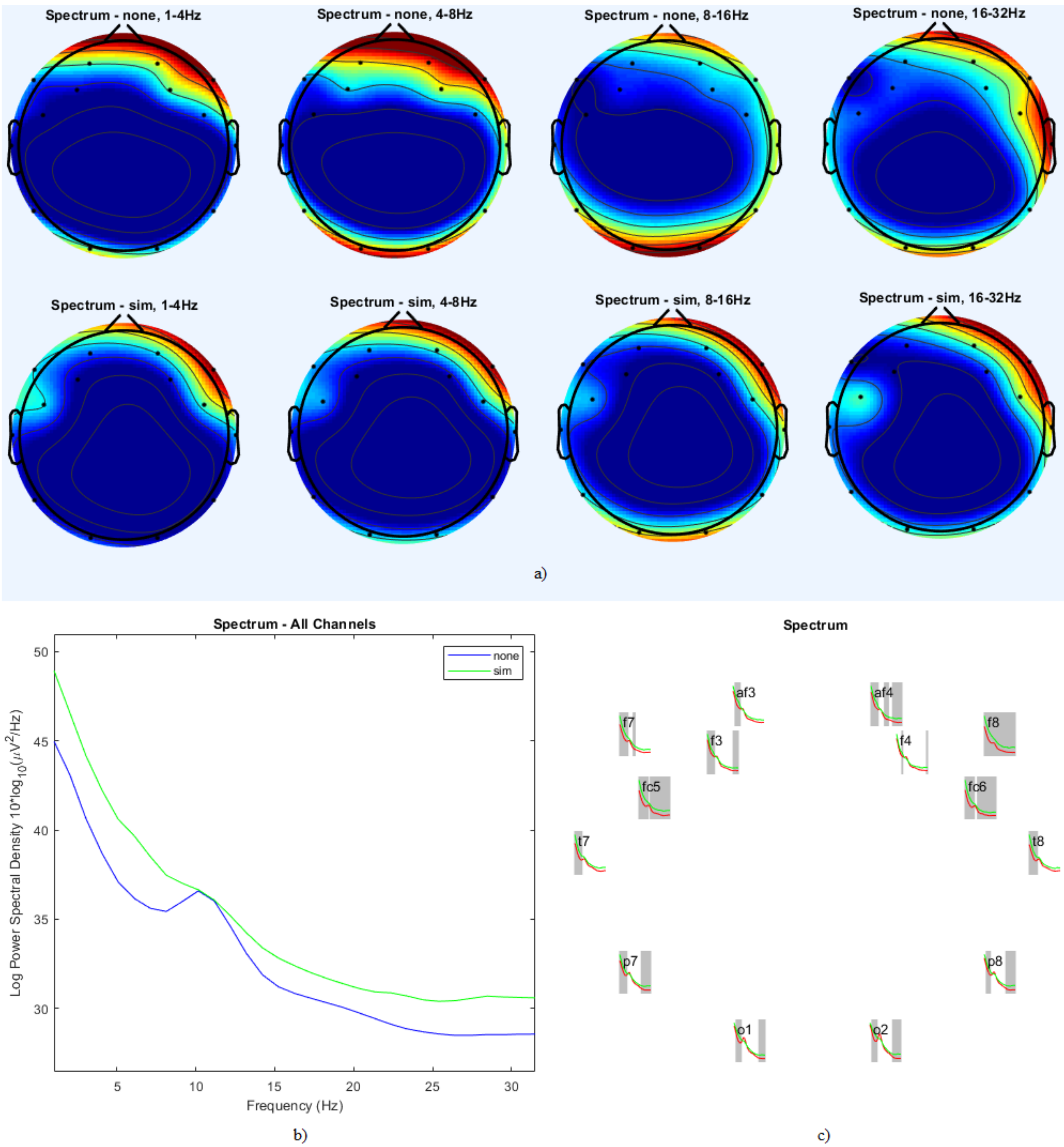

Fig. 5. Results for study between different conditions based on task a) Scalp topography for different frequency bands b) PSD for different rating conditions c) Statistically significant frequency regions for each channel

DOI:10.1109/TNSRE.2018.2872924

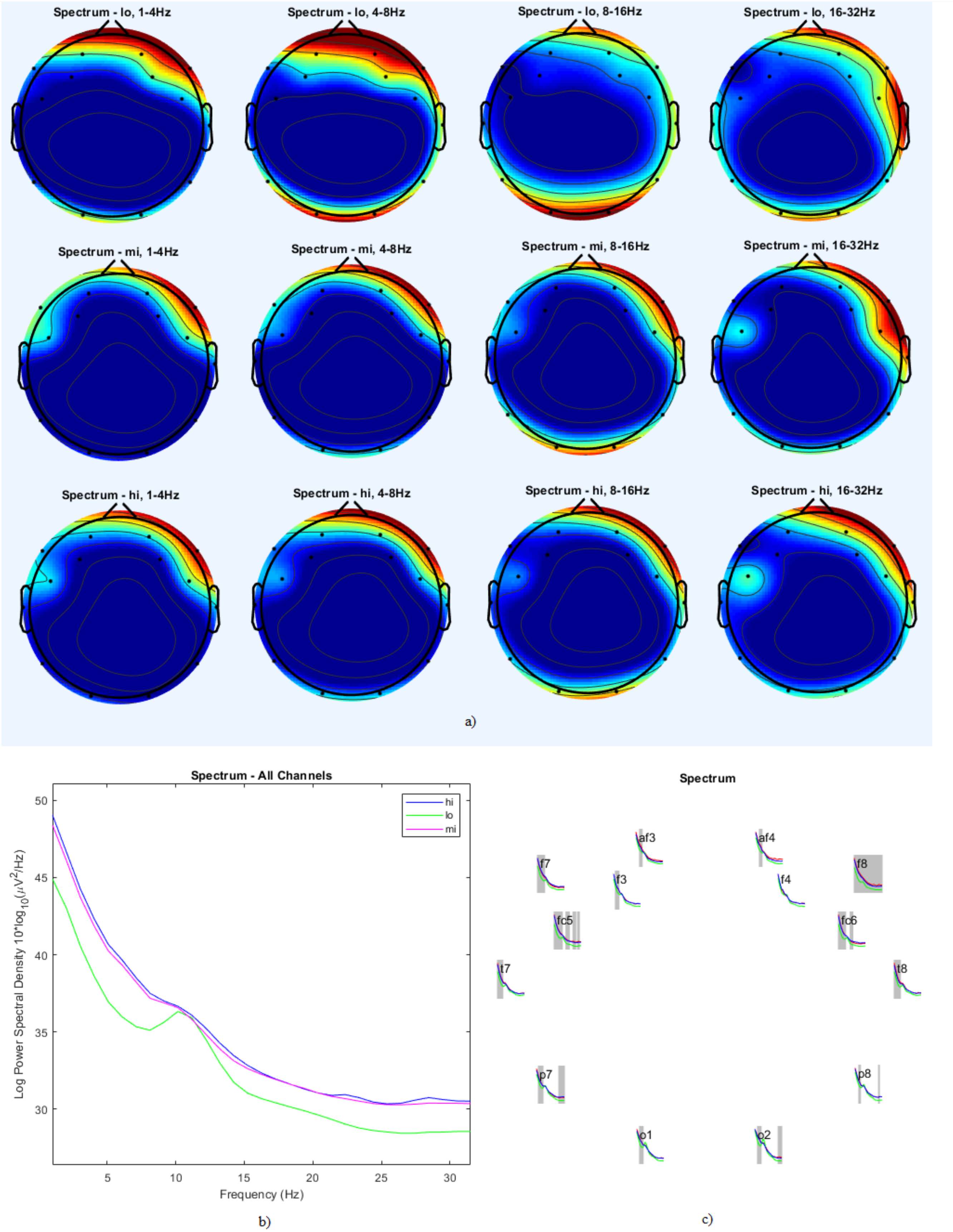

Fig. 6. Results for study between different conditions based on rating scales a) Scalp topography for different frequency bands b) PSD for different rating conditions c) Statistically significant frequency regions for each channel

DOI:10.1109/TNSRE.2018.2872924

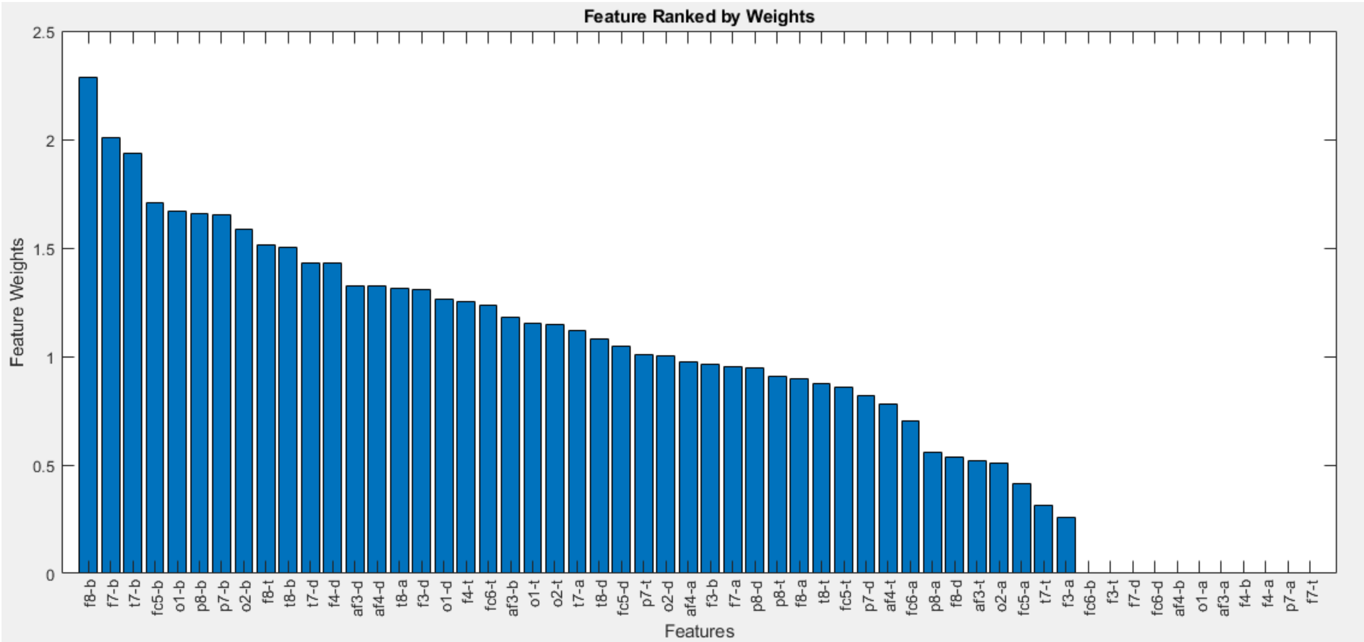

Fig. 7 Feature weights from Neighborhood Component Analysis

Fig. 8 Confusion matrix using 28 features

B. *Classification results*

NCA was performed to evaluate the weights of the 56 features, with results shown in Fig. 7. The top features accounting for at least 75% of the weights were selected to train an SVR model, which resulted in a final feature dimension of 28, a reduction of half the initial feature dimensions. The trained regression model was used to predict the rating values for the unseen 20% test data, with the predicted values converted to either “low”, “moderate” or “high” labels based on the respective range the rating values are in. A classification accuracy of 69% was achieved on the test set, with the confusion matrix shown in Fig. 8. The calculated Cohen’s kappa is 0.46 with expected random chance accuracy of 42.4%.

If all features are used to train the regression model, the resultant classification accuracy is 69.2% with kappa value of 0.47. The expected classification accuracy by random chance is 41.7%.

V. Discussion

A. *Spectral Analysis and Classification*

Spectral topographies of the different EEG frequency bands are provided for each task condition and for three possible workload classification levels based on rating scale. Delta activity is localized in frontal areas for all conditions with an average increase in PSD for a higher workload. Increase in theta PSD localized in frontal areas for higher activity was also observed, similar to results reported in [46]. Decrease in alpha activity in the occipital areas and increase in beta activity in frontal areas especially in channel location F8 was observed for increasing mental workload. A study in [47] reported similar findings.

We are also able to verify the subjective ratings of the subjects with EEG spectral activity, by observing from the PSD graph comparing “low”, “moderate” and “high” workload levels. There is a marginal positive difference between PSD values across most of the frequencies when comparing the “high” and “moderate” conditions indicating that the ratings are somewhat accurate in accounting for different workload levels. However, this slight difference also highlights the inherent weakness of subjective ratings, where subjects might not reliably report their experience after performing tasks, causing the two curves to be almost exact. As the obtained graphs are average PSD across many subjects and channels, any variation due to individual difference would complicate the prediction of workload rating levels for unseen data in the “moderate” and “high” classes.

This issue is exemplified in our classification analysis of selected PSD features. The confusion matrix shows a high

DOI:10.1109/TNSRE.2018.2872924

error rate of classification for both the "moderate" and "high" levels, at 53.8% and 68.9% respectively. While general performance of the model is acceptable, more work can be done to effectively classify the "moderate" and "high" classes.

B. Evaluation of the EEG open access dataset

The EEG MWL dataset described in this paper provides a sizable pool of 48 subjects utilizing commercial psychological multitasking test software as the stimuli. A key benefit of using a commercial test comes in the detailed support documentation provided by the company [34] if required.

The dataset has the benefit of having uniformity in terms of subject data, reducing possible individual difference arising from gender, age and education levels. The dataset is also accompanied with subjects' rating of workload, allowing the possibility for studies linking subjective and objective measures to be performed.

Furthermore, given the sizable number of subjects, it is also possible to explore approaches for both intra-subject and inter-subject classification schemes and develop algorithms for BCI applications.

However, due to the specificity in terms of subject selection, the dataset might be unable to account for an overview of EEG mental workload characteristics for the general population.

VI. Conclusion

In this paper, we have described an open access EEG database using the SIMKAP multitasking activity to obtain MWL data. Our dataset is provided open access to supplement the existing pool of MWL datasets with the double benefit of a large group of subject data with an official commercial psychological test for multitasking as stimuli. Spectral analysis and classification has been performed to illustrate the validity of the data for research, as the results obtained are similar to studies on EEG MWL data performed previously.

We hope that in providing this sizable dataset of 48 subjects, development of novel BCI and EEG data classification algorithms, particularly to account for subjective and objective data, can be facilitated. The raw dataset is available for download via: http://dx.doi.org/10.21227/44r8-ya50.

References

[1] C. Guger, S. Daban, E. Sellers, C. Holzner, G. Krausz, R. Carabalona, *et al.*, "How many people are able to control a P300-based brain–computer interface (BCI)?," *Neuroscience letters,* vol. 462, pp. 94-98, 2009.

[2] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi, "A review of classification algorithms for EEG-based brain–computer interfaces," *Journal of neural engineering,* vol. 4, p. R1, 2007.

[3] A. Khorshidtalab and M. J. E. Salami, "EEG signal classification for real-time brain-computer interface applications: A review," in *Mechatronics (ICOM), 2011 4th International Conference On*, 2011, pp. 1-7.

[4] S. Sun and J. Zhou, "A review of adaptive feature extraction and classification methods for EEG-based brain-computer interfaces," in *Neural Networks (IJCNN), 2014 International Joint Conference on*, 2014, pp. 1746-1753.

[5] J. Kevric and A. Subasi, "Comparison of signal decomposition methods in classification of EEG signals for motor-imagery BCI system," *Biomedical Signal Processing and Control,* vol. 31, pp. 398-406, 2017.

[6] E. Parvinnia, M. Sabeti, M. Z. Jahromi, and R. Boostani, "Classification of EEG Signals using adaptive weighted distance nearest neighbor algorithm," *Journal of King Saud University-Computer and Information Sciences,* vol. 26, pp. 1-6, 2014.

[7] P. Zarjam, J. Epps, and N. H. Lovell, "Beyond subjective self-rating: EEG signal classification of cognitive workload," *IEEE Transactions on Autonomous Mental Development,* vol. 7, pp. 301-310, 2015.

[8] Y. Zhang, G. Zhou, J. Jin, Q. Zhao, X. Wang, and A. Cichocki, "Sparse bayesian classification of EEG for brain–computer interface," *IEEE transactions on neural networks and learning systems,* vol. 27, pp. 2256-2267, 2016.

[9] *EEG open access datasets*. Available: http://bnci-horizon-2020.eu/database/data-sets

[10] *EEG/ ERP data available for free public download.* Available: http://sccn.ucsd.edu/~arno/fam2data/publicly_available_EEG_data.html

[11] P. Sajda, A. Gerson, K. R. Müller, B. Blankertz, and L. Parra, "A data analysis competition to evaluate machine learning algorithms for use in brain-computer interfaces," *IEEE Trans. Neural Syst. Rehabil. Eng.,* vol. 11, pp. 184-185, Jun. 2003.

[12] S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, *et al.*, "Deap: A database for emotion analysis; using physiological signals," *IEEE Transactions on Affective Computing,* vol. 3, pp. 18-31, 2012.

[13] J. Shin, A. von Luhmann, B. Blankertz, D.-W. Kim, J. Jeong, H.-J. Hwang, *et al.*, "Open Access Dataset for EEG+ NIRS Single-Trial Classification," *IEEE Transactions on Neural Systems and Rehabilitation Engineering,* 2016.

[14] M. S. Young, K. A. Brookhuis, C. D. Wickens, and P. A. Hancock, "State of science: mental workload in ergonomics," *Ergonomics,* vol. 58, pp. 1-17, 2015.

[15] J. Reason, "Human error: models and management," *BMJ: British Medical Journal,* vol. 320, p. 768, 2000.

[16] P. Aricò, G. Borghini, G. Di Flumeri, A. Colosimo, S. Bonelli, A. Golfetti, *et al.*, "Adaptive automation triggered by EEG-based mental workload index: a passive brain-computer interface application in realistic air traffic control environment," *Frontiers in human neuroscience,* vol. 10, 2016.

[17] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research," *Advances in psychology,* vol. 52, pp. 139-183, 1988.

[18] G. B. Reid and T. E. Nygren, "The subjective workload assessment technique: A scaling procedure for measuring mental workload," *Advances in psychology,* vol. 52, pp. 185-218, 1988.

[19] J. R. Comstock Jr and R. J. Arnegard, "The multi-attribute task battery for human operator workload and strategic behavior research," 1992.

[20] Y. Santiago-Espada, R. R. Myer, K. A. Latorella, and J. R. Comstock Jr, "The Multi-Attribute Task Battery II (MATB-II) Software for Human Performance and Workload Research: A User's Guide," 2011.

[21] G. F. Wilson and C. A. Russell, "Real-time assessment of mental workload using psychophysiological measures and artificial neural networks," *Human factors,* vol. 45, pp. 635-644, 2003.

[22] A. Gevins and M. E. Smith, "Neurophysiological measures of cognitive workload during human-computer interaction," *Theoretical Issues in Ergonomics Science,* vol. 4, pp. 113-131, 2003.

[23] C. Berka, D. J. Levendowski, C. K. Ramsey, G. Davis, M. N. Lumicao, K. Stanney, *et al.*, "Evaluation of an EEG workload model in an Aegis simulation environment," in *Proceedings of SPIE*, 2005, pp. 90-99.

[24] A. Holm, K. Lukander, J. Korpela, M. Sallinen, and K. M. Müller, "Estimating brain load from the EEG," *The Scientific World Journal,* vol. 9, pp. 639-651, 2009.

[25] C. A. Kothe and S. Makeig, "Estimation of task workload from EEG data: new and current tools and perspectives," in *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, 2011, pp. 6547-6551.

[26] Z. Wang, R. M. Hope, Z. Wang, Q. Ji, and W. D. Gray, "Cross-subject workload classification with a hierarchical Bayes model," *NeuroImage,* vol. 59, pp. 64-69, 2012.

[27] K. Ryu and R. Myung, "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,* vol. 35, pp. 991-1009, 2005.

[28] C. Berka, D. J. Levendowski, M. N. Lumicao, A. Yau, G. Davis, V. T. Zivkovic, *et al.*, "EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks," *Aviation, space, and environmental medicine,* vol. 78, pp. B231-B244, 2007.

DOI:10.1109/TNSRE.2018.2872924

[29] B. Rebsamen, K. Kwok, and T. B. Penney, "Evaluation of cognitive workload from EEG during a mental arithmetic task," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 2011, pp. 1342-1345.
[30] C. Mühl, C. Jeunet, and F. Lotte, "EEG-based workload estimation across affective contexts," *Frontiers in neuroscience,* vol. 8, 2014.
[31] L. J. Trejo, K. Kubitz, R. Rosipal, R. L. Kochavi, and L. D. Montgomery, "EEG-based estimation and classification of mental fatigue," *Psychology,* vol. 6, p. 572, 2015.
[32] W. K. So, S. W. Wong, J. N. Mak, and R. H. Chan, "An evaluation of mental workload with frontal EEG," *PloS one,* vol. 12, p. e0174949, 2017.
[33] O. Bratfisch and E. Hagman, "SIMKAP–Simultankapazität/Multi-Tasking," *Mödling: Schuhfried GmbH,* 2008.
[34] M. Bühner, C. J. König, M. Pick, and S. Krumm, "Working memory dimensions as differential predictors of the speed and error aspect of multitasking performance," *Human Performance,* vol. 19, pp. 253-275, 2006.
[35] Y.-H. Li and J. Shiu, "A Normative Study of Cognitive Ability Tests in Chinese-speaking Student Pilots," *J Aviat Med,* vol. 30, pp. 29-40, 2016.
[36] A. Ahmad, S. Darmoul, A. Dabwan, M. Alkahtani, and S. Samman, "Human Error in Multitasking Environments in *Proceedings of the 2016 International Conference on Industrial Engineering and Operations Management*, 2016.
[37] F. G. Paas, "Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach," *Journal of educational psychology,* vol. 84, no. 4, p. 429, 1992.
[38] T. de Jong, "Cognitive load theory, educational research, and instructional design: some food for thought," *Instructional Science,* journal article vol. 38, no. 2, pp. 105-134, March 01 2010.
[39] R. W. Homan, J. Herman, and P. Purdy, "Cerebral location of international 10–20 system electrode placement," *Electroencephalography and clinical neurophysiology,* vol. 66, pp. 376-382, 1987.
[40] A. J. Ries, J. Touryan, J. Vettel, K. McDowell, and W. D. Hairston, "A comparison of electroencephalography signals acquired from conventional and mobile systems," *Journal of Neuroscience and Neuroengineering,* vol. 3, pp. 10-20, 2014.
[41] K. Stytsenko, E. Jablonskis, and C. Prahm, "Evaluation of consumer EEG device Emotiv EPOC," in *MEi: CogSci Conference 2011, Ljubljana*, 2011.
[42] A. Delorme and S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *Journal of neuroscience methods,* vol. 134, no. 1, pp. 9-21, 2004.
[43] *Makoto's preprocessing pipeline.* Available: http://sccn.ucsd.edu/wiki/Makoto's_preprocessing_pipeline.html
[44] T. R. Mullen *et al.*, "Real-time neuroimaging and cognitive monitoring using wearable dry EEG," *IEEE Transactions on Biomedical Engineering,* vol. 62, no. 11, pp. 2553-2567, 2015.
[45] W. Yang, K. Wang, and W. Zuo, "Neighborhood Component Feature Selection for High-Dimensional Data," *JCP,* vol. 7, no. 1, pp. 161-168, 2012.
[46] K. C. Eschmann, R. Bader, and A. Mecklinger, "Topographical differences of frontal-midline theta activity reflect functional differences in cognitive control abilities," *Brain and cognition,* vol. 123, pp. 57-64, 2018.
[47] G. F. Wilson, C. R. Swain, and P. Ullsperger, "EEG power changes during a multiple level memory retention task," *International Journal of Psychophysiology,* vol. 32, no. 2, pp. 107-118, 1999.

DOI:10.1109/TNSRE.2018.2872924