Mental workload classification in *n*-back tasks based on single-trial EEG

Dai Zhongxiang[1], Bezerianos Anastasios[1], Chen Shen-Hsing Annabel[2], Sun Yu[1]

(*1. Singapore Institute for Neurotechnology (SINAPSE), Centre for Life Science, National University of Singapore, Singapore 117456, Singapore; 2. School of Humanities and Social Sciences, Nanyang Technological University, Singapore 637332, Singapore*)

Abstract: Mental workload estimation has been under extensive investigation over the years, because the capability of monitoring the cognitive workload enables the prevention of cognitive overloading and improvement of workplace safety. Electroencephalogram (EEG) signals has been found to be an objective and non-intrusive measure of mental workload. However, the evaluation of cognitive workload based on single-trial EEG data, which is an essential step towards real-time workload monitoring and brain-computer interface, has been a major challenge. Recently, a number of advanced feature extraction methods and machine learning algorithms have been employed in EEG-based mental workload assessment. In this study, we performed single-trial workload classification using the EEG data recorded during the performance of *n*-back tasks with 2 levels of difficulty (corresponding to low and high levels of workload respectively), examined the effectiveness of 3 types of feature extraction (spectral power, discrete wavelet transform and common spatial filtering), and evaluated the performance of 4 classification algorithms (support vector machine, K-nearest neighbors, random forest and gradient boosting classifiers). Our findings indicate that common spatial filtering was the best-performing individual feature extraction method for single-trial-based workload classification, and the optimal performance was achieved by combining the features from either spectral power or discrete wavelet transform with those from common spatial filtering, and adopting the random forest classifier. This study might provide some helpful guidance on the selection of feature extraction methods as well as machine learning algorithms in mental workload evaluation based on single-trial EEG data.

Keywords: electroencephalogram (EEG); single-trial; mental workload; feature extraction; classification

基于单试验脑电图的 *n*-back 任务中的脑力负荷分类

代忠祥[1], Bezerianos Anastasios[1], Chen Shen-Hsing Annabel[2], 孙　煜[1]

(1. 新加坡国立大学 生命科学中心 新加坡神经科学研究所　117456　新加坡;
2. 南洋理工大学人文社会科学学院　637332　新加坡)

摘　要:近年来,脑力负荷估计已经经历了广泛的研究,因为监测认知负荷的能力能够防止认知超负荷并且改善工作场所安全。脑电图(EEG)信号已经被发现是一种客观和非侵入性的脑力负荷的测量方式。然而,作为实时脑力负荷监测和脑机接口研究的重要一步,基于单试验 EEG 数据的认知负荷的评估一直是一个重大的挑战。最近,许多高级的特征提取方法和机器学习算法已经被采用于基于 EEG 的脑力负荷评估中。在本研究中,使用在具有 2 个难度水平的 *n*-back 任务的执行期间记录的 EEG 数据进行了单试验脑力负荷分类,测试了 3 种类型的特征提取的有效性(谱功率、离散小波变换和公共空间滤波),并评估了 4 种分类算法的性能(支持向量机、K-近邻、随机森林和梯度推进分类器)。研究结果表明,公共空间滤波是性能最好的基于单试验的脑力负荷分类的特征提取方法,而且最佳性能可以通过将来自谱功率或离散小波变换的特征与来自公共空间滤波的特征相结合,并采用随机森林分类器来实现。这项研究可能对基于单试验脑电图数据的脑力负荷评估中的特征提取方法以及机器学习算法的选择提供一些有用的指导。

关键词: 脑电图;单试;精神工作量;特征提取;分类
中图分类号: R338.8　TH7　　**文献标识码**: A　　**国家标准学科分类代码**: 416.60

收稿日期:2016-10　　Received Date: 2016-10

1 Introduction

Mental workload, which is usually evaluated by considering the interactions between the task demand and the individual's capacity [1-3], has been a major research topic in cognitive neuroscience over the years. The capability of monitoring the mental workload offers considerable benefits such as optimizing the working environment and preventing cognitive overloading [3]. Therefore, various attempts have been made to estimate the cognitive workload through different modalities, such as performance measures, subjective ratings, and physiological assessments [3-6]. Among these measures, physiological variables have been preferred mainly because of their objectiveness, capability of continuous measurement, and minimal disturbance to the operators.

Electroencephalogram (EEG) is a widely used technology for measuring electrophysiological signals from the scalp [7] has been found to be a more reliable method for workload evaluation due to various advantages including great sensitivity to variations in mental workload, high temporal resolution, and portability [3,8]. Among the various EEG studies investigating the neural correlations of mental workload, working memory (WM) tasks, most often n-back tasks, have been extensively employed to induce different levels of cognitive workload [9-11]. The widespread adoption of WM tasks mainly results from the fact that the WM system, which is a memory system that facilitates short-term storage and processing of information [12-13], has been found to be closely related to individual's mental workload capacity [9,14].

Due to the high trial-to-trial variability of EEG, traditional EEG analysis of mental workload usually apply averaging method across a number of trials to improve the prediction performance [15]. However, workload estimation based on single trials is critical for online monitoring of cognitive workload as well as the development of brain computer interface (BCI) technology [9,15]. In recent years, studies have demonstrated the feasibility of EEG-based single-trial workload classification[3] and potential of improving the prediction results through the fusion of different types of features [9] and implementation of advanced machine learning techniques.

Given the embedded rich time- and frequency-domain information in EEG signals, various feature extraction methods have been developed and employed in the evaluation of mental workload. For instance, band-specific spectral powers, particularly the powers in the theta [16-18], alpha [19-20] and beta[21-22] bands, have been consistently discovered to be highly correlated with cognitive workload. In addition, discrete wavelet transform (DWT), which decomposes the EEG signals into time-frequency representations at different scales, has also been found to be effective in workload estimation[23-25]. Furthermore, common spatial patterns (CSP) is another widely used technique for online workload classification in BCI research that transforms the multi-channel EEG data into new spatial dimensions so that the variances of the new signals are optimally discriminative [15]. Here, all three techniques, as well as their combinations, have been used to extract features from the EEG signals that were fed into the machine learning algorithms for workload classification.

Over the years, various machine learning techniques have achieved great success in the classification of mental workload using EEG. Different classifiers, including support vector machine (SVM) and K-nearest neighbors (KNN), have been adopted and proved effective in mapping the highly complex relationship between EEG signals and cognitive workload [26-29]. Recently, ensemble learning algorithms, including boosting (such as gradient boosting machine (GBM)[30]) and bagging (such as random forest (RF) classifiers[31]) algorithms, have achieved impressive performance in various contexts[32-35]. In this work, we explored the capability of these 2 types of ensemble learning algorithms in single-trial EEG workload classification, and compared their performances with more widely used conventional classification techniques including SVM and KNN.

In this study, using the EEG data recorded during n-back tasks with two levels of workload (high vs. low), 1) performed individual-based single-trial workload classification, 2) compared the effectiveness of different feature extraction methods as well as their combinations, and 3) compared the performance of different machine learning algorithms in discriminating the cognitive workload. Our findings might serve as a practical guideline for the selection of the types of features and classifiers in future studies on EEG-based single-trial workload classification.

2 Methods and materials

2.1 Subjects and experiments

In the current study, 21 university students performed n-back tasks at 2 difficulty levels: 1-back and 3-back. The study was approved by the Institutional Review Board of National University of Singapore. Written informed consent was obtained from all participants after the explanation of the experimental protocol. The participants were reimbursed S$20 for their participation. Four subjects were excluded from the following analysis due to insufficient number of correct trials.

The experimental paradigm is illustrated in Fig. 1. The subjects were instructed to watch a sequence of images, compare the currently presented image with the one n (n = 1 or 3) trials back, and make the correct response with a keyboard with 3 buttons: "1", "2" and "3". Specifically, the subjects were requested to remember both the content and the position of the image such that, compared with the image n trials back, 1) if the image (not position) was the same, they had to press "1"; 2) if there was a position (not image) match, they had to press "3"; 3) if both the image and position were the same, they had to press "2"; 4) if neither the image nor the position were the same, they do not need to press any button. There were 80 trials for both 1-back and 3-back tasks. The sequences of the n-back tasks within each session were pseudorandomized and each session lasted around 20 minutes.

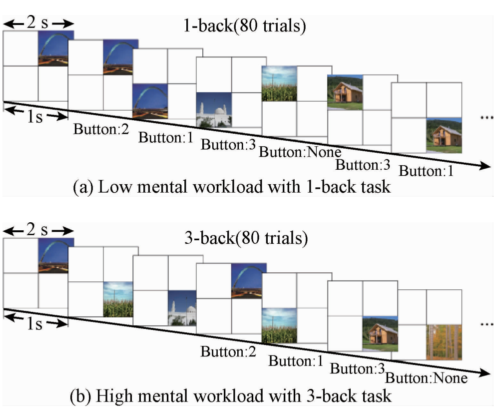


Fig. 1　Experimental paradigm of the n-back tasks

2.2 Data acquisition

During each session, 64-channel EEG data were recorded, with the positions of the electrodes based on the standard 10-20 system, with the ANT wave-guard system (ASA-Lab, ANT B. V., Netherlands). The sampling frequency was 256 Hz and the recorded signals were re-referenced to the average of the electrodes at the left and right mastoids, which resulted in 62 channels in the following analysis. The bipolar electrooculogram (EOG) signals were recorded with electrodes attached to the outer canthi (HEOG), and above and below (VEOG) the right eye. The impedances of the electrodes were maintained below 10 kΩ throughout recording. The data were band-pass filtered from 0.5 ~ 70 Hz for anti-aliasing, and main interferences were removed by applying a 50 Hz notch filter.

2.3 EEG data preprocessing

The recorded EEG signals were band-pass filtered from 1 ~ 40 Hz. Subsequently, the artefacts caused by eye blinks and eye movements were removed through independent component analysis (ICA)[36-37]. Specifically, the multi-channel EEG signals were first decomposed into independent components using the AMICA algorithm[38], following which the correlations between each of the components and the horizontal as well as vertical EOG components were calculated; next, those independent components that exhibited high correlations with the EOG signals were identified and removed, and the EEG signals were reconstructed from the remaining independent components. Afterwards, the artefact-free EEG data were segmented into individual epochs, with the beginning of each epoch marked by the stimulus onset of each trial and a time window of 1 second. In the current study, only correct trials were selected for the following analysis. All preprocessing steps were conducted using EEGLAB[39].

2.4 Feature extraction

Three types of features are extracted from the EEG signals: Spectral power, discrete wavelet transform (DWT) and common spatial patterns (CSP). Each of these feature extraction methods, and their combinations, are employed for the single-trail EEG mental workload classification.

2.4.1 Band power

The spectral power of EEG signals, mostly in the theta (4 ~ 7 Hz), alpha (8 ~ 12 Hz) and beta (12 ~ 30 Hz) frequency bands, has been consistently discovered to be reflective of mental workload[3, 40-41]. Here, the spectral powers in these 3 frequency bands of all electrodes were used as features in the classification analysis, which led to $62 \times 3 = 186$ features for each trial.

2.4.2 Discrete wavelet transform

Wavelet transform (WT) has been widely adopted in various EEG studies due to its advantage in time- and frequency-localization [23, 42-44]. In WT, the time-frequency representation of a time series is obtained through the application of time windows with variable sizes such that high-frequency information is extracted using short time windows, whereas low-frequency representations of a signal is obtained though long time windows [42]. The continuous wavelet transform (CWT) of a time series $x(t)$ is calculated by:

$$CWT(a, b) = \int_{-\infty}^{\infty} x(t) \frac{1}{\sqrt{|a|}} \varphi\left(\frac{t-a}{b}\right) \mathrm{d}t \tag{1}$$

which φ represents the wavelet function, and a and b are the scaling and shifting parameters respectively. Moreover, in order to avoid the huge computational cost of evaluating the wavelet coefficients at every possible values of a and b, discrete wavelet transform (DWT) is calculated by assigning discrete values to the scaling and shifting parameters:

$$DWT(j, k) = \int_{-\infty}^{\infty} x(t) \frac{1}{\sqrt{|2^j|}} \varphi\left(\frac{t-k2^j}{2^j}\right) \mathrm{d}t \tag{2}$$

which can be obtained from equation (1) by replacing a with 2^j and b with $k2^j$. DWT can be efficiently implemented in the following manner: the signal is passed through a quadrature mirror filter, which consists of 1) a high-pass filter that produces the detail coefficients of level 1 (D1) and 2) a low-pass filter that yields the approximation coefficients of level 1 (A1); next, A1 is passed to another quadrature mirror filter, through which the level-2 detail and approximation of the signal are obtained; the process is repeated until the desired number of levels are reached [45]. In this study, the Daubechies-4 (db4) wavelet is used as the wavelet function. Moreover, the EEG signals were decomposed into 5 levels (D1: 64 ~ 128 Hz, D2: 32 ~ 64 Hz, D3: 16 ~ 32 Hz, D4: 8 ~ 16 Hz, and D5: 4 ~ 8 Hz), and the detail coefficients in levels 3, 4 and 5 (corresponding to the beta, alpha and theta bands) were employed for extracting features for the workload classification. The energies of the detail coefficients of each channel in the 3 scales are used as features [23, 46]:

$$E_c = \frac{1}{N} \sum_{i=3,4,5} D_i^2 \tag{3}$$

which N is the number of time points in each trial ($N = 256$). Therefore, $62 \times 3 = 186$ features are produced for each trail.

2.4.3 Common spatial patterns

The common spatial patterns technique is a widely used method for extracting discriminative features in the research of Brain Computer Interface (BCI) for mental workload classification [47-48]. The CSP algorithm derives the optimal spatial filter that transforms the multi-channel EEG data to a new spatial space such that the inter-class difference in the variances of different filtered EEG signals are maximized [47]. In the current study, the variances of the 10 most discriminative CSP features are employed.

2.5 Mental workload classification

Four classifiers are employed for the single-trail workload classification in the current study, including 2 popular ensemble learning algorithms: Gradient boosting and random forest (RF) classifiers, and 2 widely used classifiers in EEG workload classification: K nearest neighbor (KNN) and support vector machine (SVM) classifiers. The Scikit-learn package in Python [49] is used for the RF, SVM and KNN classifiers, and the XGBoost Python package [32] is used for the gradient boosting classifier.

For each subject, leave-one-out (LOO) cross-validation was used to evaluate the performance of the feature extraction methods and classifiers. The LOO method has been demonstrated to produce a relatively unbiased estimation of the true performance of the model and the generalizability of the machine learning model to new testing data [50-51]. In each iteration of the LOO algorithm, one of the trials was treated as the testing sample, whereas the remaining trials were used as the training set; the percentage of correctly predicted trials were used as the classification accuracy for the particular subject. In order to avoid the curse of dimensionality when using the spectral power and DWT as features, feature selection was performed within each iteration of the LOO process by preserving the features corresponding to the top 30 F-values, which were derived from ANOVA analysis between each individual features and the class labels. In addition, in order to identify the specific power and DWT features that are important for single-trial workload discrimination across subjects, one additional feature selection was performed for every subject after the classification analysis with the ANOVA technique, and those features that were selected for at least 6 subjects were identified and presented.

In order to facilitate unbiased comparisons of the performances of different classifiers, for each type of features as well as each combination of features, the free parameters of the classifiers are tuned through grid search and the classification accuracies obtained using the best-performing parameters are presented in the following sections.

2.5.1　Gradient boosting

The gradient boosting algorithm iteratively trains a number of weak decision-tree classifiers to gradually optimize the loss function[30] and has achieved exceptional performance in many real-world classification problems[32]. Therefore, in this study, we have explored whether this algorithm can boost the results of EEG single-trial workload classification. We have adopted the eXtreme Gradient Boosting XGBoost [32] implementation of the algorithm, which provides various advantages such as fast training speed and superior performance. In order to avoid excessive computational cost, the following parameters are selected based on prior knowledge and fixed: max_*depth* = 2, *gamma* = 0, max_*delta_step* = 0, *subsample* = 0.8, *colsample_bytree* = 0.4, *alpha* = 0.5, *lambda* = 0; whereas the number of trees is chosen in the following set: {50, 100, 150, 200} and the learning rate (eta) is selected in {0.01, 0.05, 0.1}.

2.5.2　Random forest

A random forest classifier consists of a number of decision-tree classifiers, each of trees is trained on a bootstrapped subsample of the entire training dataset and their predictions are averaged to produce the final prediction[31]. RF has been previously employed in various EEG studies and demonstrated to be effective in discriminating different mental states [33-35]. In this work, we have assessed the capability of RF for classifying mental workload using single-trial EEG signals. The number of trees in the RF algorithm is tuned by applying the following values: {50, 100, 150, 200}.

2.5.3　SVM and K-nearest neighbor classifiers

Both support vector machine (SVM) and K-nearest neighbor (KNN) classifiers have been widely used in EEG studies for mental workload classification [26-29]. For SVM, the radial basis function (rbf) kernel is used and the free parameters, C and gamma, are both grid-searched using the following values: $\{e^{-6}, e^{-4}, e^{-2}, e^{2}, e^{4}, e^{6}\}$. For KNN, the number of nearest neighbors (K) used in the algorithm are selected from the following set of values: {3, 5, 10, 20}.

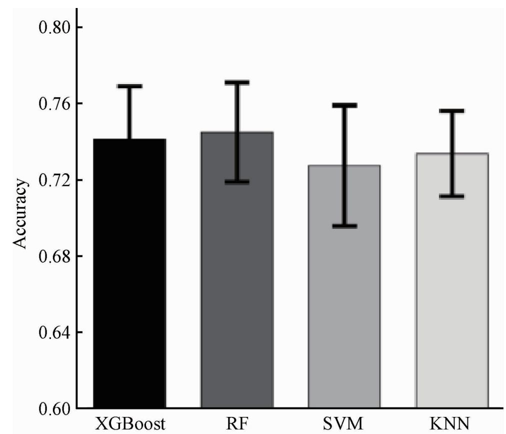
3　Results

3.1　Behavioral results

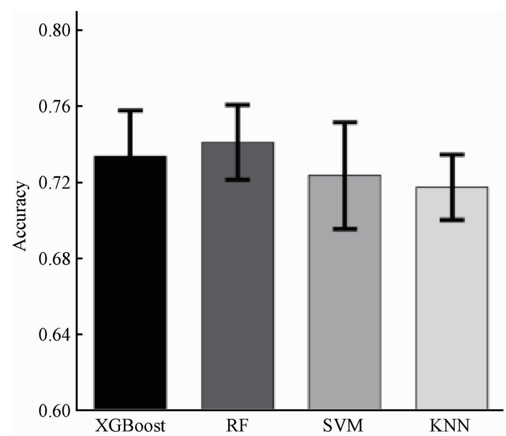
The accuracy of the 1-back task was (97.0 ± 2.7)% (mean ± S.D.), whereas the accuracy of the 3-back task was (45.5 ± 14.3)%. The significant difference ($p < 0.001$) in accuracy indicates the presence of differences in cognitive workload between these two tasks.

3.2　Workload classification

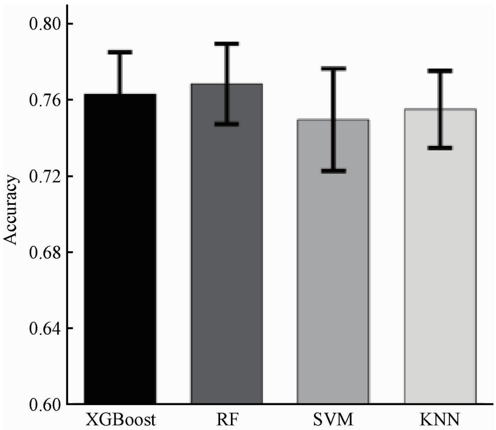
The classification accuracies of different machine learning algorithms using each of the individual feature extraction methods are shown in Fig. 2. As shown in the figure, all classifiers achieved satisfactory accuracies (higher than 70%) for all types of features; the spectral power and DWT methods produced closely comparable results, whereas the CSP technique led to significantly better accuracies; regardless of the type of features, the best classification performance was achieved by the RF classifier, followed by XGBoost.



(a) Classification accuracy using DWT features
XGBoost:learning rate=0.05;number of trees =100;
RF:number of trees=200;SVM:C=e^{-2},gamma=e^{-4};
KNN:number of neighbors=20



(b) Classification accuracy using Power features
XGBoost:learning rate=0.01;number of trees =150;
RF:number of trees=150;SVM:C=e^{-2},gamma=e^{-4};
KNN:number of neighbors=10



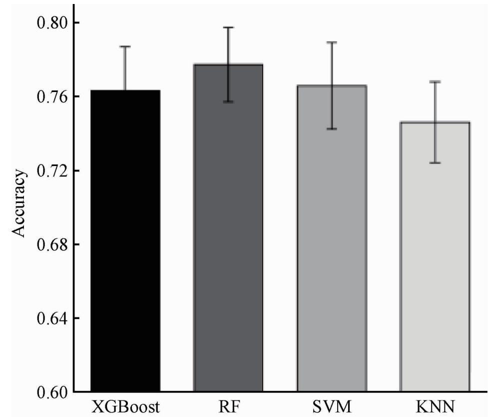
(c) Classification accuracy using CSP features
XGBoost:learning rate=0.01;number of trees =200;
RF:number of trees=100;SVM:C=e^{-2},gamma=e^{-6};
KNN:number of neighbors=5

Fig. 2 Average classification accuracy using each of the individual feature extraction methods with different classifiers and optimal hyper parameters

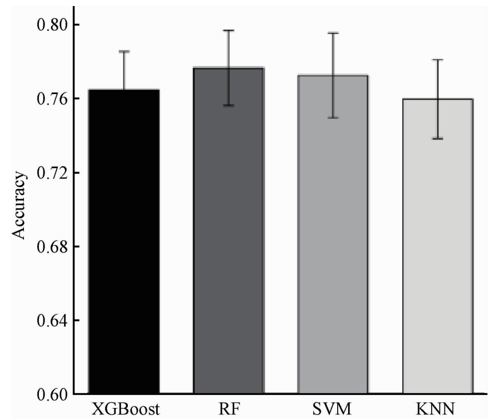
Table 1 and Fig. 3 display the performance of classifiers when different combinations of the feature extraction methods were used. Due to the comparable performance of the power and DWT features, only the combinations of each of these two methods with the CSP technique, as well as the combination of all three types of features, were explored. As shown in the figure, all 3 types of feature combinations improved the accuracies compared with the single-feature classification (Fig. 2); similar to Fig. 2, the best performance was achieved by RF, although in the combination of all 3 types of features, SVM produced closely comparable result; for the best-performing model (RF), all 3 methods of feature combination yielded similar results. Classification accuracy using the combination of different feature extraction methods with different classifiers (corresponding to Fig. 3).

Table 1 The values in the table represent mean (standard error of mean)

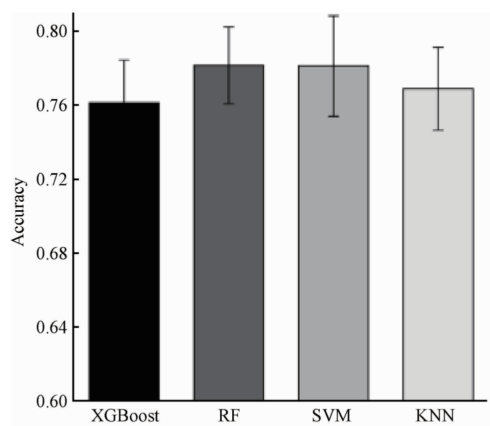
Features	XGBoost	RF	SVM	KNN
DWT + CSP	76.3 (2.3%)	77.7 (2.0%)	76.6 (2.3%)	74.6 (2.1%)
Power + CSP	76.2 (2.0%)	77.4 (2.0%)	77.0 (2.2%)	75.8 (2.0%)
DWT + Power + CSP	75.9 (2.2%)	77.8 (2.0%)	77.8 (2.6%)	76.6 (2.1%)



(a) Classification accuracy using DWT+CSP features
XGBoost:learning rate=0.01;number of trees =50;
RF:number of trees=150;SVM:C=e^{2},gamma=e^{-6};
KNN:number of neighbors=10



(b) Classification accuracy using Power+CSP features
XGBoost:learning rate=0.05;number of trees =200;
RF:number of trees=150;SVM:C=e^{2},gamma=e^{-6};
KNN:number of neighbors=5



(c) Classification accuracy using DWT+Power+CSP features
XGBoost:learning rate=0.01;number of trees =200;
RF:number of trees=50;SVM:C=e^{2},gamma=e^{-6};
KNN:number of neighbors=10

Fig. 3 Average classification accuracy using the combination of different feature extraction methods with different classifiers and optimal hyper parameters

Fig. 4 shows the power and DWT features that were selected for at least 6 subjects, which might serve as a guideline for the selection of a subset of features for single-trial workload classification. As shown in the figure, for the spectral power features, 1 channel was discriminative in all 3 frequency bands (PO5); 3 channels were identified in both alpha and beta bands (C2, C5 and TP8); in addition, 2 theta-specific, 2 alpha-specific and 4 beta-specific channels were found to possess consistently strong discriminative power. Regarding the DWT features, no theta-band discriminative features were discovered; 1 channel was identified in both alpha and beta bands; above all, a large number of channels (13 out of 14), predominantly in the frontal regions, were pinpointed in the beta band.

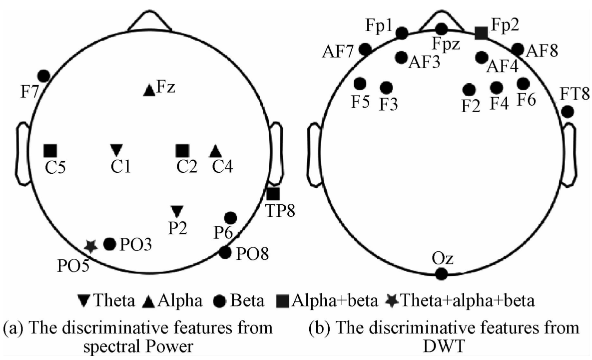


Fig. 4　The discriminative features from different frequency bands are highlighted by different shapes

4　Discussion

In this study, we performed individual-based EEG single-trial workload classification in order to discriminate between 2 working memory tasks with different workload levels. We assessed the discriminative power of 3 types of features as well as their combinations. The results suggested that comparable accuracies were produced by the features derived from spectral power and DWT, both of which were significantly outperformed by the CSP features; combining either spectral power or DWT with the CSP features further improved the classification results. In addition, we also evaluated the effectiveness of 4 different machine learning algorithms in mental workload classification based on single-trial EEG data and discovered that random forest classifier achieved the best classification accuracy regardless of the type of features.

The comparison of the performance of different types of features might suggest that despite of the widespread usage of spectral power and DWT[3, 23, 40-44] in EEG workload classification, the CSP technique might be superior in extracting informative and discriminative features for workload classification based on single-trial EEG, which is consistent with observations in BCI research [15]. The fusion of different types of features have been found to be an effective method for improving the performance of EEG mental workload classification[3, 9]. Our findings further corroborate this notion by revealing that the combination of either spectral power or DWT features with the CSP features can significantly improve the accuracy of single-trial mental workload classification.

Random forest, which has been proved to be effective in EEG mental workload classification [33-35], was found in this work to be the best-performing classifier in terms of single-trial EEG workload classification compared with the other widely used classifiers: SVM and KNN, as well as another popular ensemble learning algorithm: gradient boosting. Moreover, the exceptional performance of RF seemed robust against the type of features used in the model. Of note, gradient boosting classifier, which has gained success in various real-world applications, did not produce superior results compared with other algorithms, which might result from the small sample size for each individual or the non-exhaustive parameter tuning.

Discriminative spectral power features are identified in all frequency bands and resided in a wide range of brain regions, which is consistent with previous findings in EEG workload classification based on single trials [3]. The DWT features with consistently discriminative power are predominantly in the beta band, which has been found to reflect the cognitive demand [20]. Moreover, in agreement with the findings in previous EEG studies of cognitive workload that utilized DWT as a feature extraction method[51], the channels possessing high discriminative power are mostly located in the frontal brain regions.

5　Conclusion

To conclude, after comparing different feature extraction methods and classification algorithms, we found that the best classification result was achieved by 1) combining the spectral power or DWT features with CSP technique and 2) the random forest classifier, which might serve as a guideline for mental workload classification based on single-trial EEG data for future works.

Reference

[1] YOUNG M S, BROOKHUIS K A, WICKENS C D, et al. State of science: mental workload in ergonomics[J]. Ergonomics, 58(1): 1-17.

[2] KANTOWITZ B H. Defining and measuring pilot mental workload [C]. NASA Langley Research Center, Mental-State Estimation, 1988: 179 - 188.

[3] BROUWER A M, HOGERVORST M A, VAN ERP J B, et al. Estimating workload using EEG spectral power and ERPs in the *n*-back task [J]. Journal of Neural Engineering, 2012, 9(4): 045008.

[4] COLLET C, AVERTY P, DITTMAR A. Autonomic nervous system and subjective ratings of strain in air-traffic control [J]. Applied Ergonomics, 2009, 40(1): 23-32.

[5] ZIJLSTRA F R H. Efficiency in work behaviour: A design approach for modern tools[D]. TU Delft: Delft University of Technology, 1993.

[6] WILSON G F, EGGEMEIER F T. Psychophysiological assessment of workload in multi-task environments[M]. Damos: Multiple-task performance, 1991.

[7] GAI S, LIU J, LIU X, et al. Study on the design, preparation and test of a novel dry electrodefor long-term EEG measurement [J]. Chinese Journal of Scientific Instrument, 2016, 37(1): 144-151.

[8] YU K, PRASAD I, MIR H, et al. Cognitive workload modulation through degraded visual stimuli: A single-trial EEG study [J]. Journal of Neural Engineering, 2015, 12(4): 046020.

[9] BASHIVAN P, YEASIN M, BIDELMAN G M. Single trial prediction of normal and excessive cognitive load through EEG feature fusion [C]. IEEE Signal Processing in Medicine and Biology Symposium,2016:1-5.

[10] BERKA C, LEVENDOWSKI D J, LUMICAO M N, et al. EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks [J]. Aciation, Space, and Environmental Medicine, 2007, 78(Supp. 1): B231-B244.

[11] HOGERVORST M A, BROUWER A M, VAN ERP J B. Combining and comparing EEG, peripheral physiology and eye-related measures for the assessment of mental workload [J]. Frontiers in Neuroscience, 2015, 8(5): 322.

[12] BADDELEY A. Working memory: Theories, models, and controversies [J]. Annual Review of Psychology, 2012, 63(1): 1-29.

[13] BADDELEY A. Working memory: Looking back and looking forward [J]. Nature Reviews Neuroscience, 2003, 4(10): 829-839.

[14] SWELLER J. Cognitive load during problem solving: Effects on learning [J]. Cognitive Science, 1988, 12(2): 257-285.

[15] MÜLLER K R, TANGERMANN M, DORNHEGE G, et al. Machine learning for real-time single-trial EEG-analysis: from brain - computer interfacing to mental state monitoring [J]. Journal of Neuroscience Methods, 2008, 167(1): 82-90.

[16] GRUNWALD M, WEISS T, MUELLER S, et al. EEG changes caused by spontaneous facial self-touch may represent emotion regulating processes and working memory maintenance [J]. Brain Research, 2014, 1557(15): 111-126.

[17] LANGER N, VON BASTIAN C C, WIRZ H, et al. The effects of working memory training on functional brain network efficiency [J]. Cortex, 2013, 49 (9): 2424-2438.

[18] SAUSENG P, KLIMESCH W, SCHABUS M, et al. Fronto-parietal EEG coherence in theta and upper alpha reflect central executive functions of working memory [J]. International Journal of Psychophysiology, 2005, 57(2): 97-103.

[19] GEVINS A, MCEVOY L K, SMITH M E, et al. Long-term and within-day variability of working memory performance and EEG in individuals [J]. Clinical Neurophysiology, 2012, 123(7): 1291-1299.

[20] ROUX F, WIBRAL M, MOHR H M, et al. Gamma-band activity in human prefrontal cortex codes for the number of relevant items maintained in working memory [J]. Journal of Neuroscience, 2012, 32(36): 12411-12420.

[21] ROWLAND N, MEILE M, NICOLAIDIS S. EEG alpha activity reflects attentional demands, and beta activity reflects emotional and cognitive processes [J]. Science, 1985, 228(4700): 750-752.

[22] FERNÁNDEZ T, HARMONY T, RODRÍGUEZ M, et al. EEG activation patterns during the performance of tasks involving different components of mental calculation [J]. Electroencephalography and Clinical Neurophysiology, 1995, 94(3): 175-182.

[23] MURATA A. An attempt to evaluate mental workload using wavelet transform of EEG [J]. Human Factors, 2005, 47(3): 498-508.

[24] CHANDRA S, VERMA K L, SHARMA G, et al. EEG based cognitive workload classification during NASA MATB-II multitasking [J]. International Journal of Cognitive Research in Science: Engineering and

Education, 2015, 3(1): 35-41.

[25] AMIN H U, MALIK A S, AHMAD R F, et al. Feature extraction and classification for EEG signals using wavelet transform and machine learning techniques [J]. Australasian Physical and Engineering Science in Medicine, 2015,38(1): 139-149.

[26] YAZDANI A, EBRAHIMI T, HOFFMANN U. Classification of EEG signals using Dempster Shafer theory and a k-nearest neighbor classifier [C]. 4th International IEEE/EMBS Conference on Neural Engineering, 2009: 327-330.

[27] HEGER D, PUTZE F, SCHULTZ T. Online workload recognition from EEG data during cognitive tests and human-machine interaction [C]. Annual Conference on Artificial Intelligence, 2010: 410-417.

[28] PUTZE F, JARVIS J P, SCHULTZ T. Multimodal recognition of cognitive workload for multitasking in the car [C]. 20th International Conference on Pattern Recognition, 2010: 3748-3751.

[29] LOTTE F, CONGEDO M, LECUYER A, et al. A review of classification algorithms for EEG-based brain-computer interfaces [J]. Journal of Neural Engineering, 2007, 4(2): R1.

[30] FRIEDMAN J H. Greedy function approximation: A gradient boosting machine [J]. Annals of Statistics, 200129(5): 1189-1232.

[31] BREIMAN L. Random forests [J]. Machine Learning, 2001, 45(1): 5-32.

[32] CHEN T, GUESTRIN C. Xgboost: A scalable tree boosting system [C]. ACM Sigkdd International Conference, 2016:785-794.

[33] NOVAK D, BEYELER B, OMLIN X, et al. Workload estimation in physical human-robot interaction using physiological measurements [J]. Interacting with Computer, 2015, 27(6): 616-629.

[34] FRAIWAN L, LWEESY K, KHASAWNEH N, et al. Automated sleep stage identification system based on time-frequency analysis of a single EEG channel and random forest classifier [J]. Computer Methods & Programs in Biomedicine, 2012, 108(1): 10-19.

[35] LEHMANN C, KOENIG T, JELIC V, et al. Application and comparison of classification algorithms for recognition of Alzheimer′s disease in electrical brain activity (EEG) [J]. Journal of Neuroscience Methods, 2007, 161(2): 342-350.

[36] JAMES C J, HESSE C W. Independent component analysis for biomedical signals [J]. Physiological Measurements, 2005, 26(1): R15-R39.

[37] JUNG T P, MAKEIG S, HUMPHRIES C, et al. Removing electroencephalographic artifacts by blind source separation [J]. Psychophysiology, 2000, 37(2): 163-178.

[38] PALMER J A, MAKEIG S, KREUTZ-DELGADO K, et al. Newton method for the ICA mixture model [C]. IEEE International Confernece on Acoustics, Speech and Signal Processing, 2008:1805-1808.

[39] DELORME A, MAKEIG S. EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis [J]. Journal of Neuroscience Methods, 2004, 134(1): 9-21.

[40] KATHNER I, WRIESSNEGGER S C, MULLER-PUTZ G R, et al. Effects of mental workload and fatigue on the P300, alpha and theta band power during operation of an ERP (P300) brain-computer interface [J]. Biological Psychology, 2014, 102(5): 118-129.

[41] BORGHINI G, ASTOLFI L, VECCHIATO G, et al. Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness [J]. Neuroscience & Behavioral Review, 2014, 44(10): 58-75.

[42] OCAK H. Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy [J]. Expert Systems with Applications, 2009, 36(2): 2027-2036.

[43] ADELI H, ZHOU Z, DADMEHR N. Analysis of EEG records in an epileptic patient using wavelet transform[J]. Journal of Neuroscience Methods, 2003, 123(1): 69-87.

[44] ÜBEYLI E D. Combined neural network model employing wavelet coefficients for EEG signals classification [J]. Digital Signal Processing, 2009, 19(2): 297-308.

[45] MALLAT S G. A theory for multiresolution signal decomposition: The wavelet representation [J]. IEEE Transactions on Pattern Analysis, 1989, 11 (7): 674-693.

[46] SUBASI A. EEG signal classification using wavelet feature extraction and a mixture of expert model [J]. Expert Systems with Applications, 2007, 32 (4): 1084-1093.

[47] BLANKERTZ B, TOMIOKA R, LEMM S, et al. Optimizing spatial filters for robust EEG single-trial analysis [J]. IEEE Signal Processing Magazine, 2008, 25(1): 41-56.

[48] DIJKSTERHUIS C, DE WAARD D, BROOKHUIS K A, et al. Classifying visuomotor workload in a driving simulator using subject specific spatial brain

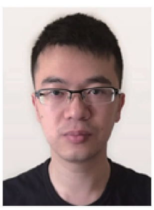
patterns [J]. Frontierns in Neuroscience, 2013, 7(8): 149.

[49] PEDREGOSA F, VAROQUAUX G, GRAMFORT A, MICHEL V, THIRION B, GRISEL O, et al. Scikit-learn: Machine learning in Python [J]. Journal of Machine Learning Research, 2011, 12 (10): 2825-2830.

[50] TEMKO A, THOMAS E, MARNANE W, et al. EEG-based neonatal seizure detection with support vector machines [J]. Clinical Neurophysiology, 2011, 122(3): 464-473.

[51] VAPNIK V N, KOTZ S. Estimation of dependences based on empirical data [M]. New York: Springer-Verlag, 1982.

Author Biography



Dai Zhongxiang received his B. Sc. degree from National University of Singapore in 2015. Now he is a research assistant in the Singapore Institute for Neurotechnology (SINAPSE). His main research interest is machine learning.

E-mail: daiz9109@gmail.com

Bezerianos Anastasios obtained from the University of Patras received his B. Sc. degree, M. Sc. degree of telecommunication from Athens University, and Ph. D. degree from the University of Patras in 1987. Currently he is a Professor at the Medical School of Patras University, Greece, and a Visiting Research Professor at the Department of Electrical and Computer Engineering at National University of Singapore, Singapore. His main research interests at large are in neuroengineering and systems medicine and bioinformatics.

E-mail: tassos.bezerianos@nus.edu.sg

Chen Shen-Hsing Annabel received her B. Sc. degree of Social Sciences with Honors from National University in 1992, and her M. Sc. and Ph. D. degrees both from Indiana University-Purdue University, Indianappolis in 1995 and 1999, respectively. She is a faculty member in School of Humanities and Social Sciences since 2008 in Nanyang Technological University and has a joint appointment at LKC Medicine since 2015. Her main research interests include investigate underlying neural substrates involved in higher cognition in the cerebellum, as well as changes in cognitive processes in healthy aging and dementia through the application of neuroimaging techniques.

E-mail: AnnabelChen@ntu.edu.sg

Sun Yu (Corresponding author) received his B. Sc. degree of Engineering from Nanjing University of Aeronautics and Astronautics in 2005, and his Ph. D. degree in Biomedical Engineering from Loughborough University, UK, in 2011, and a dual-Ph. D. degree in Shanghai Jiao Tong University in 2012. He is currently a Senior Research Fellow at the Singapore Institute for Neurotechnology (SINAPSE). His main research interest include the field of neuroengineering, multimodal neuroimaging processing with particular focus of brain connectome.

E-mail: lsisu@nus.edu.sg