Driver Fatigue Classification with Independent Component by Entropy Rate Bound Minimization Analysis in an EEG-based System

Rifai Chai, *Member, IEEE,* Ganesh R. Naik, *Senior Member, IEEE,*
Tuan N. Nguyen, *Senior Member, IEEE,* Sai Ho Ling, *Senior Member, IEEE,* Yvonne Tran,
Ashley Craig and Hung T. Nguyen, *Senior Member, IEEE*

Abstract— This paper presents a two-class electroencephalography (EEG)-based classification for classifying of driver fatigue (fatigue state vs. alert state) from 43 healthy participants. The system uses independent component by entropy rate bound minimization analysis (ERBM-ICA) for the source separation, autoregressive (AR) modeling for the features extraction and Bayesian neural network for the classification algorithm. The classification results demonstrate a sensitivity of 89.7%, a specificity of 86.8% and an accuracy of 88.2%. The combination of ERBM-ICA (source separator), AR (feature extractor) and Bayesian neural network (classifier) provides the best outcome with a *p*-value < 0.05 with the highest value of area under the receiver operating curve (AUC-ROC=0.93) against other methods such as power spectral density (PSD) as feature extractor (AUC-ROC=0.81). The results of this study suggest the method could be utilized effectively for a countermeasure device for driver fatigue identification and other adverse event applications.

Index Terms—electroencephalography (EEG), driver fatigue, autoregressive (AR) model, independent component analysis, entropy rate bound minimization, Bayesian neural network.

I. INTRODUCTION

DRIVER related fatigue is a leading factor in road accidents that can lead to serious injuries and fatalities in transportation [1]. Driver fatigue has been described as a feeling of tiredness and reduced alertness when driving which is associated with drowsiness, and impairs capability and willingness to perform the driving task [2]. The symptoms of driver fatigue include increased feelings of tiredness (yawning, sore or heavy eyes), slower reaction time and lack of concentration during driving and reduced control of speed of the vehicle [3, 4]. Fatigue is believed to contribute to 14-20% of motor vehicle accidents [5, 6], and it not only poses a risk to drivers themselves in terms of injuries and fatalities, but it could also results in injury to passengers, other vehicle drivers, cyclists and pedestrians. As a result an automated driver fatigue counter measurement/monitor system with robust and reliable fatigue classification accuracy is required as a strategy to reduce fatigue related risks on the road [6, 7, 8].

Currently, measurements of fatigue include the following: (i) psychological measurements employing psychometric questionnaires that assess an individual's self-reported fatigue [9, 10], (ii) video measurement as an indicator of performance such as facial expression, reaction time, steering errors and lane deviation [11], and (iii) physiological measurements such as electroencephalography (EEG) [4, 8, 12] for brain signal measurement, electrooculography (EOG) [13, 14] and other eye tracking systems [15] for eye movement detection, and finally electrocardiography (ECG) to detect the heart rate or heart rate variability changes associated with fatigue [16, 17].

Using psychological self-report for a fatigue counter measurement device during driving would arguably be problematic as it relies on an individual's potentially unreliable/ biased subjective feedback, and the strategy itself may distract the driver, and it requires time to validate the questionnaire of the person as an overall indicator of the fatigue symptom [1, 18]. Moreover, video recording the driver's face during driving is a non-direct measurement of detecting fatigue that may lead to privacy issues. Physiological measurements of fatigue using EOG, ECG and EEG have been explored widely [19]. For example, an increase in eye blink rates using EOG during driving may indicate impending fatigue [13]. Changes in heart rate variability (HRV) have been shown to be related to fatigue [16] while detection of the changes in brain activity using EEG has also been shown to be related to the fatigue state [4, 8]. EEG is considered to be a significant and reliable method of detecting fatigue, as it directly measures neurophysiological activity in the human brain [4]. Accordingly, this paper explores strategies for improving the fatigue vs. alert classification in an EEG-based

system.

EEG provides a high temporal resolution of brain activity in which multiple neural generators may be simultaneously active [20]. As a result, multivariate techniques can unearth more complex connections between the dependent and independent variables in EEG data. Independent components analysis (ICA) is one of these multivariate techniques and is typically used as a source separation and estimation tool [21, 22]. ICA is one of the so called blind source separation (BSS) techniques that utilize both lower and higher order statistics to estimate sets of linearly mixed variables into their independent components (ICs). In the recent past, ICA has been extensively used for EEG signal processing, especially for BCI applications [23]. One of the advantages of using ICA for EEG is that it decomposes the linearly mixed neural activities into its constituent independent components (ICs). Moreover, an advantage of using ICA-based methods for EEG analysis is that no explicit prior knowledge about brain activity is needed to estimate the source components [24, 25].

Most of the existing ICA algorithms exploit both higher order and second order statistics to minimize the non-Gaussian aspect of the sources. Recently, ICA by entropy rate bound minimization (ICA-ERBM) has emerged as an effective source separation technique. The algorithm takes both non-Gaussian property and sample correlation into account by minimizing mutual information rates. It is originally introduced as a full BSS, which results in improved general temporal structure of sources, e.g., second-order white noise, but with higher-order correlated sources [26]. This paper uses the ICA-ERBM for the separation of EEG fatigue data.

The functional basic components of EEG-based fatigue classification is similar to other EEG-based classifications, and consists of several elements including: (i) brain signal measurement and data acquisition using EEG technology, (ii) computational intelligence or algorithms such as pre-processing, features extraction and classification [27, 28]. For the features extraction in EEG analysis, power spectral density (PSD) has been used widely, especially in the study of fatigue [4]. The power spectrum estimation converts the time domain into the frequency domain of EEG data. This study explores the autoregressive (AR) model as it is an effective tool in EEG feature extraction algorithms, and will be used as a comparison to the PSD method [29, 30, 31]. For the classification algorithm, Bayesian neural networks capable of providing optimal structure [32] will be used to classify the two-state outputs classification (fatigue state vs. alert state).

The main contributions of this paper are the novel combination source separation (ICA-ERBM) and EEG feature extraction components which have not been explored previously for fatigue classification with the goal of improving classification accuracy. These components/algorithms include the use of entropy rate bound minimization as a source separation technique, the AR-modelling as the feature extraction algorithm and the Bayesian neural network for the classification algorithm.

The structure of this paper is as follows: section II covers the methodology: general structure, data collection, feature extraction methods and classification algorithms. Section III describes results, followed by section IV for discussion and section V for the conclusions.

II. BACKGROUND AND METHODOLOGY

A. General Structure

The components for the EEG-based fatigue classification system presented in this paper are shown in Fig. 1. First EEG data were collected in a simulated driver fatigue study, followed by a first signal pre-processing module for removing EEG artifact, a second signal pre-processing module as an additional source of separation processing, and third, a pre-processing module of window segmentation. The next module includes a features extraction module that transforms the signals into useful features. The features are processed into a classification algorithm that includes optimization, training and classification tasks. The classification comprises two-state outputs: fatigue state or alert (non-fatigue) state. The desired output for the fatigue state of the neural network uses value of one and the desired output for non-fatigue or alert state of the neural network uses value of zero.
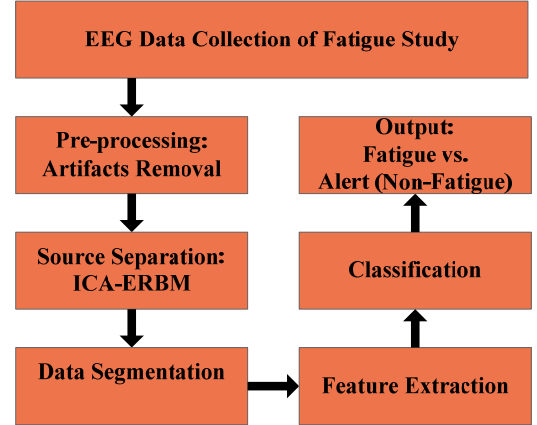

Fig. 1. Components of EEG-based fatigue classification system

B. EEG Experiment and Pre-processing

This EEG experiment used the dataset of forty-three healthy participants aged 18 to 55 years, obtained from a previous experimental study [4, 12]. Participants were given information of the study and informed consent was obtained before they commenced the experiment. This study was approved by The Human Research Ethics Committee of University of Technology Sydney (UTS). The study involved a repeated measures experimental intervention whereby baseline EEG and subjective levels of fatigue using psychometric assessment were taken, then an intervention task of monotonous simulated driving task followed by post EEG measures and post-subjective levels of fatigue [12]. The Divided Attention Steering Simulator (DASS) from Stowood Scientific Instruments was used as a driver simulator task. Experiments were conducted in a noise-, stimulus-, and temperature-controlled laboratory. Participants were asked to ensure that they kept driving at the centre of the road in the simulation task. The DASS also required participants to perform a reaction time response to a target number that

appeared in any of the four corners of the computer screen, and these were shown at random times during driving.

The experiment was terminated decrements in performance were detected, such as if they were driving off the road in the simulation driving task for greater than 15 seconds or if participants showed consistent facial signs indicating fatigue (such as head nodding and extended eyes closure) using video monitoring. The use of '15 seconds off the road' condition was to allow a reasonable time for the video monitoring assessment to check the fatigue occurrence. Although the experiment was terminated if the deviations off the road occurred for greater than 15 seconds, this may not be where the fatigue onset had occurred. In fact, fatigue was most likely to have occurred before this point. Video monitoring was used during the real time recording to check for consistent physiological signs of fatigue such as tired eyes, head nodding and extended eye closure. The video monitoring is a subjective assessment. Therefore a post recording validation is needed to verify fatigue occurrence [4, 33].

Fatigue occurrence was validated using three methods: (i) monitoring for consistent physiological signs of fatigue such as tired eyes, nodding and extended eye closure verified further with EOG analysis of blink rate and eye closure, (ii) performance decrements such as deviation off the road and (iii) subjective psychometric measures using a validated fatigue questionnaire called the Chalder Fatigue Scale and the Stanford Sleepiness scale, which measures a person's perception of how drowsy they feel [12]. The validation of fatigue in these participants has been previously reported in studies [4, 12]. The maximum time for the simulated driving was specified at 2 hours. EEG signals were recorded by attaching a 32-channel EEG system, the Active-Two system from Biosemi with the electrode positions based on the International 10-20 system. These positions are: FP1, AF3, F7, F3, FC1, FC5, T7, C3, CP1, CP5, P7, P3, PZ, PO3, O1, OZ, O2, PO4, P4, P8, CP6, CP2, C4, T8, FC6, FC2, F4, F8, AF4, FP2, FZ and CZ. The recorded EEG data was down sampled from 2048Hz to 256Hz.

For the alert group of data, the first 5 mins of EEG data when starting the driving simulation task was chosen while the fatigue group of data was selected from the last 5 mins of EEG data before the task was terminated, where consistent signs of fatigue were identified and additionally verified as fatigue using the EOG signals. Then in each group of data (alert and fatigue), 20s segment s were taken and the first 20s segments were selected for further analysis with the least movement artifact. As a result, 20s of alert state and 20s of fatigue state data were available from each participant.

The data was then processed in the signal pre-processing module. Here, the second-order blind identification (SOBI) and canonical correlation analysis (CCA) were used for removing artifact and disturbance related to sources such as eye activity, muscle activity and heart signals. The data were then fed to the source separation using ICA-ERBM. The novel combination of the source separation (ICA-ERBM) and EEG feature extraction method used for the classification is explained in detail next.

C. Source separation using ICA-ERBM

ICA by ERBM exploits both sample correlation and non-Gaussianity and takes this into account by minimizing mutual information rate. It is originally introduced as a full BSS (FBSS) algorithm [26]. Let, N statistically independent, zero mean source $s(t) = [s_1(t),\ldots,s_N(t)]^T$ be mixed with an $N \times N$ mixing matrix A such that we obtain the mixture $x(t) = [x_1(t),\ldots,x_N(t)]^T$ as $x(t) = As(t)$, where T and t denote the transpose and time index respectively. ICA separates the mixture using $y(t) = Wx(t)$ where $y(t) = [y_1(t),\ldots,y_N(t)]^T$ and W is the unmixing or separation matrix. ICA assumes the sources are independent identically distributed (IID), hence, a cost for achieving the separation of these N independent sources is the mutual information $I(y_1;\ldots;y_N)$ among N random variables y_n, $n = 1,\ldots,N$, which is:

$$I(y_1;\ldots;y_N) = \sum_{n=1}^{N} H(y_n) - \log\left|\det(W)\right| - H(x) \quad (1)$$

where $H(y_n)$ represents the entropy of the nth individual separated source, and entropy of observation $H(x)$ is constant
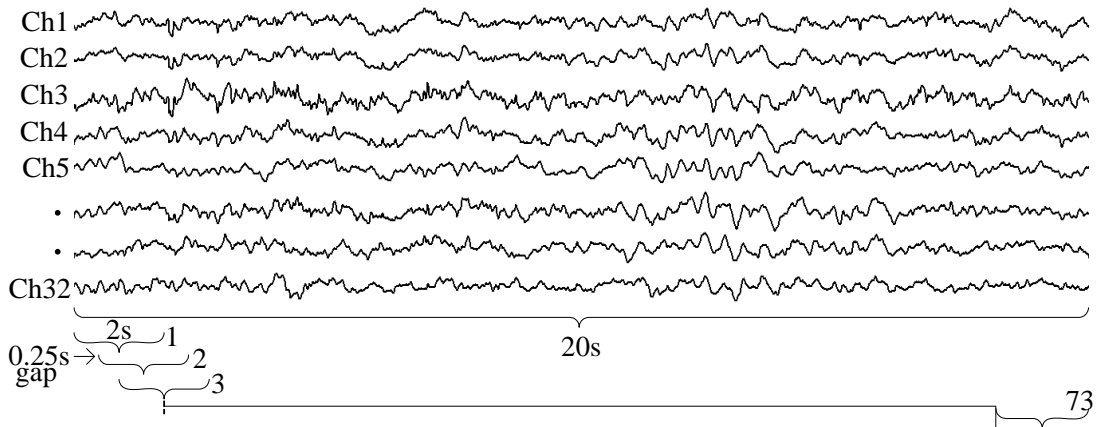

Fig. 2. Data Segmentation of Fatigue study

with respect to the unmixing matrix W. $H(x)$ is the entropy of any individual member of the separated process. However, a new cost function is needed since the equation (1) cannot obtain most of the temporal information of sources. The new cost function is therefore given as:

$$I_r(y_1;,\ldots,;y_N) = \sum_{n=1}^{N} H_r(y_n) - \log|\det(W)| - H_r(x) \quad (2)$$

where $H_r(y_n) = \lim_{t\to\infty} H[y_n(1),\ldots,y_n(t)]/t$ is the entropy rate of the nth process of y_n and entropy rate of observation $H_r(x) = \lim_{t\to\infty} H[x(1),\ldots,x(t)]/t$ of the observed vector-valued process x is constant with respect to the unmixing matrix W. $H_r(x)$ is the entropy rate of the separated process of the individual. Equation (1) is modified using the method proposed by Li and Adali [34] to obtain new entropy estimator and cost function. The new cost function is explained as:

$$J(W, p_1,\ldots,p_N) = \sum_{n=1}^{N} H(v_n) - \log|\det(W)| \quad (3)$$

where $v_n(t) = \sum_{q=0}^{p-1} a_n(q)y_n(t-q), y_n = w_n^T x$,

the n-th is separated source, and $a_n = [a_n(0),\ldots,a_n(p-1)]^T$ are the filter coefficients. Later, the algorithm is optimized to obtain a new W, which minimizes the mutual information rate. The ICs, $\hat{s}(t)$ are recovered using the equation $\hat{s}(t) = Wx(t)$, where W and $x(t)$ are the unmixing matrix and recordings (mixtures) respectively. More detail of the algorithm is provided elsewhere [26].

D. Data Segmentation and Feature Extraction

Before performing feature extraction, ICA-ERBM separated data are segmented as illustrated in Fig. 2. A moving window of 2s with overlapping 1.75s was applied to the 20s segments which provided 73 overlapping segments on each state. With the 43 participants, a total of 3139 units of datasets were formed for the alert state and another 3139 units for the fatigue state.

An autoregressive (AR) model was applied as a features extraction algorithm in combination with ICA-ERBM separated sources in this study. AR modelling has been used in EEG studies as an alternative to Fourier-based method [29, 30, 31]. The advantage of AR modelling is its inherent capacity to model the peak spectra that are characteristic of the EEG signals and it is an all-pole model making it efficient for resolving sharp changes in the spectra. The fast Fourier transform (FFT) is a widely used nonparametric approach that can provide accurate and efficient result, but which does not have spectral resolution for short data segment [35]. Further, other EEG classifications used previously have shown that the AR modelling achieved a better result [29, 36, 37].

In AR modelling, it is assumed to be a random process that is independent of the previous value of the signal. The Burg method is the most popular of the AR methods that have been used that recursively estimates the reflection coefficients of an AR lattice filter by minimizing the mean of forward and backward least squares linear prediction error. This method is used in this paper to estimate the coefficients of the AR.

AR modelling requires the selection of the model order number. The best AR model order number represents a consideration of both the signal complexity and the sampling rate. If the AR model order is too low, the whole signal cannot be captured in the model. On the other hand, if the model order is too high, then more noise is captured [38]. In this study, different AR order numbers were tested and the order providing the best classification accuracy was the chosen AR order number. The calculation of the AR modelling for ICA-ERBM separated sources is as follows:

$$\hat{s}(t) = \sum_{k=1}^{P} a(k)\hat{s}(t-k) + e(t) \quad (4)$$

where $\hat{s}(t)$ represents the ICA-ERBM separated EEG data (sources) at time (t), P is the order of the AR, $e(t)$ represents the white noise with, zero means error and finite variance, and $a(k)$ represents the AR coefficients which need to be estimated from finite samples of data $\hat{s}(1), \hat{s}(2),\ldots\hat{s}(N)$.

For comparison purposes, power spectral density (PSD), a popular feature extractor in fatigue studies, is also used in this paper [4, 8]. The PSD of the Welch spectrum is given by:

$$\hat{P}_w(f) = \frac{1}{S}\sum_{l=1}^{S} \hat{P}_l(f) \quad (5)$$

where the $\hat{P}_w(f)$ denotes the Welch PSD estimation, $\hat{P}_l(f)$ denotes the periodogram estimate of l-th segment and S denotes the number of segments.

E. Classification Algorithm

One of the crucial issues in developing a neural network is generalization, defined by how well the network can make predictions for new cases that are not in the training data. A network that is not complex enough may ignore the data, leading to "under-fitting", while a network that is too complex may fit the noise, not just training data, leading to "over-fitting". The complexity of the network is concerned with the network architecture and magnitudes of network weights and biases.

Several frameworks have been proposed to prevent MLP networks from under-fitting or over-fitting such as growing, pruning, global searches, and early stopping. However, these frameworks require intensive searching for network parameters or do not make maximum use of the available data [32, 39].

The Bayesian neural network structure uses a three layered feed-forward structure and modeled by:

$$z_k(x, w) = f\left(b_k + \sum_{j=1}^{l} w_{kj} f\left(b_j + \sum_{i=1}^{m} w_{ji}x_i\right)\right) \quad (6)$$

where $f(.)$ denotes the transfer functions and hyperbolic tangent function is used in this paper, m denotes the input

nodes number (i=1, 2, …, m), l denotes the hidden nodes number (j=1, 2, …, l), p denotes the number of output (k=1, 2, …, p), w_{ji} denotes the weight to the hidden unit y_j from input unit x_i, w_{kj} denotes the weights to output (z_k) from hidden unit (y_j), b_j and b_k denote the biases.

Bayesian regularization framework has been proposed to enhance the generalization abilities of neural networks regardless of finite and/or noisy available data [40]. The probability distribution of network parameters is considered in Bayesian learning, providing the best generalization of the trained network. Especially, this type of neural network can be trained on all of the available data. Therefore, it is suitable for applications where the data set is small.

In the Bayesian framework, the most probable model corresponding to the training data D is found in automatic fashion. Based on Gaussian probability distribution over weight values, the posterior distribution of the weights w in network H can be calculated using Bayes' theorem as shown below:

$$p(w \mid D, H) = \frac{p(D \mid w, H)\, p(w \mid H)}{p(D \mid H)} \qquad (7)$$

where $p(D \mid w, H)$ is the likelihood which contain information about weights from observations, the prior distribution $p(w \mid H)$ contains information about background weight set, and the $p(D \mid H)$ is known as the evidence of the network H.

For a multilayer perceptron neural network (MLP) depicted in Fig.3, the most probable value for the neural network weights, w^{MP}, can be found by minimizing the cost function, $F(w)$ defined as follows:

$$F(w) = \beta E_D(w) + \alpha E_W(w) \qquad (8)$$

where α and β are hyper-parameters with the ratio α / β that controls the effective complexity of the network structure, $E_D(w)$ is the error function, and $E_W(w)$ is the sum square of weight function, defined in the following equation:

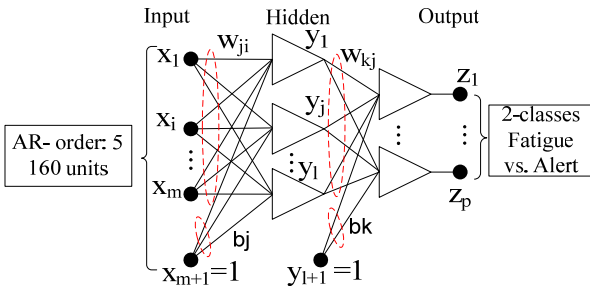$$E_W(w) = \frac{1}{2}\|w\|^2. \qquad (9)$$

Fig. 3. The ANN structure for EEG-based fatigue vs. alert classification with input of combination of ICA-ERBM and AR processes; output of 2-classes (fatigue and alert).

By introducing hyper-parameters in the cost function, neural network weights can be prevented from being too large, which would result in poor generalization for new test cases. As a result, a validation set is not required in a neural network training procedure.

The Bayesian regularization algorithm is applied to update hyperparameters as follows.

$$\alpha^{MP} = \frac{\gamma}{2 E_W(w^{MP})}; \beta^{MP} = \frac{N - \gamma}{2 E_D(w^{MP})} \qquad (10)$$

where $\gamma = n - 2\alpha^{MP} tr(H^{MP})^{-1}$ is called the effective number of parameters, n is the total number of parameters in the network, N is the total number of errors, and H is Hessian matrix of $F(w)$ at the minimum point of w^{MP}. The log evidence of model H_i, is evaluated by Bayesian framework as follows:

$$\ln p(D \mid H_i) = -\alpha_{MP} E_W^{MP} - \beta_{MP} E_W^{MP} - \frac{1}{2}\ln|A|$$
$$+ \frac{W}{2}\ln \alpha_{MP} + \frac{N}{2}\ln \beta_{MP} + \ln M \,!+ 2\ln M \qquad (11)$$
$$+ \frac{1}{2}\ln \frac{2}{\gamma} + \frac{1}{2}\ln \frac{2}{N - \gamma}$$

where W is the number of network parameters, M is the number of hidden nodes, and A is the Hessian matrix of the cost function. A network structure with highest log evidence value will be selected as the best optimal structure of that network.

For the performance measurement, the well-known performance indicators, including specificity or true negative rate (*TNR*), sensitivity or true positive rate (*TPR*) and accuracy, were used as follows:

$$Sensitivity(TPR) = \frac{TP}{TP + FN} \qquad (12)$$

$$Specificity(TNR) = \frac{TN}{TN + FP} \qquad (13)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad (14)$$

where *TP* (true positive) denotes the number of the data inputs that refer to fatigue state correctly classified as fatigue. *FP* (false positive) is the number of data inputs that refer to alert state classified as a fatigue state. *TN* (true negative) is number of the data inputs that refer to an alert state correctly classified as an alert state. *FN* (false negative) is the data inputs that refer to fatigue state classified as an alert state.

III. RESULTS

Initially, 32-channel EEG data was fed to the ICA - ERBM, which resulted in 32-channel ICA separated sources. These optimized sources were further segmented for feature extraction process. The dataset from the segmentation process comprised 3139 units of alert state datasets and another 3139 units for the fatigue state obtained from the 43 participants.

This provided a total of 6278 units for the combination of the fatigue and alert datasets. For Bayesian neural network classification, the dataset was divided into training sets with 50% of the overall sets and the remaining portion used for test sets. The Levenberg-Marquardt with Bayesian regularization algorithm was applied to train this neural network classifier, and the mean squared error function was chosen as the error function $E_D(w)$ [40].

The PSD, a popular feature extractor was used for comparison. In this paper, the Welch spectrum of the PSD was used for converting the time domain EEG data into the frequency domain. For this study the EEG bands used covering the four EEG bands included delta bands (0.5-3Hz), theta bands (3.5-7.5Hz), alpha bands (8-13Hz) and beta bands (13.5-30Hz). The total power for each EEG activity band was used for the features that were calculated using the numerical integration trapezoidal method, providing 4 units of power values. This resulted in 128 units of power for the 32 EEG channels used.
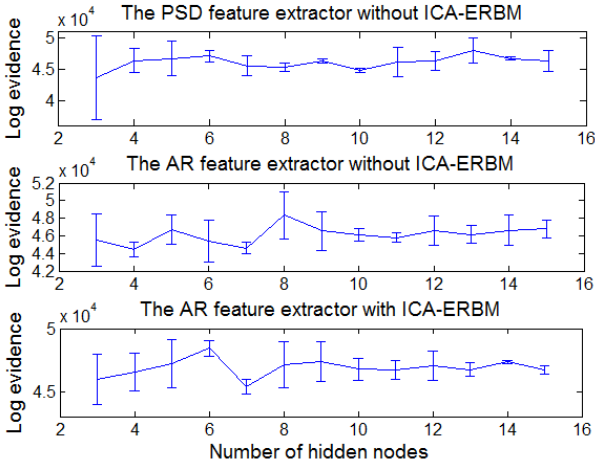

Fig. 4. Evidence framework of Bayesian inference for PSD only feature extractor, AR only feature extractor and the combination AR feature extractor with ICA-ERBM source separator.
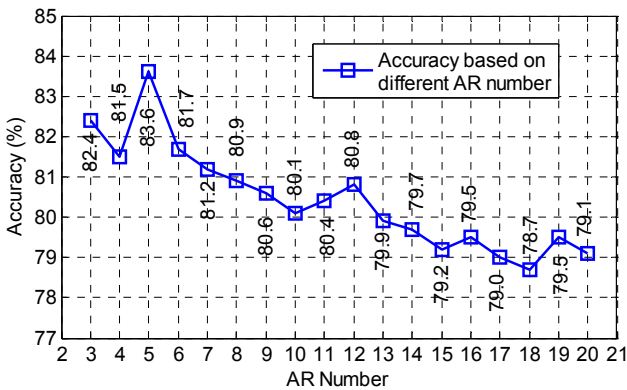

Fig. 5. Finding the best AR order number Plotting

For the AR feature extractor, different order numbers were tested from 3 to 20 units to obtain the best number with the highest classification accuracy. The size of the AR features equaled the AR order number multiplied with 32 units of EEG channels, for example, an AR order number of 5 resulted in 160 units of the AR features.

The plot of the log evidence against the optimum number of hidden neurons of the Bayesian neural network training is shown in Fig.4, which includes: (i) the PSD feature extractor without the ICA-ERBM with 13 hidden nodes resulted in the best classification evidence; (ii) using the AR feature extractor without the ICA-ERBM with 8 hidden nodes produced the best classification evidence and (iii) using the combination of ICA-ERBM source separator and AR feature extractor with 6 hidden nodes produced the best classification evidence.

Fig. 5 shows the result of the accuracy with different AR order number and plotting of the AR number versus classification accuracy. As for the purpose of finding the best AR order number, the ICA-ERBM module as the source separation is excluded. Each AR number provides a different result. The accuracy values ranged between 79.0% and 83.6%. The accuracy curve shows the order number starts at 3 with an accuracy of 82.4%. It is improved to 83.6% with an order number of 5, resulting in the highest accuracy. The remaining order numbers provided lower accuracy between 79.0% and 81.7%. As a result, the best optimal AR order number was five. For comparison purposes, the accuracy without using the ICA-ERBM pre-processing method is also provided in the manuscript.

The classification results of the fatigue states using the different feature methods and Bayesian neural networks as a classifier are shown in Table 1. First, for the features using the PSD without the inclusion of the source separator: (i) for the fatigue group, 1212 units of fatigue data were correctly classified as fatigue states (TP) in a total of 1569 units of actual fatigue dataset, resulting in a sensitivity of 77.2%. (ii) For the alert state, 1187 units of alert data were correctly classified as alert states (TN) in a total of the 1570 units of actual alert dataset, resulting in a specificity of 75.6%. This resulted in an accuracy of 76.4% of the feature using PSD without the ICA-ERBM.

TABLE I
CLASSIFICATION RESULT OF BETWEEN FATIGUE STATE AND ALERT STATE ON THE TEST SETS

Classification Results	Feature Methods		
	Source Separator: None Feature Extractor: PSD	Source Separator: None Feature Extractor: AR	Source Separator: ICA-ERBM Feature Extractor: AR
TP	1212	1322	1407
FN	357	247	162
FP	383	267	207
TN	1187	1303	1363
Sensitivity/ TPR (%)	77.2%	84.3%	**89.7%**
Specificity/ TN Rate (%)	75.6%	83.0%	**86.8%**
Accuracy (%)	76.4%	83.6%	**88.2%**

Next, compared to the PSD features extractor, improvement was achieved using the AR method (using the best AR order number =5) without the inclusion of the source separator. For the fatigue group, 1322 units of fatigue data were correctly classified as fatigue states (TP) in a total of 1569 units of actual fatigue dataset, resulting in a sensitivity of 84.3%. For the alert group, 1303 units of alert data were correctly classified as alert states (TN) in the total of 1570 units of actual alert dataset, resulting in a specificity of 83.0%. This resulted in an accuracy of 83.6% (see Table 1) of the feature using AR without the ICA-ERBM.

Further improvements were obtained by using the combination of ICA-ERBM as the source separator and AR method as the feature extractor. For the fatigue group, 1407 units of fatigue data were correctly classified as fatigue states (TP) in the total of 1569 units of actual fatigue dataset, resulting in a sensitivity of 89.7%. For the alert group, 1363 units of alert data were correctly classified as alert states (TN) in the total of 1570 units of actual alert dataset, resulting in specificity of 86.8%. This resulted in an accuracy of 88.2% (see Table 1) of the feature using AR with the source separator of ICA-ERBM.

IV. DISCUSSION

From the above results, compared to the case where PSD was used as a feature extractor but without source separation, the combination of source separation using ICA-ERBM and AR as feature extractor improved the classification performance. Results indicated sensitivity increased by 12.5% (from 77.2% to 89.7%), specificity by 11.2% (from 75.6% to 86.8%) and accuracy by 11.8% (from 76.4 to 88.2%). Furthermore, compared to the case where AR model was used as a feature extractor but without the source separation, the combination of the source separation of ICA-ERBM and AR as feature extractor, resulted in improved performance indicators for sensitivity which increased by 3.8% (from 83.0% to 86.8%), specificity by 5.4% (from 84.3% to 89.7%) and accuracy by 4.6% (from 83.6% to 88.2%).

TABLE II
STATISTICAL SIGNIFICANT (P-VALUE) OF SPECIFICITY, SENSITIVITY AND ACCURACY FROM TEST SET OF 43 PARTICIPANTS

Comparison methods : (i) Source Separation: None; Feature Extractor: PSD (ii) Source Separation: None; Feature Extraction: AR (iii) Source Separation: ICA-ERBM; Feature Extraction: AR	***p*-value of Sensitivity**	***p*-value of Specificity**	***p*-value of Accuracy**
(iii) vs. (i)	5.1×10^{-7}	4.9×10^{-7}	4.9×10^{-14}
(iii) vs. (ii)	0.0034	0.0022	0.000023

The results indicate that the use of the ICA-ERBM (source separator) and AR (feature extractor) provides the best sensitivity of 89.7%, with a specificity of 86.8% and an accuracy of 88.2%, compared to the other combination methods for the test set. Statistical significance tests are presented in Table II to allow evaluation of the comparisons of the sensitivity, specificity and accuracy from the test sets of 43 participants. The results show that the use of ICA-ERBM as source separator and AR as feature extractor (iii) compared to no source separator and PSD as feature extractor (i), and no source separator and AR as feature extractor (ii), resulted in all of the p-values of specificity, sensitivity and accuracy being less than 0.05. A p-value < 0.05 is considered statistically significant with 95% confidence level.

This indicates that EEG provides a high temporal resolution of brain activity in which multiple neural generators may be simultaneously active [20]. The combination of ICA-ERBM provides an optimal solution for extracting the complex connections between the dependent and independent variables in EEG data, converting the data into features that help to achieve the best performance in term of sensitivity, specificity and accuracy.

Moreover, to examine the performance of the proposed method with the compared methods, Fig. 6 shows the results in the receiver operating characteristic (ROC) curve analyses. The ROC curve is a plot of true positive rate (TPR) versus false positive rate (FPR) using different threshold ratios as a sweeping variable. A ROC plane has axes ranging from 0 to 1 in which FPR (1- specificity) is plotted on the X axis and TPR on the Y axis. A random performance (diagonal dotted line) of the classifier would have a straight line connecting (0, 0) to (1, 1). Any classifier that appears in the lower right triangle performs worse than random guessing. Curves that appear in the upper-left of the ROC have superior performance classifications [41, 42]. Fig.6 shows that the methods investigated provided ROC curves above random guess classifications. The method using the inclusion of the source separation of ICA-ERBM and AR feature extractor (iii) had the best upper-left ROC curve results compared to the other two methods.
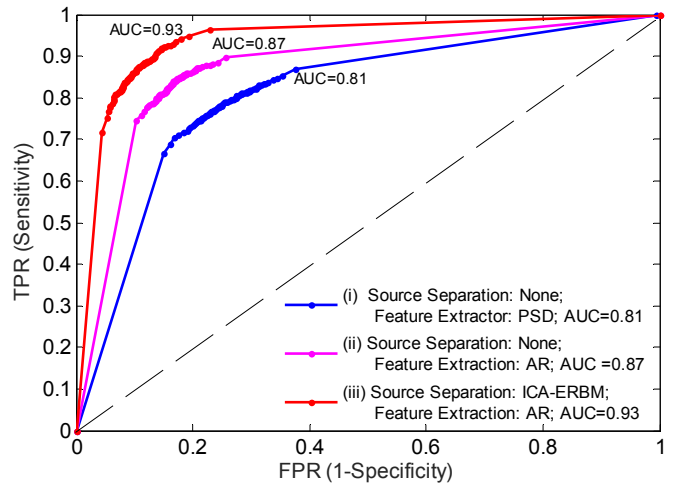

Fig. 6. ROC plot with AUC values for PSD only feature extractor (i), AR only feature extractor (ii) and the combination AR feature extractor with ICA-ERBM source separator (iii).

The areas under the curve (AUC) of ROC were also computed to evaluate quantitatively classification performance. The AUC can be calculated by using the trapezoidal areas created between each ROC point [41, 42].

The AUC value lies between 0 and 1 with a higher AUC value indicating a better classification performance. The method using the inclusion of the source separation of ICA-ERBM and AR feature extractor (iii) had the highest AUC of 0.93 compared to the method which involved no inclusion of the source separation and AR as feature extractor (ii), and which had an AUC of 0.87, while the method with no inclusion of the source separation and PSD as feature extractor (i) had an AUC of 0.81. As a result, the ROC-AUC analyses show the inclusion of the source separation of ICA-ERBM and AR feature extractor (iii) to be the best performance result compared to other methods based on highest AUC value.

The computational time of the classification algorithm was also estimated by MATLAB's built-in *tic/toc* functions whereas the *tic* function was used before the program and the *toc* afterwards (using an Intel Core i5-4570 processor 3.20GHz, 8GB RAM) to determine classification times. The result showed, in terms of the classification/execution time, the classifier was able to complete the task in 0.09 seconds or less than a second. The program developed in MATLAB takes more computational time compared to *C* language with significantly faster training time.

For the operation of real-time classification, there is therefore no necessity to perform the classifier training. The classifier only needs to compute the feed forward neural network function, based on (6), with the saved weight parameters and particular features for pattern classification system. Thus, the classification time in the runtime mode (execution) is fast, taking less than a second.

The classifier used a three-layer feed forward ANN. The network comprises of a single hidden layer, with m inputs, l hidden units and p outputs. Further, the network has $m + l + p$ nodes and $(m + p)l$ edges. As a result, the complexity of ANN is represented by notation $O(n)$ where $n = m + l + p + (m + p)l$. In this study, the number of inputs (m) was 160 units; the optimum number of hidden nodes (l) was 6 units and the number of outputs (p) was 2 units. The computational time of the classification algorithm was estimated to be 0.09 seconds using MATLAB. This is scalable depending on the related software and hardware.

V. CONCLUSIONS

In this paper, the classification of two-state outputs (fatigue vs. alert) during the simulated driving task has been applied with 43 participants. ICA-ERBM algorithm was used for the source separation method, AR model for the feature extraction method and a Bayesian neural network was used for the classification method.

The results showed that the combination of the ICA-ERBM (source separator) and AR (feature extractor) achieved the best sensitivity of 89.7%, specificity of 86.8% and accuracy of 88.2% for test set, and compared with the other combination methods, it achieved a p-value less than 0.05, which is considered statistically significant at a 95% confidence level. Using only the PSD feature extractor, without the ICA-ERBM source separator, we achieved a lower sensitivity of 77.2%,

specificity of 75.6% and accuracy of 76.4%. Using only the AR feature extractor without the ICA-ERBM source separator, we achieved a sensitivity of 84.3%, specificity of 83.0% and accuracy of 83.6%. The findings therefore suggest that a combination of the ICA-ERBM source separator and AR modelling feature extractor provides the best outcome in term of overall sensitivity, specificity and accuracy when classifying fatigue state vs. alert state. The ROC-AUC analysis results also suggest a combination of the ICA-ERBM and AR provides the best performance with highest AUC value of 0.93.

The results from this study provide optimism for the development of an operational fatigue countermeasure device that could be employed in a real driving context. Future research should focus on optimizing the above techniques for a wider pool of participants (e.g. wider age range) and also investigate the efficacy of the driver fatigue detection system in real time.

REFERENCES

[1] G. Matthews, P. A. Hancock, M. C. Neubauer, and P. A. Desmond, *The handbook of operator fatigue*: Ashgate Publishing, Ltd., 2012.

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

[3] J.-E. Lv, L.-H. Zhu, R.-S. Zheng, and Y.-F. Wei, "Effects of driver's reaction time on safe driving," *J. Transp. Syst. Eng. Inf. Technol.,* vol. 14, pp. 80-86, 2014.

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

[5] R. F. S. Job, A. Graham, C. Sakashita, and J. Hatfield, " Fatigue and Road Safety: Identifying Crash Involvement and Addressing the Problem within a Safe Systems Approach," in *The Handbook of Operator Fatigue,* vol. 1: Ashgate Publishing Limited, 2012, pp. 349-363.

[6] A. Fletcher, K. McCulloch, S. D. Baulk, and D. Dawson, "Countermeasures to driver fatigue: a review of public awareness campaigns and legal approaches," *Australian and New Zealand J. Public Health,* vol. 29, pp. 471-476, 2005.

[7] W. Vanlaar, H. Simpson, D. Mayhew, and R. Robertson, "Fatigued and drowsy driving: A survey of attitudes, opinions and behaviors," *J. Safety Res.,* vol. 39, pp. 303-309, 2008.

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

[9] H. J. Michielsen, J. De Vries, G. L. Van Heck, F. J. Van de Vijver, and K. Sijtsma, "Examination of the Dimensionality of Fatigue: The Construction of the Fatigue Assessment Scale (FAS)," *European J. Psychol. Assessment,* vol. 20, p. 39, 2004.

[10] J.-S. Lai, D. Cella, S. Choi, D. U. Junghaenel, C. Christodoulou, R. Gershon, and A. Stone, "How Item Banks and Their Application Can Influence Measurement Practice in Rehabilitation Medicine: A PROMIS Fatigue Item Bank Example," *Arch. Phys. Med. Rehabil.,* vol. 92, pp. S20-S27, 2011.

[11] L. Fletcher, N. Apostoloff, L. Petersson, and A. Zelinsky, "Vision in and out of vehicles," *IEEE Intell. Syst.,* vol. 18, pp. 12-17, 2003.

[12] A. Craig, Y. Tran, N. Wijesuriya, and P. Boord, "A controlled investigation into the psychological determinants of fatigue," *Biol. Psychol.,* vol. 72, pp. 78-87, 2006.

[13] S. Hu and G. Zheng, "Driver drowsiness detection with eyelid related parameters by Support Vector Machine," *Expert Syst. Appl.,* vol. 36, pp. 7651-7658, 2009.

[14] T. L. Morris and J. C. Miller, "Electrooculographic and performance indices of fatigue during simulated flight," *Biol. Psychol.,* vol. 42, pp. 343-360, 1996.

[15] J. Qiang, Z. Zhiwei, and P. Lan, "Real-time nonintrusive monitoring and prediction of driver fatigue," *IEEE Trans. Veh. Technol.,* vol. 53, pp. 1052-1068, 2004.

[16] Y. Tran, N. Wijesuriya, M. Tarvainen, P. Karjalainen, and A. Craig, "The relationship between spectral changes in heart rate variability and fatigue," *J. Psychophysiology,* vol. 23, pp. 143-151, 2009.

[17] M. Patel, S. Lal, D. Kavanagh, and P. Rossiter, "Applying neural network analysis on heart rate variability data to assess driver fatigue," *Expert Syst. Appl.,* vol. 38, pp. 7235-7242, 2011.

[18] C. Zhao, M. Zhao, J. Liu, and C. Zheng, "Electroencephalogram and electrocardiograph assessment of mental fatigue in a driving simulator," *Accid. Anal. Prev.,* vol. 45, pp. 83-90, 2012.

[19] S. Ye and Y. Xiong Bill, "An Innovative Nonintrusive Driver Assistance System for Vital Signal Monitoring," *IEEE J. Biomed. Health Inform.,* vol. 18, pp. 1932-1939, 2014.

[20] T. Eichele, V. D. Calhoun, M. Moosmann, K. Specht, M. L. Jongsma, R. Q. Quiroga, H. Nordby, and K. Hugdahl, "Unmixing concurrent EEG-fMRI with parallel independent component analysis," *Int. J. Psychophysiol.,* vol. 67, pp. 222-234, 2008.

[21] Y. Tran, A. Craig, P. Boord, and D. Craig, "Using independent component analysis to remove artifact from electroencephalographic measured during stuttered speech," *Med. Biol. Eng. Comput.,* vol. 42, pp. 627-633, 2004/09/01 2004.

[22] Y. Zou, V. Nathan, and R. Jafari, "Automatic Identification of Artifact-related Independent Components for Artifact Removal in EEG Recordings," *IEEE J. Biomed. Health Inform.,* vol. PP, pp. 1-1, 2014.

[23] R. Thuraisingham, Y. Tran, P. Boord, and A. Craig, "Analysis of eyes open, eye closed EEG signals using second-order difference plot," *Med. Biol. Eng. Comput.,* vol. 45, pp. 1243-1249, 2007/12/01 2007.

[24] S. Makeig, A. J. Bell, T.-P. Jung, and T. J. Sejnowski, "Independent component analysis of electroencephalographic data," *Adv. Neural Inf. Process. Syst.,* pp. 145-151, 1996.

[25] A. Delorme, T. Sejnowski, and S. Makeig, "Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis," *Neuroimage,* vol. 34, pp. 1443-1449, 2007.

[26] X.-L. Li and T. Adali, "Blind spatiotemporal separation of second and/or higher-order correlated sources by entropy rate minimization," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP),* 2010, pp. 1934-1937.

[27] J. R. Wolpaw and E. W. Wolpaw, "Brain-computer interfaces: something new under the sun," *Brain-Computer Interfaces: Principles and Practice,* p. 14, 2012.

[28] B. He, S. Gao, H. Yuan, and J. R. Wolpaw, "Brain–computer interfaces," in *Neural Engineering*: Springer, 2013, pp. 87-151.

[29] D. J. McFarland and J. R. Wolpaw, "Sensorimotor rhythm-based brain–computer interface (BCI): model order selection for autoregressive spectral analysis," *J. Neural Eng.,* vol. 5, p. 155, 2008.

[30] O. Faust, R. Acharya, A. Allen, and C. Lin, "Analysis of EEG signals during epileptic and alcoholic states using AR modeling techniques," *IRBM,* vol. 29, pp. 44-52, 2008.

[31] D. J. McFarland and J. R. Wolpaw, "Sensorimotor rhythm-based brain-computer interface (BCI): feature selection by regression improves performance," *IEEE Trans. Neural Sys. Rehabil. Eng.,* vol. 13, pp. 372-379, 2005.

[32] H. T. Nguyen, "Intelligent technologies for real-time biomedical engineering applications," *Int. J. Autom. and Control,* vol. 2, Nos.2/3, pp. 274-285, 2008.

[33] Y. Tran, N. Wijesuriya, R. A. Thuraisingham, A. Craig, and H. T. Nguyen, "Increase in regularity and decrease in variability seen in electroencephalography (EEG) signals from alert to fatigue during a driving simulated task," in *Proc. 30th Annual Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC),* 2008, pp. 1096-1099.

[34] X.-L. Li and T. Adali, "A novel entropy estimator and its application to ICA," in *Proc. IEEE Int. Workshop Mach. Learn. Signal Process. (MILSP),* 2009, pp. 1-6.

[35] N. R. Anderson, K. Wisneski, L. Eisenman, D. W. Moran, E. C. Leuthardt, and D. J. Krusienski, "An offline evaluation of the autoregressive spectrum for electrocorticography," *IEEE Trans. Biomed. Eng.,* vol. 56, pp. 913-916, 2009.

[36] C. Brunner, M. Billinger, C. Vidaurre, and C. Neuper, "A comparison of univariate, vector, bilinear autoregressive, and band power features for brain–computer interfaces," *Med. Biol. Eng. Comput.,* vol. 49, pp. 1337-1346, 2011.

[37] N.-J. Huan and R. Palaniappan, "Neural network classification of autoregressive features from electroencephalogram signals for brain–computer interface design," *J. Neural Eng.,* vol. 1, p. 142, 2004.

[38] G. Florian and G. Pfurtscheller, "Dynamic spectral analysis of event-related EEG data," *Electroencephalogr. Clin. Neurophysiol.,* vol. 95, pp. 393-396, 1995.

[39] H. B. Demuth, M. H. Beale, O. De Jess, and M. T. Hagan, "Neural network design," 2014.

[40] C. M. Bishop, *Neural networks for pattern recognition*: Oxford university press, 1995.

[41] T. Fawcett, "An introduction to ROC analysis," *Pattern Recogn. Lett.,* vol. 27, pp. 861-874, 2006.

[42] M. J. P. Castanho, L. C. Barros, A. Yamakami, and L. L. Vendite, "Fuzzy Receiver Operating Characteristic Curve: An Option to Evaluate Diagnostic Tests," *IEEE Trans. Inf. Technol. Biomed.,* vol. 11, pp. 244-250, 2007.

Rifai Chai (M'14) received the B. Eng. degree from Krida Wacana Christian University, Jakarta, Indonesia in 2000 and s PhD degree in Engineering from the University of Technology, Sydney (UTS) in 2014. From 2000 to 2010 he has worked as a Product Development Engineer; Research and Development Engineer; and Project Engineer with companies in Indonesia and Australia. Currently, he is working as an Associate Lecturer with Faculty of Engineering and Information Technology, University of Technology, Sydney, Australia. His research interests on brain-computer interfaces, biomedical signal processing, biomedical instrumentation, embedded system and computational intelligence using neural networks, fuzzy logic and evolutionary computation.

Ganesh R. Naik (M'08, SM'15) received a B.E. degree in Electronics and Communication Engineering from the University of Mysore, India, in 1997, an M.E. degree in Communication and Information Engineering from Griffith University, Brisbane, Australia, in 2002, and a PhD degree in Electronics Engineering, specialising in biomedical engineering and signal processing from RMIT University, Melbourne, Australia, in 2009.

Since 2013, he is working as a Chancellor's Post-doctoral Research Fellow in the Faculty of Engineering and Information Technology (FEIT), UTS. As an early career researcher, he has edited 10 books, authored more than 80 papers in peer reviewed journals, conferences, and book chapters over the last seven years. Dr. Naik serves as an associate editor for IEEE ACCESS and two Springer journals (*Circuits, Systems, and Signal Processing* and *Australasian Physical & Engineering Sciences in Medicine*). He is a recipient of the Baden–Württemberg Scholarship from the University of Berufsakademie, Stuttgart, Germany (2006–2007). In 2010, Dr. Naik was awarded an ISSI overseas fellowship from Skilled Institute Victoria, Australia.

Tuan Nghia Nguyen received his PhD in Engineering from the University of Technology, Sydney (UTS) in March 2010. He was awarded as a recipient of the Chancellor's List in recognition of exceptional scholarly achievement in PhD research in 2010. Currently, he is a Chancellor's Postdoctoral Research Fellow in the Centre for Health Technologies, Faculty of Engineering at the UTS. His current research interests include biomedical engineering, advanced control and artificial intelligence for health technologies. He is also a senior member of the Institution of Electrical and Electronics Engineers (IEEE) and the Institution of Engineers, Australia (IEAust).

Ashley Craig received his doctorate from the Faculty of Medicine, UNSW in 1986. He also received an Honorary Doctorate for his contribution to neurological and rehabilitation sciences in 2003 from South Western University. He is the Professor of Rehabilitation Sciences in the John Walsh Centre for Rehabilitation Research, Sydney Medical School-Northern, The University of Sydney, as well as a senior member of the Kolling Institute of Medical Research. He has been involved in research in the area of neurological injury and its impacts for many years, and has especially focused on fatigue and its association with brain activity. He has also been involved in BCI research, investigating "hands free" neural signal strategies for people with severe injury.

S.H. Ling (M'06-SM'12) received the BEng degree from the Department of Electrical Engineering, M.Phil. and Ph.D. degrees from the Department of Electronic and Information Engineering in the Hong Kong Polytechnic University in 1999, 2002 and 2007 respectively. Currently, he works in University of Technology, Sydney, Australia as Senior Lecturer. He has authored and coauthored over 150 books, international journal and conference papers on computational Intelligence and its industrial applications. His current research interests include evolution computations, fuzzy logics, neural networks, hybrid systems and biomedical applications. Currently, he serves as Co-Editors-in-Chief for Journal of Intelligent Learning Systems and Applications.

Hung T. Nguyen (SM'99) received the B.E. degree (with First Class and University Medal) and the Ph.D. degree from the University of Newcastle, Newcastle, Australia, in 1976 and 1980, respectively.

He is a Professor of electrical engineering at the University of Technology Sydney (UTS), Ultimo, Australia. He is Assistant Deputy Vice Chancellor (Innovation) and Director of the Centre for Health Technologies at the UTS. He has been involved with research in the areas of biomedical engineering, artificial intelligence, neurosciences and advanced control for more than 20 years. He has developed several medical devices and system for diabetes, disability, fatigue, cardiovascular diseases, and breast cancer. He was an Engineering Manager of Power Electronics Pty Ltd., from 1988 to 1998, the Founding and Executive Director of AIMedics Pty Ltd., from 2001 to 2006, and Dean of the Faculty of Engineering and Information Technology, from 2009 to 2014.

Prof. Nguyen is a Fellow of the Institution of Engineers, Australia, the Australian Computer Society, and the British Computer Society.

Yvonne Tran received her BSc (Hons) in Biomedical Science in 1997 and Ph.D. degree in Psychophysiology in 2001 both from the University of Technology, Sydney, Australia. In 2001, she joined the Centre of Health Technology, Faculty of Engineering and Information Technology, University of Technology, Sydney as a Postdoctoral Fellow and is currently working as a Research Associate for this centre. In 2007, she joined the Rehabilitation Studies Unit, University of Sydney, Australia as a Senior Research Officer. Her area of research include investigating neural signals for BCI use, neuropsychophysiology and cognitive associations of deafferentation in people with spinal cord injury, detection of psychophysiological signals following a fatiguing task and psychological injury following a motor vehicle accident.