Comparing Features Extractors in EEG-based Cognitive Fatigue Detection of Demanding Computer Tasks

Rifai Chai, *Member IEEE*, Mitchell R. Smith, Tuan N. Nguyen, *Member, IEEE*,
Sai Ho Ling, *Senior Member*, *IEEE*, Aaron J. Coutts and Hung T. Nguyen, *Senior Member, IEEE*

Abstract— An electroencephalography (EEG)-based classification system could be used as a tool for detecting cognitive fatigue from demanding computer tasks. The most widely used feature extractor in EEG-based fatigue classification is power spectral density (PSD). This paper investigates PSD and three alternative feature extraction methods, in order to find the best feature extractor for the classification of cognitive fatigue during cognitively demanding tasks. These compared methods are power spectral entropy (PSE), wavelet, and autoregressive (AR). Bayesian neural network was selected as the classifier in this study. The results showed that the use of PSD and PSE methods provide an average accuracy of 60% for each computer task. This finding is slightly improved using the wavelet method which has an average accuracy of 61%. The AR method is the best feature extractor compared with the PSD, PSE and wavelet in this study with accuracy of 75.95% in AX-continuous performance test (AX-CPT), 75.23% in psychomotor vigilance test (PVT) and 76.02% in Stroop task (*p*-value < 0.05).

I. INTRODUCTION

Cognitive fatigue refers to a psychobiological state, characterised by feelings of tiredness and discomfort from demanding tasks or exercise. Indeed, recent research has shown there is an impairment to exercise performance following cognitively demanding computer tasks [1, 2]. The computer task used to induce cognitive fatigue in the aforementioned research was the AX-continuous performance test (AX-CPT). Other tasks such as the psychomotor vigilance test (PVT) and the Stroop task may also induce fatigue and impair performance [3]. The AX-CPT demands vigilance, working memory and response inhibition, and requires participants to respond to a string of letters by pressing one of two buttons. The PVT is based on a simple visual reaction time to the stimuli presented at random intervals of time. The Stroop task requires response inhibition and sustained attention on controlled processes as participants respond to the colours of words rather than the meaning of the words. As a result, a practical tool for measuring and detecting cognitive fatigue measurement and detection is required.

The most common tools used to assess cognitive fatigue are the individual-report scaling questionnaires [1]. Whilst these tools are simple to use, they are indirect assessments of cognitive activity and may lack the sensitivity required to detect small but important changes in cognitive fatigue. Moreover, it may not able to detect the onset of the fatigue if the cognitive fatigue occurred before the questionnaire had been recorded. Other tools which can be used for fatigue measurement include capturing eye movements using camera systems or electroocolugraphy (EOG), detecting the heart rate variability using electrocardiography (ECG) and via brain activity using electroencephalography (EEG) [4]. EEG is considered a promising method of cognitive fatigue detection, as previous research has revealed an influence of fatigue on brain activity [5]. The EEG has also been widely used in sleep stage analysis [6]. As a result, this paper explores the use of computational intelligence with different types of feature extractors for EEG-based detection of cognitive fatigue induced by demanding computer tasks.

The most common feature extraction method for fatigue study using EEG is power spectral density (PSD). This converts the time domain data into frequency domain and groups data into EEG bands (delta, theta, alpha and beta) [7]. Other feature extraction methods in other EEG applications such as brain computer interface (BCI) include power spectral entropy (PSE) [8], wavelet as the time-frequency method [9] and autoregressive (AR) as the parametric method [10]. Linear and non-linear classifiers are the common classification method for biomedical applications [11]. Since the EEG signal is a non-linear and multi-dimensional signal, this paper uses Bayesian neural network as a classifier [12]. For practical application, a configuration with fewer EEG channels is preferable as it is time consuming to set up the system, if many channels are included. A recent study shows an association of cognitive fatigue with the frontal area in cognitively demanding tasks [13].

This paper investigates the classification of cognitive fatigue against alert states from experiments of demanding computer tasks by using different feature extraction algorithms including PSD, PSE, wavelet and AR methods with only two EEG channels used on frontal sites (F3 and F4). For the classifier, the Bayesian neural network is used in this paper which is capable of improving the generalization of neural network [12].

II. METHODOLOGY

A. Data Collection

Fourteen participants aged between 18 and 30 years participated in three cognitively demanding computer tasks which were approved by the University Research Ethics Committee. Each experimental session had duration of 45

minutes and sessions were conducted on separate days. The three computer tasks were: PVT, AX-CPT and Stroop task. The PVT task is used for testing of visual reaction time. In this task, a counter appeared in the centre of the screen until a button was pressed by the participants which stopped the counter to display the reaction time (RT). If the counter reached 500 ms without a response, the message 'missed' was displayed. If the response button was pressed before or within the first 100ms of the timer being displayed, an error was signaled by a beep sound with the displaying of a 'false start' message. In the AX-CPT task, sequences of letters were continuously presented one at a time on a computer screen. Participants were required to press one button for a target trial and a separate button for a non-target trial. A beep sound was elicited for any missed or incorrect response as a prompt to increase speed and accuracy. In the Stroop task, coloured words (red, blue, green, yellow) were presented one at a time and participants were required to indicate the colour of a word (not the word's meaning) by pressing a button on the response box corresponding to the colour of the word displayed on the screen.
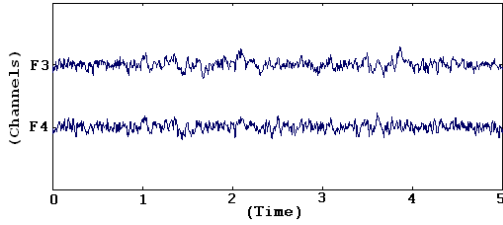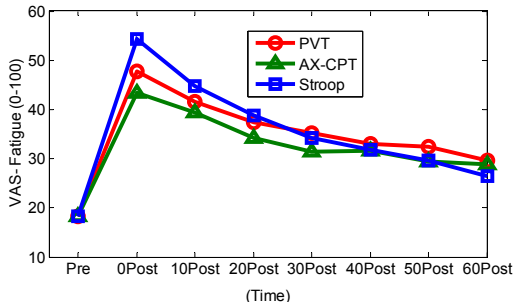

Figure 1. Two EEG channels data collection.

Figure 2. Effect of cognitive tasks on cognitive fatigue based on the VAS

An EEG system (Flexcomp Infinity, Thought Technology) was used for the recording of brain activity with the two channels positioned at frontal areas (F3 and F4) of the international 10-20 system. Brain activity was recorded at a sampling rate of 256 Hz with the signal as shown in Fig. 1. Subjective fatigue was also measured using visual analogue scales (VAS). Both EEG and VAS were measured before, and every 10 min following treatment (post-treatment) up to 60min. In each EEG measurement, a sample with duration of 3 minutes was recorded. The EEG data was divided into alert and fatigue groups. The first 3min from the pre-treatment was used for the alert group. For the fatigue group, it was taken from the EEG data of 10min post treatment of each task. Fig.2 shows the plot of

effects of cognitive tasks on cognitive fatigue using the VAS values. This validates occurrence of fatigue at post-treatment at '0 min' and '10 min' of the post treatment sessions.

B. Pre-Processing and Feature Extraction

The raw EEG data contained some artifacts such as the noise signal from the eyes, heart and muscles; as a result, the raw EEG data was imported to EEGLAB [14] for visual inspection and rejection of artifacts. The 3 min data from the pre-treatment (alert) group and from each of the 10 min post-treatment (fatigue) groups were segmented by using 2 seconds moving windows with overlapping of half of a second. This resulted in 357 segments from each 3 min of data, with 4998 data point of alert state and another 4998 data point for fatigue state of each task (AX-CPT, PVT and Stroop tasks) for 14 participants. A number of feature extraction algorithms were used to extract the features of these data. In this study, four feature extraction algorithms were used for comparison and best algorithm in term of accuracy could be found. The feature extractors for the comparison were PSD, PSE, Wavelet and AR.

The PSD is the most popular method in EEG analysis of fatigue research which is based on frequency analysis [5, 7]. In this paper, the Welch spectrum of the PSD was used. This can be calculated using the Fast Fourier transform (FFT) method. The PSD of the Welch spectrum is given by:

$$\hat{P}_w(f) = \frac{1}{S}\sum_{l=1}^{S}\hat{P}_l(f) \qquad (1)$$

where the $\hat{P}_w(f)$ denotes the Welch PSD estimation, $\hat{P}_l(f)$ denotes the periodogram estimate of lth segment and S denotes the number of segments. The time domain EEG data is converted into the frequency domain covering the four EEG bands which include delta bands (0.5-3Hz), theta (3.5-7.5Hz), alpha (8-13Hz) and beta (13.5-30Hz). The total power for each band was used for the features which were calculated using the numerical integration of trapezoidal method.

The PSE can be calculated based on the PSD value which is given by:

$$H(f) = -\frac{1}{\ln(N)}\sum_{i=1}^{N}p_i\ln(p_i) \qquad (2)$$

where p_i denotes the spectral amplitudes of frequency bin i which can be calculated as the normalized PSD at each frequency bin, N denotes the number of the frequency bin. The PSE resulted in the same size of the matrix features as the PSD.

The wavelet is the time-frequency-based feature extraction method which has been use in the brain-computer interface (BCI) as in other EEG-based studies [9] [15]. This study used discrete wavelet transform (DWT) with Daubenchies as the mother wavelet. With EEG sampling frequency 256 Hz, a total of 5 levels of the wavelet

decompositions are needed for constructing of the delta, theta, alpha and beta bands.

The AR modeling is a parametric method of feature extraction which has been used in another application of EEG-based system [10] with the calculation as follows:

$$x(m) = \sum_{k=1}^{P} a(k)x(m-k) + e(m) \qquad (3)$$

where $x(m)$ represents the signal at time m, P is the order of the AR, $a(k)$ represents the AR coefficients and $e(m)$ represents the white noise with zero means error and finite variance. The order of AR was set at 4 in this study. Each of the feature extractors (PSD, PSE, wavelet and AR) resulted in 4 units of features in each EEG channels or 8 units for two EEG channels.

C. Classification

Bayesian neural network was used in this application. This has been identified as a useful approach to handle non-linear classification problems [12]. The structure of the Bayesian neural network is shown in Fig. 3, which is a three layered feed-forward structure and modeled by:

$$z_k(x, w) = f\left(b_k + \sum_{i=1}^{l} w_{kj} f\left(b_j + \sum_{i=1}^{m} w_{ji} x_i\right)\right) \qquad (4)$$

where $f(.)$ is the transfer functions which uses the hyperbolic tangent function for this paper, m is the input nodes number ($i=1, 2, …, m$), l is the hidden nodes number ($j=1, 2, …, l$), p is the number of output ($k=1, 2, …, p$), w_{ji} is the weight to the hidden unit y_j from input unit x_i, w_{kj} is the weights to output (z_k) from hidden unit (y_j), b_j and b_k are the biases. To improve the generalization and to avoid the over-fitting of the neural network, the Bayesian framework was applied, which is able to adjust the hyper-parameters to the most probable value given by training the dataset during the process of the Bayesian neural network learning [12] [16].
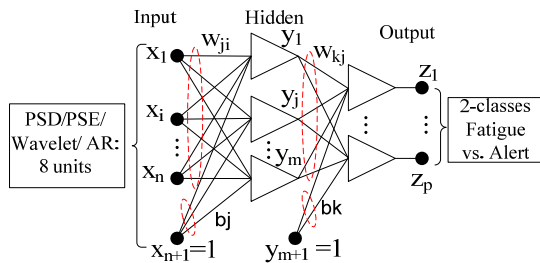

Figure 3.   The structure of neural network for fatigue vs alert classification

Based on the Bayesian inference, the posterior distribution of the weight vector w of neural network given a dataset D is given by:

$$p(w \mid D, \alpha) = \frac{p(D \mid w, \alpha)p(w \mid \alpha)}{p(D \mid \alpha)} \qquad (5)$$

where $p(w \mid \alpha)$ denotes the weight prior determined using the theory of prior, $p(w \mid D, \alpha)$ is the dataset likelihood, α is the normalization factor which is usually ignored since it is irrelevant to the first level inference, $p(D \mid \alpha)$ is the evidence for α. In Bayesian neural networks, the optimal

weights should maximize the weight posterior distribution which is equivalent of minimising the total error function. The optimal selection corresponds to the highest logarithm of the evidence.

III. RESULTS

The matrix size of the features is 8×4998 units for the alert task and another matrix with the size of 8×4998 units for the fatigue task. The overall matrix size of the combination dataset is 8×9996 units. The training sets are taken from 50% of overall sets and the remaining datasets were taken for testing sets. The Bayesian framework is applied to determine the best suitable neural network for each demanding computer task (AX-CPT, PVT and Stroop) by estimating the evidence of a set of neural network with different hidden nodes. For the performance measurement, the value of sensitivity, specificity and accuracy are used.

TABLE I.   RESULTS OF FATIGUE FROM POST-TREATMENT VS. ALERT FROM PRE-TREATMENT OF TWO EEG CHANNELS FROM 14 PARTICIPANTS.

Method	Cognitively Demanding Computer Tasks	Correctly Identified		Accuracy (%)
		Fatigue Sensitivity (%)	Alert Specificity (%)	
(i) Power Spectral Density (PSD)	AX-CPT	59.02	60.35	60.18
	PVT	60.42	59.88	60.15
	Stroop	61.47	59.75	60.61
(ii) Power Spectral Entropy (PSE)	AX-CPT	62.37	59.08	60.73
	PVT	60.36	59.95	60.15
	Stroop	60.49	61.69	61.09
(iii) Wavelet	AX-CPT	62.59	60.95	61.77
	PVT	61.40	61.55	61.45
	Stroop	63.91	60.10	62.01
(iv) Autoregressive (AR)	AX-CPT	76.20	75.69	**75.95**
	PVT	76.11	74.35	**75.23**
	Stroop	79.11	73.77	**76.02**

Table I shows the results of the classification between fatigue state (from post-treatment of the computer tasks) and alert state (from pre-treatment of the computer tasks) using two EEG channels with three cognitively demanding computer tasks (AX-CPT, PVT and Stroop). The results from four feature extraction methods are presented for comparison (PSD, PSE, Wavelet and AR) with Bayesian neural networks.

Using the PSD feature extractor, the sensitivity of each computer task is between 59% and 62%, the specificity of each computer task is between 59% and 61% and the accuracy of each computer task is around 60%. Using the PSE as the feature extractor, the sensitivity, specificity and accuracy were similar compared to the PSD feature extractor as follows: the sensitivity of each task is between 60% and 63%, the specificity of each task is between 59% and 62% and the accuracy of each task is between 60% and 61%. Using the wavelet as the feature extractor, the sensitivity, specificity and accuracy are slightly improved compared to the PSD and PSE feature extractor as follows: the sensitivity

of each task is between 61% and 64%, the specificity of each task is between 60% and 62% and the value of accuracy is between 61% and 62%. Using the AR, the sensitivity, specificity and accuracy are improved significantly compared with the PSD, PSE and wavelet methods to: the sensitivity of each task is between 76% and 80%, the specificity of each task was between 73% and 76% and the accuracy of each task is between 75% and 76%. As a result, the AR is the best feature extractor in this study which provided accuracy of 75.95% for AX-CPT task, 75.23% for PVT task and 76.02% for Stroop task.

The plots of the log evidence against the optimum number of hidden neurons of the Bayesian neural network training from each feature extractor are shown in Fig. 4. For PSD and AR feature extractors, the network with nine hidden nodes has the highest evidence. For PSE and wavelet feature extractors the network with ten hidden nodes has the highest evidence.
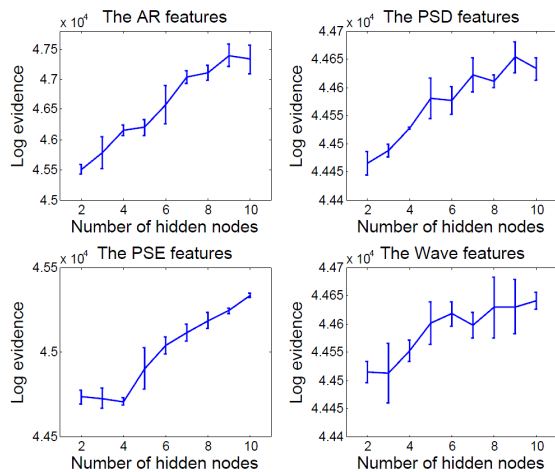

Figure 4. Log evidence vs. number of hidden nodes for each feature extractor (PSD, PSE, Wavelet and AR) from the Bayesian neural nework training

A statistical significance test was carried out in each computer task between the overall accuracy between AR feature extractor compared to PSD, PSE and wavelet feature extractors with p-value less than 0.05. The results showed that the performance of the classification using the AR feature extractor is significantly better than other feature extractors (PSD, PSE and wavelet) with a 95% confidence level.

IV. CONCLUSION

In this paper, four feature extraction methods (PSD, PSE, wavelet and AR) were investigated to find the best feature extractor for the classification of fatigue and alert from three cognitively demanding computer tasks (AX-CPT, PVT and Stroop), using two EEG channels of frontal location (F3 and F4). Bayesian neural networks were used for the classifier. The result shows that using PSD and PSE feature extractors, the average accuracy was 60%. This accuracy was improved by using the wavelet method of feature extraction with an

accuracy of approximately 61% for each computer task. A further improvement was found by using AR as it is shown to be the best feature extractor in each task compared to the PSD, PSE and wavelet methods with accuracy of 75.95% in the AX-CPT task, 75.23% in the PVT task and 76.02% in the Stroop task. The p-value between the AR method compared to PSD, PSE and wavelet is less than 0.05 or 95% confidence level. Such an improvement in accuracy using the AR method may lead to the automatic cognitive fatigue detection using the fewer EEG channels. Therefore it is recommended that this approach be used to detect fatigue during cognitively demanding computer tasks.

REFERENCES

[1] S. M. Marcora, W. Staiano, and V. Manning, "Mental fatigue impairs physical performance in humans," *Journal of Applied Physiology,* vol. 106, pp. 857-864, 2009.

[2] M. R. Smith, S. M. Marcora, and A. J. Coutts, "Mental Fatigue Impairs Intermittent Running Performance," *Medicine and science in sports and exercise,* 2014.

[3] M. R. Smith, J. Fransen, and A. J. Coutts, "The effects of three cognitively demanding tasks on psychological and performance indicators of cognitive fatigue," in *The 19th annual congress of the European College of Sports Science*, Amsterdam, 2014, p. 676.

[4] S. K. L. Lal and A. Craig, "A critical review of the psychophysiology of driver fatigue," *Biological Psychology,* vol. 55, pp. 173-194, 2001.

[5] S. K. L. Lal, A. Craig, P. Boord, L. Kirkup, and H. Nguyen, "Development of an algorithm for an EEG-based driver fatigue countermeasure," *Journal of Safety Research,* vol. 34, pp. 321-328, 2003.

[6] I. G. Campbell, "EEG recording and analysis for sleep research," *Current Protocols in Neuroscience,* pp. 10.12. 11-10.12. 19, 2009.

[7] A. Craig, Y. Tran, N. Wijesuriya, and H. Nguyen, "Regional brain wave activity changes associated with fatigue," *Psychophysiology,* vol. 49, pp. 574-582, 2012.

[8] A. Zhang, B. Yang, and L. Huang, "Feature extraction of EEG signals using power spectral entropy," in *Int. Conf. Biomed. Eng. and Inform. (BMEI)*, 2008, vol. 2, pp. 435-439.

[9] O. Rosso, M. Martin, A. Figliola, K. Keller, and A. Plastino, "EEG analysis using wavelet-based information tools," *Journal of Neuroscience Methods,* vol. 153, pp. 163-182, 2006.

[10] D. P. Burke, S. P. Kelly, P. de Chazal, R. B. Reilly, and C. Finucane, "A parametric feature extraction and classification strategy for brain-computer interfacing," *IEEE Trans. on Neural Sys. Rehabil. Eng.,* vol. 13, pp. 12-17, 2005.

[11] P. Georgieva, F. Silva, M. Milanova, and N. Kasabov, "EEG Signal Processing for Brain-Computer Interfaces," in *Springer Handbook of Bio-Neuroinformatics*: Springer, 2014, pp. 797-812.

[12] H. T. Nguyen, "Intelligent technologies for real-time biomedical engineering applications," *International Journal of Automation and Control,* vol. 2, Nos.2/3, pp. 274-285, 2008.

[13] E. Wascher, B. Rasch, J. Sänger, S. Hoffmann, D. Schneider, G. Rinkenauer, H. Heuer, and I. Gutberlet, "Frontal theta activity reflects distinct aspects of mental fatigue," *Biological Psychology,* vol. 96, pp. 57-65, 2014.

[14] 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, pp. 9-21, 2004.

[15] J. Wolpaw and E. W. Wolpaw, *Brain-computer interfaces: principles and practice*: Oxford University Press, 2011.

[16] H. T. Trieu, H. T. Nguyen, and K. Willey, "Advanced obstacle avoidance for a laser based wheelchair using optimised Bayesian neural networks," in *The 30th Ann. Int. Conf. of the IEEE Engineering in Med. Biol. Soc*, 2008, pp. 3463-3466.