Cross dataset workload classification using encoded wavelet decomposition features

Lim, Wei Lun; Sourina, Olga; Wang, Lipo

2018

Lim, W. L., Sourina, O., & Wang, L. (2018). Cross dataset workload classification using encoded wavelet decomposition features. Proceedings of the International Conference on Cyberworlds, 300-303. doi:10.1109/CW.2018.00062

https://hdl.handle.net/10356/145993

https://doi.org/10.1109/CW.2018.00062

Cross Dataset Workload Classification Using Encoded Wavelet Decomposition Features

Wei Lun Lim
Fraunhofer Singapore
Nanyang Technological University
Singapore
wlim031@e.ntu.edu.sg

Olga Sourina
Fraunhofer Singapore
Nanyang Technological University
Singapore
EOSourina@ntu.edu.sg

Lipo Wang
School of Electrical and Electronic Engineering
Nanyang Technological University
Singapore
ELPWang@e.ntu.edu.sg

***Abstract*— For practical applications, it is desirable for a trained classification system to be independent of task and/or subject. In this study, we show one-way transfer between two independent EEG workload datasets: from a large multitasking dataset with 48 subjects to a second Stroop test dataset with 18 subjects. This was achieved with a classification system trained using sparse encoded representations of the decomposed wavelets in the alpha, beta and theta power bands, which learnt a feature representation that outperformed benchmark power spectral density features by 3.5%. We also explore the possibility of enhancing performance with the utilization of domain adaptation techniques using transfer component analysis (TCA), obtaining 30.0% classification accuracy for a 4-class cross dataset problem.**

Keywords-Transfer learning; EEG workload; Wavelet decomposition; Autoencoders

I. INTRODUCTION

In workload classification applications involving electroencephalography (EEG) data as input, most would consider a subject and task specific approach. In other words, a classification system would first be trained on sample EEG data from a specific user performing a specific task. This system would then be used to classify subsequent EEG workload data generated from the same user performing the same task, with recalibration before every session. This approach by far gives the best possible performance and is widely used in practical applications. However, this system does not allow much in the way of generalization, with reduced performance when either task, subject or session is varied.

To address this issue, previous studies have provided insight on solving generalized EEG workload classification problems in a cross-subject [1, 2], cross-task [3, 4, 5] or cross-session [6] manner.

Outside of workload classification problems, transfer learning via domain adaptation has been shown as a viable method in EEG emotion classification [7, 8, 9] and fatigue studies [10]. Using large-scale data has also been shown to improve performance with transfer learning [11].

Ultimately, the aim would be to develop a system that is able to classify EEG workload data regardless of the subject or workload task in question.

With consideration to the works above, we explore possible approaches to a subject and task independent EEG workload classification problem and conduct several simulations to evaluate their performance. Formally, we shall provide analysis for the following two hypotheses: a) Learned feature representations from a large EEG workload dataset have better capability in transferring to a new, independent EEG workload dataset compared to classical features of a similar type. b) Transfer learning techniques provide additional performance compared to classical approaches for EEG workload classification problems.

II. METHODS

A. *Cross-Dataset Description*

The first dataset consists of 48 subjects selected from the post graduate university population, performing a single session of the SIMKAP multitasking test of the Vienna test system developed by Schuhfried GmbH [12].

Four levels of cognitive workload were identified from the experiment process, with 3 minutes of EEG data recorded for each level. For the first level, subjects do not perform any actions. For the second level, subjects attempted a visual matching task. For the third level, subjects performed a question and answer task with stimuli given via audio and answers chosen from a selection on screen. The final level involves multitasking where subjects perform the prior two tasks simultaneously with additional audio questions requiring them to input answers at a certain time with reference to an on-screen timer.

The second dataset consists of 18 maritime trainees who performed a four level Stroop test lasting 1 minute for each level, with EEG data being recorded.

The first level of the test had the subject observe a test screen without performing any actions. The second level required subjects to enter a correct response based on the ink color of the text displayed on screen. In this level, the ink color matches the text, e.g. the word “blue” is displayed in a blue color. The third level requires subjects to perform the same task, with a mismatch between the text description and ink color. The word “blue” might be displayed in a yellow color and the correct response would be the answer “yellow”. The final level has the subject perform the task in level three with an additional restriction of responding within 1 second of stimuli display, to encourage quicker cognitive processing of the stimuli shown.

In both datasets, questionnaires were administered for each workload level, as a subjective measure of workload. The trend from the questionnaire indicates that the subjects feel an increase in workload from level 1 to 4. Sample screens from both tests and the questionnaire can be viewed in Fig. 1 – 3.

From the above dataset description, the subjects’ background (university graduate vs maritime trainee) and tasks involved (multitasking test vs Stroop test) are sufficiently different for a cross dataset EEG cognitive workload classification analysis.

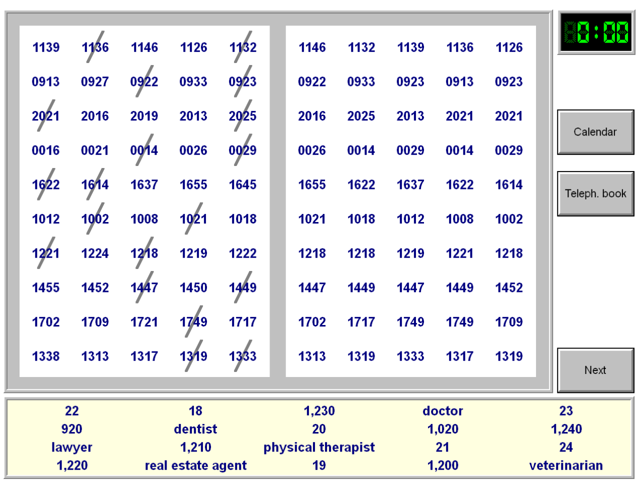

Figure 1. Screenshot of the SIMKAP multitask test. Subjects are to mark items in the right panel by matching those already crossed out on the left panel. Responses to auditory questions are completed by selecting the correct answer from the bottom panel.

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

Figure 2. Questionnaire on a 1-9 scale for rating of mental workload, which subjects were required to fill after each task.

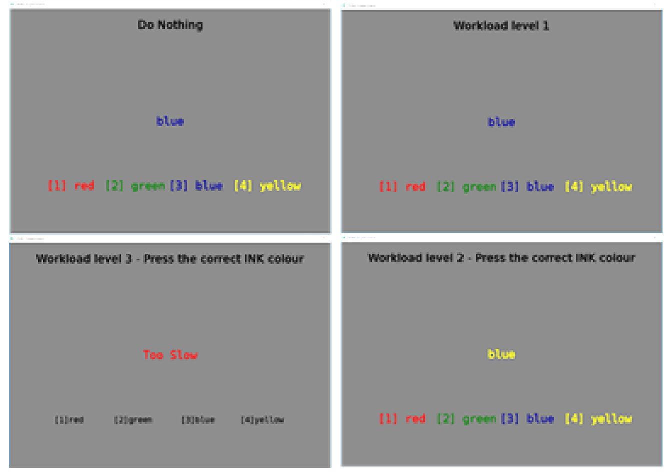

Figure 3. Stroop test for 4 levels workload calibration. Clockwise from top left: Level 0, subjects are to observe the screen but not respond. Level 1, subjects are to respond by pressing the correct key corresponding to the ink color displayed. Level 2, same task as level 1 but with mismatch between word meaning and ink color. Level 3, same task as level 2 but with response time limit of 1 second imposed after stimulus display.

B. *Data Processing*

In both datasets, the Emotiv EPOC device with sampling frequency of 128Hz was used to measure the EEG of subjects during the tasks. 14 channels at positions AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4, according to the 10-20 international system were monitored. The EEG data was processed with a non-overlapping window of 4 seconds over the theta (4-8Hz), alpha (8-16Hz) and beta (16-32Hz) bands to obtain the average power spectral density (PSD), average wavelet coefficient and encoded representation of wavelet decomposition. These three bands were specifically selected as they show correlation with EEG workload activities [4, 13]. In addition, the features under consideration were also chosen in a way that ensures a fair comparison. Being derived from the same EEG frequency bands, the physiological relationship is preserved, with the same dimensionality for each set of features.

Wavelet decomposition using symlet wavelet of order 9 (sym9) [14] was performed to obtain wavelet coefficients pertaining to the 3 bands at decomposition levels 4, 3 and 2. Average coefficient value for each stated level was taken to form a feature set.

To obtain an encoded representation, each set of wavelet coefficients were passed through separate sparse autoencoders with number of hidden neurons set to number of integer frequency values multiplied by number of channels, with logistic sigmoid as the transfer function. The resultant encoded representations were concatenated and passed through another sparse autoencoder with number of hidden neurons set to number of features multiplied by number of channels, with a positive saturating linear transfer function. This two-level encoder design tries to learn relationships within each individual band at the first level and then attempts to learn relationships between each frequency band at the second level. The encoder structure is illustrated in Fig. 4. All processing work was performed on MATLAB r2018a.

C. Classification

To establish a baseline performance of the proposed features to be studied, we apply the softmax classifier, implemented by training a fully connected neural network with 4 hidden neurons and a softmax layer. The network is trained with the SIMKAP dataset as training data and a single subject from the Stoop dataset is used as test data. This is repeated for each of the 18 subjects with classification accuracy recorded. With this baseline established, we proceed to explore domain adaptation via transfer component analysis, which learns a new representational "transfer component" subspace which minimizes data distribution between two separate domains [15]. A basic grid search was used to optimize the hyperparameters for each feature set. In our study, the first domain will be the entire SIMKAP dataset while the second domain will be a subject from the Stroop dataset.

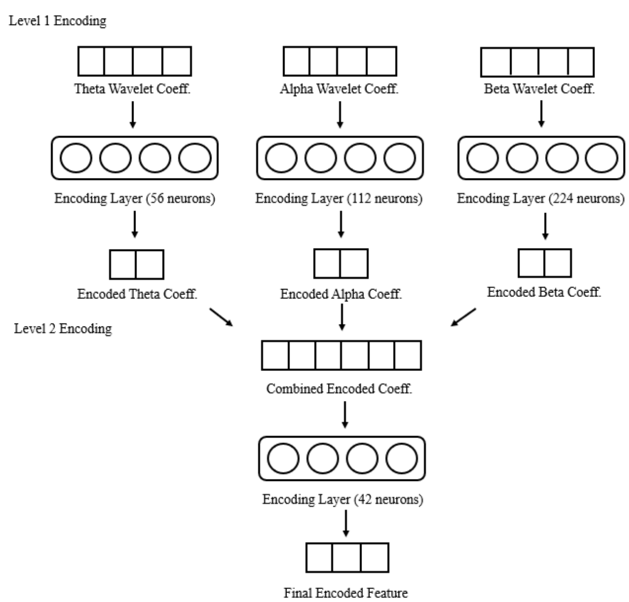

Figure 4. Two level encoding structure for wavelet coefficient representation.

III. RESULTS

We first describe the performance in terms of average accuracy obtained for the different sets of features studied, using the softmax classifier. An average classification accuracy of 27.2% was obtained using the encoded wavelet features, compared to a 23.7% using PSD features and 23.5% using average wavelet coefficient features. Statistical testing with a two tailed Wilcoxon's signed rank test comparing treatments between the different features obtains a significant difference, with at least $p < 0.1$, for the encoded wavelet representation with respect to PSD or average wavelet coefficient. No significant difference was obtained when comparing between PSD and wavelet coefficient features.

When TCA was tested, the average numerical performance of all input features showed an increase. Statistical significance, however, was not observed either between feature sets or when comparing between results obtained by utilizing TCA and that of the softmax classifier.

Table 1 summarizes the main results described above while Fig. 5 and 6 shows a plot of accuracy for the subjects with the three studied features, with plots for the two different classification methods.

TABLE I. CLASSIFICATION ACCURACY OF DIFFERENT FEATURES USING SOFTMAX AND TCA

Features	Accuracy	
	Softmax	***TCA***
PSD	23.7 % (8.2)	26.9 % (5.3)
Wavelet Coeff.	23.5 % (7.0)	26.8 % (5.0)
Encoded Wavelet Coeff.	27.2 % (5.2)	30.0 % (7.5)

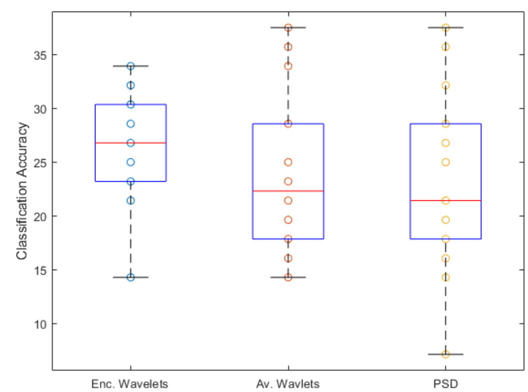

Figure 5. Box plot of 4 class classification accuracy for studied features with softmax classifier.

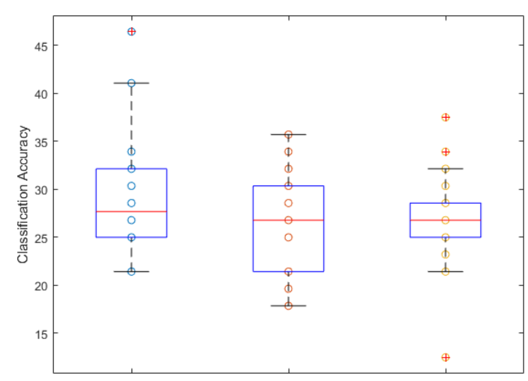

Figure 6. Box plot of 4 class classification accuracy for studied features with domain adaptation via TCA.

IV. Discussion

In this study, we seek out to verify the following two hypotheses, that a) Learned feature representations from a large EEG workload dataset have better capability in transferring to a new, independent EEG workload dataset compared to classical features of a similar type, and b) Transfer learning techniques provide additional performance compared to classical approaches for EEG workload classification problems.

From the results, we observed that there is a significance difference, at a $p = 0.1$ level, between the median values of the treatments when encoded wavelet coefficients were used as features. This finding suggests that implementing an encoding scheme on features from a large dataset might provide more generalization power when performing workload classification on a separate dataset. When domain adaptation was implemented, a significant difference was not observed in both cases comparing between studied features and between classification schemes.

While a conclusion based on statistical significance for the second hypothesis cannot be reached, we observe that the numerical average classification accuracy of all studied features shows an increase of around 3%. This incentivizes the usage of domain adaptation techniques such as TCA as we see a stable improvement in performance across features.

Although our analysis shows that cross dataset problems can benefit from the proposed methodology, the current performance is still too low to be utilized in any real application. For a 4-class problem, with chance levels at 25%, using basic feature sets such as PSD or average wavelet coefficients in a cross-dataset manner leads to performance below 25%. Even with encoding and TCA, while average accuracy has improved to above chance levels, it is still too low for practical implementations.

The formulation of our problem statement was aimed to address cross dataset performance as opposed to previous works on cross subject or cross tasks. A second degree of independence from both an entirely new subject set and task seems to greatly affect performance, hence there is a need to study possible ways to mitigate its effect. With reference to the results in this study, future development in cross dataset analysis is needed and can develop further in any of these three areas: 1) feature development, 2) improved representation learning algorithms and 3) improved domain adaptation algorithms. These areas hold the potential for progress in achieving the goal of obtaining useful levels of classification performance for a cross dataset problem.

Acknowledgment

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.

References

[1] Zhang, J., Wang, Y., and Li, S.: “Cross-subject mental workload classification using kernel spectral regression and transfer learning techniques”, Cognition, Technology & Work, 2017, 19, (4), pp. 587-605.

[2] Spüler, M., Walter, C., Rosenstiel, W., Gerjets, P., Moeller, K., and Klein, E.: “EEG-based prediction of cognitive workload induced by arithmetic: a step towards online adaptation in numerical learning”, ZDM, 2016, 48, (3), pp. 267-278.

[3] Zhao, G., Liu, Y.-J., and Shi, Y.: “Real-Time Assessment of the Cross-Task Mental Workload Using Physiological Measures During Anomaly Detection”, IEEE Transactions on Human-Machine Systems, 2018, 48, (2), pp. 149-160.

[4] Dimitrakopoulos, G.N., Kakkos, I., Dai, Z., Lim, J., Bezerianos, A., and Sun, Y.: “Task-Independent Mental Workload Classification Based Upon Common Multiband EEG Cortical Connectivity”, IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2017, 25, (11), pp. 1940-1949.

[5] Roy, R.N., Charbonnier, S., Campagne, A., and Bonnet, S.: “Efficient mental workload estimation using task-independent EEG features”, Journal of neural engineering, 2016, 13, (2), pp. 1-11.

[6] Yin, Z., and Zhang, J.: “Cross-session classification of mental workload levels using EEG and an adaptive deep learning model”, Biomedical Signal Processing and Control, 2017, 33, pp. 30-47.

[7] Chai, X., Wang, Q., Zhao, Y., Liu, X., Bai, O., and Li, Y.: “Unsupervised domain adaptation techniques based on auto-encoder for non-stationary EEG-based emotion recognition”, Computers in biology and medicine, 2016, 79, pp. 205-214.

[8] Lan, Z., Sourina, O., Wang, L., Scherer, R., and Müller-Putz, G.R.: “Domain Adaptation Techniques for EEG-based Emotion Recognition: A Comparative Study on Two Public Datasets”, IEEE Transactions on Cognitive and Developmental Systems, 2018.

[9] Zheng, W.-L., Zhang, Y.-Q., Zhu, J.-Y., and Lu, B.-L.: “Transfer components between subjects for EEG-based emotion recognition”, In International Conference on Affective Computing and Intelligent Interaction (ACII), 2015, pp. 917-922.

[10] Zhang, Y.-Q., Zheng, W.-L., and Lu, B.-L.: ‘Transfer Components Between Subjects for EEG-based Driving Fatigue Detection’ In International Conference on Neural Information Processing (Springer International Publishing, 2015, edn.), pp. 61-68.

[11] Wei, C.-S., Lin, Y.-P., Wang, Y.-T., Lin, C.-T., and Jung, T.-P.: “Transfer learning with large-scale data in brain-computer interfaces”, In IEEE 38th Annual International Conference of the Engineering in Medicine and Biology Society (EMBC), 2016, pp. 4666-4669.

[12] O. Bratfisch and E. Hagman, "SIMKAP–Simultankapazität/Multi-Tasking," Mödling: Schuhfried GmbH, 2008.

[13] 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.

[14] Al-Qazzaz, N.K., Ali, S., Ahmad, S.A., Islam, M.S., and Ariff, M.I.: “Selection of mother wavelets thresholding methods in denoising multi-channel EEG signals during working memory task”, In IEEE Conference on Biomedical Engineering and Sciences (IECBES), 2014, pp. 214-219.

[15] Pan, S.J., Tsang, I.W., Kwok, J.T., and Yang, Q.: ‘Domain adaptation via transfer component analysis’, IEEE Transactions on Neural Networks, 2011, 22, (2), pp. 199-210.