Inter-Subject Transfer Learning for EEG-based Mental Fatigue Recognition

Yisi Liu
Fraunhofer Singapore
Singapore, Singapore
LiuYS@ntu.edu.sg

Zirui Lan
Fraunhofer Singapore
Nanyang Technological University
Singapore, Singapore
LanZ0001@e.ntu.edu.sg

Jian Cui
Fraunhofer Singapore
Nanyang Technological University
Singapore, Singapore
Cuijian@ntu.edu.sg

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

Wolfgang Müller-Wittig
Fraunhofer Singapore
Nanyang Technological University
Singapore, Singapore
Wolfgang.Mueller-wittig@fraunhofer.sg

***Abstract*—Mental fatigue is one of the major factors leading to human errors. To avoid failures caused by mental fatigue, researchers are working on ways to detect/monitor fatigue using different types of signals. Electroencephalography (EEG) signal is one of the most popular methods to recognize mental fatigue since it directly measures the neurophysiological activities in the brain. Current EEG-based fatigue recognition algorithms are usually subject-specific, which means a classifier needs to be trained per subject. However, as fatigue may need a relatively long period to induce, collecting training data from each new user could be time-consuming and troublesome. Calibration-free methods are desired but also challenging since significant variability of physiological signals exists among different subjects. In this paper, we proposed algorithms using inter-subject transfer learning for EEG-based mental fatigue recognition, which did not need a calibration. To explore the influence of the number of EEG channels on the algorithms' accuracy, we also compared the cases of using one channel only and multiple channels. Random forest was applied to choose the channel that has the most distinguishable features. A public EEG fatigue dataset recorded during driving was used to validate the algorithms. EEG data from 11 subjects were selected from the dataset and leave-one-subject-out cross-validation was employed. The channel from the occipital lobe is selected when only one channel is desired. The proposed transfer learning-based algorithms using Maximum Independence Domain Adaptation (MIDA) achieved an accuracy of 73.01% with all thirty channels, and using Transfer Component Analysis (TCA) achieved 68.00% with the selected one channel.**

Keywords- transfer learning; EEG; mental fatigue recognition

I. INTRODUCTION

Resulting from long working hours, poor sleep, or night shifts, mental fatigue is a common problem nowadays in workplaces and it can hamper productivity and overall cognitive functions. As revealed by the NSC (National Safety Council) [1], the health-related loss of productivity caused by fatigue is estimated to cost employers $1,200 to $3,100 per employee annually. In addition, the decline in the abilities of perception, recognition, and control caused by fatigue is dangerous for certain jobs requiring continuous concentration, such as pilots, vehicle drivers, and helmsmen. For example, according to National Highway Traffic Safety Administration (NHTSA), there were over 72,000 police-reported crashes involving fatigued drivers from 2009 to 2013. The number of fatalities involving a drowsy driver was 775 (2.1 percent of total fatalities) in 2018 [2].

To deal with such problems caused by mental fatigue, an automatic recognition of mental fatigue is crucial. Different types of physiological signals, including eye activities, electroencephalography (EEG), and electrocardiography (ECG), etc. are explored by researchers to track fatigue states. It is found that the increase of eye blink rates [3], the percentage of eye closure [4], the pupil size [5], the change of heart rate [6], and the change of EEG band power [7] could be indicators of fatigue. Among these physiological signals, EEG is one of the most favored ways to detect fatigue since it can directly reflect the neural activities in the brain. However, as there is large variability of the EEG signals by both physical differences (e.g., electrode displacements, environmental noises, skin-electrode impedance) and biological differences (e.g., different head shapes, size, and brain activity patterns), the major inconvenience when using EEG for cognitive state recognition is that the current EEG-based algorithms are subject-dependent which means every user has to conduct a calibration before the recognition. It would be impractical especially for mental fatigue recognition to conduct the calibration since eliciting fatigue could be a relatively long and gradual process up to hours. Thus, it is desired to build a universal classifier which is trained with data pooled from existing subjects and does not require a calibration for new users. Classic machine learning methods relied on feature selection but still failed to adapt accurately to new subjects since there is always a remarkable mismatch of the selected features among different subjects.

Recent progress in transfer learning has proved its usefulness over classic machine learning methods. It addresses the distribution dissimilarity among different subjects by domain adaptation. It has been successfully used in the field of text classification [8], WiFi location classification [8], gas sensor drift correction [9], breath data analysis [9], etc. In our previous work [10], we have shown that the transfer learning-enabled classifier outperformed the other classifiers such as traditional logistic regression, and deep-learning based classifiers when all channels from the EEG signals were used.

In this paper, we further improved the transfer learning-based algorithms by using Maximum Independence Domain Adaptation (MIDA), which could achieve higher accuracy compared with our previous work which used Transfer Component Analysis (TCA) in [10]. Besides, we investigated the influence of the number of EEG channels on the recognition accuracy. Though there is some existing work on fatigue recognition using a single EEG channel, those algorithms are not subject-independent. In this work, we showed that it is possible to reduce the number of channels to just one with adequate recognition accuracy. With less channel, the time to set up the device is reduced and the comfort level of the users is improved.

The paper is organized as follows. Related works are reviewed in Section II. The dataset used to validate the proposed algorithms is described in Section III. The methodology for channel selection, the transfer learning-based algorithms, and baseline algorithms for comparison are given in Section IV. The results are presented in Section V, followed by the conclusion in Section VI.

II. RELATED WORKS

A. Mental fatigue analysis by frequency components of EEG signals

EEG is an electrophysiological monitoring method which measures the voltage fluctuations on the scalp through a set of electrodes. The recording of each electrode reflects neural activities of underlying cortical sources. The frequency bands of EEG signals are commonly classified into delta (1-4 Hz), theta (4-8 Hz), alpha (8-12 Hz), beta (12-30 Hz), low gamma (30-70 Hz), and high gamma (70-150 Hz).

Much early research has found the association between mental fatigue and the oscillation patterns under different experimental conditions, which provides the foundation of the field using EEG signals to predict mental fatigue. For example, Akerstedt et. al. [11] compared the EEG patterns of industrial shift workers at different times of a day. They found that the alpha power density was significantly higher during night shift than in the afternoon. By further analysis of the data, they showed that an increase in alpha power preceded sleep by a few minutes, while theta activity increased majorly during the sleep period. In another experiment [12], the EEG data after sleep deprivation were investigated. It was found that sleep deprivation causes higher absolute power of upper alpha (9.77-12.45 Hz) and beta bands.

Driver fatigue was also extensively studied in real or simulated environments. Torsvall et. al. [13] conducted an experiment on 11 drivers and compared the EEG patterns between day and night journeys. They witnessed increases in all delta, theta and alpha bands power during night driving while significant increase was observed from the theta and alpha bands. Similar results were also reported in [14], where an experiment on 35 subjects in a driving simulator was conducted. It was found that significant increase in delta, theta, alpha and beta band power occurred during the fatigue phases caused by sleep deprivation, compared to the alert period. Eoh et al. [15] studied the pattern of EEG signals from drivers in the morning after a whole-night sleep deprivation. The EEG signals were measured twice before and twice after the driving task. It was found that alpha power and the (alpha+theta)/beta ratio showed a gradually increasing tendency, while beta and beta/alpha showed gradually decreasing tendency when fatigue increases. In [16], the increase in lower alpha power was found to occur when subjects were asked to sustain awake while they were sleepy. If the subjects were allowed to fall asleep, a decrease in alpha and increase in theta power appears. In [17], alpha spindles (defined by short narrowband bursts in the alpha band) were proposed to assess fatigue in real driving conditions. In comparison to alpha band power, alpha spindle is claimed to be more reliable and sensitive to identify mental fatigue.

In summary, results from existing works reveal that mental fatigue can be reflected by the change in different EEG frequency bands.

B. Mental Fatigue recognition by EEG-based machine learning methods

Usually, the supervised machine learning method is the most widely used way to recognize mental fatigue from EEG signals. The common practice is to extract a set of features that can maximumly distinguish the EEG signals of the fatigue state from normal state. The extracted features are used to train a classifier which will identify unseen data later. As mentioned in the last section, usually the frequency components or their transformations are used as features. For example, Yeo et. al. [18] extracted dominant frequency, average power of dominant peak, center of gravity frequency (CGF), and frequency variability (FV) from the power spectrum density of the delta, theta, alpha and beta band, and used Supported Vector Machine (SVM) for classification. In addition to frequency domain features, time domain features have also been found to be useful for mental fatigue recognition. Hu et. al. [19] considered four types of entropy features, namely sample entropy, fuzzy entropy, approximate entropy, and spectral entropy, together with gradient-boosting decision tree as the classifier to detect driver mental fatigue. Luo et. al. [20] proposed an adaptive multi-scale entropy feature extraction algorithm to obtain the scale factor of the signal. Wang et. al. [21] advocated to use wavelet entropy features since the wavelet transform can intrinsically incorporate information from both time and frequency domain of EEG signals.

Another emerging direction in this field uses deep learning infrastructure for EEG signal classification, since it can directly learn a cascade of representations from raw EEG data and jointly train the classifier by adjusting the parameters through back propagation. Lawhern et. al. [22] proposed a compact convolutional neural network named EEGNet for EEG-based brain-computer interface (BCI). Inspired by the filter-bank common spatial pattern (FBCSP) algorithm [23], they proposed two blocks of depthwise and separable convolutions to extract spatial and temporal features from the raw EEG signals. Fahimi et. al. [24] propose a convolutional neural network to detect the attentive mental state from single-channel raw EEG data. A detailed review on current progress of deep learning based EEG analysis can be found in [25]. As for mental fatigue recognition, initial attempts have been made to use CNN [26] and residual networks [27] for driver fatigue recognition. However, these existing works on EEG-based mental fatigue recognition remain on the intra-subject level.

C. Mental Fatigue recognition by transfer learning

Intra-subject fatigue recognition has the benefit of better recognition accuracy. However, we stress that it is impractical to induce fatigue for calibration purpose in actual applications, as the induction of fatigue is usually a long and gradual process. In this context, inter-subject fatigue recognition could potentially be more practical, which considers using existing fatigue EEG data from other subjects to calibrate the classifier. However, inter-subject recognition poses another challenge since neural patterns of fatigue are subject-specific—EEG data and features tend to distribute dissimilarly among different subjects. Classic machine learning approaches assume that the training data and test data follow the same distribution, which can hardly be satisfied in inter-subject fatigue recognition. Domain adaptation [28], a branch of transfer learning, addresses the concern of distribution mismatch between training data and test data. In a typical unsupervised domain adaptation problem, we have source domain data $D_S = \{(x_{S_1}, y_{S_1}), (x_{S_2}, y_{S_2}), \ldots, (x_{S_{n_1}}, y_{S_{n_1}})\}$ with labels, and target domain data $D_T = \{x_{T_1}, x_{T_2}, \ldots, x_{T_{n_2}}\}$ needing to be classified. Let $P(X_s)$ and $Q(X_T)$ be the marginal distributions of X_S and X_T from the source and target domain, respectively. It is assumed that $P \neq Q$, but there exists a transformation ϕ such that $P(\phi(X_S)) \approx P(\phi(X_T))$ and $P(Y_S|\phi(X_S)) \approx P(Y_T|\phi(X_T))$ [8]. Using this mapping, classic machine learning methods can be applied, where we can train a classifier on $\phi(X_s)$ with Y_S, and predict the class labels for $\phi(X_T)$.

D. Channel selection for mental fatigue recognition

The EEG devices usually have a number of electrodes to record signals from different lobes. However, researchers are working on reducing the number of channels needed in the EEG-based brain states recognition algorithms to a) reduce the computational complexity, b) lower the risk of overfitting, and c) ease the work to set up the device [29].

Some existing research has been done on fatigue recognition using one channel. For example, Lin et. al. [30] developed an unsupervised algorithm with one channel placed at the occipital area for drowsiness detection from EEG signals. As the data size was not balanced, instead of accuracy, 76.9% positive predictive value and 88.7% sensitivity on 10 subjects were reported. Silveira et. al. [31] proposed a method using best m-term approximation combined with discrete wavelet transform. The method was tested on a set of 6 hours and 50 minutes EEG drowsiness signal from the prefrontal lobe, yielding an overall sensitivity of 84.98% and 98.65% of precision. Ogino et. al. [32] used a prefrontal single-channel EEG device to capture EEG data and achieved a classification accuracy of 72.7% using the power spectral density (PSD) with stepwise linear discriminant analysis (SWLDA) and SVM. Albalawi et. al. [33] proposed to detect drowsiness using powers of 8 EEG frequency bands and SVM. An accuracy of 83.36% was achieved using one channel and data from 16 subjects. Belakhdar et. al. [34] proposed a method based on features extracted from alpha band and a single channel. An accuracy of 88% was obtained. However, all the above-mentioned works using one channel are not subject-independent and there is no channel selection to determine which channel to be used in the algorithms.

A lot of efforts have been put into channel selection algorithms for EEG signal processing. It is summarized in [29] that filtering, wrapping, embedded, hybrid, and human-based techniques were applied for channel selection in the literature. The filtering techniques do not require any classifiers to participate. Instead, they usually measure distance, information, dependency, and consistency from different channels, etc. As they are fast, classifier independent and scalable, these techniques are widely used in channel selection. For example, Pal et al. [35] performed an experiment on driver fatigue analysis and found that alpha band at channel Oz showed the best correlation with the subject's drowsiness state, based on which they proposed an unsupervised method for subject-independent drowsiness detection. Jap et. al. [7] investigated four different combinations using powers features, namely (alpha+theta)/beta, alpha/beta, (alpha+theta)/(alpha+beta), and theta/beta, of 52 subjects during a monotonous driving session. Results showed all these four combinations had increasing tendencies during fatigue states, while the most significant tendency was observed at the temporal lobe channels. Chai et. al. [36] used the independent component analysis (ICA) method to select the best ones out of 32 EEG channels for EEG-based driver fatigue classification. The top 3 channels were P8, O1 and T8.

Compared with the filtering techniques, the wrapping techniques consider classifiers to find the best channel subsets. The embedded techniques select the channel during the learning process of the classifiers since the channel selection criteria is embedded in such classifiers. Hybrid techniques combine both filtering and wrapping techniques and human-based techniques are more dependent on the expertise from human experts. Recently, random forest belonging to embedded techniques became popular in the computational biology area, which has advantages such as being able to deal with small sample size, high-dimensional feature space, and complex data structures [37].

In conclusion, EEG channels at occipital, temporal and parietal lobe areas of the scalp could contain useful information for classifying fatigue EEG signals based on the review. In this paper, we applied the random forest method for channel selection.

III. Data preparation

A. Dataset description

In order to test cross-subject fatigue recognition, experiments are conducted on an open dataset. The dataset was published by Cao et al. [38], which were collected during the period from 2005 to 2012. In the experiment, fatigue and drowsiness were induced by a 90-minute sustained-attention night-time driving task in an immersive driving simulator. The participants were tasked to drive and maintain the car in the center of the lane. Lane-departure events were randomly induced which makes the car drift to the left or right from the lane, and participants were asked to move back as quickly as possible by steering the wheel. The next event happened in 5-10 seconds after the car returned to the center lane. The participants needed to sustain their attention to the random lane departure events throughout the whole experiment, and their reactions were timed. The reaction time provides a gauge of the subject's fatigue level.

Twenty-seven participants with age between 22 and 28 were invited to the experiment. The participants were students or staff from the National Chiao Tung University. EEG signals were recorded during the whole 90-minute experiment using Quik-Cap (NeuroScan) with 30 valid channels plus 2 reference channels based on a modified international 10–20 system. The sampling rate is 500 Hz. The raw dataset contains 18.21 GB files which are saved in ".set" format. The dataset was released recently and accessible from [39].

B. Data preparation

The preprocessed version of the dataset available from [40] was used in this paper. As described by the authors [38], the raw EEG signals were filtered by 1-Hz high-pass and 50-Hz low-pass finite impulse response (FIR) filters. Apparent eye blinks that contaminate the EEG signals were manually removed through visual inspection by the authors of the dataset. Ocular and muscular artefacts were removed by the Automatic Artifact Removal (AAR) plug-in of EEGLAB [41]. The processed data were finally downsampled to 128 Hz.

As for epoch extraction, we follow the procedures the authors used on the same dataset in their previous paper [42]. Specifically, 3-second long EEG data (epoch) prior to the onset of the lane-departure events were extracted. The fatigue state in this duration was quantitatively estimated based on the reaction time (RT), which was the duration between the onset of the lane-departure event and the onset of the counter steering event. The RT in each lane-departure event was termed as local RT. Additionally, global RT was defined as the average of local RTs across all epochs within a 90-second window before the onset of the deviation event. An alert RT was defined as the 5th percentile of all local RTs in the entire session. Then, the EEG epochs were labeled as such [42]: epochs with both local RT and global RT less than 1.5 × alert RT were non-fatigue epochs; epochs with both local RT and global RT greater than 2.5 × alert RT were fatigue epochs. To ensure enough epochs for training and testing, we consider subjects that have at least 50 epochs for both states. Finally, the non-fatigue and fatigue epochs were balanced for each participant by subsampling the major class. In this way, totally 837 non-fatigue epochs and 837 fatigue epochs from 11 different participants were extracted, and the size of each epoch was 30 (channels) × 384 (sample points). Sixteen subjects were excluded due to not having enough epochs for fatigue/non-fatigue states. The epoch numbers for each eligible subject are shown in Table 1.

Table 1 Epoch numbers for each eligible subject

Subject ID	**Number of epochs**	
	Fatigue	**Non-fatigue**
1	94	94
5	66	66
22	75	75
31	74	74
35	85	85
41	83	83
42	51	51
43	70	70
44	72	72
45	54	54

53	113	113
Total	837	837

IV. Methodology

A. Feature extraction

We used the spectral band power as fatigue features in this study, which were widely used in existing EEG-based fatigue-related studies [7, 11-17]. The spectral band power was computed via Fast Fourier Transform on each EEG epoch from these four spectral bands: delta (1 – 4 Hz), theta (4 – 8 Hz), alpha (8 – 12 Hz) and beta (12 – 30Hz). The final feature vector was a concatenation of spectral powers extracted from four bands and all available channels. In this study, the final feature vector was of $4 \times 30 = 120$ dimensions.

B. Random forest-based channel selection

We used random forest as the channel selection method in this study, which is an ensemble learning method that constructs multiple decision trees during training and outputs the prediction based on decisions by individual trees. The random forest method is one of the most widely used channel selection methods due to its good predictive performance and easy interpretability. The importance score for each feature can be obtained by averaging the difference in out-of-bag error before and after the permutation over all trees. In our case, we fed all the labeled samples from all subjects into a random forest classifier for training. The random forest method from scikit-learn library [43] and the default parameters (number of estimators is 100, minimum number of samples to split an internal node is 2, minimum number of samples at leaf node is 1, and no restriction on the maximum depth of the tree) implemented in the model were used. The training stopped when the impurity value of the tree model reached 1e-7. After training, we got an importance value for each of the 120 features. We averaged the importance values of four band power features for each channel and used it as a channel importance score. The channel with the top importance score was selected.

C. Transfer learning methods

a) Transfer Component Analysis

Transfer Component Analysis (TCA) was proposed by Pan et al. [8] to mitigate the mismatch of distributions between source data and target data, which causes degraded classification accuracy. It seeks a projection to a latent subspace, where the projected source data and target data achieve a reduced Maximum Mean Discrepancy (MMD) in a reproducing kernel Hilbert space (RKHS) [44], which measures the distance between the empirical means of two distributions. It has proven that MMD will asymptotically approach zero if and only if the two distributions are identical when the RKHS is universal [45]. Using the kernel trick, the MMD of source data X_s and target data X_T in the resultant latent subspace evaluates to

$$\mathrm{MMD}(X_S, X_T) = \mathrm{tr}(W^\top KLKW), \tag{1}$$

where $tr(\cdot)$ is the trace operator, $W \in R^{(n_1+n_2)\times h}$ is the projection matrix, $n_1$and n_2 are the number of examples in X_S and X_T, respectively, h is the dimension of the latent space, $K = [k_{ij}]$ is the kernel matrix defined on $X = [X_S;\ X_T]$, and $L = [L_{ij}]$ where

$$L_{ij} = \begin{cases} \frac{1}{n_1^2}\ if\ x_i, x_j \in X_s \\ \frac{1}{n_2^2}\ if\ x_i, x_j \in X_T \\ -\frac{1}{n_1 n_2}\ otherwise \end{cases} \tag{2}$$

TCA seeks to minimize the MMD plus a regularization term subjecting to a variance constraint:

$$\min_{W} \ \operatorname{tr}(W^\top KLKW) + \mu \operatorname{tr}(W^\top W),$$

$$\text{s. t. } W^\top KHKW = I, \tag{3}$$

where $H = I - n^{-1}\mathbf{1}_n\mathbf{1}_n^\top$ is a centering matrix, $n = n_1 + n_2$, $\mathbf{1}_n$ is an all-one vector of dimension n, and μ is the trade-off parameter. The solution to W is the h eigenvectors of $(KLK + \mu I)^{-1}KHK$ corresponding to the h leading eigenvalues [8].

b) Maximum independence domain adaptation

Maximum independence domain adaptation (MIDA) [9] projects the data of different domains to a latent, domain-invariant space, where the projected samples are independent of the domain features. Domain feature $\boldsymbol{d}$ denotes which domain a sample belongs to and is defined using a one-hot coding scheme: $d_i = 1$ if the sample comes from subject i and 0 otherwise. In the context of inter-subject transfer, each subject constitutes a domain by himself, and the domain feature denotes which subject a sample belongs to. Feature vectors are augmented with their respective domain features by concatenation,$\hat{X} = \begin{bmatrix} X \\ D \end{bmatrix}$. We project $\hat{X}$ to the desired latent subspace by applying a mapping ϕ followed by a linear transformation matrix $\widehat{W}$ to $\hat{X}$, expressed as $X' = \widehat{W}^\top \phi(\hat{X})$. Similar to other kernel dimensionality reduction methods [9], the key idea is to construct $\widehat{W}$ as a linear combination of $\phi(\hat{X})$, namely $\widehat{W} = \phi(\hat{X})W$. It follows that $X' = W^\top \phi(\hat{X})^\top \phi(\hat{X})$. $\phi(\hat{X})^\top \phi(\hat{X})$ can be efficiently computed by way of kernel trick and a proper kernel function. Let $K_{\hat{X}} = \phi(\hat{X})^\top \phi(\hat{X}) = [k_{ij}]$, $k_{ij} = ker(\hat{X}_{:i}, \hat{X}_{:j})$, where $ker(\cdot)$ can take the form of any valid kernel function such as linear function, polynomial function, and radial basis function, etc.

The desired latent subspace should achieve maximum independence between X' and D. Intuitively, when X' is independent of D, we cannot distinguish which subject a specific sample belongs to, thus minimizing the difference of distribution of samples among different subjects. The Hilbert–Schmidt independence criterion (HSIC) is used to evaluate the independence between X' and D, which, dropping the scalar, is empirically evaluated to [9]

$$\mathrm{HSIC}(X', D) = \operatorname{tr}(W^\top K_{\hat{X}} H K_D H K_{\hat{X}} W), \tag{4}$$

where $K_D = D^\top D$ is the kernel matrix on domain feature D.

Besides minimizing $\mathrm{HSIC}(X', D)$, it is also important to preserve the statistical property of the data in the latent subspace, such as the variance [9]. This is achieved by maximizing the trace of the covariance matrix of X'. The final objective function is

$$\max_{W} \ -\operatorname{tr}(W^\top K_{\hat{X}} H K_D K_{\hat{X}} W) + \mu \operatorname{tr}(W^\top K_{\hat{X}} H K_{\hat{X}} W)$$

$$\text{s. t. } W^\top W = I, \tag{5}$$

which consists of the HSIC term and the variance term and is subject to an orthonormal constraint on W. The solution of W is the eigenvectors of $K_{\hat{X}}(-HK_DH + \mu H)K_{\hat{X}}$ corresponding to the h largest eigenvalues [9].

D. Baseline methods for comparison

We used logistic regression (LR) as a baseline method. LR is a simple yet effective binary classifier. Given a feature example x, the probability of predicting x as class 1 is estimated by

$$P(y = 1 | x;\ W, b) = \frac{1}{(1 + e^{-(W^\top x + b)})} \tag{6}$$

where W and b are the model parameters that need to be fit on the training data and are usually found by gradient-descent based optimization.

The state-of-the-art deep learning model EEGNet [22] was used for comparison. It was designed for a general EEG signal classification purpose and has been tested for both intra-subject and cross-subject classification on several different EEG datasets. The model takes raw EEG signals as input instead of handcrafted features. The EEGNet structure contains two-step convolution—depthwise convolution and spatial convolution.

E. Classification

To simulate cross-subject fatigue recognition, we carried out leave-one-subject-out cross-validation on the dataset. In a leave-one-subject-out cross-validation setting, the data from one subject were held out from the dataset and reserved as test data, and the data from remaining subjects were pooled together and used as the training data. The recognition accuracy was then evaluated on the held-out subject. The process was repeated until each subject had served as a test subject once. For TCA and MIDA, we followed the practice by their authors [8, 9] and set the trade-off parameter μ to 1. For the deep learning model, we chose 2 structures - EEGNet 8,2 and EEGNet 4,2. Adam method was used as the optimizer and the default parameters (learning_rate=0.001, beta_1=0.9, beta_2=0.999, epsilon=1e-07) implemented in TensorFlow was used. Considering the neural networks are stochastic so the results are not consistent for each training, we repeated the evaluation of each model on each test subject for 10 times and reported the average accuracy.

V. Result

A. Channel selection

As mentioned in Section IV A, we applied random forest-based channel selection. The obtained importance score for each channel is shown in Table 2. The results show that the occipital channels have the highest importance scores for classifying fatigue EEG signals, which is also consistent with the literature review. Therefore, we choose O1 when only one channel is desired.

Table 2 Channel Importance by Random Forest Method (×100)

Channel	Score	Channel	Score	Channel	Score
Fp1	1.17	FC4	0.55	CP4	0.56
Fp2	0.84	FT8	0.68	TP8	0.54
F7	0.84	T3	0.68	T5	0.59
F3	0.44	C3	0.87	P3	0.89
Fz	0.75	Cz	1.05	Pz	1.88
F4	0.42	C4	0.37	P4	0.82
F8	0.47	T4	0.50	T6	0.82
FT7	0.77	TP7	0.52	O1	1.89
FC3	0.86	CP3	1.01	Oz	1.83
FCz	0.64	CPz	0.68	O2	1.08

B. Mental fatigue recognition

The classification results using leave-one-subject-out cross-validation with 30 EEG channels are presented in Table 3. The MIDA-based algorithm achieved the highest accuracy of 73.01%, which is slightly higher than TCA by 0.31%, significantly higher than baseline, EEGNet-8,2 ($t(10) = 3.47$, $p < 0.01$), and EEGNet-4,2 ($t(10) = 5.95$, p <0.01). From the accuracy of each individual subject, 7 out of 11 subjects had equal or higher accuracy when using MIDA than TCA. The corresponding sensitivity and specificity for the best performing methods, TCA and MIDA, are presented in Table 4. The average (sensitivity, specificity) for TCA is (71.28%, 75.54%), while the average (sensitivity, specificity) for MIDA is (73.30%, 73.84%). The ROC is depicted in Fig. 1 and it shows that the performance of both TCA (AUC=0.7758) and MIDA (AUC=0.7792) is close to each other and above random guess (AUC=0.5).

Table 3 Leave-one-subject-out fatigue recognition accuracy using all channels

Subject	Logistic Regression	EEGNet-8,2	EEGNet-4,2	TCA	MIDA
1	65.96%	65.96%	57.98%	72.87%	65.96%
2	53.79%	63.64%	50.00%	62.12%	62.12%
3	42.67%	48.67%	49.33%	58.67%	61.33%
4	64.19%	52.03%	52.70%	68.92%	72.30%
5	59.41%	76.47%	66.47%	75.29%	80.59%
6	62.65%	48.80%	64.46%	75.30%	66.87%
7	77.45%	63.73%	58.82%	64.71%	70.59%
8	65.71%	71.43%	60.71%	82.14%	71.43%
9	78.47%	79.17%	73.61%	85.42%	81.94%
10	66.67%	75.93%	75.93%	87.04%	90.74%
11	62.83%	57.96%	48.23%	67.26%	79.20%
Average	63.62%	63.98%	59.84%	72.70%	73.01%
Std	9.93%	11.10%	9.55%	9.42%	9.17%

Table 4 Accuracy, sensitivity, and specificity using all channels

Subject	TCA			MIDA		
	Accuracy	Sensitivity (True non-fatigue rate)	Specificity (True fatigue rate)	Accuracy	Sensitivity (True non-fatigue rate)	Specificity (True fatigue rate)
1	72.87%	70.87%	75.29%	65.96%	64.15%	68.29%
2	62.12%	66.67%	59.52%	62.12%	61.43%	62.90%
3	58.67%	57.14%	61.02%	61.33%	61.64%	61.04%
4	68.92%	69.44%	68.42%	72.30%	72.60%	72.00%
5	75.29%	69.37%	86.44%	80.59%	76.53%	86.11%
6	75.30%	73.86%	76.92%	66.87%	66.67%	67.07%
7	64.71%	61.54%	70.27%	70.59%	81.82%	65.22%
8	82.14%	83.58%	80.82%	71.43%	70.83%	72.06%
9	85.42%	85.92%	84.93%	81.94%	78.05%	87.10%

10	87.04%	80.30%	97.62%	90.74%	86.67%	95.83%
11	67.26%	65.35%	69.70%	79.20%	85.87%	74.63%
Average	72.70%	71.28%	75.54%	73.01%	73.30%	73.84%
Std	9.42%	9.01%	11.42%	9.17%	9.24%	11.19%

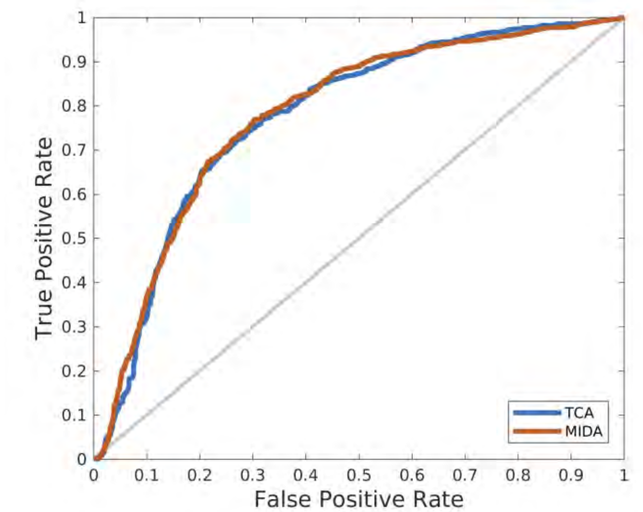

Fig. 1. ROC for TCA-based and MIDA-based fatigue/non-fatigue classification using 30 EEG channels and aggregated from 11 subjects. Positive class = non-fatigue. Number of positive class = 837. Number of negative class = 837. TCA AUC = 0.7758. MIDA AUC = 0.7792.

The classification results using leave-one-subject-out cross-validation with only one EEG channel are presented in Table 5. The TCA-based algorithm achieved the highest accuracy of 68.00%, which is slightly higher than MIDA by 0.11%, significantly higher than EEGNet-8,2 ($t(10) = 3.01, p < 0.05$) and EEGNet-4,2 ($t(10) = 2.29, p < 0.05$). The corresponding sensitivity and specificity are presented in Table 6. The average (sensitivity, specificity) for TCA is (68.20%, 70.57%), while the average (sensitivity, specificity) for MIDA is (64.05%, 70.57%). The ROC is depicted in Fig. 2 and it shows that the performance of MIDA (AUC=0.7396) is slightly better than TCA (AUC= 0.6931) and above random guess (AUC=0.5).

Compared to the results using all 30 channels, the accuracy drops by 4.70% and 5.12% for TCA and MIDA respectively. However, it also proves the possibility of using just one channel to recognize the state of fatigue.

Table 5 Leave-one-subject-out fatigue recognition accuracy using one channel

Subject	Logistic Regression	EEGNet-8,2	EEGNet-4,2	TCA	MIDA
1	83.51%	76.33%	76.22%	83.51%	78.19%
2	41.67%	42.35%	39.77%	40.15%	48.48%
3	49.33%	49.27%	50.00%	49.33%	37.33%
4	61.49%	64.66%	66.55%	60.81%	60.81%
5	71.18%	60.12%	68.06%	73.53%	78.82%
6	66.87%	65.30%	61.63%	66.87%	71.08%

7	73.53%	58.92%	65.20%	74.51%	73.53%
8	67.86%	59.64%	67.43%	68.57%	64.29%
9	86.11%	68.26%	69.51%	86.81%	83.33%
10	82.41%	73.98%	81.30%	82.41%	81.48%
11	61.95%	56.06%	56.42%	61.50%	69.47%
Average	67.81%	61.35%	63.83%	68.00%	67.89%
Std	13.92%	10.02%	11.67%	14.47%	14.38%

Table 6 Accuracy, sensitivity, and specificity using one channel

Subject	TCA			MIDA		
	Accuracy	Sensitivity (True non-fatigue rate)	Specificity (True fatigue rate)	Accuracy	Sensitivity (True non-fatigue rate)	Specificity (True fatigue rate)
1	83.51%	89.87%	78.90%	78.19%	71.20%	92.06%
2	40.15%	36.17%	42.35%	48.48%	48.65%	48.28%
3	49.33%	49.66%	0.00%	37.33%	28.89%	40.95%
4	60.81%	58.70%	64.29%	60.81%	60.81%	60.81%
5	73.53%	67.24%	87.04%	78.82%	73.33%	87.69%
6	66.87%	60.29%	96.67%	71.08%	67.33%	76.92%
7	74.51%	83.78%	69.23%	73.53%	73.08%	74.00%
8	68.57%	84.21%	62.75%	64.29%	59.09%	83.33%
9	86.81%	89.55%	84.42%	83.33%	78.57%	90.00%
10	82.41%	73.97%	100.00%	81.48%	73.61%	97.22%
11	61.50%	56.70%	90.63%	69.47%	70.00%	68.97%
Average	68.00%	68.20%	70.57%	67.89%	64.05%	74.57%
Std	14.47%	17.65%	28.91%	14.38%	14.41%	18.31%

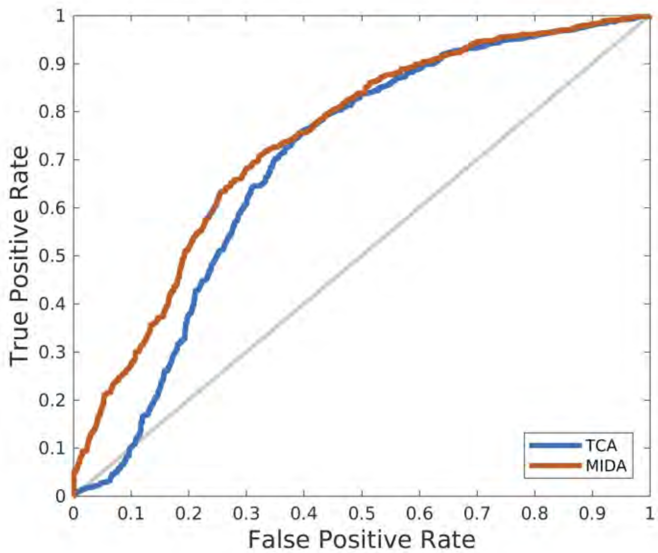

Fig. 2. ROC for TCA-based and MIDA-based fatigue/non-fatigue classification using 1 EEG channel and aggregated from 11 subjects. Positive class = non-fatigue. Number of positive class = 837. Number of negative class = 837. TCA AUC = 0.6931. MIDA AUC = 0.7396.

VI. CONCLUSION

Mental fatigue is a common problem nowadays in workplaces and it can hamper productivity and overall cognitive functions. To deal with such problems, an automatic recognition of mental fatigue is crucial. EEG measures the neural activities in the brain and has the potential to recognize the states of mental fatigue. However, as there is large variability of the EEG signals by both physical differences and biological differences, the major inconvenience when using EEG for cognitive state recognition is that the current EEG-based algorithms are subject-dependent which means every user has to conduct a calibration before the recognition. We stress that it would be impractical especially for mental fatigue recognition to conduct the calibration since eliciting fatigue could be a relatively long and gradual process up to hours. Thus, it is desired to build a universal classifier which is trained with data pooled from existing subjects and does not require a calibration for new users. In this paper, we investigate inter-subject mental fatigue recognition and the methods to improve the recognition accuracy. We found that domain adaptation methods MIDA and TCA can effectively enhance the recognition accuracy to 73.01% and 72.70%, respectively, in the leave-one-subject-out cross-validation experiment on a public dataset, significantly outperforming the baseline method or deep learning-based methods. We have also explored the possibility of using one single electrode for the recognition of mental fatigue. Our experiments show that fatigue can be recognized with the best accuracy of 68.00% with just one electrode placed at O1.

ACKNOWLEDGEMENT

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] *Fatigue – You're More Than Just Tired*, National Safety Council, Available: https://www.nsc.org/work-safety/safety-topics/fatigue, Accessed on: 1 Mar 2020.

[2] NHTSA, "2018 Fatal motor vehicle crashes: overview," in "Traffic Safety Facts Resesarch Note," NHTSA's National Center for Statistics and Analysis, Washington, DOT HS 812 826, Oct 2019, Available: https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812826, Accessed on: 1 Mar 2020.

[3] S. Hu and G. Zheng, "Driver drowsiness detection with eyelid related parameters by Support Vector Machine," *Expert Systems with Applications,* vol. 36, no. 4, pp. 7651-7658, 2009.

[4] W. W. Wierwille, S. Wreggit, C. Kirn, L. Ellsworth, and R. Fairbanks, "Research on vehicle-based driver status/performance monitoring; development, validation, and refinement of algorithms for detection of driver drowsiness," National Highway Traffic Safety Administration, Washington, DOT HS 808 247, Oct 1994, Available: https://rosap.ntl.bts.gov/view/dot/2578, Accessed on: 1 Mar 2020.

[5] Y. Morad, H. Lemberg, N. Yofe, and Y. Dagan, "Pupillography as an objective indicator of fatigue," *Current eye research,* vol. 21, no. 1, pp. 535-542, 2000.

[6] Y. Tran, N. Wijesuriya, M. Tarvainen, P. Karjalainen, and A. Craig, "The relationship between spectral changes in heart rate variability and fatigue," *Journal of Psychophysiology,* vol. 23, no. 3, pp. 143-151, 2009.

[7] B. T. Jap, S. Lal, P. Fischer, and E. Bekiaris, "Using EEG spectral components to assess algorithms for detecting fatigue," *Expert Systems with Applications,* vol. 36, no. 2, pp. 2352-2359, 2009.

[8] S. J. Pan, I. W. Tsang, J. T. Kwok, and Q. Yang, "Domain adaptation via transfer component analysis," *IEEE Transactions on Neural Networks,* vol. 22, no. 2, pp. 199-210, 2010.

[9] K. Yan, L. Kou, and D. Zhang, "Learning Domain-Invariant Subspace Using Domain Features and Independence Maximization," *IEEE transactions on cybernetics,* vol. 48, no. 1, pp. 288-299, 2017.

[10] Y. Liu, Z. Lan, J. Cui, O. Sourina, and W. Müller-Wittig, "EEG-Based Cross-Subject Mental Fatigue Recognition," in *2019 International Conference on Cyberworlds (CW)*, 2019, pp. 247-252: IEEE.

[11] T. Åkerstedt, G. Kecklund, and A. Knutsson, "Manifest sleepiness and the spectral content of the EEG during shift work," *Sleep,* vol. 14, no. 3, pp. 221-225, 1991.

[12] M. Corsi-Cabrera, J. Ramos, C. Arce, M. Guevara, M. Ponce-de Leon, and I. Lorenzo, "Changes in the waking EEG as a consequence of sleep and sleep deprivation," *Sleep,* vol. 15, no. 6, pp. 550-555, 1992.

[13] L. Torsvall, "Sleepiness on the job: continuously measured EEG changes in train drivers," *Electroencephalography and clinical Neurophysiology,* vol. 66, no. 6, pp. 502-511, 1987.

[14] S. K. Lal and A. Craig, "Driver fatigue: electroencephalography and psychological assessment," *Psychophysiology,* vol. 39, no. 3, pp. 313-321, 2002.

[15] H. J. Eoh, M. K. Chung, and S.-H. Kim, "Electroencephalographic study of drowsiness in simulated driving with sleep deprivation," *International Journal of Industrial Ergonomics,* vol. 35, no. 4, pp. 307-320, 2005.

[16] W. Klimesch, "EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis," *Brain research reviews,* vol. 29, no. 2-3, pp. 169-195, 1999.

[17] M. Simon *et al.*, "EEG alpha spindle measures as indicators of driver fatigue under real traffic conditions," *Clinical Neurophysiology,* vol. 122, no. 6, pp. 1168-1178, 2011.

[18] M. V. Yeo, X. Li, K. Shen, and E. P. Wilder-Smith, "Can SVM be used for automatic EEG detection of drowsiness during car driving?," *Safety Science,* vol. 47, no. 1, pp. 115-124, 2009.

[19] J. Hu and J. Min, "Automated detection of driver fatigue based on EEG signals using gradient boosting decision tree model," *Cognitive neurodynamics,* vol. 12, no. 4, pp. 431-440, 2018.

[20] H. Luo, T. Qiu, C. Liu, and P. Huang, "Research on fatigue driving detection using forehead EEG based on adaptive multi-scale entropy," *Biomedical Signal Processing and Control,* vol. 51, pp. 50-58, 2019.

[21] Q. Wang, Y. Li, and X. Liu, "Analysis of feature fatigue EEG signals based on wavelet entropy," *International Journal of Pattern Recognition and Artificial Intelligence,* vol. 32, no. 08, pp. 1-15, 2018.

[22] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, "EEGNet: a compact convolutional neural network for EEG-based brain–computer interfaces," *Journal of neural engineering,* vol. 15, no. 5, pp. 1-17, 2018.

[23] K. K. Ang, Z. Y. Chin, C. Wang, C. Guan, and H. Zhang, "Filter bank common spatial pattern algorithm on BCI competition IV datasets 2a and 2b," *Frontiers in neuroscience,* vol. 6, pp. 1-9, 2012.

[24] F. Fahimi, Z. Zhang, W. B. Goh, T.-S. Lee, K. K. Ang, and C. Guan, "Inter-subject transfer learning with an end-to-end deep convolutional neural network for EEG-based BCI," *Journal of neural engineering,* vol. 16, no. 2, pp. 1-12, 2019.

[25] Y. Roy, H. Banville, I. Albuquerque, A. Gramfort, T. H. Falk, and J. Faubert, "Deep learning-based electroencephalography analysis: a systematic review," *Journal of neural engineering,* vol. 16, no. 5, pp. 1-37, 2019.

[26] Z. Gao *et al.*, "EEG-based spatio–temporal convolutional neural network for driver fatigue evaluation," *IEEE transactions on neural networks and learning systems,* vol. 30, no. 9, pp. 2755-2763, 2019.

[27] H. Zeng, C. Yang, G. Dai, F. Qin, J. Zhang, and W. Kong, "EEG classification of driver mental states by deep learning," *Cognitive neurodynamics,* vol. 12, no. 6, pp. 597-606, 2018.

[28] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on knowledge and data engineering,* vol. 22, no. 10, pp. 1345-1359, 2010.

[29] T. Alotaiby, F. E. A. El-Samie, S. A. Alshebeili, and I. Ahmad, "A review of channel selection algorithms for EEG signal processing," *EURASIP Journal on Advances in Signal Processing,* vol. 2015, no. 1, pp. 1-21, 2015.

[30] C.-T. Lin, C.-J. Chang, B.-S. Lin, S.-H. Hung, C.-F. Chao, and I.-J. Wang, "A real-time wireless brain–computer interface system for drowsiness detection," *IEEE transactions on biomedical circuits and systems,* vol. 4, no. 4, pp. 214-222, 2010.

[31] T. d. Silveira, A. d. J. Kozakevicius, and C. R. Rodrigues, "Drowsiness detection for single channel EEG by DWT best m-term approximation," *Research on Biomedical Engineering,* vol. 31, no. 2, pp. 107-115, 2015.

[32] M. Ogino and Y. Mitsukura, "Portable drowsiness detection through use of a prefrontal single-channel electroencephalogram," *Sensors,* vol. 18, no. 12, pp. 1-19, 2018.

[33] H. Albalawi and X. Li, "Single-Channel Real-Time Drowsiness Detection Based on Electroencephalography," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2018, pp. 98-101: IEEE.

[34] I. Belakhdar, W. Kaaniche, R. Djemal, and B. Ouni, "Single-channel-based automatic drowsiness detection architecture with a reduced number of EEG features," *Microprocessors and Microsystems,* vol. 58, pp. 13-23, 2018.

[35] N. R. Pal *et al.*, "EEG-based subject-and session-independent drowsiness detection: an unsupervised approach," *EURASIP Journal on Advances in Signal Processing,* vol. 2008, no. 1, pp. 1-11, 2008.

[36] R. Chai, G. R. Naik, S. H. Ling, Y. Tran, A. Craig, and H. T. Nguyen, "Channels selection using independent component analysis and scalp map projection for EEG-based driver fatigue classification," in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2017, pp. 1808-1811: IEEE.

[37] M. Fernández-Delgado, E. Cernadas, S. Barro, and D. Amorim, "Do we need hundreds of classifiers to solve real world classification problems?," The journal of machine learning research, vol. 15, pp. 3133-3181, 2014.

[38] Z. Cao, C.-H. Chuang, J.-K. King, and C.-T. Lin, "Multi-channel EEG recordings during a sustained-attention driving task," *Scientific data,* vol. 6, no. 1, pp. 1-8, 2019.

[39] *Multi-channel EEG recordings during a sustained-attention driving task, (raw dataset)*, Available: https://figshare.com/articles/Multichannel_%20EEG_recordings_during_a_sustainedattention_driving_task/6427334/5, Accessed on: 1 Mar 2019.

[40] *Multi-channel EEG recordings during a sustained-attention driving task (pre-processed dataset)*, Available: https://figshare.com/articles/Multichannel_EEG_recordings_during_a_sustainedattention_driving_task_preprocessed_dataset_/7666055, Accessed on: 1 Mar 2019.

[41] A. Delorme and S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *Journal of neuroscience methods,* vol. 134, no. 1, pp. 9-21, 2004.

[42] C.-S. Wei, Y.-T. Wang, C.-T. Lin, and T.-P. Jung, "Toward drowsiness detection using non-hair-bearing EEG-based brain-computer interfaces," *IEEE transactions on neural systems and rehabilitation engineering,* vol. 26, no. 2, pp. 400-406, 2018.

[43] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *Journal of machine learning research,* vol. 12, no. Oct, pp. 2825-2830, 2011.

[44] A. Gretton, K. Borgwardt, M. Rasch, B. Schölkopf, and A. J. Smola, "A kernel method for the two-sample-problem," in *Advances in neural information processing systems*, 2007, pp. 513-520.

[45] A. Smola, A. Gretton, L. Song, and B. Schölkopf, "A Hilbert space embedding for distributions," in *International Conference on Algorithmic Learning Theory*, 2007, pp. 13-31.