Assessing cognitive mental workload via EEG signals and an ensemble deep learning classifier based on denoising autoencoders

Shuo Yang[a,b], Zhong Yin[a,b,*], Yagang Wang[a,b], Wei Zhang[a,b], Yongxiong Wang[a,b], Jianhua Zhang[c]

[a] School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai, 200093, PR China
[b] 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
[c] OsloMet Artificial Intelligence Lab, Department of Computer Science, Oslo Metropolitan University, Oslo, N-0130, Norway

ABSTRACT

To estimate the reliability and cognitive states of operator performance in a human-machine collaborative environment, we propose a novel human mental workload (MW) recognizer based on deep learning principles and utilizing the features of the electroencephalogram (EEG). To determine personalized properties in high dimensional EEG indicators, we introduce a feature mapping layer in stacked denoising autoencoder (SDAE) that is capable of preserving the local information in EEG dynamics. The ensemble classifier is then built via the subject-specific integrated deep learning committee, and adapts to the cognitive properties of a specific human operator and alleviates inter-subject feature variations. We validate our algorithms and the ensemble SDAE classifier with local information preservation (denoted by EL-SDAE) on an EEG database collected during the execution of complex human-machine tasks. The classification performance indicates that the EL-SDAE outperforms several classical MW estimators when its optimal network architecture has been identified.

1. Introduction

Due to inadequate progress in artificial intelligence, a group of sophisticated automation technologies needs to be implemented alongside human operations. It is widely recognized that human operators become an important component in such human-machine (HM) systems [1]. In a HM operational environment, human mental workload (MW) monitoring is crucial for safety-critical tasks such as driving and operating flights, nuclear power plants, and manned-space flights [2]. While there is no universally accepted definition of MW, it is known as the transient human cognitive capacity or resource required to process a specific cognitive task [3]. The degree of MW is also linked to vigilance and situational awareness to indicate the capability to handle emergencies [4].

A high level of MW represents the operator's high-risk operation state, and implies a high possibility of leading to serious HM system error. Aim at maximizing the HM collaboration efficiency, a moderate level of MW is necessary to maintain optimal operator functional states (OFSs) [5]. During certain mental tasks, the required cognitive resources attempt to match the amount of mental effort [6]. Unlike machines and/or computers, the OFS of an operator does not always meet the requirements of the task because of limited physiological experience

and working memory [4]. Since increased demand for mental resources results in performance degradation [7] or even catastrophic accidents, MW assessment is a significant requirement in human-centered HM systems.

The approaches for assessing MW mainly include subjective ratings, human performance and physiological measures. In a recently reported work, physiological measures were deemed attractive because they can be implemented with the least interference to the main task [8]. Among the various physiological measures, electroencephalogram (EEG) is one of the most commonly used techniques for recording brain activity [9]. Electroencephalogram signals have the capability to estimate human cognitive states in real time. To classify EEG signals into accurate MW levels, the machine learning-based approach has received much attention. However, the high dimensionality of EEG features may reduce the training stability of a data-driven model. To tackle this problem, we introduce deep neural networks to identify useful intermediate EEG features that indicate distinguishable cognitive states. Different from the feed-forward neural network with a single hidden layer, a deep network is constructed by using multiple hidden layers aimed at the abstracted combination of the salient EEG indicators.

The main purpose of this study is to find a personalized deep network mapping between the psychophysiological features of, and the

---

inherent variations in, MW levels. However, there are two issues that limit the generalization capability for implementing a deep learning model. (1) The noise induced by movements and muscular activities contaminate the useful information in EEG features. (2) The data distribution of the EEG indicators varies across different individuals. To remove the noise components in multidimensional features more successfully, we introduce the stacked denoising autoencoder (SDAE) [10] as a feature filter to discover the stable EEG variables at higher levels. Moreover, individual-specific MW classifiers are developed via the ensemble learning principle to alleviate cross-subject feature variations. Thus, the personalized ensemble deep learning committee is built to adapt cognitive properties to a certain human operator.

To improve the accuracy of classification further, we redesign the shallow hidden layer of the member SDAE to a feature mapping layer with the capability of preserving the local information in EEG dynamics. Motivated by Refs. [11,12], the individual-specific properties in EEG are expected to be better extracted and fused to reflect the MW levels. Finally, a new MW classification framework, an ensemble SDAE classifier with local information preservation (denoted by EL-SDAE), is proposed and validated in this work. Aiming at MW classification, we also investigate and determine the optimal number of hidden layers and hidden neurons for the EL-SDAE. The final results computed on the basis of EEG signals collected from real-world HM tasks are compared against several competitive deep and shallow MW classifiers to demonstrate the effectiveness of the EL-SDAE.

This paper is structured as follows. Section 2 briefly reviews related works on physiological signal-based MW recognition. In Section 3, the proposed EL-SDAE and details of the EEG databases used for validating the MW classifier are presented. In Sections 4 and 5, we present the results and corresponding discussions, respectively. Finally, conclusions are presented in Section 6.

2. Related works

Human physiological signals, such as EEG, electrocardiogram (ECG), and electrooculogram, can reflect hidden information related to MW patterns [13]. Specifically, multi-channel EEG signals indicate advanced cognitive functions such as reasoning and judgment. As the corresponding brain network can become very active during cognitive tasks, EEG has been implemented as an important tool for assessing MW [14]. Based on the entire EEG power spectrum, the electrophysiological activity of the cortex can be interpreted into several frequency bands, which are known as delta, theta, alpha, beta, and gamma [15].

Among the related works, Orlandi et al. employed power spectral densities obtained from EEG signals to investigate the effects of ship handling maneuvers on MW in 10 marine pilots [16]. Fallahi et al. utilized the subjective workload assessment technique accompanied by EEG to analyze the MW of traffic control operators. The results indicated that operators experience greater MW at high traffic density than at lower levels [17]. Wang et al. showed that the wireless acquisition of EEG signals can be utilized to classify different memory workload levels [18]. Liu et al. conducted an experiment using a virtual ship bridge simulator. The EEG signals were rerecorded to study the MWs of the crew members during cooperative tasks [19]. Based on EEG data and therapist's evaluation during driving skill training, Fan et al. established a group-level classification model to distinguish the emotional states and MW of patients in the autism spectrum disorder during driving skill training [20]. Chen et al. demonstrated the validity of EEG signals in evaluating the MW of construction workers and described the development of a wearable EEG safety helmet prototype [21]. Puma et al. used EEG signals to assess the increase in MW in multitasking environments. The results showed that the consistent difference between the power levels in the alpha and theta bands is related to the level of task performance [22].

Machine-learning based pattern recognition techniques can abstract massive physiological data into a structural model automatically. Such

a model is capable of discriminating between different MW levels [4]. In related works, Cui et al. oversampled EEG data in minority class and classified MW using a dynamic multilayer perceptron [23]. Zarjam et al. investigated workload discrimination using EEG signals and achieved high accuracy in subject-independent multichannel classification [24]. Wei et al. introduced EEG-based MW recognition related to multitasking [25].

Machine learning has gained popularity in various industrial, medical, and commercial domains while deep learning is one of the most competitive branches [10,26]. Among the popular primitives of deep learning approaches, SDAE has performed well in a wide range of domains, such as fault diagnosis and natural language analysis [27,28]. Each hidden layer of a SDAE model is an autoencoder, i.e., a symmetric neural network with a single hidden layer, while the output and input neurons are the same. The autoencoder adopts the unsupervised learning principle wherein the outputs approximate to the inputs [29]. A special form of the autoencoder is called the denoising autoencoder (DAE) [30]. It minimizes the reconstruction loss of corrupted input features and learns robust representation of the noise inputs. The training procedure is accomplished by using the back-propagation (BP) algorithm. Following greedy layer-wise training for SDAE, a deep network can be generated based on supervised fine-tuning [31].

3. Methods

3.1. EEG database for MW assessment

We employed an EEG database for validating the MW classifiers built in our previous work [4]. In the laboratory environment, participants were instructed to perform a simulated HM task based on automatic enhanced cabin air management system (ACAMS) [32,33] with physiological signals simultaneously recorded under different task loads. The ACAMS platform provided a micro-world in a spaceflight. When a fault occurred in any parameter of the automatic control system, the operator was required to fix the error manually. A total of eight healthy subjects (A, B, C, D, E, F, G and H) were involved. Two sessions of the same experimental task were carried out with each subject.

The variations in the difficulty levels of each session were designed via a cyclic-loading schedule of eight consecutive phases. The first and the last phases were the baseline conditions of 5 min each. The six consecutive task-load conditions were programmed with 1, 3, 4, 4, 3, and 1 failures (denoted as the number of failed subsystems i.e., NTF) in ACAMS. Each task load condition lasted for 15 min. In this study, the control conditions NTF = 1 and NTF = 4 were chosen for analysis and labeled as low and high MW levels (denoted by LMW and HMW), respectively.

Eleven channels of simultaneous EEG signals were recorded at a sampling frequency of 500 Hz at the positions of F3, F4, Fz, C3, C4, Cz, P3, P4, Pz, O1, and O2 [34]. A third-order Butterworth filter with a low-pass cutoff frequency of 40 Hz was used to eliminate high-frequential artifacts from the raw data. Independent component analysis was employed to correct the ocular artifacts. Fast Fourier transform with a frequential resolution of 0.5 Hz was adopted to calculate the power spectral density (PSD) in each 2 s, non-overlapping EEG segment [35].

For each EEG segment in all the channels, four PSD features, i.e., the average power in the theta (4–8 Hz), alpha (8–13 Hz), beta (14–30 Hz), and gamma (31–40 Hz) bands, were computed. The power differences between four channel pairs —F3-F4, C3-C4, P3-P4, and O1-O2 — were also elicited. In addition, we extracted EEG temporal features including mean, variance, zero crossing rate, Shannon entropy, spectral entropy, kurtosis, and skewness. The data matrix of each session is of the size 1800×137, where 1800 and 137 denote the number of EEG feature vectors and the number of features, respectively. It is noted the first and the last 450 instances (control condition NTF = 1) were labeled as the LMW class while the remaining were labeled as the HMW class (control
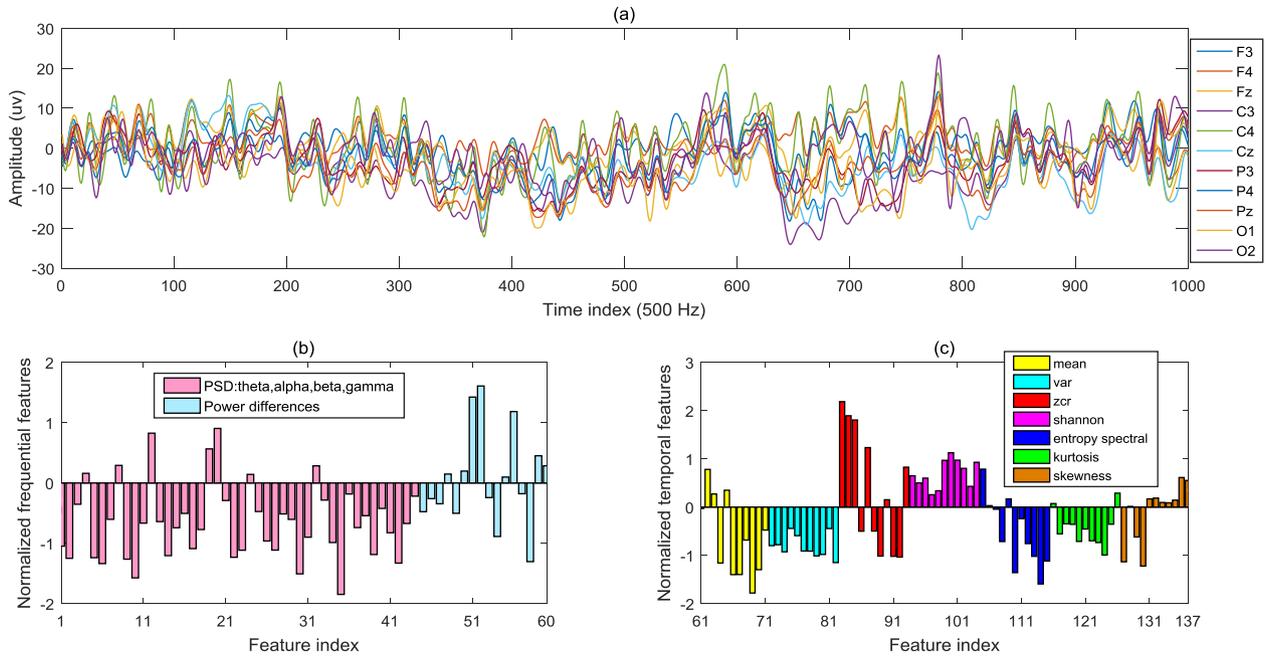
Fig. 1. Filtered EEG data in (a) a 2 s segment of 11 channels, (b) the corresponding 60 frequential features and (c) the corresponding 77 temporal features. Power spectral density, variance, and zero crossing rate are denoted by PSD, var, and zcr in the legend, respectively.

condition NTF = 4). For each matrix, each column of EEG data was separately standardized (z-scored) to ensure that all the EEG features were of equal significance. Fig. 1 depicts a 2 s segment of filtered EEG signals of subject B and the corresponding EEG features.

3.2. Ensemble SDAE with local information preservation

3.2.1. Stacked denoising autoencoder

The competence of the SDAE benefits from its deep structure-based multiple hidden layers [36]. The outline of the SDAE for EEG modeling is illustrated in Fig. 2, wherein the numbers of input and output neurons

Fig. 2. Architecture of SDAE for EEG modeling.

are 137 and 2, respectively. The basic element of SDAE is the auto-encoder (AE) [37] for nonlinear feature dimensionality reduction. In particular, an AE model is indeed a single-hidden-layer neural network that shares the same inputs and outputs. Considering the issue of MW classification, let us define the input and hidden activations of an AE as $\mathbf{x} \in R^D$ and $\mathbf{x}_h \in R^d$, respectively. Here, $\mathbf{x}$ and $\mathbf{x}_h$ are the input EEG features, their abstractions with D and d being the corresponding dimensionalities. The transformation from $\mathbf{x}$ to $\mathbf{x}_h$ is achieved by

$$\mathbf{x}_h = s(\mathbf{W}\mathbf{x} + \theta), \tag{1}$$

where $\mathbf{W} \in R^{D \times d}$ and $\theta \in R^d$ denote the input weight matrix and bias vector, respectively. The function $f(\cdot)$ is the logistic sigmoid function denoted by $s(\mathbf{z}) = 1/(1 + e^{-\mathbf{z}})$. The reconstructed output $\mathbf{x}_o \in R^D$ is generated by the following mapping from hidden to output neurons, i.e.,

$$\mathbf{x}_o = s(\mathbf{W}^T\mathbf{x}_h + \tilde{\theta}) = \frac{1}{1 + e^{-(\mathbf{W}^T\mathbf{x}_h + \tilde{\theta})}} \tag{2}$$

The BP algorithm for optimizing the squared-error cost function is used to train the AE parameters $\mathbf{W}$, θ, and $\tilde{\theta}$.

$$L(\mathbf{x}, \mathbf{x}_o) = \sum_{i=1}^{D} |x^{(i)} - x_o^{(i)}|^2. \tag{3}$$

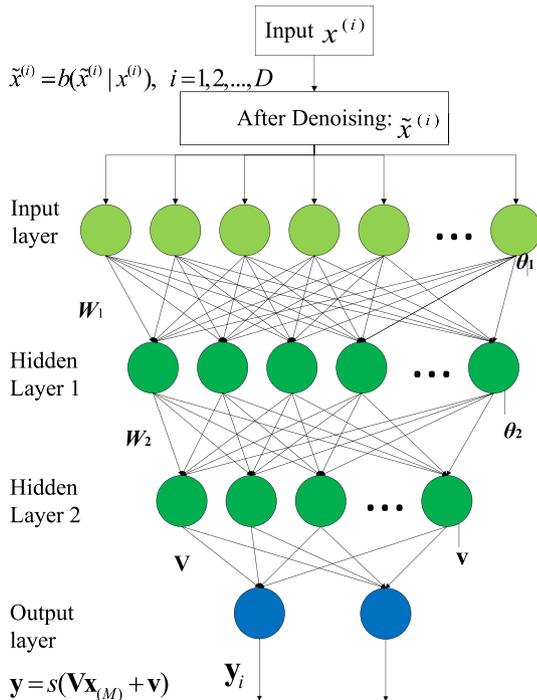The trained parameters are defined as

$$\{\mathbf{W}^*, \theta^*, \tilde{\theta}^*\} = \arg\min \frac{1}{N} \sum_{i=1}^{N} L\{\mathbf{x}_i, s[\mathbf{W}^T s(\mathbf{W}\mathbf{x}_i + \theta) + \tilde{\theta}]\}. \tag{4}$$

In Eqn. (4), N represents the size of training EEG instances. The trained AE model consists of its input and hidden layers, $\mathbf{x}_{(1)} = s(\mathbf{W}^*\mathbf{x} + \theta^*)$. Let us denote $\mathbf{x}_{(1)}$ as the input of another AE. High-level representations of the EEG feature abstractions $\mathbf{x}_{(m)}$ could be hierarchically extracted by stacking AE (SAE) structures.

$$\mathbf{x}_{(M)} = s(\mathbf{W}_M^* \dots s(\mathbf{W}_2^* s(\mathbf{W}_1^* \mathbf{x} + \theta_1^*) + \theta_2^*) \dots + \theta_M^*). \tag{5}$$

By adding an output layer with two nodes that correspond to binary MW levels, the deep SAE network is derived as

$$\mathbf{y} = s(\mathbf{V}\mathbf{x}_{(M)} + \mathbf{v}), \quad \mathbf{y} = s(\mathbf{W}\mathbf{x}_{(M)} + \theta) \tag{6}$$

Table 1
Pseudo codes for training a SDAE.

	Function:	SDAE_train
	Inputs:	Training datasets $\{\mathbf{X}, \mathbf{Y}\} \mathbf{x} \in R^D$, $\mathbf{x}_h \in R^d$
		Number of the hidden layers M
	Outputs:	SDAE parameters $\mathbf{W}_m^*, \theta_m^*, \mathbf{V}^*, \mathbf{v}^*$
1		Randomly assigned $\mathbf{W}_m, \theta_m, \mathbf{V}, \mathbf{v}$
2		**For** $m = 1: M$
3		**For** $i = 1: D$
4		$\tilde{x}_{(m)}^{(i)} = b\left(\tilde{x}_{(m)}^{(i)} \mid x_{(m)}^{(i)}\right)$
5		**End For**
6		Compute $\mathbf{x}_{(m)} = s[\mathbf{W}_m \mathbf{x}_{(m-1)} + \theta]$
7		Compute $\mathbf{W}_m^*, \theta_m^*$ according to Eqn. (4)
8		**End For**
9		Compute $\mathbf{y} = s(\mathbf{V}\mathbf{x}_{(M)} + \mathbf{v})$
10		Compute $\mathbf{W}_m^*, \theta_m^*, \mathbf{V}^*, \mathbf{v}^*$ according to Eqn. (4)
11		**Return** $\mathbf{W}_m^*, \theta_m^*, \mathbf{V}^*, \mathbf{v}^*$

Table 2
Pseudo codes for training the EL-SDAE.

	Function:	EL-SDAE
	Inputs:	Dataset
		Number of weak learners
	Outputs:	Strong classifier $C_T(X)$C(X).
1		Compute α via $\mathbf{X}\mathbf{L}\mathbf{X}^T\alpha = \lambda \mathbf{X}\mathbf{B}\mathbf{X}^T\alpha$
2		Generate mapping $\mathbf{A}$ from α
3		**For** $t = 1: T$.
4		Generate $S_t = \{\mathbf{X}_t, \mathbf{Y}_t\}$ from S via bootstrap
5		Randomly assigned $\mathbf{W}_{1,t}, \theta_{1,t}, \mathbf{W}_{2,t}, \theta_{2,t}\mathbf{V}_t, \mathbf{v}_t$
6		Compute $\tilde{\mathbf{x}}_o = s[\mathbf{W}_{1,t} \cdot b(\tilde{\mathbf{x}} \mid \mathbf{A}\mathbf{x}) + \theta_{1,t}]$
7		Compute $L(\mathbf{A}\mathbf{x}, \tilde{\mathbf{x}}_o) = \|\mathbf{A}\mathbf{x} - \hat{\mathbf{x}}_o\|_2^2$
8		Determine $\mathbf{W}_{1,t}^*$ and $\theta_{1,t}^*$ via Eqn. (16)
9		Compute $\mathbf{h}_{1,t} = s[\mathbf{W}_{1,t}^* \cdot \mathbf{A}\mathbf{x} + \theta_{1,t}^*]$
10		**Call SDAE_train** on $\mathbf{h}_{1,t}$ with $M = 1$
11		**Return** $\mathbf{W}_{2,t}^*, \theta_{2,t}^*, \mathbf{V}^*, \mathbf{v}^*$
12		Build $C_t(\cdot)$ on all $\mathbf{W}_t^*, \theta_t^*, \mathbf{V}_t^*, \mathbf{v}_t^*$
13		**End for**
14		Return $C_T(\mathbf{x}) = \text{sign}[\sum_{t=1}^T C_t(\mathbf{x})]$

where $\mathbf{y} = [1 \ \ 0]^T$, $\mathbf{y} = [0 \ \ 1]^T$, $\mathbf{V}$, and $\mathbf{v}$ represent the LMW class, HMW class, output weight matrix, and bias vector, respectively.

To elicit stable feature abstractions, a small portion of the input features could be artificially corrupted by a random mapping.

$$\tilde{x}^{(i)} = b(\tilde{x}^{(i)} \mid x^{(i)}), \quad i = 1,2, \ldots, D \tag{7}$$

where the probability from a uniform distribution is used to set $x^{(i)}$ to 0. By introducing Eqn. (7), the SDAE is generated to elicit noise-free EEG feature representations for MW assessment. The pseudo codes of the SDAE training algorithm are summarized in Table 1, wherein the pre-training and fine-tuning are performed in lines 7 and 10, respectively.

3.2.2. Ensemble SDAE with local information preservation

To improve the MW classification accuracy of participant-specific EEG features further, we proposed the EL-SDAE framework shown in Fig. 3. The EL-SDAE employs the principle of ensemble learning that integrates multiple individual learners into a model committee to perform complex classification tasks [38]. The base learner is designed based on SDAE and shares a homogeneous network structure while Bagging is used to construct the classification committee. It is noted that the SDAE mapping only has the ability to find stable hidden variables among the global feature distribution. To improve the discrimination capability of the weak model further, we incorporate an additional hidden layer into the member SDAE to preserve salient local information by using locality preserving projection (LPP) [39]. A three-layer deep SDAE (denoted by local SDAE) is combined with the Bagging ensemble method to form the final EL-SDAE MW classifier.

In the first hidden layer of the base learner, the input EEG features $\mathbf{X} = \{\mathbf{x}_i \in R^{D_1}, i = 1, 2, \ldots, N\}$ are represented via a low dimensional embedding by using a linear mapping.

$$\mathbf{a}_i = \mathbf{A}\mathbf{x}_i, i = 1,2, \ldots, N \tag{8}$$

To find the transform matrix $\mathbf{A}$, we define a neighborhood graph H on the feature set $\mathbf{X}$. To preserve the local information of the hidden
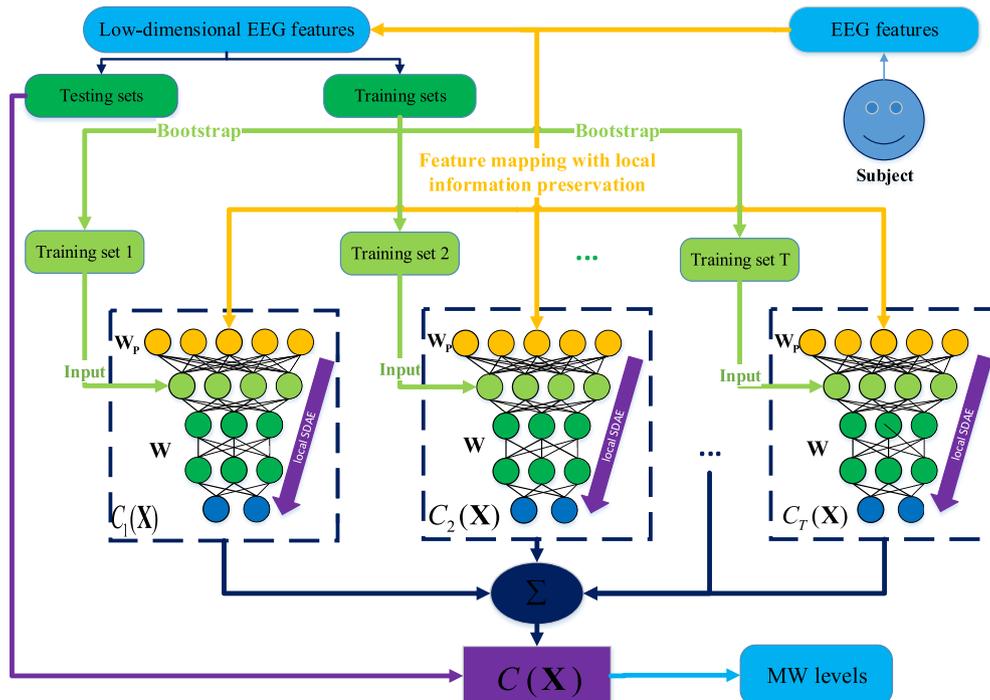

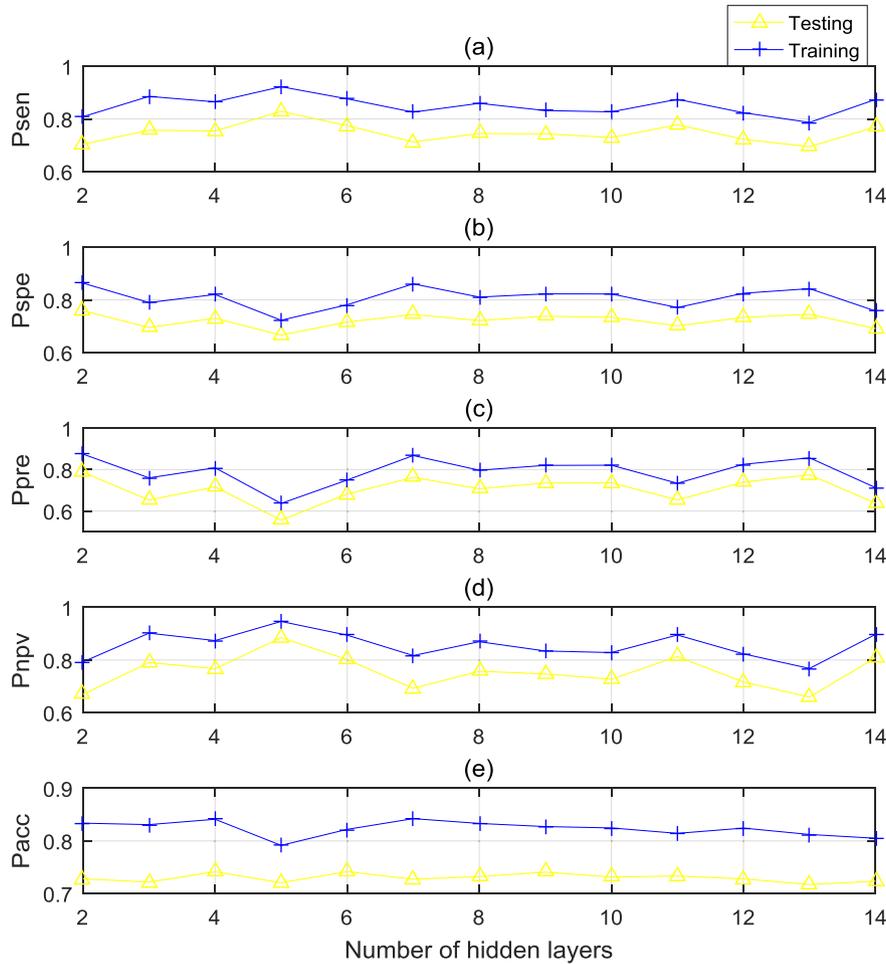**Fig. 3.** Architecture of the EL-SDAE model.

Fig. 4. Participant-average MW classification performance vs. the number of hidden neurons in a SDAE model with 70 hidden neurons in each hidden layer.

distribution in $\mathbf{X}$, we use S_{ij} to weight the edge of each of the two EEG feature vectors, $\mathbf{x}_i$ and $\mathbf{x}_j$. The value of $S_{ij} = 1$ (or $S_{ij} = 0$) indicates that the two feature vectors are (or are not) neighbors, i.e., belonging to the same class. The following objective function of the low dimensional embedding is then optimized [39].

$$\min \Phi = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \|\mathbf{z}_i - \mathbf{z}_j\|_2^2 S_{ij} \tag{9}$$

In Eqn. (9), if the neighboring points $\mathbf{x}_i$ and $\mathbf{x}_j$ are mapped far apart, the weighting function will impose a heavy penalty to ensure that the mapping of two nearest neighbors in the original space is such that they are also close in the projection space.

The objective function shown in Eqn. (9) can be rearranged as

$$\Phi = \sum_{i=1}^{N} \mathbf{A}\mathbf{x}_i B_{ii} \mathbf{x}_i^T \mathbf{A}^T - \sum_{i=1}^{N} \sum_{j=1}^{N} \mathbf{A}\mathbf{x}_i P_{ij} \mathbf{x}_j^T \mathbf{A}^T = \mathbf{A}\mathbf{X}\mathbf{L}\mathbf{X}^T \mathbf{A}^T. \tag{10}$$

In Eqn. (10), $\mathbf{B}$ is a diagonal matrix whose elements B_{ii} are the sum of the columns of $\mathbf{S}$, and $\mathbf{L} = \mathbf{B} - \mathbf{P}$ is the Laplacian matrix with the constraint $\mathbf{A}\mathbf{X}\mathbf{B}\mathbf{X}^T\mathbf{A}^T = 1$. By solving the generalized eigenvalue problem, the transform matrix $\mathbf{A}$ can be computed via Eqn. (11).

$$\mathbf{X}\mathbf{L}\mathbf{X}^T\boldsymbol{\alpha} = \lambda \mathbf{X}\mathbf{B}\mathbf{X}^T\boldsymbol{\alpha} \tag{11}$$

In Eqn. (11), λ indicates the eigenvalue and the corresponding eigenvector [11,40]. Let the column vectors $\boldsymbol{\alpha}_0, \boldsymbol{\alpha}_1, ..., \boldsymbol{\alpha}_{l-1}$ be the solution of Eqn. (11), so that the input weights of the three-hidden-layer SDAE are given by $\mathbf{A} = [\boldsymbol{\alpha}_0, \boldsymbol{\alpha}_1, ..., \boldsymbol{\alpha}_{l-1}]$.

According to Eqn. (6), the activation of the last hidden layer $\mathbf{h}_t$ in each local SDAE C_t can be expressed as Eqn. (12).

$$\mathbf{h}_t = s\{\mathbf{W}_{2,t}^* \cdot s[\mathbf{W}_{1,t}^* \cdot b(\tilde{\mathbf{x}}|\mathbf{A}\mathbf{x}) + \boldsymbol{\theta}_{1,t}^*] + \boldsymbol{\theta}_{2,t}^*\} \tag{12}$$

By adding an output layer, the MW class decision of C_t is computed as Eqn. (13).

$$\tilde{\mathbf{y}}_t = C_t(\mathbf{h}_t) = s(\mathbf{V}_t^* \cdot \mathbf{h}_t + \mathbf{v}_t^*) \tag{13}$$

The procedure of determining $\mathbf{W}_{1,t}^*$ and $\boldsymbol{\theta}_{1,t}^*$ is similar to that in the case of the standard SDAE, and the squared-error cost function is used.

$$L(\mathbf{A}\mathbf{x}, \tilde{\mathbf{x}}_o) = \|\mathbf{A}\mathbf{x} - \hat{\mathbf{x}}_o\|_2^2 \tag{14}$$

It is noted that different local SDAEs can also have different training sets. We can define the summation of the training error as

$$E_t = \frac{1}{\tilde{N}} \sum_{i=1}^{\tilde{N}} L(\mathbf{A}\mathbf{x}_i, \hat{\mathbf{x}}_{i,o}), \quad i = 1, 2, ..., \tilde{N}, \tag{15}$$

where $\tilde{N}$ is the sample size of the training subset. Finally, the BP algorithm can be implemented to compute the updating value δ_t for the network weights.

$$\delta_t = \frac{\partial E_t}{\partial \mathbf{W}_{1,t}} = \frac{\partial E_t}{\partial L} \frac{\partial L}{\partial s(\mathbf{W}_{1,t} \cdot \tilde{\mathbf{x}})} \frac{\partial s(\mathbf{W}_{1,t} \cdot \tilde{\mathbf{x}})}{\partial \mathbf{W}_{1,t} \cdot \tilde{\mathbf{x}}} \frac{\partial \mathbf{W}_{1,t} \cdot \tilde{\mathbf{x}}}{\partial \mathbf{W}_{1,t}} \tag{16}$$

In Eqn. (16), $\boldsymbol{\theta}_{1,t}$ is absorbed into $\mathbf{W}_{1,t}$ with the index of zero and $\tilde{x}_0 = 1$. Eventually, $\mathbf{W}_{2,t}^*$ and $\boldsymbol{\theta}_{2,t}^*$ are similarly computed via the layer-wise pre-training while $\mathbf{V}_t$ and $\mathbf{v}_t$ are determined by fine-tuning.

To fuse the outputs from all the base learners C_t, we implement Bagging approach [41] in the EL-SDAE. The Bagging ensemble framework randomly selects a subset from the entire EEG training data with replacement sampling via the bootstrap method [42]. Approximately $P_{choose} = 63\%$ training EEG instances are employed for modeling a weak
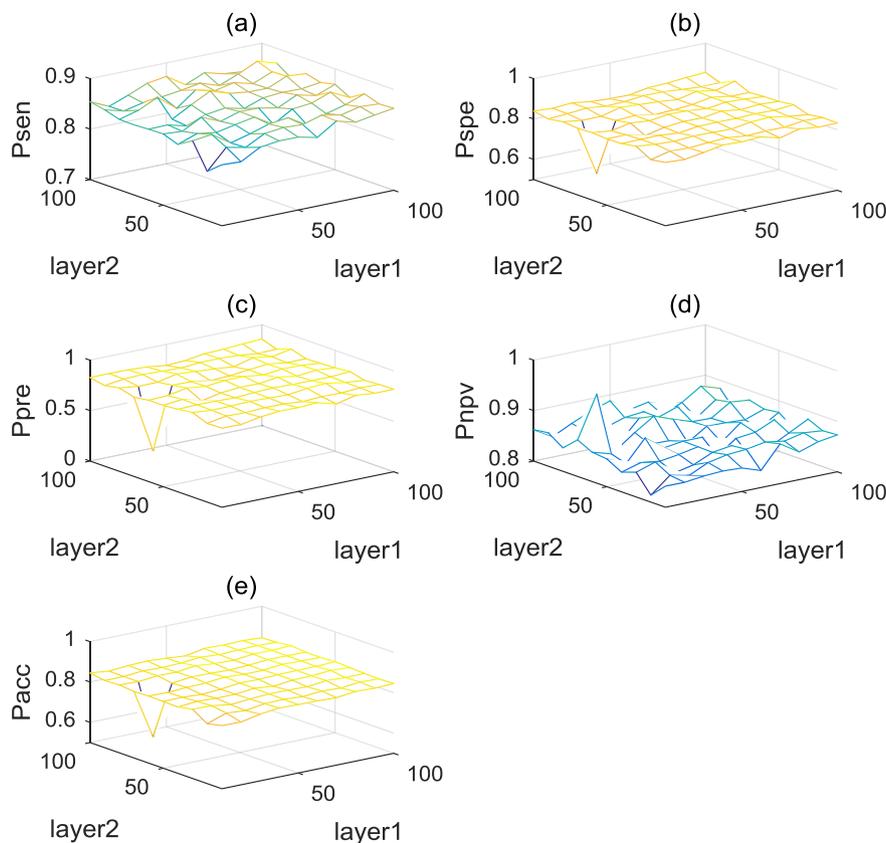
Fig. 5. Model selection of the member SDAE with a two-hidden-layer deep structure.

Table 3
Optimal neuron size of the first and second hidden layers.

Metrics	1st hidden layer	2nd hidden layer	Maximum values
P_{sen}	90	90	0.8822
P_{spe}	100	80	0.8627
P_{pre}	100	80	0.8606
P_{npv}	20	70	0.9553
P_{acc}	100	80	0.8555

learner to ensure diversity in different models.

$$P_{choose} = \lim_{l \to +\infty} [1 - (1 - 1/l)^l] = 1 - 1/e \approx 0.63 \qquad (17)$$

In Eqn. (12), l denotes the total times for sampling. Following the repetition of T-round bootstrap, T different training subsets are generated, wherein each EEG feature vector may not appear in certain subsets. Therefore, we can separately train T different member EEG classifiers $C_t(\mathbf{x})$, while the estimated MW level for an unseen EEG vector is fused from T outputs by majority voting. Assuming that T is odd, through majority voting combined with T base classifiers, whether more than half of the base classifiers would be of the HMW (or LMW) class can be predicted based on the positive (or negative) values of $C_T(\mathbf{x})$.

$$\bar{\mathbf{y}} = C_T(\mathbf{x}) = \text{sign}\left(\sum_{t=1}^{T} C_t(\mathbf{x})\right) \qquad (18)$$

The pseudo codes of the EL-SDAE training algorithm are shown in Table 2.

To evaluate the performance of the proposed and classical MW classifiers, the predicted MW levels are compared with the target labels eliciting five metrics, i.e., sensitivity, specificity, precision, negative predictive value and accuracy. These are denoted by P_{sen}, P_{spe}, P_{pre}, P_{npv},

and P_{acc}, respectively. Let us denote the positive class as LMW and negative class as HMW. True positive (TP) and true negative (TN) values are the number of correctly predicted EEG feature vectors of LMW and HMW classes, respectively. False positive (FP) is the number of overall cases wherein the HMW features are wrongly classified as LMW, while the false negative (FN) value corresponds to the contrary case.

4. Results

To verify the effectiveness of the proposed EL-SDAE in MW classification, two data splitting paradigms are designed. The subject-generic paradigm is used to identify the optimal model structure of the EL-SDAE, wherein the deep model is trained and tested on all 28800 EEG instances of the eight subjects. The classification results are derived by using 7/9 data for training and the remaining 2/9 for testing (it is noted that the 10-fold cross validation is applied in Section 4.7). On the other hand, the subject-specific paradigm is adopted to compare the classification performance across the different MW recognition models. The EEG data of 3600 instances of each participant are divided into training and testing sets in the same proportion, i.e., the sample sizes of the non-overlapped training and testing sets for each subject are 2800 and 800, respectively.

4.1. Model selection of member SDAE classifier

To improve the generalization capability of the base learner in the EL-SDAE framework, we examined the classifier's performance under different combinations of two parameters, i.e., the number of hidden layers and the number of neurons in the hidden layers. The number of hidden layers in member SDAEs is determined from the candidate parameter set of {2,3, …,14} with 13 choices. To reduce computational cost, the number of hidden neurons is fixed at 70. It is noted that classifier accuracy is reduced to lower than 0.5 when more than 14

Table 4
Testing confusion matrices of eight participants for SDAE and EL-SDAE.

Subject index	Predicted	SDAE		EL-SDAE	
		Target LMW	Target HMW	Target LMW	Target HMW
A	Low	326	44	**373**	18
	High	74	356	27	**382**
B	Low	373	10	**384**	1
	High	27	390	16	**399**
C	Low	369	63	**388**	23
	High	31	337	12	**377**
D	Low	303	133	**317**	91
	High	97	267	83	**309**
E	Low	357	84	**364**	47
	High	43	316	36	**353**
F	Low	342	75	**372**	37
	High	58	325	28	**363**
G	Low	395	36	**396**	17
	High	5	364	4	**383**
H	Low	322	68	**357**	28
	High	78	332	43	**372**

Note: The largest number of correctly predicted instances for each subject is in boldface.

Table 5
Testing classification performance of SDAE and EL-SDAE classifiers.

Methods	Subject	P_{sen}	P_{spe}	P_{pre}	P_{npv}	P_{acc}
SDAE	A	0.8811	0.8279	0.8150	0.8900	0.8525
	B	0.9739	0.9353	0.9325	0.9750	0.9538
	C	0.8542	0.9158	0.9225	0.8425	0.8825
	D	0.6950	0.7335	0.7575	0.6675	0.7125
	E	0.8095	0.8802	0.8925	0.7900	0.8413
	F	0.8201	0.8486	0.8550	0.8125	0.8338
	G	0.9165	0.9864	0.9875	0.9100	0.9488
	H	0.8256	0.8098	0.8050	0.8300	0.8175
	Mean	0.8470	0.8672	0.8709	0.8397	0.8553
EL-SDAE	A	**0.9540**	**0.9340**	**0.9325**	**0.9550**	**0.9438**
	B	**0.9974**	**0.9614**	**0.9600**	**0.9975**	**0.9788**
	C	**0.9440**	**0.9692**	**0.9700**	**0.9425**	**0.9563**
	D	**0.7770**	**0.7883**	**0.7925**	**0.7725**	**0.7825**
	E	**0.8856**	**0.9075**	**0.9100**	**0.8825**	**0.8963**
	F	**0.9095**	**0.9284**	**0.9300**	**0.9075**	**0.9188**
	G	**0.9588**	**0.9897**	**0.9900**	**0.9575**	**0.9738**
	H	**0.9273**	**0.8964**	**0.8925**	**0.9300**	**0.9113**
	Mean	**0.9192**	**0.9218**	**0.9222**	**0.9181**	**0.9202**

Note: The largest value for each participant and the performance index are in boldface.

member classifiers are adopted. The variations in the performance metrics of the training and testing EEG sets are illustrated in Fig. 4. The results show that the generalization capability of the SDAE does not increase in a deeper network structure. Therefore, we employ the member SDAE with the structure of two hidden layers to reduce computational costs.

To find the optimal structure of a four-layer SDAE (with two hidden layers), the classification performance indices, i.e., P_{sen}, P_{spe}, P_{pre}, P_{npv} and P_{acc} are investigated and shown in Fig. 5. We search for the optimal number of neurons within a finite parameter set {10,20, ...,100}.Specifically, the size of the hidden neurons across two layers can differ, and thus there are 100 model structures required to be examined. The classification performance under different combinations of neuron sizes

is shown via mesh plots. The results indicate that P_{sen} fluctuates within a small interval of 0.1. Except in the case of the parameter combination of (20,70) (i.e., when the first and second hidden layers possess 20 and 70 neurons, respectively), the value of P_{npv} reaches its maximum while the smaller P_{spe} and P_{pre} indicate the difficulty in recognizing low MW level. Table 3 lists the optimal neuron sizes at which each of the five classification indices reaches its maximum value. It is noted that P_{spe}, P_{pre}, and P_{acc} all reach their highest values when the hidden layer contain 100 and 80 neurons while P_{sen} and P_{npv} are also close to their maximum values under these conditions. Hence, we select the SDAE with 100 and 80 neurons in the hidden layers as the optimal structure of the base model.

4.2. EL-SDAE performance for MW classification

Table 4 summarizes the testing confusion matrices for the eight subjects under the optimal performance of the SDAE and EL-SDAE. For the EL-SDAE classifier, the output dimensionality of the local information preservation mapping is fixed at 40 for all feature sets. The number of hidden neurons is selected at around 100 and 80 to improve the diversity of the local SDAEs. Citing subject A as an example, the number of LMW levels that are correctly classified is 326 while the number of LMW levels erroneously classified as HMW levels is 44. Moreover, the number of HMW levels that are correctly classified is 356 and that of instances incorrectly classified as LMW levels is 74. Therefore, we can compute P_{sen} = 326/(326 + 44) = 0.8811, P_{spe} = 356/(356 + 74) = 0.8279, P_{pre} = 326/(326 + 44) = 0.8150, and P_{npv} = 356/(356 + 44) = 0.8900. The overall classification rate P_{acc} is thus 0.8525. The five performance metrics of the SDAE and EL-SDAE for the testing EEG feature sets are summarized in Table 5. It is shown that the EL-SDAE classifier outperforms the standard SDAE for all participants, and the average accuracy of binary MW classification is improved by 6.5%.

4.3. Performance comparison with shallow MW classifiers

In Fig. 6, the performance of the proposed EL-SDAE is computed based on the optimal SDAE structure and compared with those of four shallow learning machines, i.e., logistic regression (LR), naive Bayesian model (NB), extreme learning machine (ELM), K-nearest neighbor (KNN). The LR model is a classical binary classifier trained via the maximum likelihood method and has no hyperparameter that needs to be tuned [43]. For the NB classifier, we examine the accuracy across different prior probability combinations {(0.1, 0.9), (0.2, 0.8), ...(0.9, 0.1)}, and the optimal values (0.2, 0.8) of the highest average accuracy are selected. In addition, ELM is a single-hidden-layer neural network with random input weights [44]. We investigate its hidden neuron size across {10,20, ...100} and the optimal value is 90. Moreover, KNN is an instance-based learning method and we set 100 nearest neighbors for the classification model [45]. It is found from the figure that the EL-SDAE achieves the highest accuracy for the eight subjects among all classifiers.

4.4. Performance comparison with different ensemble classifiers

In Fig. 7, we compare different ensemble learning methods with the EL-SDAE on the basis of five MW classification metrics. We first introduce a Bagging-based ensemble model via standard SDAE (denoted by B-SDAE), wherein 17 random sample sets are drawn from the EEG training data. Moreover, two special SDAE-based ensemble models with the same number of base learners are introduced to verify the effectiveness of the temporal and frequential EEG features. These two models generated via the Bagging method are denoted by B-Time and B-Frequency. It can be seen that the EL-SDAE model achieves the highest average performance in terms of all five classification metrics. It is found that when the EEG feature selection is carried out in the B-Frequency and B-Time models, the MW classification performance
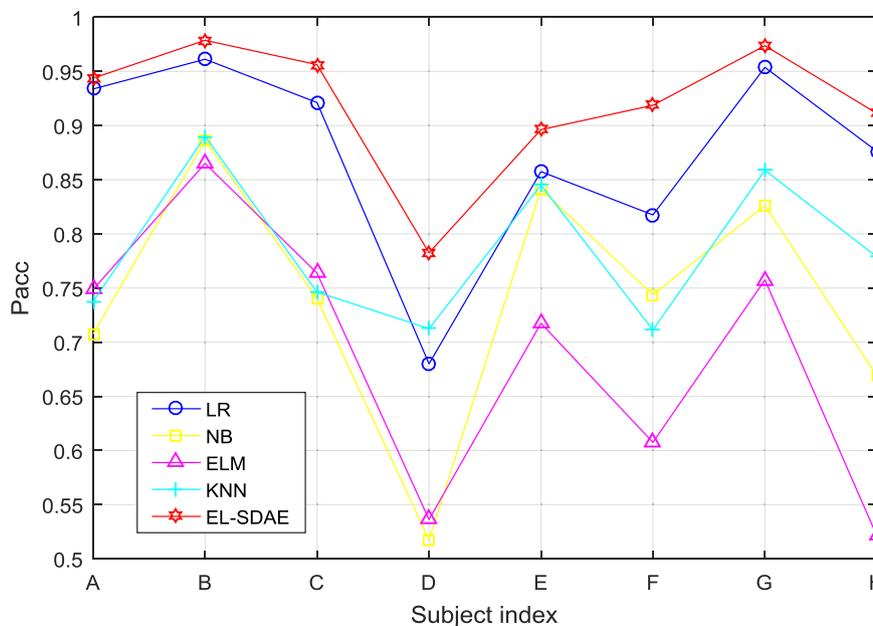
Fig. 6. Classification accuracy comparison between EL-SDAE and shallow MW classifiers.
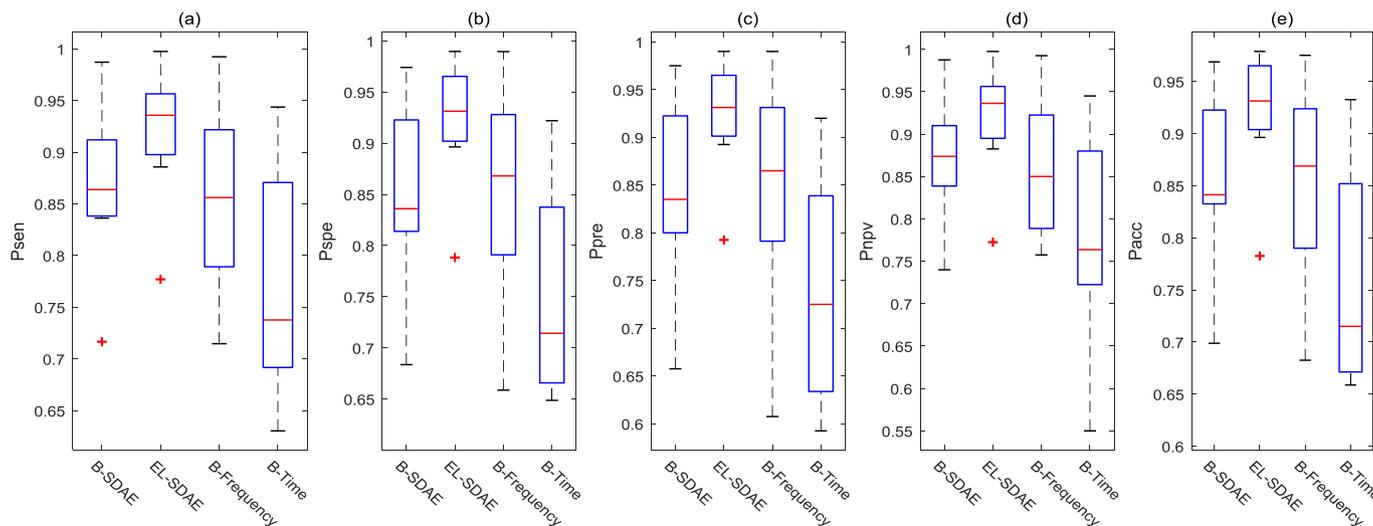

Fig. 7. Box-whisker plots of the MW classification performance indices derived from different ensemble classifiers.

significantly decreases.

4.5. Analysis of member classifiers in EL-SDAE

In this section, we examine the performance of different base learners in the EL-SDAE and the results are shown in Fig. 8. It can be seen that the testing MW classification rates are lower than those of the training state. By obtaining the average value of the eight subjects for each performance metric, we observe that the performance of the DAE is slightly higher than that of the SDAE. Moreover, after employing linear mapping with local information preservation as an additional hidden layer in the local SDAE model, salient psychophysiological features can be obtained through feature representations. Significant improvements are found in all the testing values of P_{spe}, P_{pre} and P_{acc}. In the case of P_{sen} and P_{npv}, the corresponding values of only one subject's EEG features (i.e., subject D) are slightly lower than those in the DAE and SDAE. According to the average values of the classification metrics of eight subjects, the base learner constructed via the local SDAE possesses better interclass discriminating capability under both training and testing conditions. It indicates the effectiveness of implementing

local SDAE as the member classifier to build the final MW recognition committee.

4.6. Computational cost analysis

Table 6 compares the average CPU time of 10 runs of training and testing all eleven MW classifiers on the EEG feature set from a single subject. According to the table, we find that the EL-SDAE and KNN achieve the highest and lowest training computational cost, respectively. This can be attributed to two aspects. Construction of the local information preservation mapping induces additional computational cost pertaining to feature dimensionality reduction, i.e., the training time cost for the base learner using local SDAE is higher than that employing standard SDAE. Moreover, the EL-SDAE method introduces multiple weak classifiers, and the outputs of each weak classifier need to be voted on produce the final decision. This results in higher computational burden on the Bagging-based models, i.e., EL-SDAE, B-SDAE, B-Frequency, and B-Time. However, the testing time cost of the EL-SDAE remains comparable to that of the standard SDAE but is significantly lower than that of the KNN. The observation shows the
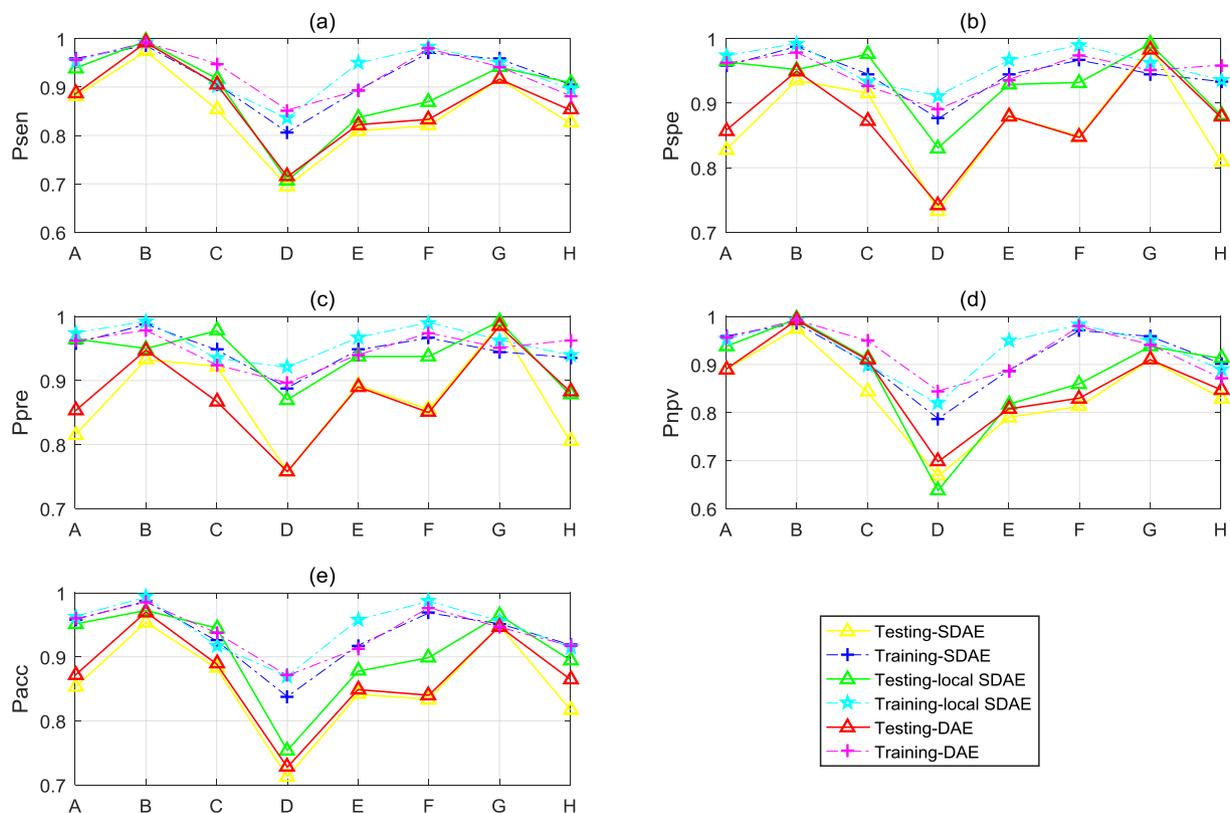
Fig. 8. Comparison of training and testing classification performance indices of different base learners in EL-SDAE.

Table 6
Average CPU time (in seconds) for training and testing a MW classifier with EEG feature set for a subject's data. The standard deviation (denoted by s.d.) and mean values are calculated on the basis of 10 runs of trials. The average testing classification accuracy for all subjects are also listed for the purpose of comparison.

MW Classifier	Average testing P_{acc}	Training		Testing	
		Mean	s.d.	Mean	s.d.
ELM	0.6897	0.1209	0.0447	0.0254	0.0185
NB	0.7414	1.8155	0.1378	0.2574	**0.0703**
KNN	0.7850	0.1190	0.0138	**1.0901**	0.0685
LR	0.8752	6.0197	0.5487	0.0429	0.0238
DAE	0.8698	11.1892	0.8581	0.1385	0.0100
SDAE	0.8553	10.4677	0.6923	0.1014	0.0264
Local SDAE	0.9073	12.2753	**1.4970**	0.1034	0.0333
B-SDAE	0.8575	13.7963	0.9189	0.0936	0.0186
B-Frequency	0.8528	12.4840	0.7497	0.1092	0.0194
B-Time	0.7584	12.0784	0.6396	0.1092	0.0264
EL-SDAE	**0.9202**	**15.9082**	0.6070	0.1014	0.0212

Note: The minimum and maximum values in each column are underlined and marked in boldface, respectively.

potential of the EL-SDAE for online testing.

4.7. Cross validation and statistical analysis

In Table 7, we present the classification accuracy derived by the 10-fold cross-validation on the entire dataset of training and testing instances of each subject [46]. That is, the EEG data of two sessions (3600 instances) of each subject are randomly divided into 10 subsets (each subset has 360 instances). For each iteration, nine subsets are used for training the MW classifier and the remaining subset is used for testing. After all subsets are tested, the average accuracy is computed and presented in the table. In total, the 10-fold cross-validation accuracy of

11 classifiers mentioned in the previous sections is computed. From Table 7, it can be seen that the proposed EL-SDAE achieves the highest performance on subjects A, B, C, D, E, F, and H while the local-SDAE achieves the highest accuracy on subject G. This indicates that the EL-SDAE outperforms the other 10 MW classifiers on the basis of 10-fold cross-validation accuracy of all subjects.

To investigate the robustness of the classification algorithms, we compute the standard deviation of the 10 values of the accuracy that are derived on each subset in the process of the cross-validation. The results of each classifier and each subject are presented in Table 8. We observe that the minimum value of the standard deviation for subjects B, C, and E and the lowest average standard deviation are achieved by the EL-SDAE. Thus, the generalization capability and the stability of the EL-SDAE are competitive for the task of the subject-specific MW classification.

We also validate the obtained accuracy and standard deviation of the 10-fold cross validation of all subjects based on two-tailed t-test. The performance of the EL-SDAE is compared with those of the remaining 10 MW classifiers. All the accuracy and standard deviation values for statistical tests are on the basis of Tables 7 and 8 The p values and the statistics of t-test method are presented in Table 9. Since there are 10 hypotheses that are tested on the same data source, we apply the Holm–Bonferroni method to correct the significance level for different cases. The elicited p values are ordered from the lowest to the highest and the correct significance level is $\alpha/(10 + 1 - k)$ wherein k is the ranking of the certain p value. We observe that the mean accuracy of the EL-SDAE is significantly higher than those of the other 10 MW classifiers. Moreover, the standard deviation of the accuracy of the EL-SDAE is significantly lower than that of the DAE, ELM, KNN, NB, SDAE, B-Time, B-SDAE, and is comparable with that of the LR, B-Frequency, and local-SDAE. The t-test results partially validated the robustness of the EL-SDAE for the generalization capability across multiple subjects.

Table 7
Subject-specific classification accuracy derived by the 10-fold cross-validation of each MW classifier. MACC denotes the mean accuracy.

MW Classifier	A	B	C	D	E	F	G	H	MACC
DAE	0.9167	0.9547	0.8797	0.7475	0.8278	0.9114	0.9081	0.8608	0.8758
ELM	0.8647	0.9131	0.8414	0.6997	0.8008	0.8667	0.8750	0.7961	0.8322
KNN	0.8661	0.9089	0.7372	0.6922	0.8233	0.8250	0.8808	0.8072	0.8176
LR	0.9381	0.9653	0.9047	0.8206	0.9181	0.9494	0.9361	0.8969	0.9161
NB	0.8222	0.8825	0.7092	0.6294	0.8308	0.8414	0.8564	0.6433	0.7769
SDAE	0.8997	0.9392	0.8592	0.6981	0.8303	0.8786	0.8956	0.8275	0.8535
B-Frequency	0.9267	0.9753	0.8733	0.7417	0.8827	0.9110	0.9130	0.8563	0.8850
B-Time	0.7283	0.8970	0.7313	0.6513	0.6703	0.8333	0.8740	0.7663	0.7690
B-SDAE	0.9053	0.9530	0.8717	0.7117	0.8357	0.8930	0.9033	0.8530	0.8658
Local-SDAE	0.9217	0.9797	0.8806	0.7978	0.8961	0.9556	**0.9378**	0.9000	0.9086
EL-SDAE	**0.9493**	**0.9810**	**0.9237**	**0.8483**	**0.9353**	**0.9597**	0.9363	**0.9013**	**0.9294**

Note: The minimum and maximum values in each column are underlined and in boldface, respectively. The subject index is shown by A, B, C, D, E, F, G, and H.

Table 8
Standard deviations (SDs) of the 10-fold cross validation accuracy for each MW classifier tested.

MW Classifier	A	B	C	D	E	F	G	H	SD
DAE	0.0231	0.0159	0.0135	0.0262	0.0200	0.0148	0.0163	0.0165	0.0183
ELM	0.0240	0.0133	0.0145	0.0267	0.0265	0.0192	0.0188	0.0200	0.0204
KNN	0.0305	0.0134	0.0158	0.0269	0.0297	0.0189	**0.0231**	0.0143	0.0216
LR	0.0107	0.0103	0.0134	0.0177	0.0127	0.0126	0.0125	0.0129	0.0129
NB	**0.0420**	0.0153	0.0256	**0.0343**	0.0228	0.0102	0.0185	0.0348	0.0254
SDAE	0.0195	0.0151	0.0186	0.0220	0.0216	0.0087	0.0119	0.0156	0.0166
B-Frequency	0.0220	0.0092	0.0156	0.0234	0.0235	**0.0201**	0.0090	0.0197	0.0178
B-Time	0.0251	**0.0208**	0.0248	0.0329	0.0329	0.0132	0.0195	**0.0364**	0.0257
B-SDAE	0.0252	0.0184	0.0221	**0.0343**	0.0236	0.0157	0.0145	0.0286	0.0228
local-SDAE	0.0249	0.0119	**0.0273**	0.0330	**0.0656**	0.0079	0.0088	0.0315	**0.0264**
EL-SDAE	0.0113	0.0047	0.0101	0.0229	0.0077	0.0101	0.0100	0.0138	0.0113

Note: The minimum and maximum values in each column are underlined and in boldface, respectively. The subject index is shown by A, B, C, D, E, F, G, and H.

Table 9
Results of two-tailed t-test on the 10-fold validation performance between the EL-SDAE and the other 10 MW classifiers.

MW Classifier	Mean		s.d.	
	p value	t value	p value	t value
EL-SDAE vs. DAE	**0.0022**	4.7020	**0.0020**	−4.7685
EL-SDAE vs. ELM	**< 0.0001**	8.9254	**0.0012**	−5.2335
EL-SDAE vs. KNN	**0.0002**	7.0861	**0.0059**	−3.8956
EL-SDAE vs. LR	**0.0035**	4.3120	0.2734	−1.1886
EL-SDAE vs. NB	**0.0004**	6.4100	**0.0031**	−4.4019
EL-SDAE vs. SDAE	**0.0007**	5.7861	**0.0339**	−2.6301
EL-SDAE vs. B-Frequency	**0.0043**	4.1458	**0.0130**	−3.3084
EL-SDAE vs. B-Time	**0.0003**	6.4865	**0.0008**	−5.6756
EL-SDAE vs. B-SDAE	**0.0019**	4.8520	**0.0001**	−7.6596
EL-SDAE vs. local-SDAE	**0.0307**	2.6991	0.0592	−2.2502

Note: The significant cases are in boldface. The significance level is 0.05 and is corrected by the Holm–Bonferroni method for different cases.

5. Discussion

In the present study, we introduce a novel ensemble deep learning framework denoted by EL-SDAE for recognizing MW levels via EEG signals. In particular, the LPP-based local information preservation mapping is used to learn subject-specific, intermediate EEG abstractions. After identifying the optimal mapping architecture, we combine it with the SDAE to generate a local deep SDAE network as the base learner. The EL-SDAE employs the Bagging paradigm to build the classifier ensemble for two reasons. On the one hand, the sizes of the instances in each training subset can be enlarged by random sampling with replacement. This mitigates the problem of inadequate training of the deep learning model since the deep architecture involves more tunable parameters. On the other hand, the Bagging approach takes only a part of the EEG examples for constructing each member model

and enables the classification committee to inhibit the noised feature distribution in the entire training set.

The superiority observed in the case of both the local SDAE and EL-SDAE over the classical SDAE under the paradigm of subject-specific MW classification indicates that local information preservation mapping has the capability to find the personalized EEG feature representations. It also verifies the usefulness of the LPP in supervised physiological feature fusion [1]. Consistent with the results from Wu et al. [41], the Bagging-based approach further increase the classification rate by integrating the outputs of multiple weak models. Specifically, EEG features fed into the EL-SDAE adopt comprehensive frequency ranges in the theta, alpha, beta, and gamma bands, which are consistent with most existing work on MW classification using EEG power features. For instance, Ke et al. inferred that there are workload-related EEG features that are at least partially unaffected by task types and time effects on alpha and theta bands. Baldwin et al. indicated that a wide range of task-sensitive features exist in the alpha, theta, and beta bands in the frontal, central and parietal areas [35,47,48].

Another contribution of the study is the optimal structure of the two-hidden-layer member SDAE that has been identified as 100 and 80 hidden neurons. As suggested in Ref. [49], both the SDAE and AE can be utilized to abstract neurophysiological signals. The implementation of the AE on EEG signals for the issue of sleep state identification was reported with a classification accuracy of 80.4% [50]. In the EL-SDAE framework, the number of hidden neurons is set around the optimal value to ensure the required diversity of the base learners. It is also noted that a large number of hidden neurons may not improve the generalization capability. In Fig. 9, we implement an AE model with the number of hidden neurons selected within a candidate set {10,20, …,300}of the EEG feature sets from eight participants. It is shown that the testing results do not improve significantly with increase in the number of neurons in the hidden layer.

On the other hand, the limitations and future work of the present study include the following three aspects. (1) Despite the high accuracy
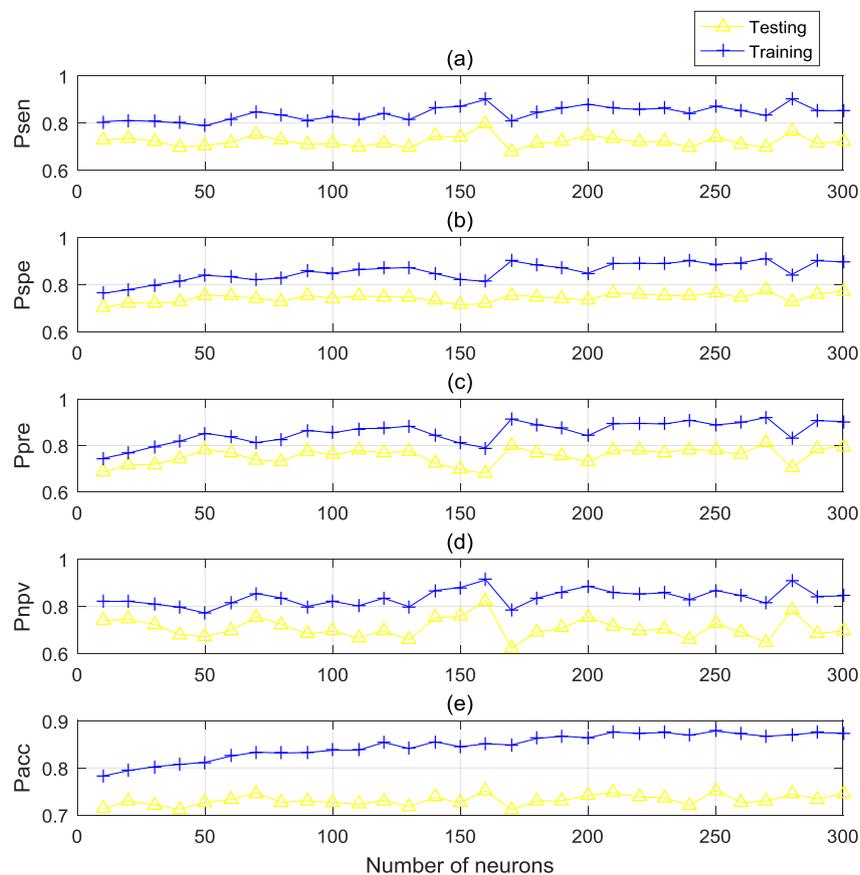
Fig. 9. Participant-average MW classification performance vs. the number of hidden neurons in a SDAE model.

of the EL-SDAE, the average values of the two classification indices, P_{spe} and P_{spe}, are still slightly lower than those of the local SDAE. The potential reason is that the effect of Bagging is not obvious in a stable deep model with minimal deviation. (2) To mitigate over-fitting, we employ only three hidden layers of the local SDAE in the ensemble classifier. In future work, we will enrich the diversity of the SDAE architecture with deeper structures to extract more salient feature representations. (3) Du et al. proposed a multimodal detection method combining EEG and ECG signals, and the results indicated that the average correlation coefficient and root mean square error are 0.85 and 0.09, respectively [51]. The algorithm based on AE achieves high classification performance in fetal-ECG signal reconstruction [52]. However, the input of the EL-SDAE was only fixed with EEG features. It will be worth considering other modalities of peripheral physiological signals in future research.

6. Conclusion

In this paper, an ensemble deep learning model EL-SDAE is proposed for binary MW classification by abstracting high-dimensional EEG features. The EEG data is collected on eight subjects with two-session experiments under the ACAMS platform in complex HM collaboration environments. In the EL-SDAE framework, base models are constructed via local SDAE of three hidden layers, and this learns useful EEG feature representation with local information preservation mapping. By employing the Bagging method, multiple local SDAEs are fused into personalized EEG classification committees to improve subject-specific MW recognition performance. The subject-average classification achieved an accuracy of 92%, which outperforms several classical shallow and deep classifiers. It is noted the high performance of the EL-SDAE entails cost of the increased computational burden and additional model parameters. Our future work will focus on using of the multiple modalities of physiological data with adaptable network depth.

Conflicts of interest

None declared.

References

[1] J. Zhang, Z. Yin, R. Wang, Recognition of mental workload levels under complex human–machine collaboration by using physiological features and adaptive support vector machines, IEEE Trans. Human-Mach Syst. 45 (2) (2015) 200–214.

[2] H.J. Foy, P. Chapman, Mental workload is reflected in driver behaviour, physiology, eye movements and prefrontal cortex activation, Appl. Ergon. 73 (2018) 90–99.

[3] T. Heine, G. Lenis, P. Reichensperger, T. Beran, O. Doessel, B. Deml, Electrocardiographic features for the measurement of drivers' mental workload, Appl. Ergon. 61 (2017) 31–43.

[4] Z. Yin, J. Zhang, Cross-session classification of mental workload levels using EEG and an adaptive deep learning model, Biomed. Signal Process. Control 33 (2017) 30–47.

[5] J.H. Zhang, J.J. Xia, J.M. Garibaldi, P.P. Groumpos, R.B. Wang, Modeling and control of operator functional state in a unified framework of fuzzy inference petri nets, Comput. Methods Progr. Biomed. 144 (2017) 147–163.

[6] X. Hou, Y. Liu, O. Sourina, W. Mueller-Wittig, CogniMeter: EEG-based emotion, mental workload and stress visual monitoring, International Conference on Cyberworlds, 2015, pp. 153–160 Visby, Sweden.

[7] P. Arico, G. Borghini, I. Graziani, F. Taya, Towards a multimodal bioelectrical framework for the online mental workload evaluation, International Conference of the IEEE Engineering in Medicine and Biology Society, 2014, pp. 3001–3004 Chicago, USA.

[8] M.K. Choi, S.M. Lee, J.S. Ha, P.H. Seong, Development of an EEG-based workload measurement method in nuclear power plants, Ann. Nucl. Energy 111 (2018)

595–607.

[9] J. Luo, Z. Feng, J. Zhang, N. Lu, Dynamic frequency feature selection based approach for classification of motor imageries, Comput. Biol. Med. 75 (2016) 45–53.

[10] T. Hwang, M. Kim, M. Hwangbo, E. Oh, Comparative analysis of cognitive tasks for modeling mental workload with electroencephalogram, International Conference of the IEEE Engineering in Medicine and Biology Society, 2014, pp. 2661–2665 Chicago, USA.

[11] G. Shikkenawis, S.K. Mitra, On some variants of locality preserving projection, Neurocomputing 173 (2016) 196–211.

[12] A. Özçift, Random forests ensemble classifier trained with data resampling strategy to improve cardiac arrhythmia diagnosis, Comput, Biol. Med. 41 (5) (2011) 265–271.

[13] Y.H. Ayaz, P.A. Shewokis, S. Bunce, K. Izzetoglu, B. Willems, B. Onaral, Optical brain monitoring for operator training and mental workload assessment, Neuroimage 59 (1) (2012) 36–47.

[14] H. Aghajani, A. Omurtag, Assessment of mental workload by EEG + FNIRS, International Conference of the IEEE Engineering in Medicine and Biology Society, 2016, p. 3773 Orlando, USA.

[15] C. Cimpanu, F. Ungureanu, V.I. Manta, T. Dumitriu, A comparative study on classification of working memory tasks using EEG signals, International Conference on Control Systems and Computer Science, 2017, pp. 245–251 Bucharest, Romania.

[16] L. Orlandi, B. Brooks, Measuring mental workload and physiological reactions in marine pilots: building bridges towards redlines of performance, Appl. Ergon. 69 (2018) 74–92.

[17] M. Fallahi, M. Motamedzade, R. Heidarimoghadam, A.R. Soltanian, M. Farhadian, S. Miyake, Analysis of the mental workload of city traffic control operators while monitoring traffic density: a field study, Int. J. Ind. Ergon. 54 (2016) 170–177.

[18] S. Wang, J. Gwizdka, W.A. Chaovalitwongse, Using wireless EEG signals to assess memory workload in the n-Back task, IEEE Trans. Human-Mach Syst. 46 (3) (2016) 424–435.

[19] Y. Liu, O. Sourina, S.H.P. Liew, G. Krishnan, D. Konovessis, H.E. Ang, EEG-based mental workload and stress recognition of crew members in maritime virtual simulator: a case study, International Conference on Cyberworlds, 2017, pp. 64–71 Chester, UK.

[20] J. Fan, J.W. Wade, A.P. Key, Z. Warren, N. Sarkar, EEG-based affect and workload recognition in a virtual driving environment for ASD intervention, IEEE Trans. Biomed. Eng. 99 (2017) 1-1.

[21] J. Chen, X. Song, Z. Lin, Revealing the "Invisible Gorilla" in construction: estimating construction safety through mental workload assessment, Automatic, Constr. Met. (CTICM) 63 (2016) 173–183.

[22] S. Puma, N. Matton, P.V. Paubel, É. Raufaste, R. El-Yagoubi, Using theta and alpha band power to assess cognitive workload in multitasking environments, Int. J. Psychophysiol. 123 (2018) 111–120.

[23] X. Cui, J. Zhang, R. Wang, Identification of mental workload using imbalanced EEG data and DySMOTE-based neural network approach, Ifac Papersonline 49 (19) (2016) 567–572.

[24] P. Zarjam, J. Epps, N. Lovell, Beyond subjective self-rating: EEG signal classification of cognitive workload, IEEE Trans. Auton. Men. De. 7 (4) (2015) 301–310.

[25] L.L. Wei, O. Sourina, Y. Liu, L. Wang, EEG-based mental workload recognition related to multitasking, International Conference on Information, Communications and Signal Processing, 2016, pp. 1–4 Singapore, Singapore.

[26] A. Gonczarek, J.M. Tomczak, S. Zaręba, J. Kaczmar, P. Dąbrowski, M.J. Walczak, Interaction prediction in structure-based virtual screening using deep learning, Comput. Biol. Med. 100 (2018) 253–258.

[27] S. Luo, L. Zhu, K. Althoefer, H. Liu, Knock-knock: acoustic object recognition by using stacked denoising autoencoders, Neurocomputing 267 (2017) 18–24.

[28] C. Lu, Z.Y. Wang, W.L. Qin, J. Ma, Fault diagnosis of rotary machinery components using a stacked denoising autoencoder-based health state identification, Signal Process. 130 (2017) 377–388.

[29] Y.J. Fan, Autoencoder node saliency: selecting relevant latent representations, Pattern Recogn. 88 (2019) 643–653.

[30] J. Li, Z. Struzik, L. Zhang, A. Cichocki, Feature learning from incomplete EEG with denoising autoencoder, Neurocomputing 165 (2015) 23–31.

[31] Z. Wang, J. Wang, Y. Wang, An intelligent diagnosis scheme based on generative adversarial learning deep neural networks and its application to planetary gearbox fault pattern recognition, Neurocomputing 310 (2018) 213–222.

[32] B. Lorenz, F. Nocera, S. Röttger, R. Parasuraman, Automated fault-management in a simulated spaceflight micro-world, Aviat. Space Environ. Med. 73 (2002) 886–897.

[33] J. Sauer, D.G. Wastell, G.R.J. Hockey, A conceptual framework for designing micro-worlds for complex work domains: a case study of the Cabin Air Management System, Comput. Hum. Behav. 16 (1) (2000) 45–58.

[34] H.H. Jasper, Report of to the committee on method of clinical examination in electroencephalography. Appendix: the ten twenty electrode system of the international federation, Clin. Neurophysiol. 10 (1958) 371–375.

[35] 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 electro-encephalography based on feature selection and support vector machine regression, Int. J. Psychophysiol. 98 (2) (2015) 157–166.

[36] Y. Zhao, J. Li, L. Yu, Multimodal video classification with stacked contractive autoencoders, Signal Process. 120 (2016) 761–766.

[37] Y. Zhao, J. Li, L. Yu, A deep learning ensemble approach for crude oil price forecasting, Energy Econ. 66 (2017) 9–16.

[38] Y. Wang, S. Huang, Z. Liu, H. Wang, D. Liu, Locality preserving projection based on endmember extraction for hyperspectral image dimensionality reduction and target detection, Appl. Spectrosc. 70 (9) (2016) 1573.

[39] G.F. Lu, Y. Wang, J. Zou, Z. Wang, Matrix exponential based discriminant locality preserving projections for feature extraction, Neural Network. 97 (2018) 127–136.

[40] R. Jiang, W. Fu, L. Wen, S. Hao, R. Hong, Dimensionality reduction on anchorgraph with an efficient locality preserving projection, Neurocomputing 187 (2016) 109–118.

[41] Z. Wu, N. Li, J. Peng, H. Cui, P. Liu, H. Li, X. Li, Using an ensemble machine learning methodology-Bagging to predict occupants' thermal comfort in buildings, Energy Build. 173 (2018) 117–127.

[42] A.R. Hassan, S. Siuly, Y. Zhang, Epileptic seizure detection in EEG signals using tunable-Q factor wavelet transform and bootstrap aggregating, Comput. Methods Progr. Biomed. 137 (2016) 247–259.

[43] F. Zou, Y. Wang, Y. Yang, K. Zhou, Y. Chen, J. Song, Supervised feature learning via l 2 -norm regularized logistic regression for 3D object recognition, Neurocomputing 151 (2015) 603–611.

[44] W. Zhang, Q. Jiang, L. Chen, C. Li, Two-stage ELM for phishing Web pages detection using hybrid features, World Wide Web 20 (4) (2017) 797–813.

[45] S.B. Chen, Y.L. Xu, C.H.Q. Ding, B. Luo, A nonnegative locally linear KNN model for image recognition, Pattern Recogn. 83 (2018) 78–90.

[46] L. Xu, O. Hu, Y. Guo, M. Zhang, D. Lu, C.B. Cai, S. Xie, M. Goodarzi, H.Y. Fu, Y.B. She, Representative splitting cross validation, Chemometr Intell. Lab. 183 (2018) 29–35.

[47] C.L. Baldwin, B.N. Penaranda, Adaptive training using an artificial neural network and EEG metrics for within- and cross-task workload classification, Neuroimage 59 (1) (2012) 48–56.

[48] D. Chris, D.W. Dick, K.A. Brookhuis, B.L.J.M. Mulder, D.J. Ritske, Classifying visuomotor workload in a driving simulator using subject specific spatial brain patterns, Front. Neurosci. 7 (2013) 149.

[49] 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 Progr. Biomed. 161 (2018) 1–13.

[50] L. Fraiwan, K. Lweesy, Neonatal sleep state identification using deep learning autoencoders, IEEE International Colloquium on Signal Processing and its Applications, Batu, Malaysia, 2017, pp. 228–231.

[51] L.H. Du, W. Liu, W.L. Zheng, B.L. Lu, Detecting driving fatigue with multimodal deep learning, International IEEE/EMBS Conference on Neural Engineering, 2017, pp. 74–77 Shanghai, China.

[52] P.R. Muduli, R.R. Gunukula, A. Mukherjee, A deep learning approach to fetal-ECG signal reconstruction, National Conference on Communication, 2016, pp. 1–6 Guwahati, India.