Neurocomputing 347 (2019) 212–229

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Physiological-signal-based mental workload estimation via transfer dynamical autoencoders in a deep learning framework

Zhong Yin[a,*], Mengyuan Zhao[b], Wei Zhang[a], Yongxiong Wang[a], Yagang Wang[a,*], Jianhua Zhang[c]

[a] Engineering Research Center of Optical Instrument and System, Ministry of Education, Shanghai Key Lab of Modern Optical System, University of Shanghai for Science and Technology, Shanghai 200093, PR China
[b] School of Marxism, University of Shanghai for Science and Technology, Shanghai 200093, PR China
[c] OsloMet Artificial Intelligence Lab, Department of Computer Science, Oslo Metropolitan University, Oslo N-0130, Norway

ARTICLE INFO

Article history:
Received 25 June 2018
Revised 14 January 2019
Accepted 19 February 2019
Available online 11 April 2019

Communicated by Dr Zhu Yu

Keywords:
Mental workload
Electroencephalograph
Deep learning
Stacked autoencoder
Operator functional state

ABSTRACT

Evaluating operator mental workload (MW) in human-machine systems via neurophysiological signals is crucial for preventing unpredicted operator performance degradation. However, the feature of physiological signals is associated with the historical values at the previous time steps and its statistical properties vary across individuals and types of mental tasks.

In this study, we propose a new transfer dynamical autoencoder (TDAE) to capture the dynamical properties of electroencephalograph (EEG) features and the individual differences. The TDAE consists of three consecutively-connected modules, which are termed as feature filter, abstraction filter, and transferred MW classifier. The feature and abstraction filters introduce dynamical deep network to abstract the EEG features across adjacent time steps to salient MW indicators. Transferred MW classifier exploits large volume EEG data from an source-domain EEG database recorded under emotional stimuli to improve the model training stability. We tested our algorithms on two target EEG databases. The classification performance shows TDAE significantly outperforms existing shallow and deep MW classification models. We also investigated how to select TDAE hyper-parameters and found its superiority in accuracy can be achieved with proper filter orders.

© 2019 Elsevier B.V. All rights reserved.

1. Introduction

The increased complexity of the control and automation systems served to generate attention in understanding the locus of the human-machine interaction. Unlike the unmanned, fully-automated agents, the task performance in a human-machine collaboration environment is closely linked to the functional states of human operators [1]. To meet the task demands, human cognitive resources are partially occupied from the overall capacity accounting for information processing functionality of the brain [2]. When assigning improper amount or type of tasks to the operator, the high mental workload (MW) may arise leading to the potential degradation of operator performance. Assessing the MW levels is thus considered highly crucial aiming at reducing the possibility of accidents caused by human factors.

Considering the fact that electrical neurophysiological signals can reflect the summation of spontaneous activity of the human cortex, MW states were recognized via electroencephalograph (EEG) in well-documented works [3–5]. However, building a generic EEG model is problematic since EEG signals are known as stochastic processes with nonstationarity and individual difference. By employing the advantages of the machine learning principles, interpreting EEG into different cognitive states is obtained by various data-driven modeling techniques [4–10]. For instance, the multi-layer perceptron [4], support vector machine (SVM) [5], extreme learning machine (ELM) [6], random forest [7], stacked autoencoder (SAE) [8], deep belief networks (DBN) [9] and multi-layer convolutional neural network (CNN) [10]. Among them, the deep learning classifiers received high attention in the newest works due to its higher classification accuracy benefiting from the hierarchical feature abstractors [8–10].

In our primary investigations on deep learning approaches [11–12], feature vectors are computed from non-overlapped EEG segments and the current MW output is solely determined via the feature vector of the current time step. However, inherent cortical states are known as continuous-time process. The current MW level is also associated with the historical feature vectors

* Corresponding authors.
E-mail addresses: yinzhong@usst.edu.cn (Z. Yin), ygwang@usst.edu.cn (Y. Wang).

https://doi.org/10.1016/j.neucom.2019.02.061
0925-2312/© 2019 Elsevier B.V. All rights reserved.

Nomenclature

Acronym	*Description*
AE	Autoencoder
ANN	Single-hidden-layer feedfoward neural network
AutoCAMS	Automation-enhanced cabin air management system
CNN	Convolutional neural network
DBN	Deep belief network
DEAP	Database for emotion analysis using physiological signals
DCE	Deep collaborative embedding
ECG	Electrocardiogram
EEG	Electroencephalograph
EOG	Electrooculogram
ELM	Extreme learning machine
KNN	*K*-nearest neighbor
LSSVM	Least square support vector machine
LR	Logistic regression
MATB	Multiple-Attribute Task Battery
MW	Mental workload
NB	Naive Bayesian model
NPV	Negative predicting value
PSD	Power spectral density
PLS	Partial least square
PCA	Principal component analysis
RBF	Radial basis function
RSSL	Robust structured subspace learning
SD	Source domain dataset
SAE	Stacked autoencoder
SVM	Support vector machine
TD	Target domain dataset
TDAE	Transfer dynamical autoencoder

computed from the EEG recordings at the previous time steps. Moreover, we found the deep learning models possess outstanding generalization capability only if the huge amount psychophysiological data are available for training [13]. Since the statistical properties of the physiological signals vary across individuals and types of mental tasks [14–15], the EEG features extracted from different domains cannot be directly fused as the training database for a specific MW-assessing environment. The reason behind is that most machine learning methods follow the assumption that the training and testing instances are independently drawn from identical distribution. It then leads to a high burden in collecting personal EEG data for each operator.

To overcome the two shortcomings mentioned above, the primary objective of the present study is to introduce dynamical deep network to abstract the EEG features across adjacent time steps to salient MW indicators. Specifically, we compress the EEG features via a two-layer SAE where the input neurons are linked to both of the current and historical feature vectors corresponding to multiple recording time steps. Then, the pre-trained SAE can be considered as a feature filter aiming at finding the dynamical properties for ongoing EEG processes. To validate the effectiveness of the feature abstractor, two EEG databases collected in our previous works are employed, where the operators were engaged with a simulated process control system [11,16].

The secondary objective is to employ transfer learning principle to improve fine-tuning performance in the deep learning framework. Transfer learning emphasizes sharing knowledge from a source domain to a target domain [17]. When achieving domain-adaptation across multiple training data sources, it could facilitate MW classifier to be sufficiently trained via EEG feature sets extracted from a different but similar task environment. Here, we adopt the EEG recordings of DEAP, the Database for Emotion Analysis using Physiological signals, that was originally utilized for emotion recognition [18]. We then designed two dependent SAE networks for knowledge transferring by comparing the similarity of the high-level EEG feature abstractions between the target MW databases and source DEAP database for consistent data distribution. Finally, a new transfer dynamical autoencoders (TDAE) in the deep learning framework is proposed by successively linking the filtering SAEs and the transferred SAE classifier.

The rest of the paper is structured as follows. The related works are briefly reviewed in Section 2. In Section 3, we first introduce all adopted EEG databases and how we prepare the consistent EEG features. Then, details of the proposed TDAE MW classifier are presented. The MW classification performance and the accuracy comparisons are shown in Section 4. In Section 5, we provide some useful discussions on the present algorithms. The contribution of the study is concluded in Section 6.

2. Related works

The ultimate goal for MW assessment is to achieve adaptive systems that can automatically allocate the tasks between human and machine agents. Byrne and Parasuraman then proposed the concept of *adaptive automation* [19]. The type of human-machine collaboration can be defined via the *level of automation* according to Kaber et al., which is intended to design a human-centered working environment [20]. Lin et al. applied Kaber's framework on a simulated control room of the nuclear power plant to enhance the operator performance [21]. Parallel work on air traffic control task can be noted in [22]. In the literature, MW states were measured via the secondary task performance. For instance, Parasuraman et al. used the reaction time of radio communications in the Multiple-Attribute Task Battery (MATB) platform to assess the MW level for the flight crew members [23].

Considering the fact that introducing secondary task may lead to a complex dual-task mode and impairs the main task performance, physiological measure received more attention in the past a few years because of its superiority in nonintrusiveness, temporal resolution and diagnostic capability. In early works, Pope et al. proposed the engagement index for assessing MW, which is computed by the ratio of the EEG power spectral densities [24]. Freeman et al. investigated four different EEG power ratios by statistical analysis [25]. Besides the EEG feature, the electrocardiograms (ECG) were discussed in [26] by Haarmann et al.. Then, the middle- and high-frequency components (0.07–0.5 Hz) of the heart rate variability, i.e., the fast Fourier transformation of the ECG signal, were used to measure MW in [27] by Hoover et al. Moreover, the blink rate captured from the electrooculogram (EOG) was adopted by Noel et al. to estimate the MW level of the trainer pilot [28]. Compared to the ECG signal, EEG features are directly generated from the central nervous system and are closely related to the working memory management manipulated in the prefrontal cortex. The multichannel EEG recordings also possess better spatial resolution than physiological signals collected on the peripheral nervous system.

The implementation of the machine learning approaches as cognitive state classifiers were widely documented. The single-hidden-layer feedfoward neural network (ANN) was used by Russell et al. to classify MW levels [29]. Kumar et al. combined the fuzzy C-means clustering and the fuzzy maximum likelihood estimation to classify EEG into different sleep states [30]. Guler and Ubeyli fused the discrete wavelet transformation and the adaptive neuron fuzzy inference system as a multiclass EEG classifier [31]. In a newly reported work [32], the classical *k*-nearest neighbor method also showed high performance. In our previous works [16,33], we applied the least square support vector machine

(LSSVM) and the locally linear embedding to classify operator MW. Recently, Faust et al. reviewed the popular deep learning methods applied in this domain [34], where the DBN, SAE and CNN were listed as attractive deep models.

In particular, Wang et al. indicated the perfect MW classifier should accurately predict the MW levels under any given day, training subjects and human-machine tasks [14]. In their work, a cross-subject MW classifier based on the hierarchical Bayes model was proposed. Cross-day MW classifiers built by Christensen et al. with LSSVM, ANN and the linear discriminant analysis approaches can be found in [35]. Baldwin et al. reported their attempt to reach the cross-task classification [15]. Recently, a task-independent MW recognition framework using the SVM and the cortical connectivity features was reported by Dimitrakopoulos et al. [36]. In our previous works [11–12,37], we developed an adaptive SAE, switching DBN, and deep ELM model to tackle the cross-day, individual-independent and task-generic MW classification problems. From the above related works, we found dynamical properties of the data time courses and a limited size of training samples may impair the classification performance when applying deep learning methods for assessing mental workload via EEG features.

For the shallow feedforward ANN applied in [29], a single hidden layer can accurately approximate a mapping across input EEG features and class labels. However, the potential overfitting may arise when too many hidden neurons are predefined. The SVM implemented in [33] avoided this problem since it introduces a regularization term to minimize the norm of the weight vector in the decision hyper-plane that achieves max margin classification. In [37], the input weight of an ELM network is completely random while its output weight is computed by minimizing the least square error and the weight norm. Since only a part of weight parameters need to be tuned, the training speed of the ELM is much higher than ANN and SVM methods. It is noted the learning capability is limited for above shallow learning machines on the MW classification issue when raw EEG indicators were employed [12]. For the neural network structure in deep SAE, DBN, and CNN models, there are several intermediate layers for abstracting high level feature representation. It provides a wider range of possibilities for the feature fusion and improves the efficiency for the physiological feature engineering.

It is noted that latent subspace learning techniques have shown promising performance for automatic feature engineering. In particular, Li et al. [38] proposed the robust structured subspace learning (RSSL) method to capture complex geometric structures of image data. They designed a new objective function for feature mapping by evaluating the local and global structure consistency when minimizing the least square cost and the regularization term. Their results showed RSSL outperformed a wide range of manifold learning methods when they were combined with a rigid regression classifier. In a newly-reported work, Li et al. [39] proposed the deep collaborative embedding (DCE) algorithm aiming at finding the optimal latent feature representations for social images. The DCE integrated the collaborative factor analysis with the CNN training algorithm to generate an end-to-end classification framework. The statistical analysis indicated a significant improvement was achieved by the DCE. Both of the works above show the intermediate feature representation module is critically important for a pattern classification system.

Motivated by the results shown in [33–38], in this study we employ a dynamical deep learning architecture to uncover proper latent EEG variables targeting the mental workload levels. Across multiple hidden layers, the original EEG feature can be automatically fused to generate high level feature representations. Note that in [39] the outlier problem was tackled via introducing a new term in the objective function that evaluated the consistency between the local and global data structures. Here, we controlled the overfitting of deep model by applying the transfer learning principle to tackle the problem for the limited size of the training dataset. That is, multiple data sources from different domains are integrated to learn the proper weights. The adequate size of the training sample is collected and used for network fine-tuning. In the end, we characterize the dynamical properties for both of EEG features and abstractions in the proposed deep learning framework.

3. Materials and methods

3.1. Neurophysiological databases

In this section, we provide descriptions for three EEG databases involved in modeling the proposed deep MW classifier. Regarding the transfer learning principle, the first two databases are target domain sets, where subjects were engaged with process control tasks under different task demands. For the third database, however, EEG signals were recorded along with emotional stimuli and are defined as source domain database for knowledge transferring.

3.1.1. Target domain EEG sets

Two target domain EEG sets (denoted as TD1 and TD2 hereinafter) were collected in our previous works [11,33]. When subjects performed tasks on automation-enhanced cabin air management system (AutoCAMS), we simultaneously measured the EEG signals from the their scalps. Details of AutoCAMS software were noted in [40], where operators were instructed to manually perform a safety-critical task with multiple subsystems that control the air quality of a space cabin. Task demand of AutoCAMS was programmed via the number of failed subsystems and the actuator sensitivity. For TD1 [33] and TD2 [11], eight and six healthy participants were engaged, respectively. Each of them accomplished two sessions of experiments.

In TD1, four-phase EEG signals in each session were selected with each phase lasting 15 min. In phases #1 and #4, subjects interacted with AutoCAMS under low task demand while only one failed subsystem was to be manipulated. In phases #2 and #3, high task demand was activated via four failed subsystem. In TD2, we selected seven phases with each of 5 min from each session, where phases #1, #3, #5, and #7 were baseline conditions with zero-task-demand. In the remaining phases #2, #4, and #6, subjects operated 2, 3, and 4 failed subsystems under high actuator sensitivity, respectively. For facilitating supervised learning, we then set the EEG instances corresponding to low- and zero-task-demand phases as *low* MW class while the remaining phases as *high* MW class. The variation of the task demand for TD1 and TD2 was illustrated in Fig. 1. It is noted we removed transient phases to improve the data separability between low and high MW classes.

3.1.2. Source domain EEG set

Considering the data volume in TD1 and TD2 was relatively limited, we further employ DEAP as the source domain database (denoted as SD hereinafter). The DEAP was publicly accessible for investigating human physiological responses stimulated by music videos and details were noted the literature [18]. The SD was constructed by selecting EEG data from the DEAP that contains 32 healthy subjects. For each subject, 40-trial experiment was carried out. In each trial, 1-min EEG signals induced by the video stimulus were selected. That is, 40-min EEG data were available for each subject while the total data length was 1280-min for all 32 subjects.

3.2. Consistent EEG feature extraction

For SD, TD1 and TD2, 32, 11, and 15 channels of EEG were recorded according to 10–20 international system, respectively. To

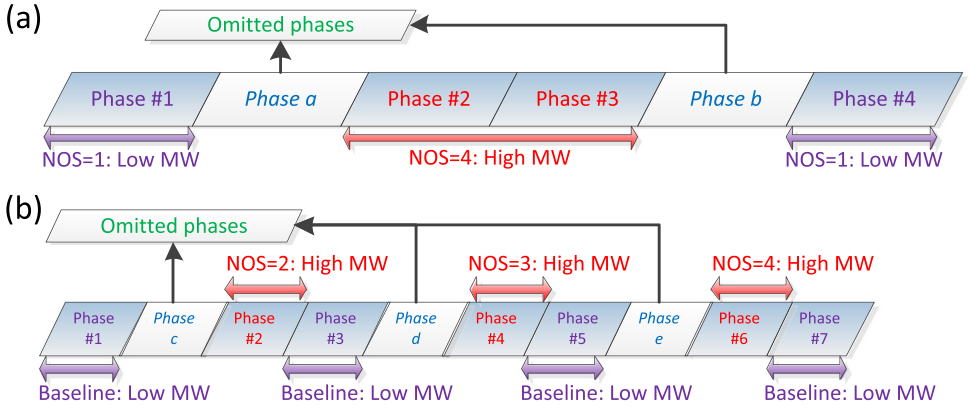

Fig. 1. Task demand variation for a session of experiment in (a) TD1 and (b) TD2. Number of the failed subsystem is denoted by NOS. Note that EEG data for transient phases *a*, *b*, *c*, *d*, and *e* were omitted. The details follow our previous work [11,33].

transfer knowledge from the source domain to the target domain, extracted EEG features from three databases must be exactly consistent possessing same dimensionality and mathematical representations. Therefore, we only choose 11 shared channels, i.e., F3, F4, Fz, C3, C4, Cz, P3, P4, Pz, O1, and O2. The EEG signals were filtered and preprocessed via independent component analysis following a framework proposed in [11]. For TD1 and TD2, EEG signals were sampled at 500 Hz and split to non-overlapped segments with each of 2-second-length. Since we removed the first 5 segments in each phase of TD2 due to muscular noise [16], 1800 and 1015 EEG segments were elicited for each session of TD1 and TD2, respectively. For the SD database, EEG signals were downsampled to 128 Hz. We split the signals in each trial into 30 non-overlapped segments with the invariable 2-second-length. In total, there were 30×40×32=38,400 segments of SD.

For each data segment, 137 EEG features were extracted. The EEG features with the index #1 to #44 were the average power spectral density (PSD) within frequency bands of theta (4–8 Hz), alpha (8–12 Hz), beta (12–30 Hz) and gamma (30–40 Hz) across 11 channels. That is, for each channel four power features were computed via fast Fourier transform. #45 to #60 features were PSD differences between right and left scalps on the channel pairs F3-F4, P3-P4, C3-C4, and O1-O2. For each channel pair, four difference features from the same frequency bands were employed. The remaining 77 features were the mean, variation, zero crossing rate, Shannon entropy, spectral entropy, kurtosis, and skewness across 11 channels.

In Fig. 2(a), we visualize the extracted features by displaying the Spearman correlation coefficient between EEG features and target MW levels. Each value in the line plot is the coefficient averaged in all sessions and subjects in each database between the feature time course and the corresponding MW levels. The high correlation was discovered for both of the features in frequency and time domains, e.g., PSD features, variation, zero crossing rate, and Shannon entropy. In Fig. 2(b) and (c), 3D-scatterplots were shown via EEG features #1-#3 and #35-#37, respectively. We selected 20 instances of each MW class in TD1 (Fig. 2(b)), TD2 (Fig. 2(c)) and 40 instances from SD. Both of the subfigures indicate two distinguishable clusters for low and high MW levels. Such inter-class margin insures the session-specific classification performance. It is important to note that the cluster of the unlabeled instances from SD was located near the manifold constructed by the target domain databases, which facilitate the knowledge transferring from SD to TD1 and TD2.

3.3. Deep transfer dynamical autoencoder for MW classification

A deep-SAE-based MW classifier hypothesizes that the EEG features can be hierarchically represented as high-level abstractions characterizing inherent cortex states for distinguishing MW levels. This section reviews the architecture of the classical SAE and introduces the formularized TDAE framework for binary MW classification.

3.3.1. Stacked autoencoder

The architecture of a SAE-based deep learning model is generated by stacking multiple autoencoders [41]. An autoencoder is equivalent to a single-hidden-layer feed-forward neural network with identical inputs and outputs. Its architecture is shown in Fig. 3(a). Given a training set of N EEG feature vectors,

$$S_{tr} = \{(\mathbf{x}_k, \mathbf{y}_k) | \mathbf{x}_k \in R^m, \mathbf{y}_k \in R^2, k = 1, 2, ..., N\}, \tag{1}$$

the function signal for m-dimensional EEG instance $\mathbf{x}_i$ propagates in the feed-forward path and leads to hidden neuron activations $\mathbf{g}$. For ith hidden neuron, the ith element of $\mathbf{g}$ is computed as,

$$g_i = f(u_i) = f\left(\sum_{j=1}^{m} w_{ij}x_j + b_i\right), i = 1, 2, ..., \tilde{m}, \tag{2}$$

where u_i, w_{ij}, b, and $\tilde{m}$ denote the induced local field, synaptic weight, bias, and number of hidden neurons, respectively. In this study, the activation function $f(\cdot)$ employs the logistic sigmoid function $f(u) = 1/(1+e^{-au})$ with $a = 1$.

In the reconstruction layer, the identical features $\tilde{\mathbf{x}} = [\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_m]^T$ are approximated via,

$$\tilde{x}_i = f(\tilde{u}_i) = f\left(\sum_{j=1}^{\tilde{m}} \tilde{w}_{ij}g_j + \tilde{b}_i\right), i = 1, 2, ..., m. \tag{3}$$

The error signal in the feedback path employs the squared error function ε,

$$\varepsilon_i(x_i, \tilde{x}_i) = \frac{1}{2}(x_i - \tilde{x}_i)^2. \tag{4}$$

The correction of output weights $\Delta\tilde{w}_{ij}$ can be computed through Back-Propagation (BP) algorithm via

$$\frac{\partial \varepsilon_i}{\partial \tilde{w}_{ij}} = \frac{\partial \varepsilon_i}{\partial \tilde{x}_i} \cdot \frac{\partial \tilde{x}_i}{\partial \tilde{u}_i} \cdot \frac{\partial \tilde{x}_i}{\partial \tilde{w}_{ij}}. \tag{5}$$

For input weights, Δw_{ij} is derived by summing the feedback ε_i of all nodes in the reconstruction layer, $1/2\sum_{i=1}^{m}(x_i - \tilde{x}_i)^2$. Note that b_i and $\tilde{b}_i$ are rewritten as w_{i0} and $\tilde{w}_{i0}$ when applying Eq. (5). Then, the optimal parameters $\theta = \{\mathbf{W}^*, \tilde{\mathbf{W}}^*, \mathbf{b}^*, \tilde{\mathbf{b}}^*\}$ are learned as,

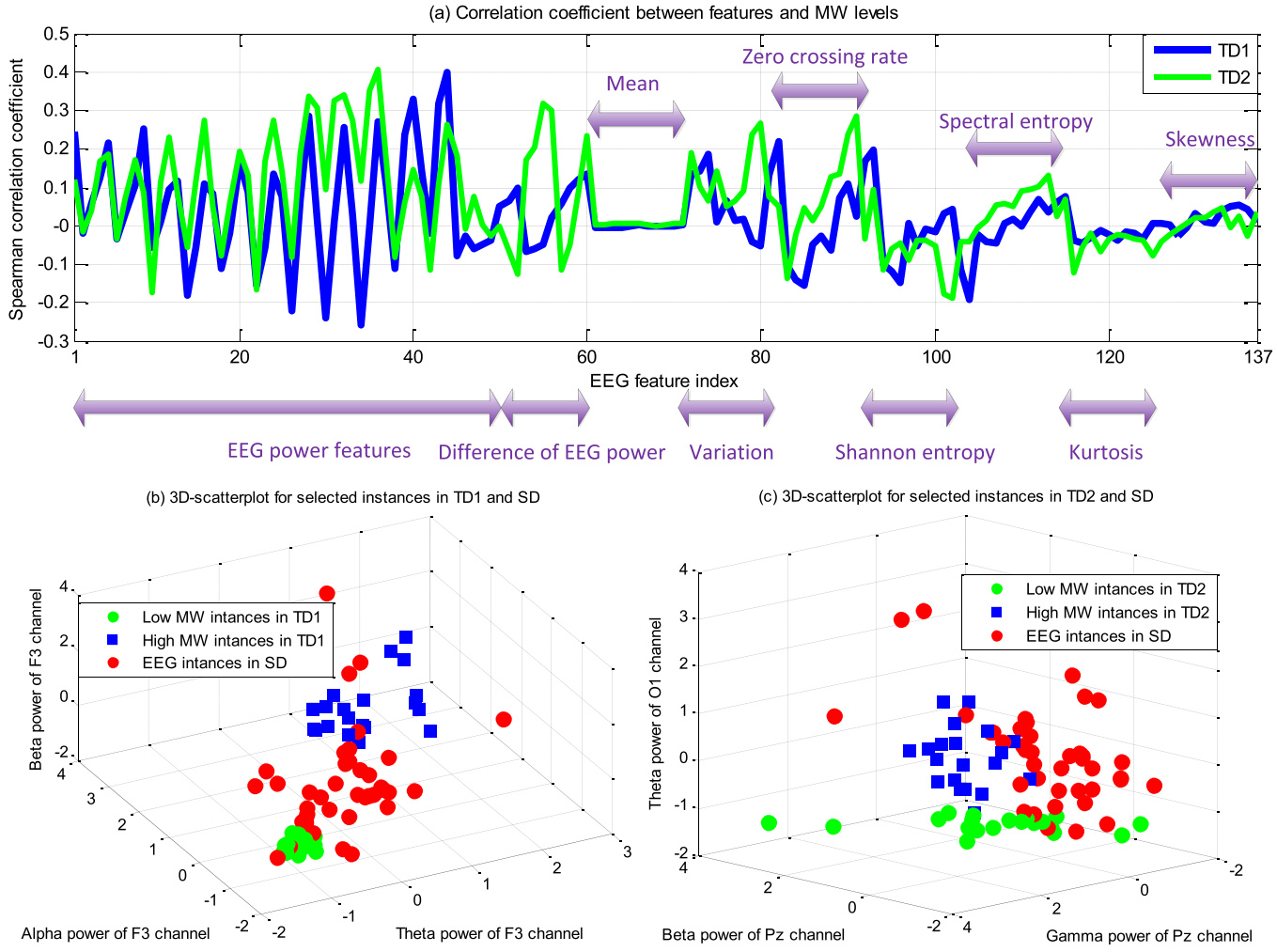

Fig. 2. Visualization of the extracted EEG features. (a) Spearman correlation coefficient between time course of 137 EEG features and the target MW level in TD1 and TD2. (b) 3D-scatterplot for selected EEG feature vectors corresponding to low, high MW classes in TD1 and instances in SD without class labels. (c) 3D-scatterplot for selected EEG feature vectors corresponding to low, high MW classes in TD2 and instances in SD without class labels.

$$\boldsymbol{\theta} = \arg\min \frac{1}{N}\sum_{k=1}^{N}\sum_{j=1}^{m}\varepsilon_{j,k}\left\{x_j(k), f\left[\sum_{i=1}^{\tilde{m}}\tilde{w}_{ji} \cdot f\left(\sum_{j=1}^{m} w_{ij}x_j(k)+b_i\right)+\tilde{b}_j\right]\right\}. \quad (6)$$

In Eq. (6), w_{ij}, $\tilde{w}_{ji}$, b_i, and $\tilde{b}_j$ are corresponding entries for the matrices $\mathbf{W}^*$, $\tilde{\mathbf{W}}^*$, $\mathbf{b}^*$, and $\tilde{\mathbf{b}}^*$, respectively.

The architecture of a SAE model is illustrated in Fig. 3(b), where the high-level feature representation $\mathbf{g}^{(l)}$ is hierarchically elicited by stacking multiple autoencoders with reconstruction layer removed,

$$\mathbf{g}^{(L)} = f(\mathbf{W}^{(L)}...f(\mathbf{W}^{(2)}f(\mathbf{W}^{(1)}\mathbf{x}+\mathbf{b}^{(1)})+\mathbf{b}^{(2)})...+\mathbf{b}^{(L)}). \quad (7)$$

By denoting $s_\theta^{(L)}(\mathbf{x}) = \mathbf{g}^{(L)}$, the SAE parameters $\boldsymbol{\theta}$ are pre-trained in a greedy, layerwise manner. Then, an output layer $\mathbf{y}$ of two nodes corresponding to low and high MW levels is added to $s_\theta^{(L)}(\mathbf{x})$,

$$\mathbf{y} = f(\mathbf{V}\mathbf{g}^{(L)}+\mathbf{b}_v) = f[\mathbf{V}s_\theta^{(L)}(\mathbf{x})+\mathbf{b}_v], \quad (8)$$

where $\mathbf{V}$ and $\mathbf{b}_v$ are the output weights and bias. In the end, $\boldsymbol{\theta}$, $\mathbf{V}$ and $\mathbf{b}_v$ are finely-tuned based on their pre-trained values to achieve optimal $\boldsymbol{\theta}^*$, $\mathbf{V}^*$ and $\mathbf{b}_v{}^*$ via the BP algorithm.

3.3.2. Transfer dynamical autoencoders

The proposed TDAE framework for MW classification is shown in Fig. 4. The TDAE consists of three consecutively-connected modules, which are termed as feature filter, abstraction filter, and transferred MW classifier. All modules are designed based on the SAE architecture shown in the previous section. The function of feature and abstraction filters is to discover hidden patterns underlying the dynamics of the EEG features and high-level representations. Transferred MW classifier exploits large volume EEG data from the SD database and is expected to find generalizable synaptic weights in TDAE deeper layers. The details for each module are presented as follows.

Feature filter module: As shown in Fig. 4, we design m feature filters to process m-dimensional EEG features, where inputs of each filter are the values of the same feature at current and historical time steps,

$$h_{i\hat{j}}^{(1)}(k) = f\left[\sum_{j=1}^{D_1} w_{\hat{j}j,i}^{(1)}x_i(k-j+1)+b_{\hat{j}}^{(1)}\right], i = 1, 2, ..., m. \quad (9)$$

In Eq. (9), the indices of the input and hidden nodes for filter i are denoted by j and $\hat{j}$, respectively. The number of the hidden nodes is set to be equal to that of the input dimensionality D_1. The value of D_1 can be considered as the order of an weighted-moving-average filter, where the filter parameters are the synaptic weights

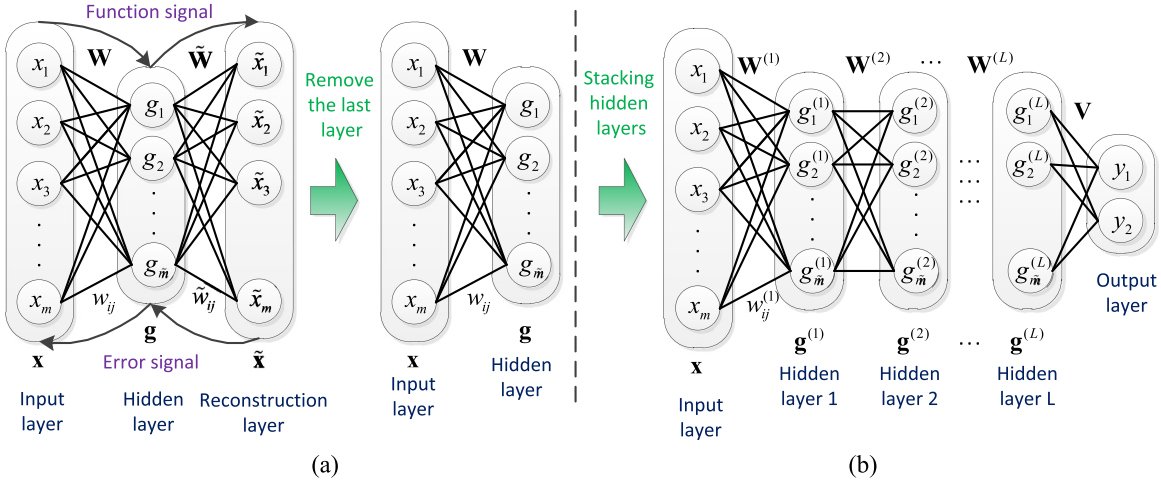

Fig. 3. Architectures of (a) an autoencoder and (b) an SAE.

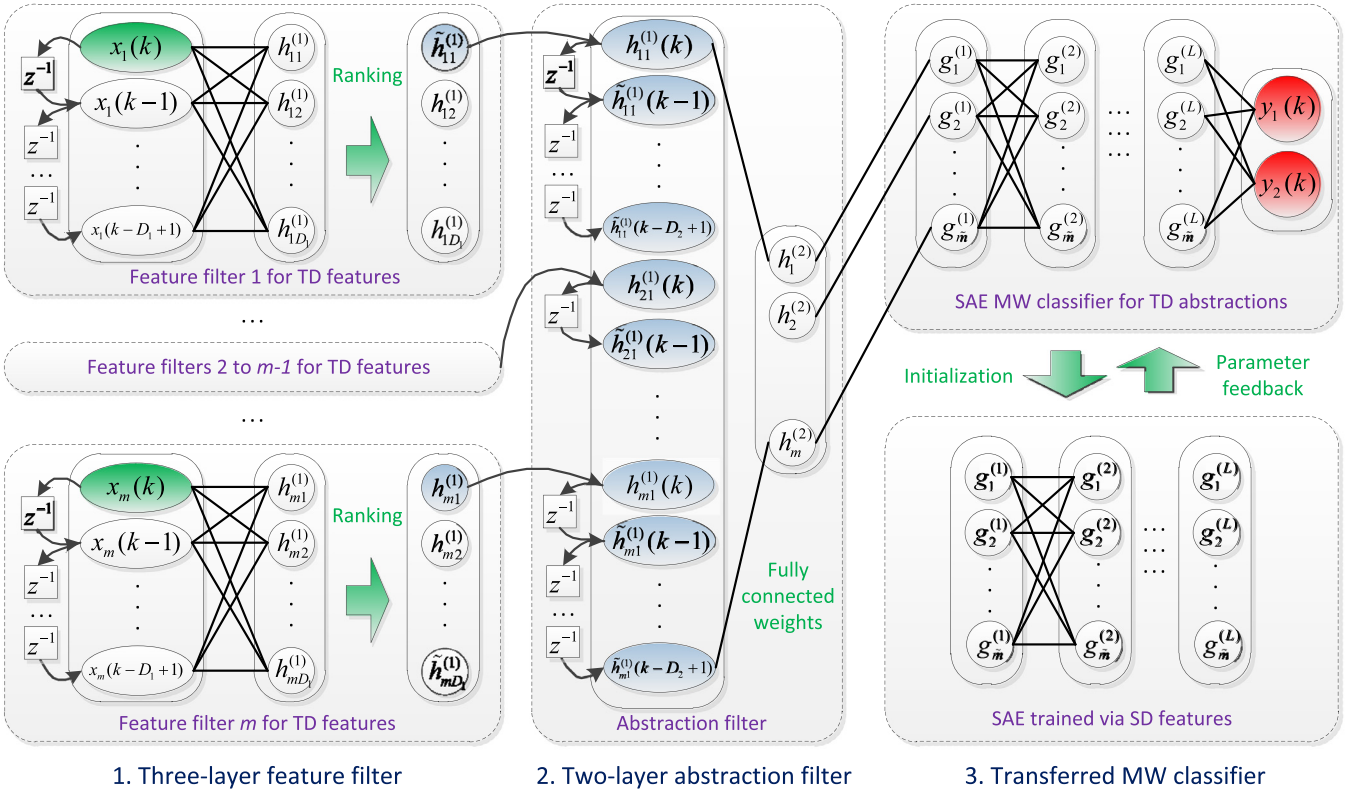

Fig. 4. Architecture of the proposed TDAE for binary MW classification. For feature and abstraction filters, $x_1(k)$ denotes the value of #1 EEG feature at time step k, Z^{-1} represents one-step backward shifting operator, h and g are hidden neuron activations of the SAEs. For the transferred MW classifier in the right section, $y_1(k)$ and $y_2(k)$ denote low and high MW level at time step k, respectively

of an autoencoder. For simplicity, we employ the same value of the hyper-parameter D_1 for all filters.

The filter weights $w_{\hat{j}j,i}$ are computed by repeatedly applying the BP algorithm noted in Eqs. (5)–(6). Then, $h_{i\hat{j}}^{(1)}(k)$ is recognized as the value of a hidden variable at time step k that integrates dynamical properties of EEG feature $x_i(k-j)$. Since positive weights are required for data smooth, we rescaled $w_{\hat{j}j,i}^{(1)}$ as follows,

$$\hat{w}_{\hat{j}j,i}^{(1)} = w_{\hat{j}j,i}^{(1)} / \sum_{j=1}^{D_1} w_{\hat{j}j,i}^{(1)} + z_i, i = 1, 2, ..., m. \quad (10)$$

where z_i is a constant that controls the range of $\hat{w}_{\hat{j}j,i}^{(1)}$ values. For simplicity, we employ the same z_i for all filters.

After all hidden neuron activations are elicited via the pre-trained, rescaled $\tilde{w}_{\hat{j}j,i}$, the abstraction dimension increases to $m \times D_1$. To avoid the curse of the dimensionality, we only select one neuron for each filter linking to the next hidden layer according to the minimum Shannon entropy,

$$\hat{j}^* = \arg\min_{\hat{j}} \; -\frac{1}{N}\sum_{k=1}^{N} h_{i\hat{j}}^{(1)}(k) \cdot \log_2[h_{i\hat{j}}^{(1)}(k)], i = 1, 2, ..., m. \quad (11)$$

Table 1
Pseudo codes for building feature filters in TDAE, the notations are consistent with Eqs. (1)–(11).

Algorithm 1: Building feature filters.

Set the training set $S_{tr}^{(1)}$
$S_{tr}^{(1)} = \{\mathbf{x}(k)|\mathbf{x}(k) \in R^m, k = 1, 2, ..., N\}$
for $i = 1 : m$
  Initialize all$w^{(1)}, b^{(1)}$ with random values
  for $\hat{j} = 1 : D_1$
    $h_{i\hat{j}}^{(1)}(k) = f\left[\sum_{j=1}^{D_1} w_{\hat{j}j,i}^{(1)} x_i(k-j+1) + b_{\hat{j}}^{(1)}\right]$
  Endfor
  $\underset{w^{(1)},\tilde{w}^{(1)},b^{(1)},\tilde{b}^{(1)}}{\arg\min} \frac{1}{N}\sum_{k=1}^{N}\sum_{j=1}^{D_1} \varepsilon_{j,k}\left\{x_j(k-j+1), f\left[\sum_{\hat{j}=1}^{D_1} \tilde{w}_{\hat{j}j,i}^{(1)} \cdot h_{i\hat{j}}^{(1)}(k) + \tilde{b}_{\hat{j}}^{(1)}\right]\right\}$
  for $\hat{j} = 1 : D_1$
    $\hat{w}_{\hat{j}j,i}^{(1)} = w_{\hat{j}j,i}^{(1)} / \sum_{j=1}^{D_1} w_{\hat{j}j,i}^{(1)} + z_i$
  Endfor
  $\hat{j}_i^* = \underset{\hat{j}}{\arg\min} \; -\frac{1}{N}\sum_{k=1}^{N} h_{i\hat{j}}^{(1)}(k) \cdot \log_2[h_{i\hat{j}}^{(1)}(k)]$
  $\tilde{h}_{i1}^{(1)} = h_{i\hat{j}^*}^{(1)}$
Endfor
Return $\hat{w}^{(1)}, b^{(1)}, \tilde{h}^{(1)}$, **as the feature filter parameters and outputs**

Targeting the hidden abstraction $h_{i\hat{j}}^{(1)}(k)$ with minimum entropy is aimed to discover the most informative representation of the EEG feature dynamics. In the end, the 1st ranking abstraction is selected, i.e., $\tilde{h}_{i1}^{(1)}(k) = h_{i\hat{j}^*}^{(1)}(k), i = 1, 2, ..., m$. The algorithm for building the feature filters in TDAE is summarized in Table 1.

Abstraction filter module: The output $\tilde{h}_{i1}^{(1)}(k)$ of each feature filter is the value of the smoothed feature i at time step k under the lowest entropy. To further discover the data structure across current and historical time steps for these shallow representations, we add a shared SAE layer next to all EEG feature filters to generate an abstraction filter. The corresponding hidden activation $h_{\hat{j}}^{(2)}(k)$ is derived as follows,

$$h_{\hat{j}}^{(2)}(k) = f\left[\sum_{i=1}^{m}\sum_{j=1}^{D_2} w_{\hat{j}j,i}^{(2)} \tilde{h}_{i1}^{(1)}(k-j+1) + b_{\hat{j}}^{(2)}\right]. \tag{12}$$

In Eq. (12), the value of D_2 is a predefined constant indicating the order of the abstraction filter, i, j and $\hat{j}$ are indices for the EEG feature, time steps and hidden nodes, respectively. To reduce the model complexity, the number of the hidden neurons is simply equal to the EEG feature dimension, i.e., $\hat{j} = 1, 2, ..., m$.

The computation of the synaptic weights $w_{\hat{j}j,i}^{(2)}$ and bias $b_{\hat{j}}^{(2)}$ is achieved by implementing Eqs. (4) and (5) to minimize the squared loss function,

$$\min_{w^{(2)},\tilde{w}^{(2)},b^{(2)},\tilde{b}^{(2)}} \frac{1}{N}\sum_{j=1}^{D_2} \times \sum_{i=1}^{m}\sum_{k=1}^{N} \varepsilon_{ij,k}\left\{\tilde{h}_{i1}^{(1)}(k-j+1), f\left[\sum_{\hat{j}=1}^{D_2} \tilde{w}_{\hat{j}j,i}^{(2)} \cdot h_{\hat{j}}^{(1)}(k) + \tilde{b}_{ij}^{(2)}\right]\right\} \tag{13}$$

where $\tilde{w}_{\hat{j}j,i}^{(2)}$ and $\tilde{b}_{ij}^{(2)}$ denoting the reconstruction weights and bias, respectively. The algorithm for building the abstraction filter of the TDAE is listed in Table 2.

Transferred MW classifier: The transferred MW classifier consists of two SAE networks as shown in Fig. 4. We first employ a SAE to examine the geometric similarity of the feature abstractions across SD and TD databases. Given a training set from the TD with two MW levels indicated, the pre-trained, finely-tuned SAE parameters

Table 2
Pseudo codes for building the abstraction filter in TDAE, the notations are consistent with Eqs. (12)–(13).

Algorithm 2: Building the abstraction filter.

Call Algorithm 1
Return $\hat{w}^{(1)}, b^{(1)}, \tilde{h}^{(1)}$
Set the training set $S_{tr}^{(2)}$
$S_{tr}^{(2)} = \{\tilde{\mathbf{h}}^{(1)}(k)|\tilde{\mathbf{h}}^{(1)}(k) \in R^m, k = 1, 2, ..., N\}$
  Initialize all $w^{(2)}, b^{(2)}$ with random values
  for $\hat{j} = 1 : m$
    $h_{\hat{j}}^{(2)}(k) = f\left[\sum_{i=1}^{m}\sum_{j=1}^{D_2} w_{\hat{j}j,i}^{(2)} \tilde{h}_{i1}^{(1)}(k-j+1) + b_{\hat{j}}^{(2)}\right]$.
  End for
  $\underset{w^{(2)},\tilde{w}^{(2)},b^{(2)},\tilde{b}^{(2)}}{\arg\min} \frac{1}{N}\sum_{j=1}^{D_2}\sum_{i=1}^{m}\sum_{k=1}^{N} \varepsilon_{ij,k}\left\{\tilde{h}_{i1}^{(1)}(k-j+1), f\left[\sum_{\hat{j}=1}^{D_2} \tilde{w}_{\hat{j}j,i}^{(2)} \cdot h_{\hat{j}}^{(1)}(k) + \tilde{b}_{ij}^{(2)}\right]\right\}$
Return $w^{(2)}, b^{(2)}, h^{(2)}$, **as the abstraction filter parameters and outputs**

with L hidden layers are computed according to Eq. (6),

$$\boldsymbol{\theta}^* = \{\mathbf{W}^{(1)*}, \mathbf{W}^{(2)*}, ..., \mathbf{W}^{(L)*}, \mathbf{b}^{(1)*}, \mathbf{b}^{(2)*}, ..., \mathbf{b}^{(L)*}\}. \tag{14}$$

Then, we map EEG features from the input space to the abstraction space. High-level feature representations for the TD databases are elicited as neuron activations of the last hidden layer according to Eq. (7),

$$s_{\theta^*}^{(L)}[\mathbf{x}_{TD}(k)] = f(\mathbf{W}^{(L)*}...f(\mathbf{W}^{(2)*}f(\mathbf{W}^{(1)*}\mathbf{x}_{TD}(k) + \mathbf{b}^{(1)*}) + \mathbf{b}^{(2)*})... + \mathbf{b}^{(L)*}). \tag{15}$$

In Eq. (15), $\mathbf{x}_{TD}(k)$ and $s_{\theta^*}^{(L)}[\mathbf{x}_{TD}(k)]$ denote the kth EEG feature and abstraction vectors in the TD database, respectively. By implementing exactly the same SAE, we then compute feature abstractions from SD database as $s_{\boldsymbol{\theta}^*}^{(L)}[\mathbf{x}_{SD}(k)]$.

The domain adaptation is achieved by minimizing the *L2* norm of the MW and emotion EEG abstractions,

$$\min_j \varphi[\mathbf{x}_{SD}(j)] = \left\| \frac{1}{N_{\mathbf{y}}} \sum_{i=1}^{N_{\mathbf{y}}} s_{\theta^*}^{(L)}[\mathbf{x}_{TD}(i)] - s_{\theta^*}^{(L)}[\mathbf{x}_{SD}(j)] \right\|_2. \tag{16}$$

In Eq. (16), $N_{\mathbf{y}}$ is the number of training instances of class $\mathbf{y}$. By applying Eq. (16), the EEG feature vector $\mathbf{x}_{SD}$ of the index j^* with the minimum Euclidean distance to the class center is determined. Then, we generate a set of adaptive training instances $\mathbf{x}_{SD}(j^*)$ that are iteratively drawn from the SD database. Finally, the adaptive training instances are used in the fine-tuning stage for the other SAE as the MW classifier. The training algorithm for building such transferred MW classifier is summarized in Table 3.

4. Results

4.1. Model selection of MW classifiers

For each subject in TD1, the 3600 instances were split into training, validating, and testing sets with each of 1200 instances. That is, the overall size of the above three sets is $1200 \times 8 = 9600$ instances. For each subject in TD2, there are 2030 EEG vectors while the sizes for training, validating and testing sets are 677, 677, and 676 instances, respectively. Accordingly, the overall sizes of the three sets from six subjects in TD2 are 4062, 4062, and 4056 instances. It is noted that only a single trial of the classification experiment was performed. To make the results repeatable, the initial states of all model parameters are all fixed as the same. To guarantee the consistency of the temporal sequence and the class balance, the training, validating and testing data were evenly sampled

Table 3
Pseudo codes for building the transferred MW classifier in TDAE, the notations are consistent with Eqs. (14)–(16).

Algorithm 3: Building the transferred MW classifier.

Set the training set $S_{tr}^{(1)}$ from TD
$S_{tr,TD}^{(1)} = \{\mathbf{x}_{TD}(k)|\mathbf{x}_{TD}(k) \in R^m, \mathbf{y}_{TD}(k) \in R^2, k = 1, 2, ..., N_{TD}\}$
Set the SD dataset $S_{sd}^{(1)}$
$S_{SD}^{(1)} = \{\mathbf{x}_{SD}(k)|\mathbf{x}_{SD}(k) \in R^m, \mathbf{y}_{SD}(k) \in R^2, k = 1, 2, ..., N_{SD}\}$
Pre - train and finely - tune a SAE $s_1^{(L)}$ via $S_{tr,TD}^{(1)}$
for $k = 1 : N_{TD}$
    $\mathbf{a}_{TD}(k) = s_1^{(L)}[\mathbf{x}_{TD}(k)]$
End for
for $k = 1 : N_{SD}$
    $\mathbf{a}_{SD}(k) = s_1^{(L)}[\mathbf{x}_{SD}(k)]$
End for
Set adaptive training set $S_{tr,AD}^{(1)} = \emptyset$
for $j = 1 : N_{AD}$
    $j^* = \arg\min_j \ \varphi(j) = \| \frac{1}{N_y} \sum_{i=1}^{N_y} \mathbf{a}_{TD}(k) - \mathbf{a}_{SD}(k) \|_2.$
    $S_{tr,AD}^{(1)} = S_{tr,AD}^{(1)} \cup \{\mathbf{x}_{SD}(j^*), \mathbf{y}(j^*)\}$
    $S_{SD}^{(1)} = S_{SD}^{(1)} \backslash \{\mathbf{x}_{SD}(j^*), \mathbf{y}(j^*)\}$
End for
Call Algorithm 1 and 2 on $S_{tr,TD}^{(1)}$
Return $\hat{w}^{(1)}, b^{(1)}, w^{(2)}, b^{(2)}, h^{(2)}$
Pre - train a SAE $s_2^{(L)}$ via $h^{(2)}$
Finely - tune $s_2^{(L)}$ via $h^{(2)}$ and $S_{tr,AD}^{(1)}$
Return $s_2^{(L)}$ **as the transferred MW classifier**

Note: N_{TD}, N_{SD}, and N_{AD} denote the number of the training instances in TD, SD and adaptive training set, respectively.

from the overall EEG feature set of each subject. That is, in every three consecutive instances of a session, the 1st, 2nd, and 3rd EEG feature vectors were used for training (with the temporal index of 1, 4, 7 …), validating (with the temporal index of 2, 5, 8, …), and testing (with the temporal index of 3, 6, 9, …), respectively.

In particular, the validating set is employed for model selection to predetermine the optimal hyper-parameters of the learning machine. Since the conventional SAE is repeatedly used in the transferred MW classifier module of TDAE, we carry out a grid search on different numbers of hidden neurons (denoted by $\tilde{m}$ in Fig. 4) and training epochs. During the grid search, the learning rate and number of hidden layers (denoted by L in Fig. 4) are fixed to 1 and 4, respectively, according to our previous work [33].

Validating classification accuracy of SAE on different hyper-parameter combinations is illustrated in Fig. 5(a) and (b) for TD1 and TD2, respectively. The candidate values of the hidden neuron and epoch numbers are set to {5, 10, 15, ..., 100} and {1, 2, 3, ..., 20}. Among all 400 cases, the range of validating accuracy varies from 0.4 to 0.74 with the optimal values achieved by 0.7297 and 0.7450 for two cases. Then, we employ 95 hidden neurons for the transferred MW classifier in TDAE while the training epoch number is selected at 20 and 17 for two cases.

To facilitate the classification performance comparison, we also apply two powerful machine learning methods, LSSVM and ELM on the same model selection procedure. The validating accuracies are shown in Fig. 5(c)–(f). For LSSVM, Gaussian radial basis function (RBF) kernel is implemented. A grid search is conducted in a space of the kernel width and the regularization parameter. For ELM, we employ the logistic sigmoid function $1/(1 + \rho \cdot e^{-x})$ with a tunable parameter ρ. Then, the grid search exhausts different values of ρ and hidden neuron number. From the figure, we found both of the optimal LSSVM and ELM models outperform standard SAE regarding the optimal validating accuracy.

In the proposed TDAE framework for mental workload classification, we integrated multiple autoencoder layers for feature representation. In the fine tuning stage, TDAE predicted the validating EEG feature vectors via an additional ANN layer as shown in Fig. 4. Here, we adopted the ELM and the LSSVM independent from TDAE for the comparison purpose. Due to their shallow structures, the corresponding validating performance is superior to that of the classical SAE when the size of the training set is limited.

For the low validating performance of the deep SAE shown in Fig. 5(a), the possible reason is the potential overfitting of the EEG features for particular network structure. When the number of the hidden neurons were set around 50, the high level feature abstractions may be ill represented with the growth of the number of the training epochs. For Fig. 5(b), the poor performance can be caused by the inadequate size of the training set since there are much fewer instances in TD2 than TD1. For instance, when only two or three epochs were applied for training, the SAE parameters in large number of hidden neurons cannot be perfectly tuned. Other failed cases can be interpreted as the result of the algorithm instability for specific values of the initial weights. For Fig. 5(c)–(d), the wrong combinations of the regularization and kernel parameters of LSSVM yield that all EEG instances were assigned to the same mental workload class. Then, lowest performances of 0.5–0.55 are achieved. For ELM classifiers shown in Fig. 5(e)–(f), the validating performance is closely linked to the hidden neuron number. The low accuracy corresponds to those simple ELM models of low fitting capability, e.g., the validating accuracy is only around 0.6 with 10–30 hidden neurons implemented.

The details of the classification experimental settings used in the following sections are listed in Table 4. In total, ten different classification methods are adopted and analyzed, i.e., naive Bayesian model (NB), logistic regression (LR), k-nearest neighbor (KNN), ANN, ELM, LSSVM, SAE, DBN, CNN, and the proposed TDAE. In the table, the selected hyper-parameter values on TD1 and TD2 are presented along with the size and the dimensionality of the EEG databases.

4.2. MW classification performance of TDAE

After selecting the optimal hyper-parameters via the validating set, MW classification is carried out in a subject-generic manner by training one classifier that is generalizable for all individuals in each target-domain database. For TDAE, the structure of the transferred MW classifier module is consistent with the optimized SAE shown in Fig. 5. The orders of the feature and abstraction filters (D_1 and D_2 in Fig. 4) are set to 19 and 3, respectively. Since the sizes of the training set in TD1 and TD2 are different, we employ 10,000 and 20,000 adaptive training instances (N_{AD} in Table 3) for two databases.

By denoting the low (or high) MW level as the positive (or negative) class, the performance metrics of accuracy, sensitivity, specificity, F1-score, precision, and negative predicting value (NPV) of TDAE, SAE, LSSVM, and ELM are computed and shown in Figs. 6 and 7 for TD1 and TD2, respectively. We find TDAE achieves the highest classification accuracy for 11 subjects, i.e., A, C, D, E, F, G, I, J, K, M and N (as shown in Figs. 6(a) and 7(a)). In Fig. 6(b) and (f), the substantial improvement of the sensitivity and NPV of TD1 indicates most of low-MW EEG vectors are correctly predicted by TDAE. It is also shown the values of all metrics of TDAE increase in TD2. For other classifiers, LSSVM outperforms both of ELM and SAE while SAE achieves the lowest accuracy. The finding is in line with the rank of the optimal validating accuracy illustrated in Fig. 5.

The classification confusion matrices corresponding to different MW classifiers are given in Table 5. Each matrix is the summation of testing results for all subjects in each database. Numbers of testing EEG vectors of TD1 and TD2 are 9600 and 4056, respectively. When implementing TDAE on TD1, 4534 low-MW and 3744 high-MW instances are correctly predicted. The classifier misclassifies 266 low-MW instances as high-MW while 1056 high-MW

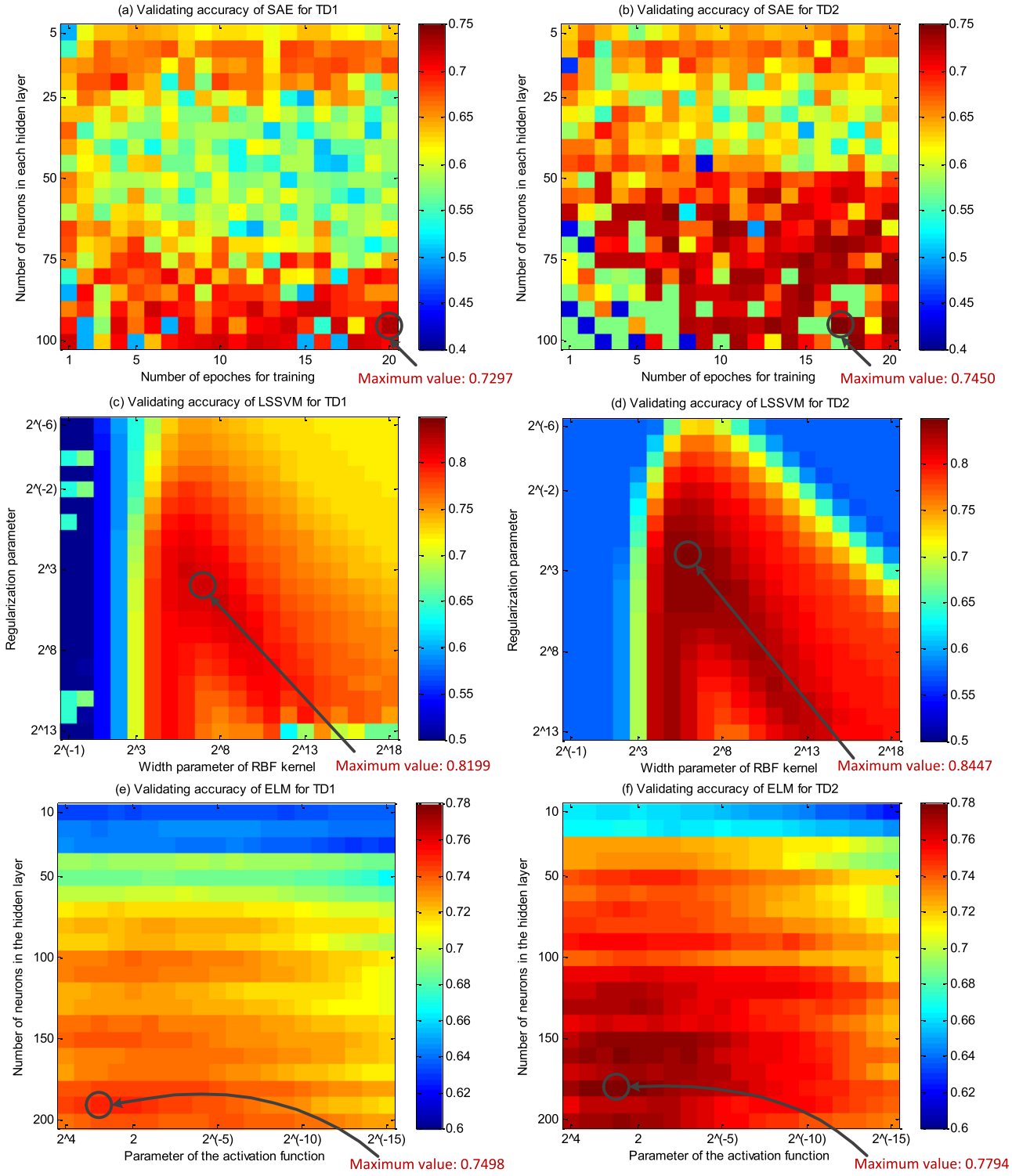

Fig. 5. Model selection results via validating sets for (a)–(b) SAE, (c)–(d) LSSVM and (e)–(f) ELM.

instances are wrongly predicted as low MW level. Among all four methods, TDAE possesses the highest accuracy, sensitivity, and NPV on TD1. For TD2, TDAE is superior to others on all five metrics.

4.3. Accuracy comparison for TDAE against shallow and deep classifiers

In this section, we implement several popular shallow and deep classifiers on TD1 and TD2. The corresponding MW classification performance is compared with that of TDAE shown in the previous section.

In Figs. 8 and 9, the performance metrics of TDAE and 12 different shallow classifiers is illustrated in error-bar plots. Subfigures provide the subject-average accuracy, F1-score, sensitivity, and specificity on each database. The results of TDAE, LSSVM and ELM are consistent with those in the previous section. After carefully selecting model hyper-parameters, we newly introduce NB, LR, KNN, and ANN. By incorporating principal component analysis

Table 4
Experimental settings for mental workload classification using different methods.

Classification methods	Hyper-parameter settings for TD1	Hyper-parameter settings for TD2	Number of EEG features	Number of EEG instances
NB	Prior probabilities: 0.5, 0.5	Prior probabilities: 0.43, 0.57	137	TD1: 28,800 TD2: 12,180
LR	Distribution model: binomial	Distribution model: binomial	137	TD1: 28,800 TD2: 12,180
KNN	*k* value: 51	*k* value: 31	137	TD1: 28,800 TD2: 12,180
ANN	Number of hidden nodes: 190	Number of hidden nodes: 180	137	TD1: 28,800 TD2: 12,180
ELM	Number of hidden nodes:190 Activation function parameter: 4	Number of hidden nodes:180 Activation function parameter: 2	137	TD1: 28,800 TD2: 12,180
LSSVM	Regularization and kernel parameters: 16, 128	Regularization and kernel parameters: 4, 64	137	TD1: 28,800 TD2: 12,180
SAE	Numbers of hidden nodes, layers, and training epochs: 90, 4, 20	Numbers of hidden nodes, layers, and training epochs: 90, 4, 17	137	TD1: 28,800 TD2: 12,180
DBN	Numbers of hidden nodes, layers, and training epochs: 90, 2, 20	Numbers of hidden nodes, layers, and training epochs: 90, 2, 17	137	TD1: 28,800 TD2: 12,180
CNN	Numbers of hidden nodes, layers, and training epochs: 144, 4, 20	Numbers of hidden nodes, layers, and training epochs: 144, 4, 20	137	TD1: 28,800 TD2: 12,180
TDAE	Numbers of hidden nodes, layers, and training epochs: 90, 4, 20	Numbers of hidden nodes, layers, and training epochs: 90, 4, 20	137	TD1: 28,800 TD2: 12,180

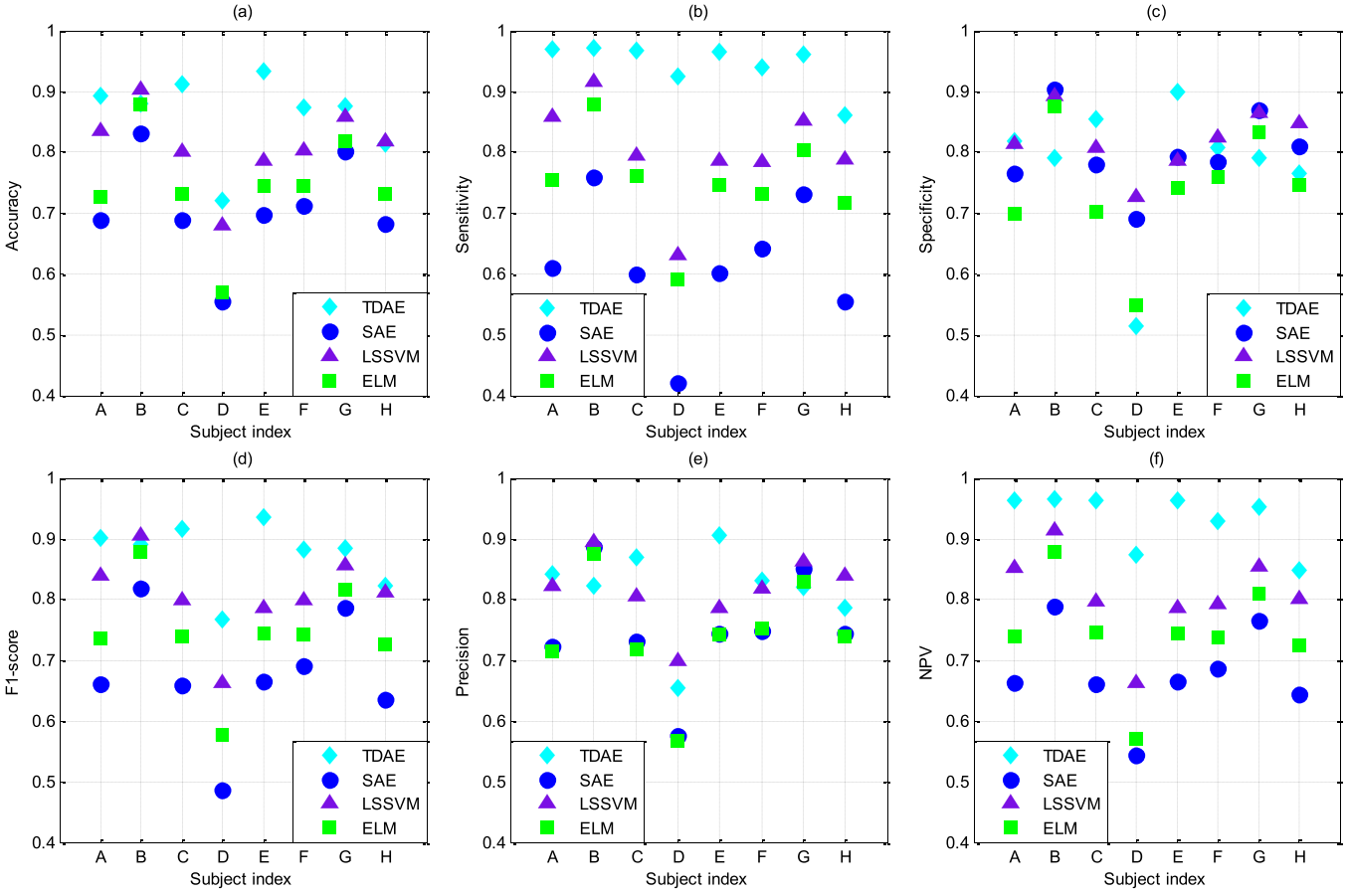

Fig. 6. Testing classification performance of 8 subjects (denoted by A–H) in TD1 computed by TDAE, SAE, LSSVM, and ELM. The hyper-parameters of all classifiers are selected according to Fig. 5. NPV denotes the negative predicting value.

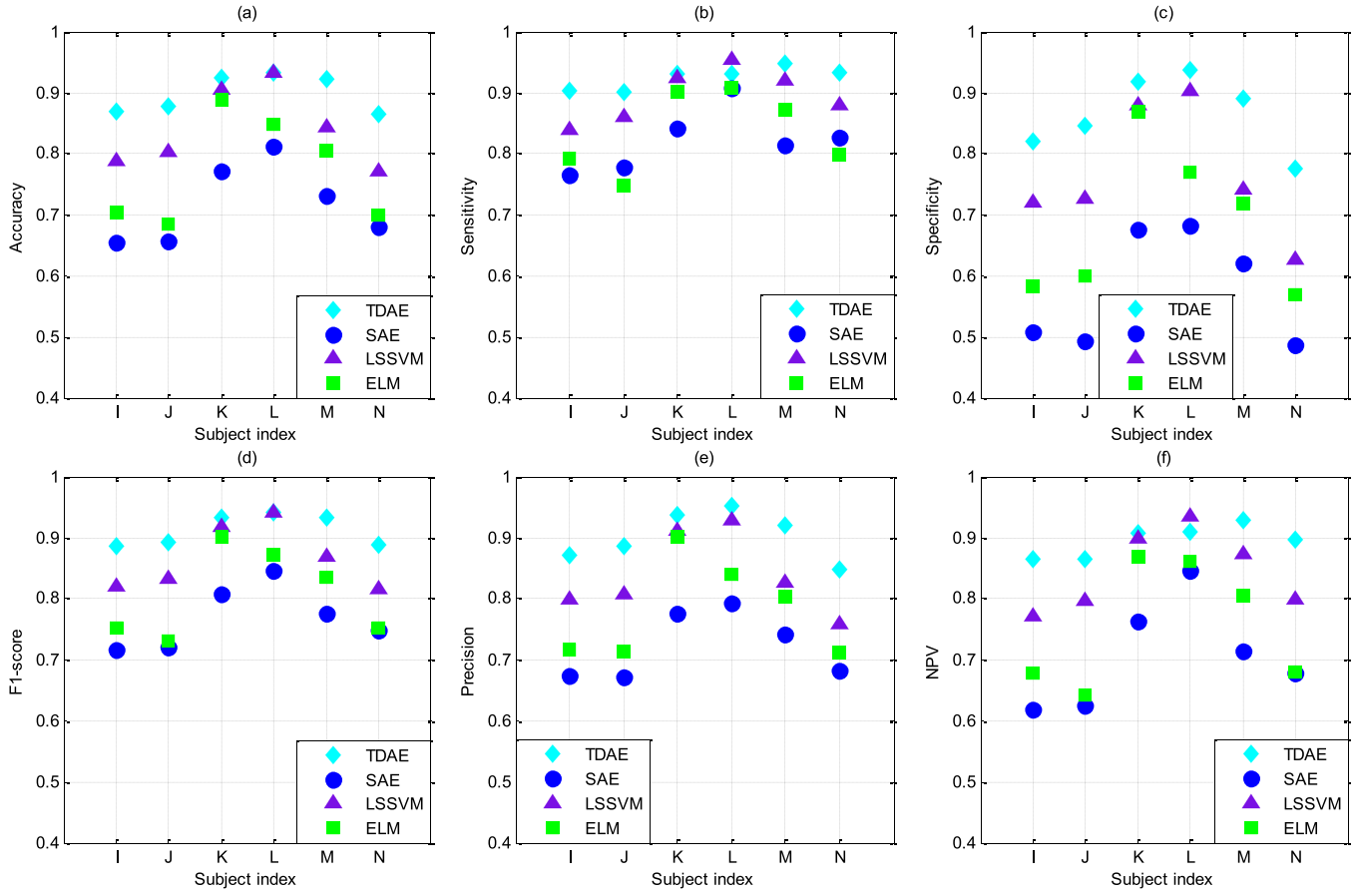

Fig. 7. Testing classification performance of 6 subjects (denoted by I–N) in TD2 computed by TDAE, SAE, LSSVM, and ELM. The hyper-parameters of all classifiers are selected according to Fig. 5.

(PCA) for feature dimensionality reduction, six hybrid classifiers are generated and defined as PCA-LSSVM, PCA-ELM, PCA-NB, PCA-LR, PCA-KNN and PCA-ANN, where the principal components are determined by a threshold of 0.9 for the total variance.

For TD1 database, TDAE achieves the highest value of the accuracy, F1-socore, and sensitivity among all 13 MW classifiers while the optimum specificity is reached by KNN. PCA-NB elicits the worst overall performance regarding the accuracy and F1-score. It is noted both of LSSVM and PCA-LSSVM are competitive. The results of ELM, LR, and ANN are comparable and outperform KNN model. The PCA preprocessing does not significantly improve the classification performance on all shallow classifiers. For TD2 database, TDAE is superior to other shallow classifiers on all performance metrics. Competitive results from LSSVM and LR methods are found. Similarly, NB and PCA approaches are not effective. By implementing a two-sided Wilcoxon signed-rank test, we notice the classification improvement of TDAE on the average accuracy and F1-score is statistically significant across all 14 subjects ($p<0.005$).

Furthermore, we compare TDAE against three deep learning primitives, i.e., SAE, DBN, and CNN. For the DBN network, severe overfitting is struck for three or more hidden layers so that we adopt a two-hidden-layer structure for feature abstraction. The CNN model is constructed with four hidden layers of two blocks. Each block consists of a convolution layer and a consecutive subsampling layer. Since CNN only accepts 2-dimensional input, we reshape 137 EEG features into a 12-by-12 pixel matrix. The value of seven empty pixels (i.e., 144–137=7) is set to 1 for the training, validating and testing datasets. Subject-average error-bar plots are shown in Fig. 10, where the performance of TDAE and SAE is consistent with that in Section 4.1. From the figure, we observe TDAE is superior to other classifiers except for the specificity derived on TD1 while CNN achieves the lowest performance. By using the two-sided Wilcoxon signed-rank test, we find the classification improvement on the average accuracy and F1-score is statistically significant ($p<0.005$).

4.4. Performance of TDAE using different hyper-parameter settings

There are three important hyper-parameters of TDAE that affect its generalization capability, i.e., the size of the adaptive training set (N_{AD}), the order of feature filters (D_1), and the order of the abstraction filter (D_2). The value of N_{AD} stands for how many instances from SD has been properly selected and added to TD training sets. In Fig. 11, we examine the testing performance of TDAE under different N_{AD} values while D_1 and D_2 are separately fixed to 19 and 3 as given in Section 4.2. For most cases the accuracy can be higher than the optimum listed in Table 5 (i.e., 0.8623 for TD1 and 0.8987 in TD2). Moreover, the accuracy can be improved to 0.9030 and 0.9127 when adding 14,000 and 31,000 instances on TD1 and TD2, respectively. The observation indicates transferring knowledge from the EEG features stimulated by emotional clues via deep networks is quite helpful for further enhancing the performance of TDAE for MW recognition tasks. Here, a moderate size, e.g., around 30,000 instances, of the adaptive training set is recommended since too small or too large N_{AD} may impair the stability of the classifier performance.

We also investigate how D_1 and D_2 affect the TDAE performance in Fig. 12. Different combinations of two hyper-parameters are examined while the values of D_1 and D_2 vary from 1 to 5 and

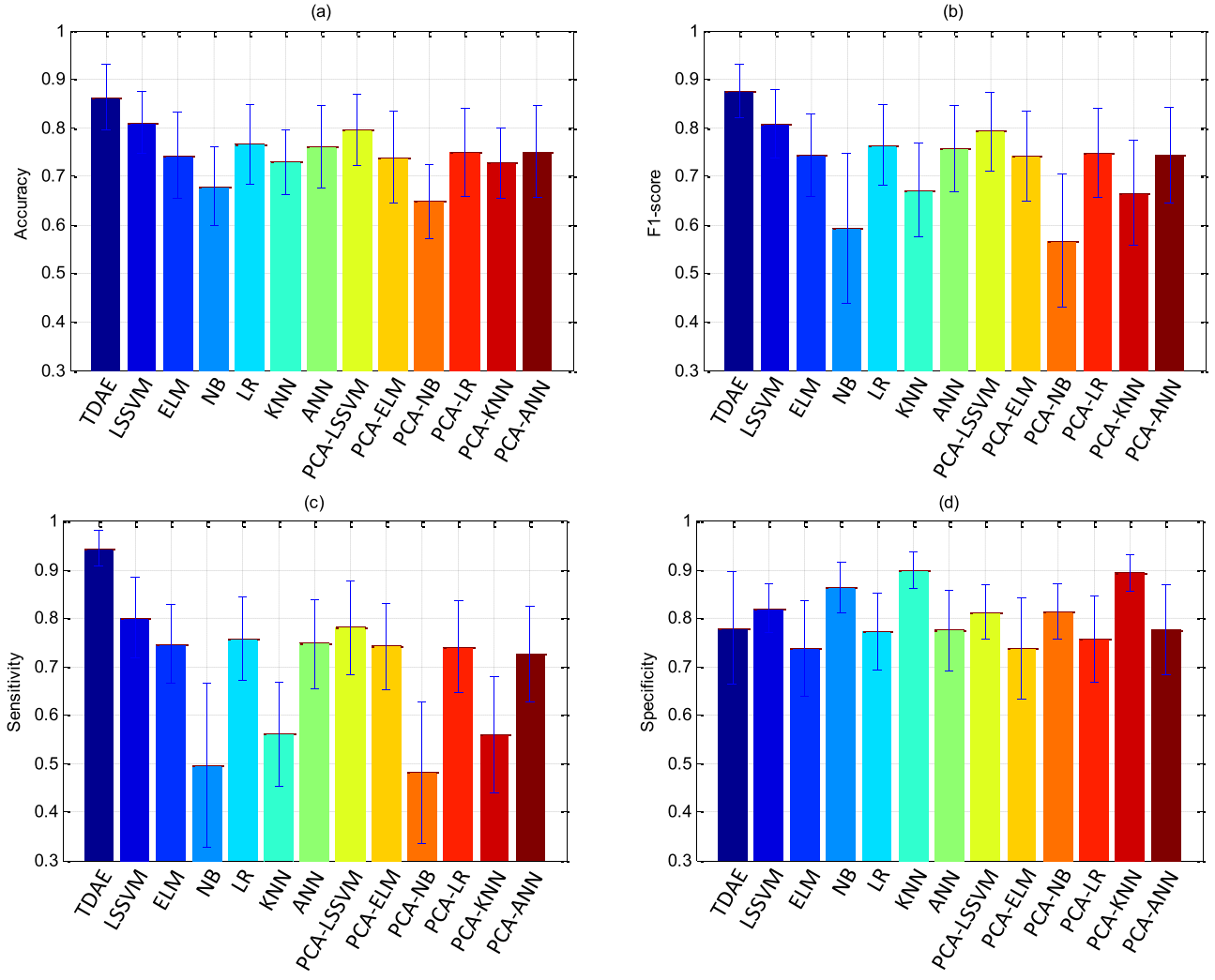

Fig. 8. Testing MW classification performance on TD1 computed by TDAE and 12 shallow classifiers. In each column data, the subject-average (subjects A–H) performance metric is shown. The error bar marks the standard deviation.

1 to 20, respectively. When both of them are set to 1, the architecture of TDAE is equivalent to SAE. Given $D_1 = 1$, moderate D_2, i.e., four, can improve the accuracy by approximately 7%. In general, larger D_1 is more effective for improving the classification performance. The underlying reason is that the EEG features can be strongly smoothed when using SAE as a local filter across a wider range of historical time instants. However, combining too large D_1 and D_2 may lead to overfitting. For instance, the accuracy degrades with $D_1 = 19$ and $D_2 = 5$ for TD1 and $D_1 = 16$ and $D_2 = 5$ for TD2. Here, we recommend a large (15–19) and a moderate (2–4) value for D_1 and D_2 to be selected, respectively.

4.5. Computational cost and performance stability of different MW classifiers

The computational cost of all shallow and deep classifiers for two target-domain databases is shown in Table 6. For each method, the CPU time for training and testing a subject-generic MW recognition model is recorded. The table lists the average and standard deviation of the time duration for ten same trials. All algorithms were programmed and implemented via Matlab® 2011b software and tested by a laptop computer with Windows 7® operation system, Intel®i7 CPU 2.5GHZ and 8GRAM configurations. It is noted LSSVM is implemented by using a function toolbox developed by Suykens et al. [42]. The ELM classifier is constructed via the codes from Huang et al. [43]. The reason for employing ELM for comparison is that it possesses very fast training speed.

From the table, we find the lowest computational cost is obtained by the NB classifier according to the smallest average CPU time. Similarly, ELM and LR based approaches were also fast trained and implemented. By using PCA for feature dimension reduction, the time cost for PCA-LSSVM, PCA-KNN, and PCA-ANN decreases. Neural network based approaches, i.e., ANN, SAE, CNN, DBN and TDAE, show higher training and testing time cost. In particular, TDAE possesses the highest computational burden. The reason behind is that a large number of feature and abstraction filters have to be trained via autoencoders with gradient-based optimization algorithms. Moreover, as the SD database is utilized for knowledge transferring via an adaptive training set, additional computational burden is induced by an enlarged training sample. However, taking into account the fact that the high computational cost of TDAE corresponds to the significantly-improved accuracy, the proposed method is still competitive against conventional MW classifiers.

To investigate the stability of the proposed algorithm, we repeatedly performed the mental workload classification on 10 trials. In each trial, the model parameters are randomly initialized. The mean testing accuracy along with the standard deviation on for each classifier is summarized in Table 7. It is shown the results are consistent with that illustrated in Figs. 8-10, i.e., TDAE

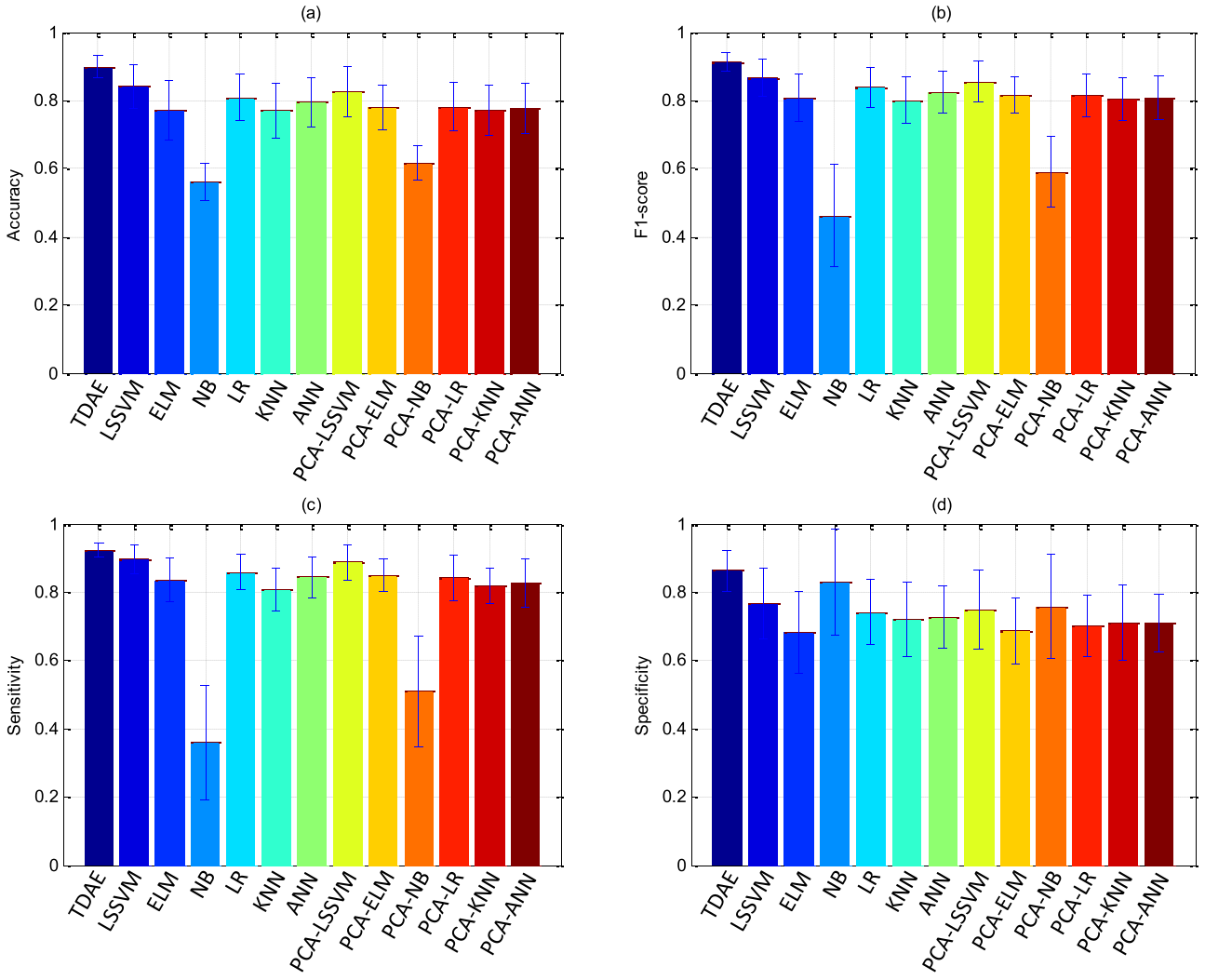

Fig. 9. Testing MW classification performance on TD2 computed by TDAE and 12 shallow classifiers. In each column data, the subject-average (subjects I–M) performance metric is shown.

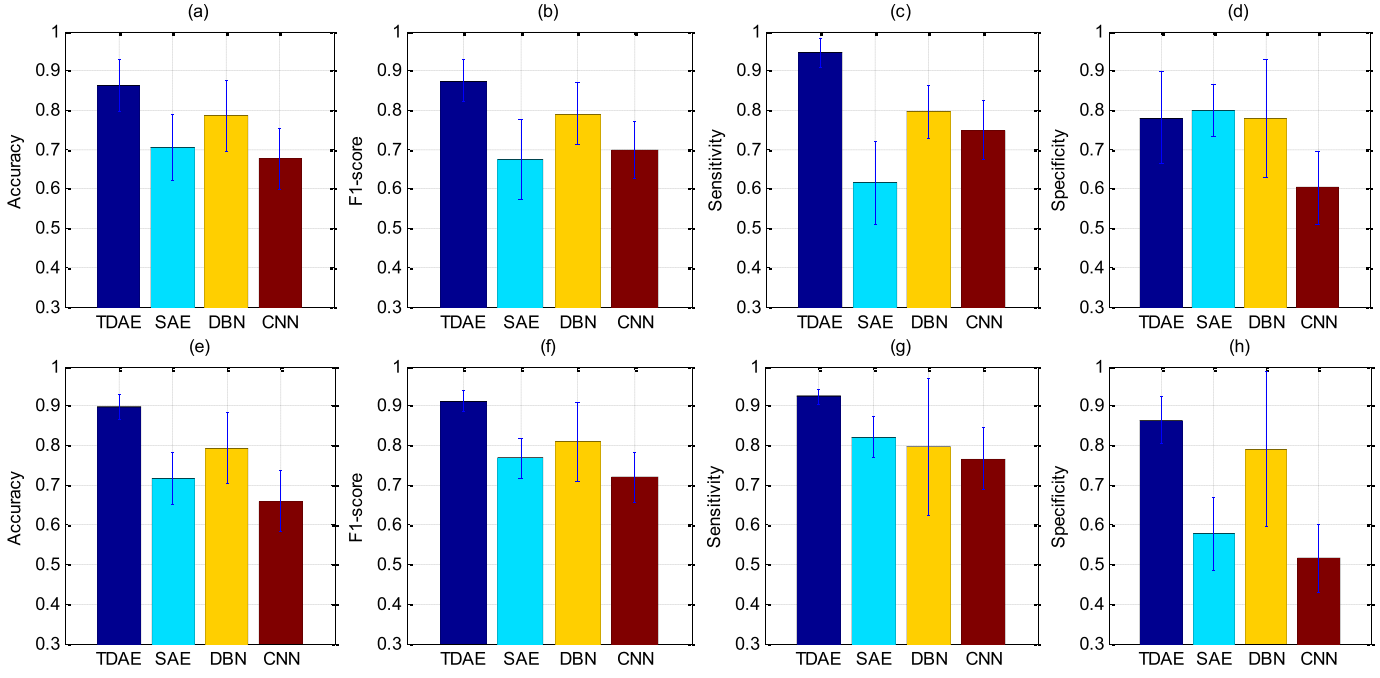

Fig. 10. Testing MW classification performance on (a)–(d) TD1 and (e)–(f) TD2 computed by TDAE and three deep classifiers.

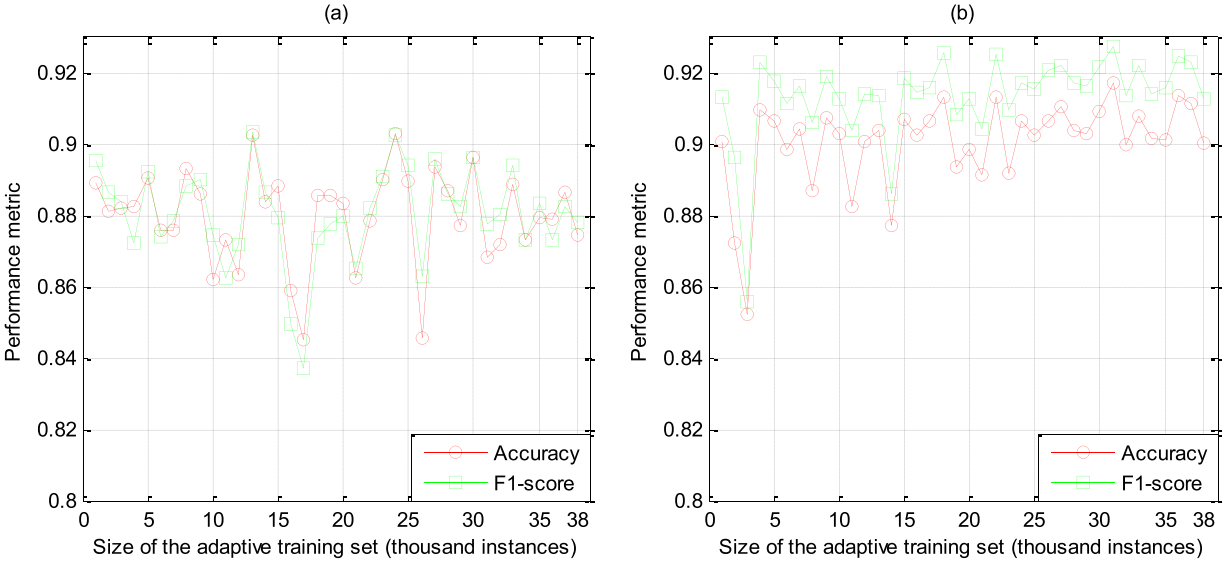

Fig. 11. Testing MW classification performance of TDAE vs. the size of the adaptive training set from SD. The results of TD1 and TD2 are shown in subfigure (a) and (b), respectively.

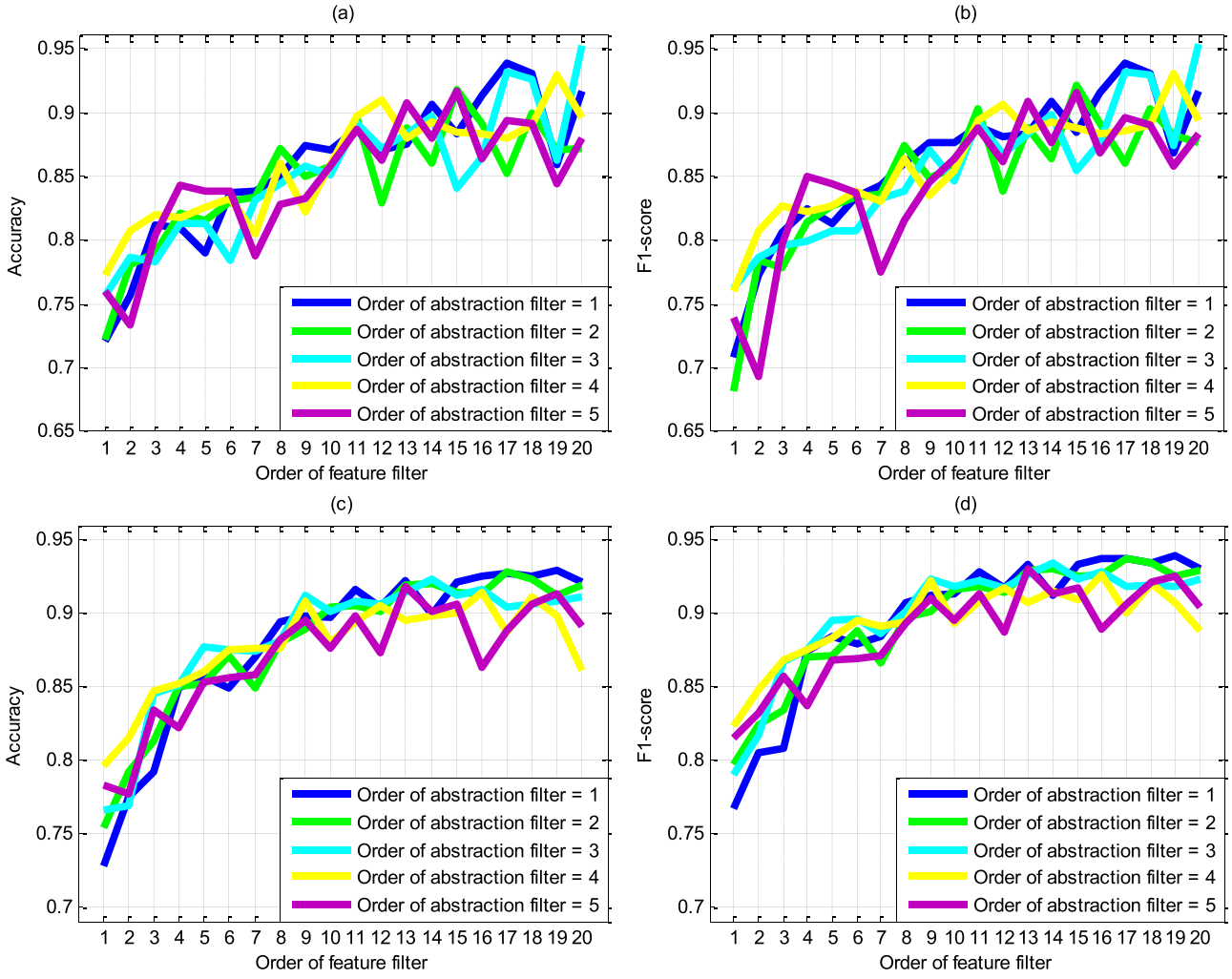

Fig. 12. Testing MW classification performance of TDAE vs. orders of feature and abstraction filters on (a)–(b) TD1 and (c)–(d) TD2.

Table 5
Mental MW confusion matrices summated over all subjects in each database.

Classifier-Database	Predicted MW level	Target MW level		Classification performance
		Low	High	
TDAE-TD1		P_{pre}=0.8111	P_{npv}=**0.9337**	P_{acc}=**0.8623**
	Low	4534	1056	P_{sen}= **0.9446**
	High	266	3744	P_{spe}=0.7800
TDAE-TD2		P_{pre}=**0.9007**	P_{npv}=**0.8958**	P_{acc}=**0.8987**
	Low	2141	236	P_{sen}=**0.9244**
	High	175	1504	P_{spe}=**0.8644**
SAE-TD1		P_{pre}=0.7536	P_{npv}=0.6745	P_{acc}=0.7068
	Low	2949	964	P_{sen}=0.6144
	High	1851	3836	P_{spe}=0.7992
SAE-TD2		P_{pre}=0.7214	P_{npv}=0.7087	P_{acc}=0.7170
	Low	1903	735	P_{sen}=0.8217
	High	413	1005	P_{spe}=0.5776
LSSVM-TD1		P_{pre}=**0.8170**	P_{npv}=0.8052	P_{acc}=0.8109
	Low	3847	862	P_{sen}=0.8015
	High	953	3938	P_{spe}=**0.8204**
LSSVM-TD2		P_{pre}=0.8364	P_{npv}=0.8475	P_{acc}=0.8407
	Low	2076	406	P_{sen}=0.8964
	High	240	1334	P_{spe}=0.7667
ELM-TD1		P_{pre}=0.7403	P_{npv}=0.7447	P_{acc}=0.7425
	Low	3586	1258	P_{sen}=0.7471
	High	1214	3542	P_{spe}=0.7379
ELM-TD2		P_{pre}=0.7792	P_{npv}=0.7586	P_{acc}=0.7712
	Low	1937	549	P_{sen}=0.8364
	High	379	1191	P_{spe}=0.6845

Note: The optimal values of the results on each database are marked in bold. P_{acc}, P_{sen}, P_{spe}, P_{pre}, and P_{npv} denote the values of accuracy, sensitivity, specificity, precision, and NPV, respectively.

Table 6
Average CPU time (in sec) for training and testing a MW classifier using EEG features from each database. The mean and standard deviation (s.d.) are computed for 10 repeated trials.

MW Classifier	TD1		TD2	
	Mean	s.d.	Mean	s.d.
LSSVM	36.22	0.56	5.38	0.19
ELM	1.47	0.09	0.66	0.04
NB	**0.52**	**0.05**	**0.23**	**0.03**
LR	1.54	0.08	0.80	0.04
KNN	17.88	0.18	3.31	0.02
ANN	145.98	*17.07*	51.79	2.79
PCA-LSSVM	35.73	0.57	5.02	0.12
PCA-ELM	1.53	0.06	0.71	0.07
PCA-NB	0.64	0.08	0.32	0.06
PCA-LR	0.84	0.06	0.43	0.07
PCA-KNN	7.32	0.11	1.38	0.04
PCA-ANN	59.18	0.33	22.38	0.27
SAE	193.74	0.34	101.47	0.80
DBN	75.26	1.29	38.50	0.42
CNN	150.16	3.67	89.32	1.08
TDAE	*496.35*	4.04	*394.80*	*8.67*

Note: The smallest and largest values in each column are marked in bold and italic styles, respectively.

achieves the highest mean value and outperforms the 2nd rank classifier, LSSVM, by 0.06 for both of TD1 and TD2. Note that the testing performance of LSSVM, NB, KNN, and LR based methods is unique since initial values are unnecessary. For the remaining neural networks, the gradient descent approach has made the trained model be related to the initial weights. In general, the *s.d.* value indicates the stability of the TDAE is superior to classical deep learning methods but worse than shallow classifiers except for ANN.

Table 7
Mean value of the subject-average testing accuracy for each MW classifier on two EEG databases. The mean and standard deviation (s.d.) are computed for 10 repeated trials. In each experiment trial, the initial parameters of the classifiers are randomized.

MW Classifier	TD1		TD2	
	Mean	s.d.	Mean	s.d.
LSSVM	0.8109	**0**	0.8407	**0**
ELM	0.7440	0.003	0.7759	0.006
NB	0.6799	**0**	0.5624	**0**
LR	0.7660	**0**	0.8092	**0**
KNN	0.7303	**0**	0.7707	**0**
ANN	0.7514	0.0125	0.7657	0.0693
PCA-LSSVM	0.7971	**0**	0.8277	**0**
PCA-ELM	0.7420	0.0023	0.7705	0.0060
PCA-NB	0.6484	**0**	0.6166	**0**
PCA-LR	0.7497	**0**	0,7818	**0**
PCA-KNN	0.7266	**0**	0.7724	**0**
PCA-ANN	0.7519	0.0016	0.7584	0.0659
SAE	0.6953	0.0460	0.6616	0.1047
DBN	0.7672	0.0255	0.7680	0.0298
CNN	0.6547	0.0215	0.6552	0.0214
TDAE	**0.8774**	0.0092	**0.9071**	0.0173

Note: The optimal value in each column is marked in bold.

Note that the *s.d.* value is also related to the size of the training database. That is, for TD2, the stability of all training algorithms are undermined (except for CNN) because of fewer training EEG instances.

4.6. TDAE compared with linear feature mapping approaches

To investigate the difference between TDAE and linear feature engineering methods, we compare TDAE against PCA and partial least square (PLS) approaches. For PCA and PLS based models, two linear mapping layers (i.e., the weight matrix is computed via PCA or PLS) are employed as feature and abstraction extractors, then we adopted two autoencoder (AE) layers for classifying mental workload levels because of the reduction of the dimensionalities. The corresponding two frameworks are denoted as PCAAE, and PLSAE, respectively. For each PCA layer, the number of the principle components are determined by a threshold of 0.9 for the total variance. For each PLS layer, the number of feature representations are selected via the optimal validating performance. Each algorithm was carried out for 10 trials under randomized initial weights. We illustrate the testing classification accuracy of three cases via box-whisker plots in Fig. 13. It is observed that the feature mapping of TDAE is superior to that of PLS and PCA. The potential reason lies in two as aspects. The linear feature mapping is very sensitive to the outliers and the mapping quality can be impaired by the noise in EEG features. Moreover, the employed EEG features are heterogeneous from multiple domains and are known to be nonlinearly correlated with the mental workload classes.

5. Discussions

The results shown in Section 4 indicate the proposed TDAE has the capability to accurately classify EEG feature vectors into binary MW levels. To give an insight into the mechanism of how TDAE improves the classification performance, we illustrate an input feature, a feature filter output and an abstraction filter output in Fig. 14. The EEG feature is selected as the theta power of subject A from the 1st session experiment. From the figure, it is clearly shown the feature filter output could track the variation of the original EEG indicator with outliers properly smoothed. The underlying reason is the feature values at both current and historical time steps are nonlinearly fused. Therefore, the drastic variation is weighted and averaged in a wider time interval.

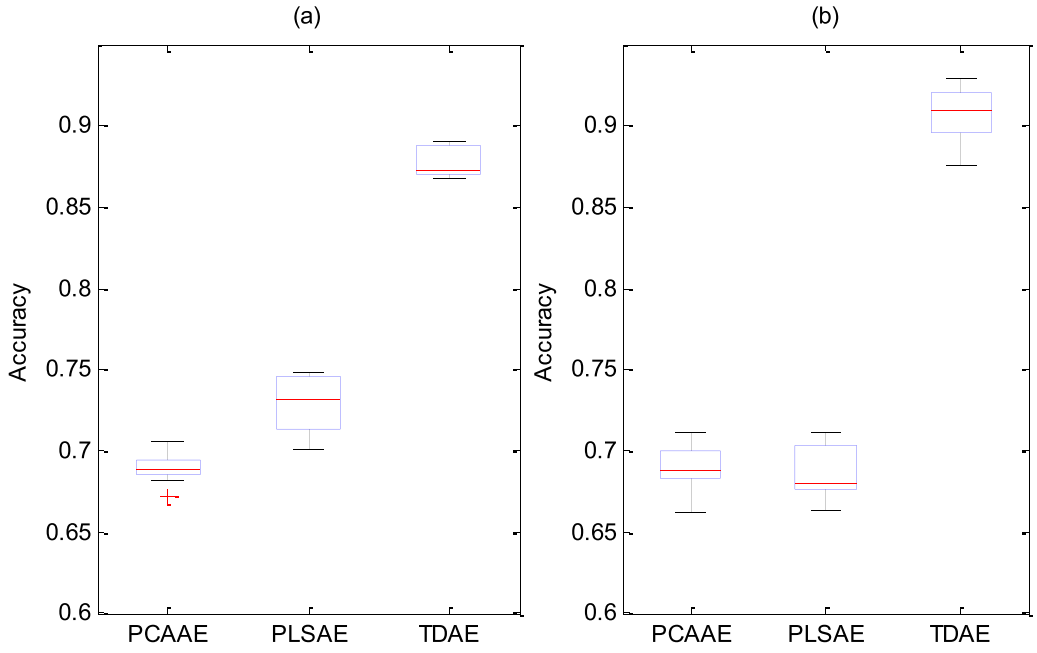

Fig. 13. Average testing classification accuracy of TDAE and the linear feature mapping based deep AE classifiers, i.e., PCAAE and PLSAE. For each column data, 10 trials of the experiment was performed with randomized initial weights.

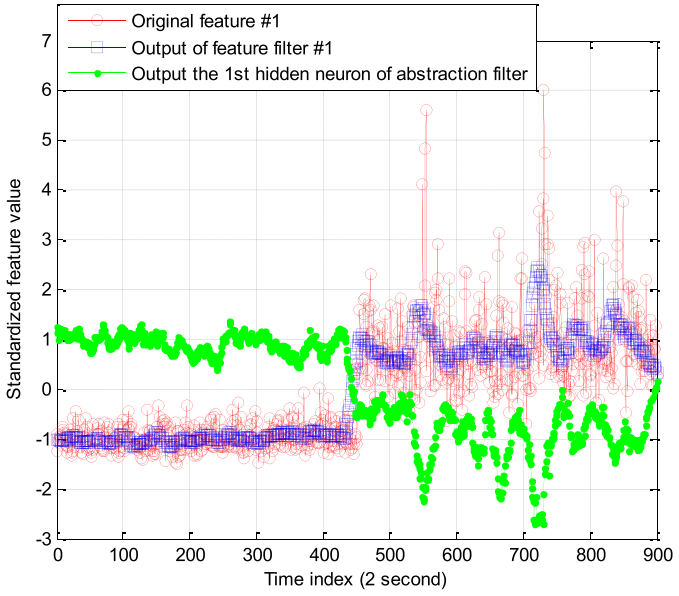

Fig. 14. Illustration of a TDAE input (original feature #1, the theta power of F3 channel), the output of the 1st feature filter and the activation of the 1st output neuron in the abstraction filter. The 30-min data of subject A from the 1st session are shown. The first 450 instances are from the low-MW level while the remaining corresponds to the high-MW class.

For the NB classifier, the prior probability of each class was computed via the ratio of the instances between two mental workload classes. However, low classification accuracy was achieved because of the potential noise in the EEG time sequences. Similar observation can be found for the KNN classifier, where the selected nearest neighbors can be contaminated by outliers. It is noticed the performance of ELM and ANN is comparable since the same network structure was employed. Among all shallow classifiers, LSSVM possesses the highest accuracy and F1-score. It is because the maximum margin classification is more generalizable against the noise for designing the decision hyper-plane. The limitation of the LSSVM in this study is its single layer structure that is inadequate for modeling the dynamical properties in EEG feature sequences.

Compared to the feature dimensionality (i.e., 137), the sizes of the training sets for TD2 (4062) is relatively inadequate. It leads to unsatisfactory performance of the classical deep learning methods as shown in Fig. 10. Note that the CNN model achieves the lowest performance since its convolution filers are not suitable for capturing local information across heterogeneous EEG indicators. For shallow models, their performance on TD1 (with 9600 training instances) is generally lower than that of TD2 (Figs. 8 and 9). It may be due to that the simple model structure without intermediate feature representations cannot perfectly tackle raw EEG feature vectors. For both of the TDAE and the compared mental workload classifiers, the choice of the hyper-parameters is critically important. In general, for a smaller database of TD2, a simpler model structure is preferable, e.g., the low value of k for KNN and the small number of the hidden nodes in ELM are more suitable as shown in Table 4.

Another advantage of TDAE is that the fine tuning stage of the classifier employed transfer learning principles. Depending on the results show in Figs. 5, 9 and 10, the generalization capability of the conventional deep classifiers (e.g., SAE and CNN) is not comparable with the shallow classifier like LSSVM. The reason behind is the limited number of training instances as well as the outlier problem cause the deep model to be easily overfitted in the supervised learning. The results in Fig. 11 imply it is possible to improve the model performance via EEG features of a different mental task. As the size of the training sample in TD2 is smaller than TD1, transferring EEG features is more effectual for the latter according to Fig. 11(b). It should be noted such scheme possesses the risk of inducing accuracy impairment if improper EEG data were selected. In Fig. 11(a), it is shown the classification rate was only achieved to 0.84 when 17,000 data points from SD were selected.

In our previous work, the average classification rate can reach 0.87 when an adaptive SAE was used for binary MW estimation [11]. In this study, optimal accuracies elicited by TDAE for TD1 and TD2 are 0.8623 and 0.8987, respectively. It is shown the average performance of the proposed method is competitive against our

previous one. Such MW recognition accuracy is also comparable with the newest work reported in [32], where 0.86 workload classification rate is achieved (note that the environment for implementing classifier is different in two works). However, the subject-specific classification paradigm is used in [11] while we trained subject-generic classifiers for TDAE. The latter can be more challenging since individual differences in the EEG data were to be properly tackled. Moreover, when the orders of feature and abstraction filters were exhausted by using the grid search, the TDAE performance can be further improved to 0.92 according to Fig. 12.

The main limitation of the proposed TDAE deep learning framework for MW recognition lies in two aspects. The computational cost for training the entire network is significantly higher than classical shallow and deep classifiers. It leads to high time cost in selection optimal hyper-parameters of the model. Therefore, we employed the same value of the feature filter order to reduce the computational burden. However, it is no doubt that the filer order should feature-specific. Moreover, there exists a prerequisite for knowledge transferring across two mental-task domains. That is, we need to select exactly the same EEG channels for data preprocessing and it leads to a possibility that useful MW indicators are excluded. In future work, we will further investigate the deep learning methods for MW assessment on these two aspects.

6. Conclusion

In the proposed TDAE framework for tackling the MW recognition issue, we first implement a group of autoencoder-based models to design multiple feature filters. The input of the feature filter consists of the EEG feature value of both current and historical time steps. Then, the optimal hidden activation with the minimum entropy is selected and fed to an abstraction filter for representing high-level dynamical properties of filtered EEG features. Finally, two SAEs were employed to transfer knowledge from a source-domain EEG database recorded under emotional stimuli to MW-related databases collected with varied operation complexity. The results show TDAE can achieve binary classification accuracies of 0.86 and 0.90 for two cases. The statistical analysis indicates the improvement is significant compared to conventional shallow and deep models with optimal hyper-parameters determined. The main contributions of TDAE is its effectiveness of the dynamical deep structure in exploiting EEG data from different tasks and subjects for MW recognition. The limitations of the study are the high computational cost and the restriction of the same paradigm for EEG channel settings when applying knowledge transferring.

Conflict of interest

None declared.

Acknowledgments

This work is sponsored by the National Natural Science Foundation of China under Grant No. 61703277 and the Shanghai Sailing Program (17YF1427000). This work is partially supported by the National Natural Science Foundation of China under Grant No. 61673276, No. 11502145 and the Shanghai Philosophy and Social Sciences Fund (2017EZX008).

References

[1] G. Hockey, Operator Functional State as a Framework for the Assessment of Performance Degradation, IOS press, G. Hockey, Italy II Ciocco, 2003.
[2] N. Hollender, C. Hofmann, M. Deneke, B. Schmitz, Integrating cognitive load theory and concepts of human-computer interaction, Comput. Hum. Behav. 26 (2010) 1278–1288.
[3] G. Borghini, L. Astolfi, G. Vecchiato, D. Mattia, F. Babiloni, Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness, Neurosci. Biobehav. Rev. 44 (2014) 58–75.
[4] E.P. Shaw, J.C. Rietschel, B.D. Hendershot, A.L. Pruziner, M.W. Miller, B.D. Hatfield, R.J. Gentili, Measurement of attentional reserve and mental effort for cognitive workload assessment under various task demands during dual-task walking, Biol. Psychol. 134 (2018) 39–51.
[5] Y. Ke, H. Qi, L. Zhang, S. Chen, X. Jiao, P. Zhou, X. Zhao, B. Wan, D. Ming, Towards an effective cross-task mental workload recognition model using electroencephalography based on feature selection and support vector machine regression, Int. J. Psychophysiol. 98 (2015) 157–166.
[6] Y. Zhang, Y. Wang, G. Zhou, J. Jin, B. Wang, A.Cichocki X., Multi-kernel extreme learning machine for EEG classification in brain-computer interfaces, Expert Syst. Appl. 96 (2018) 302–310.
[7] M.N. Anastasiadou, M. Christodoulakis, E.S. Papathanasiou, S.S. Papacostas, G.D. Mitsis, Unsupervised detection and removal of muscle artifacts from scalp EEG recordings using canonical correlation analysis, wavelets and random forests, Clin. Neurophysiol. 128 (2017) 1755–1769.
[8] J. Li, Z. Struzik, L. Zhang, A. Cichocki, Feature learning from incomplete EEG with denoising autoencoder, Neurocomputing 165 (2015) 23–31.
[9] S.-K. Kim, H.-B. Kang, An analysis of smartphone overuse recognition in terms of emotions using brainwaves and deep learning, Neurocomputing 275 (2018) 1393–1406.
[10] Z. Jiao, X. Gao, Y. Wang, J. Li, H. Xu, Deep convolutional neural networks for mental load classification based on EEG data, Pattern Recogn. 76 (2018) 582–595.
[11] Z. Yin, J. Zhang, Cross-session classification of mental workload levels using EEG and an adaptive deep learning model, Biomed. Signal Process. 33 (2017) 30–47.
[12] Z. Yin, J. Zhang, Task-generic mental fatigue recognition based on neurophysiological signals and dynamical deep extreme learning machine, Neurocomputing 283 (2018) 266–281.
[13] Z. Yin, M. Zhao, Y. Wang, J. Yang, J. Zhang, Recognition of emotions using multimodal physiological signals and an ensemble deep learning model, Comput. Methods Progr. Biomed. 140 (2017) 93–110.
[14] Z. Wang, R.M. Hope, Z. Wang, Q. Ji, W.D. Gray, Cross-subject workload classification with a hierarchical Bayes model, NeuroImage 59 (2012) 64–69.
[15] C. Baldwin, B. Penaranda, Adaptive training using an artificial neural network and EEG metrics for within and cross-task workload classification, NeuroImage 59 (2012) 48–56.
[16] Z. Yin, J. Zhang, Identification of temporal variations in mental workload using locally-linear-embedding-based EEG feature reduction and support-vector-machine-based clustering and classification techniques, Comput. Methods Progr. Biomed. 115 (2014) 119–134.
[17] L. Zhang, J. Yang, D. Zhang, Domain class consistency based transfer learning for image classification across domains, Inform. Sci. 418-419 (2017) 242–257.
[18] S. Koelstra, C. Muehl, M. Soleymani, J.-S. Lee, A. Yazdani, T.E. Ebrahimi, T. Pun, A. Nijholt, I.Y. Patras, DEAP: a database for emotion analysis using physiological signals, IEEE Trans. Affect. Comput. 3 (2012) 18–31.
[19] E. Byrne, R. Parasuraman, Psychophysiology and adaptive automation, Biol. Psychol. 42 (1996) 249–268.
[20] D. Kaber, R. Endsley, Level of automation effects on performance, situation awareness and workload in a dynamic control task, Ergonomics 42 (1999) 462–492.
[21] C. Lin, T. Yenn, C. Yang, Automation design in advanced control rooms of the modernized nuclear power plants, Saf. Sci. 48 (2010) 63–71.
[22] D. Kaber, C. Perry, N. Segall, C. McClernonc, L. Prinzel, Situation awareness implications of adaptive automation for information processing in an air traffic control-related task, Int. J. Ind. Ergonom. 36 (2006) 447–462.
[23] R. Parasuraman, M. Mouloua, R. Molloy, Effects of adaptive task allocation on monitoring of automated systems, Hum. Factor 38 (1996) 665–679.
[24] A. Pope, E. Bogart, D. Bartolome, Biocybernetic system evaluates indices of operator engagement in automated task, Biol. Psychol. 40 (1995) 187–195.
[25] F. Freeman, P. Mikulka, L. Prinzel, M. Scerbo, Evaluation of an adaptive automation system using three EEG indices with a visual tracking task, Biol. Psychol. 50 (1999) 61–76.
[26] A. Haarmann, W. Boucsein, F. Schaefer, Combining electrodermal responses and cardiovascular measures for probing adaptive automation during simulated flight, Appl. Ergon. 40 (2009) 1026–1040.
[27] A. Hoover, A. Singh, S. Fishel-Brown, E. Muth, Real-time detection of workload changes using heart rate variability, Biomed. Signal Process. 7 (2012) 333–341.
[28] J. Noel, K. Bauer, J. Lanning, Improving pilot mental workload classification through feature exploitation and combination: a feasibility study, Comput. Oper. Res. 32 (2005) 2713–2730.
[29] C. Russell, G. Wilson, Applications of Artificial Neural Networks for Air Traffic Controller Functional State Classification, Wright-Patterson AFB: United States Air Force Research Laboratory, Dayton, Ohio, 2001.
[30] A. Kumar, A real-time system for pattern recognition of human sleep stages by fuzzy system analysis, Pattern Recogn. 9 (1977) 43–46.
[31] I. Guler, E.D. Ubeyli, Adaptive neuro-fuzzy inference system for classification of EEG signals using wavelet coefficients, J. Neurosci. Methods 148 (2005) 113–121.
[32] J. Fan, J.W. Wade, A.P. Key, Z.E. Warren, N. Sarkar, EEG-based affect and workload recognition in a virtual driving environment for ASD intervention, IEEE Trans. Biomed. Eng. 65 (2018) 43–51.

[33] Z. Yin, J. Zhang, Operator functional state classification using least-square support vector machine based recursive feature elimination technique, Comput. Methods Prog. Bio. 113 (2014) 101–115.
[34] O. Faust, Y. Hagiwara, T.J. Hong, O.S. Lih, U.R. Acharya, Deep learning for healthcare applications based on physiological signals: a review, Comput. Methods Prog. Bio. 161 (2018) 1–13.
[35] J. Christensen, J. Estepp, G. Wilson, C. Russell, The effects of day-to-day variability of physiological data on operator functional state classification, NeuroImage 59 (2012) 57–63.
[36] G.N. Dimitrakopoulos, I. Kakkos, Z. Dai, J. Lim, J.J. deSouza, A. Bezerianos, Y. Sun, Task-independent mental workload classification based upon common multiband EEG cortical connectivity, IEEE Trans. Neur. Syst. Rehabil. 25 (2017) 1940–1949.
[37] Z. Yin, J. Zhang, Cross-subject recognition of operator functional states via EEG and switching deep belief networks with adaptive weights, Neurocomputing 260 (2017) 349–366.
[38] Z. Li, J. Liu, J. Tang, H. Lu, Robust structured subspace learning for data representation, IEEE Trans. Pattern Anal. 37 (2015) 2085–2098.
[39] Z. Li, J. Tang, T. Mei, Deep collaborative embedding for social image understanding, IEEE Trans. Pattern Anal. (2018) in press, doi:10.1109/TPAMI.2018.2852750.
[40] J. Sauer, P. Nickel, D. Wastell, Designing automation for complex work environments under different levels of stress, Appl. Ergon. 44 (2013) 119–127.
[41] G.E. Hinton, Reducing the dimensionality of data with neural networks, Science 313 (2006) 504–507.
[42] J.A.K. Suykens, J. Vandewalle, Least squares support vector machine classifiers, Neural Process. Lett. 9 (1999) 293–300.
[43] G.-B. Huang, Q.-Y. Zhu, C.-K. Siew, Extreme learning machine: theory and applications, Neurocomputing 70 (2006) 489–501.

Zhong Yin received his Ph.D. in control science and engineering from the East China University of Science and Technology. He has been a Lecture at School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, China, from 2015 to 2017. He has been an Associate Professor at School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, China, since 2018. His research interests include intelligent human-machine systems, biomedical signal processing and pattern recognition.

Mengyuan Zhao received her Ph.D. in logic from Zhejiang University. She has been a Lecturer at School of Marxism, University of Shanghai for Science and Technology, China, since 2015. Her research interests include logic of language, philosophy of mind and cognitive science.

Wei Zhang received the B.S. degree from the Department of Mathematics, Shaanxi Normal University, Xi'an, China, in 2003, the M.S. degree in operation research and cybernetics from Northwestern University, Shenyang, China in 2008, and the Ph.D. degree in control theory and engineering from Shanghai Jiao Tong University, Shanghai, China, in 2014. He is currently an Associate Professor in School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai, China. His research interests focus on robust control, optimal control, decoupling control, and their applications in industry.

Yongxiong Wang received his Ph.D. in Department of Automation, Shanghai JiaoTong University in 2012. He is a professor in School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology now. His research interests include the intelligent robot and robot vision. His research interests include the design, control of specialized robot, and robot vision.

Yagang Wang received his B.E. degree in electrical engineering from China University of Mining and Technology, China, in 1988, the M.E. degree in control theory and application from Taiyuan University of Technology, China, in 1991, and the Ph.D. degree in control theory and control engineering from Shanghai Jiao Tong University, China, in 2000. He worked as a Research Fellow in Nanyang Echnological University, Singapore, from 2000 to 2002. He worked as Senior Engineer in Honeywell (China) Co., Ltd. from 2002 to 2007. He is currently a Professor in University of Shanghai for Science and Technology at Shanghai, China. His research interests include process automation, system identification, wireless sensor network.

Jianhua Zhang received his Ph.D. in Electrical Engineering from Ruhr-Universität Bochum, Germany in 2005. Between 2005 and 2006 he was a Research Associate at Intelligent Systems Research Laboratory, The University of Sheffield, UK. From 2007 to 2017 he was a Professor and Head of Intelligent Systems Group at School of Information Science and Engineering, East China University of Science and Technology (ECUST), Shanghai, China. In 2017 he joined VEKIA (Lille, France) as R&D Scientific Director and Head of Machine Learning Research Lab. Since Sep. 2018 he has been a Professor of AI at Applied AI Group, Department of Computer Science, Oslo Metropolitan University (OsloMet), Norway. He also held research positions of Guest Scientist at Technische Universität Dresden, Germany from 2002 to 2003, invited Adjunct Professor at Institute for Cognitive Neurodynamics of ECUST from 2009 to 2017, and Visiting Professor at Technische Universität Berlin and Max Planck Institute for Dynamics of Complex Technical Systems, Germany, between 2008 and 2015. His research interests are in computational intelligence, machine learning, intelligent systems and control, intelligent data modeling and analysis, pattern recognition, biomedical signal processing, modeling and control of complex engineering and biomedical systems, human-machine systems, brain-machine interaction, computational neuroergonomics, and cognitive neuroengineering. So far he has published one research monograph, three translated books, and more than 110 referred journal and conference proceedings papers.