

Instantaneous Mental Workload Recognition Using Wavelet-Packet Decomposition and Semi-Supervised Learning*

Jianhua Zhang
OsloMet AI Lab

Department of Computer Science
Oslo Metropolitan University
Oslo, Norway
Email: jianhuaz@oslomet.no

Jianrong Li
School of Information Science and
Engineering
East China University of Science and
Technology
Shanghai, China

Stefano Nichele
OsloMet AI Lab
Department of Computer Science
Oslo Metropolitan University
Oslo, Norway

Abstract—The real-time monitoring of human operator's mental workload (MWL) is crucial for development of adaptive/intelligent human-machine cooperative systems in various safety/mission-critical application fields. Although data-driven approach has shown promise in MWL recognition, its major difficulty lies in how to acquire sufficient labeled data to train the model. This paper applies semi-supervised extreme learning machine (SS-ELM) to the problem of MWL classification based only on a small number of labeled data. The experimental data analysis results have shown that the proposed SS-ELM paradigm can effectively improve the accuracy and efficiency of MWL classification. The proposed semi-supervised learning paradigm may provide an alternative data-driven machine learning approach to effectively utilize a large number of unlabeled data, which can be readily collected under naturalistic (operational) task environments in many real-world applications.

Keywords—Mental workload; Operator functional state; Physiological signals; Feature engineering; Semi-supervised learning; Extreme learning machine

I. INTRODUCTION

Automation, automatic control system, and artificial intelligence (AI) techniques have been widely applied to various fields, but there is still a long way for the current development of automation and AI technologies to achieve fully-automated control for many real-world complex and uncertain systems. In this connection, Human-Machine Systems (HMS) are still ubiquitous in practice in most safety-critical application domains [1]. Compared with machines, human operators are more susceptible to external disturbances or the impact of their own psychophysiological fluctuations [2]. Therefore, it is not surprising that human factors play a significant role in the achievement of desired performance of HMS. In recent years, researchers from multiple disciplines have focused on the research of how to maintain the optimum Operator Functional State (OFS) to ensure the successful completion of the tasks in the HMS context [3].

The operator's mental workload (MWL) is an essential dimension of the multi-dimensional construct of OFS. The

MWL can be considered as a candidate variable for measuring mental status of human operator, which reflects the mental demand for operators to accomplish the tasks [4]. For operators, too high or too low psychological load is detrimental to the performance of HMS. In order to mitigate this problem, researchers conceived Adaptive Automation (AA) strategy. The AA system can adaptively allocate the tasks between operators and the machines based on the estimated levels of operators' MWL. MWL measurement/assessment/evaluation approaches can be roughly divided into three categories [5]: (1) subjective assessment; (2) task performance measures; and (3) physiological data based assessment. Compared with the former two approaches, the last approach is featured by continuous on-line measurement. ElectroEncephaloGram (EEG), ElectroCardioGram (ECG) and ElectroOculoGram (EOG) have been widely used for MWL recognition [6-8]. In this paper we evaluate the operators' MWL by using multi-modal psychophysiological signals and examine the potential and efficacy of semi-supervised learning technique for enhancing the accuracy and efficiency of high-risk MWL detection.

II. SEMI-SUPERVISED EXTREME LEARNING MACHINE

The SS-ELM is a semi-supervised learning algorithm based on ELM theory and manifold regularization framework, which can take advantage of the unlabeled data to improve the classification accuracy when labeled data are scarce [9]. It determines the output weights by minimizing the squared sum of the empirical training error of labeled data, the norm of the output weights, as well as the manifold regularization term based on both labeled and unlabeled data.

We have the SS-ELM formulation:

$$\begin{aligned} \min_{W \in \mathbb{R}^{L \times c}} & \frac{1}{2} \|W\|^2 + \frac{1}{2} \sum_{i=1}^l C_{y_i} \|\xi_i\|^2 + \frac{\lambda}{2} \text{Tr}(Y^T L Y) \\ \text{s.t.} & \mathbf{h}(\mathbf{x}_i)W = \mathbf{y}_i^T - \xi_i^T, \quad i = 1, 2, \dots, l \\ & \mathbf{y}_i = \mathbf{h}(\mathbf{x}_i)W, \quad i = 1, 2, \dots, n \end{aligned} \quad (1)$$

This work was supported in part by the OsloMet Faculty TKD Strategic (Lighthouse) R&D Project [Grant No. 201369-100].

Where λ is a tradeoff parameter, C_{y_i} is a penalty factor for training error of data from class y_i , $L \in \mathbb{R}^{n \times n}$ is the graph Laplacian built from both labeled and unlabeled data, and $Y \in \mathbb{R}^{n \times c}$ is the output matrix of the network with its i -th row equal to \mathbf{y}_i .

Note that similar to the weighted ELM (W-ELM) algorithm, here we assign different penalty factor C_{y_i} to the prediction errors w.r.t. samples from different classes because when the data is unbalanced, i.e., some classes have significantly more samples than other classes, traditional ELM tend to fit the majority classes well, but fits minority classes poorly. This usually results in poor generalization to the testing set. Therefore, in order to cope with the possibly imbalanced classification problem, we reweigh examples from different classes. Suppose that \mathbf{x}_i belongs to class \mathbf{y}_i which has N_{y_i} training samples, then we assign ξ_i with a penalty of $C_{y_i} = \frac{C_0}{N_{y_i}}$, where C_0 is a user-defined parameter as in

traditional ELM and N_{y_i} is the number of training samples in the class \mathbf{y}_i . In this way, the samples from the dominant classes will not be overfitted by the algorithm and the samples from a class with less samples will not be ignored.

Substituting the constraints into the objective function yields the new formulation in matrix form:

$$\min_{W \in \mathbb{R}^{L \times c}} \left[\frac{1}{2} \|W\|^2 + \frac{1}{2} \|C^{\frac{1}{2}}(\tilde{Y} - HW)\|^2 + \frac{\lambda}{2} \text{Tr}(W^T H^T LHW) \right] \quad (2)$$

where $H = [\mathbf{h}(\mathbf{x}_1)^T, \mathbf{h}(\mathbf{x}_2)^T, \dots, \mathbf{h}(\mathbf{x}_l)^T]^T \in \mathbb{R}^{l \times L}$, $\tilde{Y} \in \mathbb{R}^{n \times c}$ is the augmented training target with its first l rows equal to Y_l and the rest equal to 0, and $C \in \mathbb{R}^{n \times n}$ is a (penalty) diagonal matrix with its first l diagonal elements $c_{ii} = C_i$ and the rest equal to 0.

Now let us solve the above optimization problem. We first compute the gradient of the objective function w.r.t. W and then by setting the gradient to zero, we obtain the optimal output weights (i.e., the SS-ELM solution) if $L \leq l$:

$$W^* = (\mathbf{I}_L + H^T C H + \lambda H^T L H)^{-1} H^T C \tilde{Y} \quad (3)$$

where \mathbf{I}_L is an identity matrix of dimension L .

If $L > l$ (common in SSL), the optimal output weights can be solved by the alternative form:

$$W^* = H^T (\mathbf{I}_n + C H H^T + \lambda L H H^T)^{-1} C \tilde{Y} \quad (4)$$

where \mathbf{I}_n is an identity matrix of dimension n .

In summary, SS-ELM training algorithm consists of two key steps:

Step 1: Initialize an ELM network of L hidden neurons with random input weights and biases, and calculate the output matrix of the hidden neurons $H \in \mathbb{R}^{n \times L}$.

Step 2: Use (3) or (4) to compute the output weights W .

III. FEATURE EXTRACTION ALGORITHM

Usually EEG feature extraction algorithms can be divided into time-domain, frequency-domain, time-frequency, and nonlinear dynamics analysis. This paper uses the discrete wavelet-packet transform [10] to extract the time-frequency features of the physiological signals.

Wavelet transform is a multi-scale signal analysis method. The method can characterize the local features of the signal in time and scale domain, so it is very suitable for the analysis of transient characteristics and time-frequency characteristics of non-stationary EEG signal.

Wavelet Packet Decomposition (WPD) is a generalization of wavelet decomposition. In the wavelet analysis, the approximation part is decomposed into the approximation part and detail part at another level. This process is repeated until the maximal number of decomposition levels is reached. However, in the WPD details are also decomposed. WPD has multi-scale characteristics and provides great choice for time-frequency analysis. In the multiresolution wavelet analysis, the Hilbert space $L^2(\mathbb{R})$ is decomposed into the sum of all orthogonal wavelet subspaces W_j (scale factor $j \in \mathbb{Z}$):

$$L^2(\mathbb{R}) = \bigoplus_{j \in \mathbb{Z}} W_j \quad (5)$$

WPD continues to dichotomize W_j ($j = 1, 2, \dots$), as shown in

Fig. 1, where U_j^n is the wavelet-packet space of the scale j and its orthogonal basis $u_{j,k}^n(t) = 2^{-j/2} u^n(2^{-j}t - k)$ (k is the translation factor) satisfies:

$$u_{j,0}^n(t) = \begin{cases} \sum_k h_0(k) u_{j-1,k}^i, & \text{if } n \text{ is even} \\ \sum_k h_1(k) u_{j-1,k}^j, & \text{otherwise} \end{cases} \quad (6)$$

where $j, k \in \mathbb{Z}, n = 0, 1, \dots, 2^j - 1, h_0(k)$ and $h_1(k)$ are a pair of orthogonal mirror filters with the relationship $h_1(k) = (-1)^{1-k} \cdot h_0(1-k)$.

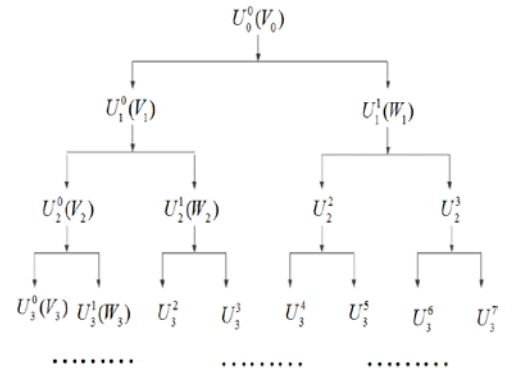


Fig. 1. Schematic of the spatial wavelet packet decomposition.

If $f(t)$ is a function in the Hilbert space $L^2(\mathbb{R})$, when the scale is small enough we can approximate the coefficient $d_0^0(k)$ of the space U_0^0 by the sampling sequence $f(k\Delta t)$ or the normalized $f(k)$. According to fast algorithm of WPD, the wavelet-packet coefficient of the j -th scale and k -th node can be expressed by:

$$d_j^n(k) = \begin{cases} \sum_m h_0(m-2k)d_{j-1}^{n/2}(m), & \text{if } n \text{ is even} \\ \sum_m h_1(m-2k)d_{j-1}^{(n-1)/2}(m), & \text{otherwise} \end{cases} \quad (7)$$

In this way, we can get wavelet-packet coefficients of a signal at all scales. It is known that the EEG signals relevant to MWL are in the frequency band of [0-50] Hz. The 17-channel electrophysiological signals are decomposed into five levels. Using (7), we extract the spectral power of the first six nodes ([0-7.8], [7.8-15.6], [15.6-23.4], [23.4-31.2], [31.2-39], [39-46.8]) as the features of the EEG signal.

IV. DATA COLLECTION EXPERIMENTS AND DATA PREPROCESSING

A. Subjects

Six subjects (22-24 y/o, male; coded by A, B, C, D, E, and F) participated in the experiments. All subjects were healthy, had normal vision and dextrmanual. Before the experiments, all subjects were informed of goals and procedure of the experiment and were trained for more than 10 hrs on aCAMS-based task operations.

B. Experimental Task Environment

The simulated task platform used in our experiments is automated-enhanced Cabin Air Management System (aCAMS), which consists of four subsystems: concentration of oxygen (O₂), air pressure (P), concentration of carbon dioxide (CO₂), and temperature (T). In the experiment, we used the aCAMS to simulate the task environment in a closed cabin. The operator's MWL is mainly affected by the Number Of Subsystems (NOS) assigned to him for manual control and the Actuator Sensitivity (AS) in the manual control systems. The aCAMS simulation platform constitutes a complex human-machine cooperative task environment. Nihon Kohden® measurement system was used to measure physiological signals at a sampling rate of 500 Hz.

C. Experimental Procedure and Data Acquisition

The aCAMS system has four subsystems, each having two control modes: automatic or manual control. The two modes of control can be switched arbitrarily. The control objective of the experiment is to maintain the output variables of the four subsystems within their target ranges by automatic control by automation systems, manual control by human operator, or a mixture of both modes. For manual control, there are two levels of actuator sensitivity (AS): Low or High. The sensitivity of the control variable under High AS is higher than Low AS [11].

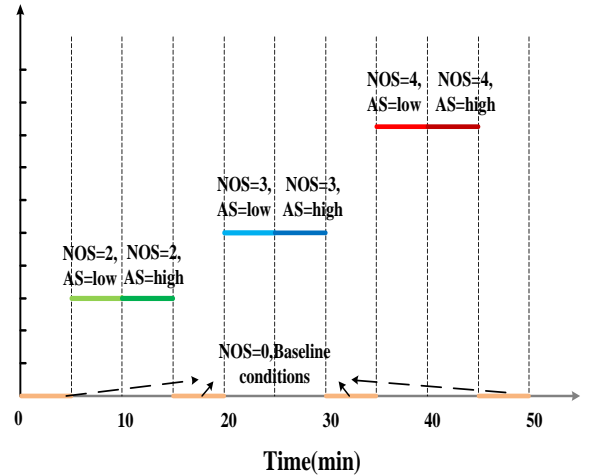


Fig. 2. The 10 task-load conditions in an experimental session.

Each session lasts for 50 min. and consists of 10 different task-load conditions. The conditions #1, 4, 7, and 10 are under automatic control mode. Operator manually controls two subsystems (O₂ and P) in the conditions #2 and 3, the only difference between the two conditions is that the AS is different. Fig. 2 illustrates the 10 task-load conditions in a session of experiment. During the last 10 s of each condition, the operator performs self-assessment of his performance in that condition, so we only consider the measured data of 290 s per condition.

The EEG, ECG and EOG signals for each subject were collected during the aCAMS operation by using a signal acquisition instrument (sampling rate: 500 Hz). The instrument has the function of removing the disturbance of the power frequency on the electrophysiological signals. In the international standard 10-20 EEG electrode placement system [12], 15 electrodes that are most relevant to the MWL variations were selected, namely F3, F4, C3, C4, P3, P4, O1, O2, Fz, Cz, CPz, Pz, AFz, POz, and Oz [13,14]. In addition, the potential difference between the upper middle part of the clavicle and the lower middle part of the left rib was recorded as ECG signal. The EOG signal was measured by the potential between the electrodes above and below the left eye. The recorded raw signals is filtered by a Butterworth band-pass filter (0-40 Hz) and the coherent method is used to remove the eye artifacts.

D. Data Labeling

The preprocessed data is divided by the sliding time window with length of 1 s (with no overlapping), then each load condition contains 290 sample data. In addition to physiological data, the experiment also records the task performance data, i.e., the output of the subsystems under control. Performance data for subject A is shown in Fig. 3, where the area between the red lines is the target range and the area between the pink lines is the safe range.

In order to quantify the MWL level, we define the Mental Workload Index (MWLI):

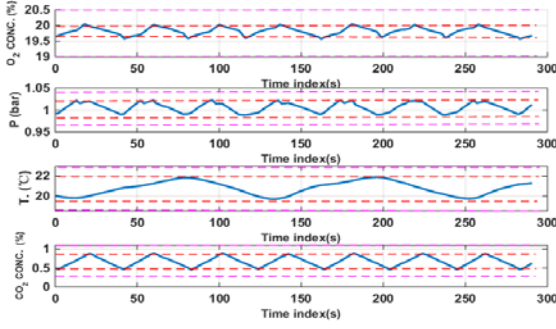


Fig. 3. The performance measures (i.e., four output time trajectories of human-machine cooperative control system under study) (subject A).

$$MWLI = w_{o_2} r_{o_2}(t) + w_p r_p(t) + w_{co_2} r_{co_2}(t) + w_T r_T(t) \quad (8)$$

where $r(t)$ with different subscript is Boolean variable of the corresponding subsystem (when the output of the corresponding subsystem is in target range at time t , $r(t) = 0$; otherwise $r(t) = 1$) and w represents the weight of the corresponding subsystem that can be determined by:

$$w = w_1 w_2 w_3 \quad (9)$$

Where w_1 represents the control weight of the corresponding subsystem (when the subsystem is under manual control, $w_1 = 1$; otherwise $w_1 = 0$), w_2 represents the difficulty level of the corresponding subsystem among four subsystems, and w_3 denotes the difficulty level of control of the corresponding subsystem with different level of AS. The values of w_2 and w_3 are empirically determined. The basic idea of entropy method is to determine the weight according to the indicator variability. In general, the smaller the information entropy of an indicator, the greater the variation in the indicator, the greater the amount of information provided, the greater the weight. By using (11) and (12), we obtain the second-to-second MWLI variations, as shown in Fig. 4. We can see that there exists individual difference across 6 subjects, but the overall trend of change is similar, for example, condition #9 has the peak (highest) level of MWL. The MWL level is higher in the conditions #3, 6 and 8, while the MWL level in the condition #2 and 5 is lower. The baseline conditions #1, 4, 7, and 10 are under automatic control, thus the MWL level in those 4 baseline conditions is zero (under-loaded). Based on those observations, we will classify the MWL into four classes (baseline, low, medium, high).

V. CLASSIFICATION RESULTS AND DISCUSSION

In this section, we present the SSL performance across the six subjects. We use the WPD algorithm to extract the relevant features. The effectiveness of SSL algorithm for MWL classification problem is validated by the pertinent empirical results. In addition, we examine the effect of the number of training data and the number of unlabeled data on the performance of SSL algorithm. Finally, we compare the performance of SSL algorithm and the commonly used supervised learning algorithms for OFS analysis.

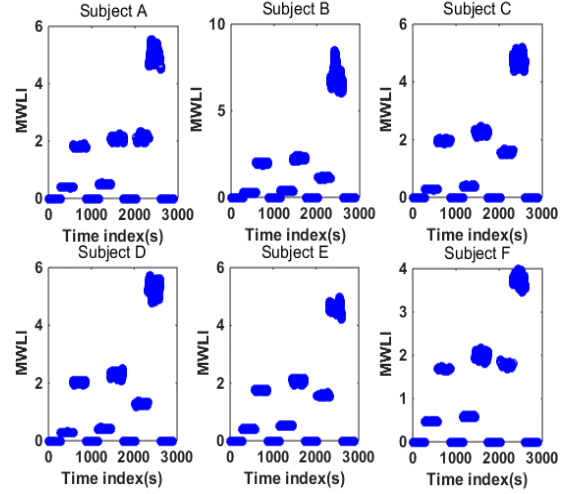


Fig. 4. The change of MWLI over time for each subject.

In this study we utilized a windowing approach with a sliding time window with length of 1 s. The Daubechies wavelet (db4) function was used to decompose the EEG signals by using 5-level WPD. In this way, we can obtain a dataset of 2900 feature data with a feature dimensionality of 102 (= 6 features/channel * 17 channels).

To avoid the impact of the smaller test set on the ability of the SSL algorithm to effectively utilize labeled and unlabeled data (i.e., unbalanced data classification problem), we divide the dataset into labeled data and unlabeled data at the rate of 1:4. The labeled data is equally divided into training and testing data. Finally, the number of training samples, test samples and unlabeled samples are 290, 290, and 2320, respectively.

A. Classification Results

In this section, the SS-ELM is applied to classify MWL. we give the 4-class classification confusion matrix for each subject in Figs. 5 and 6. Overall, using wavelet-packet-based features, SSL algorithm leads to promising classification performance.

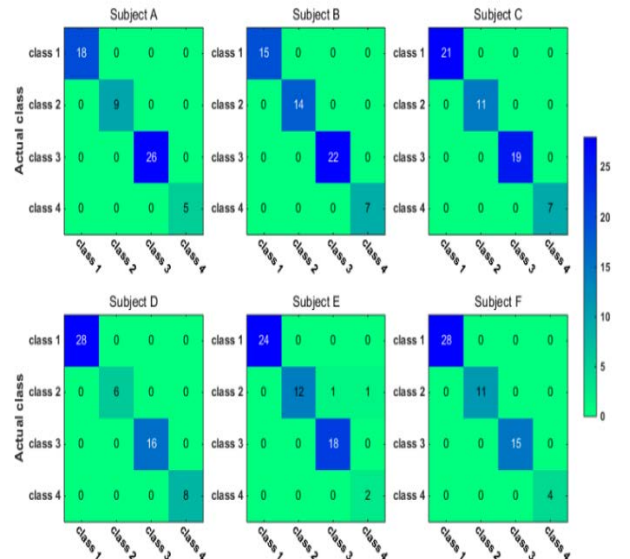


Fig. 5. Testing classification confusion matrix using wavelet packet features.

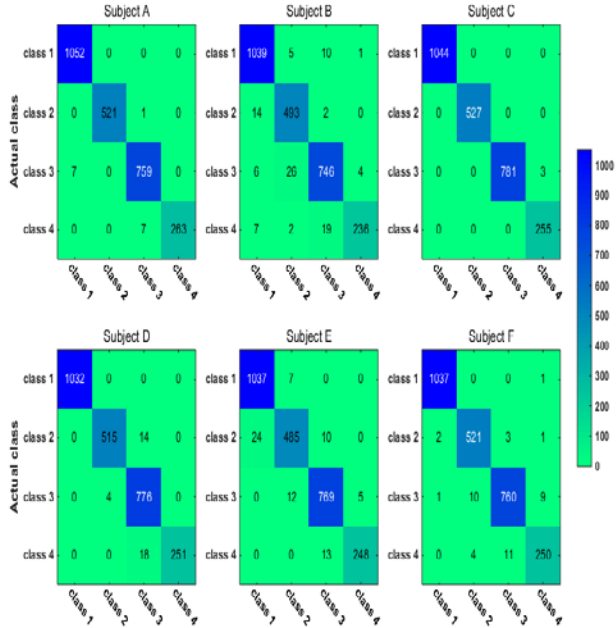


Fig. 6. The classification confusion matrix on unlabeled sample set using wavelet packet features.

B. Discussions

1) Effect of size of labeled dataset

To test the effect of the number of labeled training data on the performance of SS-ELM algorithm, we gradually increased the size of training set, while fixing the size of both the unlabeled and testing set to 2630. The training and test accuracy as well as the accuracy calculated on unlabeled data for each subject are shown in Fig. 7 (mean \pm standard deviation (s.d.)). We can see that except for subject C, the classification performance for other five subjects improves with the increase of the number of labeled data. For subject C, when the size of training set is 29, the accuracy already approaches 100%, so with the increase of the number of training samples, there is little room for further improvement of the accuracy. Therefore, we may conclude that satisfactory classification results can be obtained by using only a small number of labeled data. For other subjects, if training samples are scarce/sparse, the increase of the number of training samples has a great impact on the accuracy of the algorithm; However, if the training set is larger, the accuracy of the algorithm would improve little or stops improving with continued increase of the number of training samples. In summary, the benefit of SSL algorithm is reflected the best in the situations where only little labeled data is available.

2) Effect of size of unlabeled dataset

To test the capacity of the graph-based SSL algorithm in utilizing unlabeled data, we gradually increase the number of unlabeled data, while fixing the size of labeled set to 29. The corresponding classification accuracy is compared in Fig. 8 (mean \pm s.d.). We can see that except for subject C, the classification accuracy for other five subjects is improved with the increase of the number of unlabeled data.

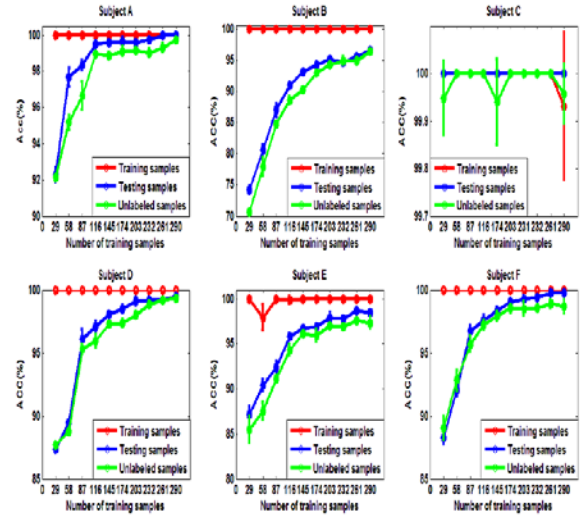


Fig. 7. The classification accuracy of using training sets of different size for each subject.

Does this observation really indicate that the more unlabeled data, the better the classification performance? To answer this question, we gradually increase the number of the unlabeled data while fixing the size of the labeled set to 290. The corresponding classification accuracy for each subject is shown in Fig. 9. It can be seen that when the number of training samples is 290, increasing the number of unlabeled samples has little effect on the classification performance. Therefore, when the labeled data are sufficiently extensive to characterize the data manifold, increasing the unlabeled data does not have much effect on the performance improvement. The fundamental advantage of the SSL algorithm for risky MWL detection is that if the labeled set is smaller, it has outstanding advantages over supervised learning; conversely, if the labeled set is large, its performance is comparable to that of supervised learning algorithm.

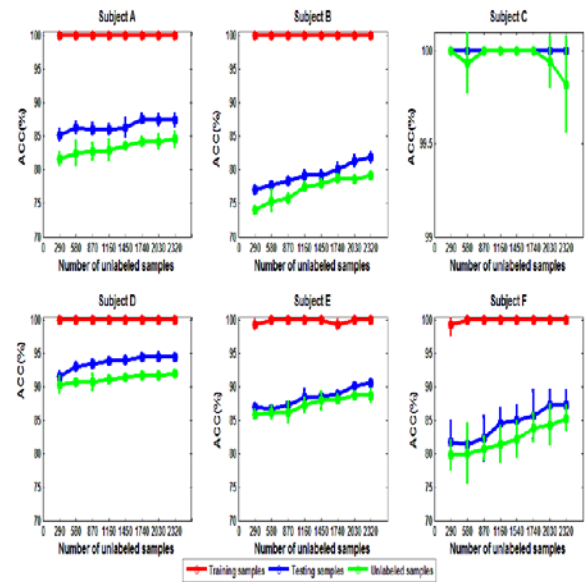


Fig. 8. The classification accuracy of using unlabeled sets of different size for each subject (size of labeled set: 29).

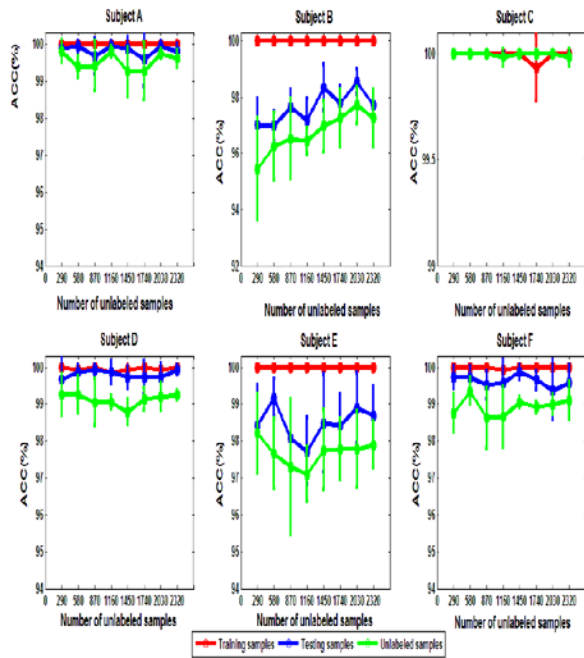


Fig. 9. The classification accuracy of using unlabeled sets of different size for each subject (size of labeled set: 290).

3) Performance comparison with supervised learning

In order to further verify the potential of the SS-ELM method for MWL classification, we compare it with four classical supervised learning algorithms, namely Naive Bayesian (NB), Random Forest (RF), Support Vector Machines (SVM), and ELM. The comparative results are shown in Fig. 10.

Since the SSL algorithm takes full advantages of a large number of unlabeled data, its classification accuracy is shown to be superior to that of the four major supervising learning algorithms, but the improvement of accuracy depends on the size of training set. When the number of labeled data is small, the performance enhancement of the SS-ELM method is most significant compared with supervised learning algorithms. On the contrary, when the number of labeled samples is large, the difference in classification performance between them is only marginal. Consequently, the SSL algorithm would be more applicable to the special scenarios in which the labeled data is difficult or expensive to collect.

VI. CONCLUSION

Although supervised learning techniques have shown promising performance in model-based MWL detection, a practical limitation of ML methods is the lack of sufficient number of labeled training data. Labeling massive data can be expensive or even erroneous given the little known domain knowledge about OFS state available. As the SSL method only requires small amount of labeled data, in this study the SSL paradigm is applied to real-time detection of high-risk MWL state using physiological data. We use SSL to exploit unlabeled data and improve the accuracy of high-risk MWL state detection.

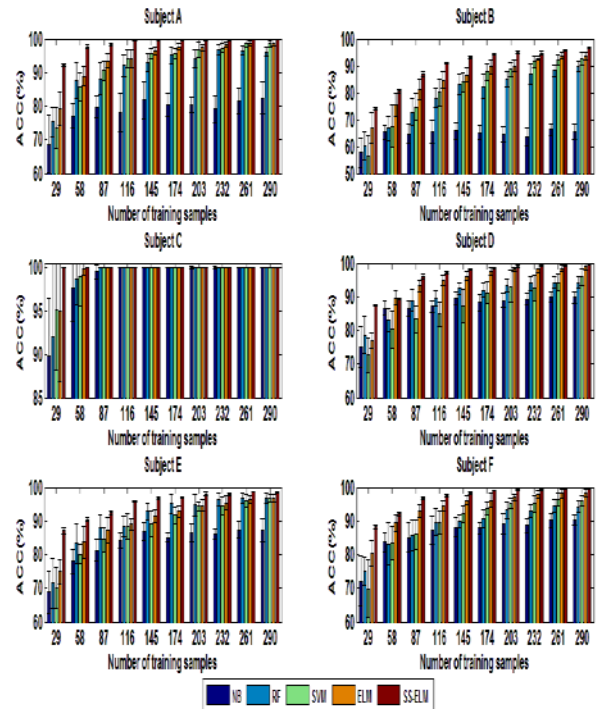


Fig. 10. The comparison of classification accuracy with gradual increase of size of the labeled (training) set for five different classifiers.

The results presented show that the proposed SSL approach is a promising alternative for risky MWL detection based on physiological signals. By exploiting the information contained in unlabeled data, the graphic semi-supervised learning method can reduce the computational cost and at the same time improve the detection accuracy. It was shown that even perfect classification accuracy can be achieved sometimes by SS-ELM. Furthermore, more can be gained by using SSL method with the increase of the size of unlabeled dataset. This result suggests that by exploring the structure of those unlabeled data, we can exploit additional information to improve the performance of MWL detection.

REFERENCES

- [1] S.K.L. Lal and A. Craig, "A critical review of the psychophysiology of driver fatigue," *Biological psychology*, vol. 55(3), pp. 173-194, 2001.
- [2] N. Bobko et al., "The mental performance of shiftworkers in nuclear and heat power plants of Ukraine," *Int. J. of Industrial Ergonomics*, vol. 21(3-4), pp. 333-340, 1998.
- [3] N. Hollender et al. "Integrating cognitive load theory and concepts of human-computer interaction," *Computers in Human Behavior*, vol. 26(6), pp. 1278-1288, 2010.
- [4] B. Cain, "A review of the mental workload literature," Defense Research and Development Toronto (Canada), 2007.
- [5] M. Mahfouf, J. Zhang, et al., "Adaptive fuzzy approaches to modelling operator functional states in a human-machine process control system," in *Proc. of IEEE Int. Conf. on Fuzzy Systems (FUZZ-IEEE 2007)*, 23-26 July 2007, London, UK, pp. 1-6.
- [6] J. Zhang, H. Liu, X. Peng, J. Raisch, and R. Wang, "Classifying human operator functional state based on electrophysiological and performance measures and fuzzy clustering method," *Cognitive Neurodynamics*, vol. 7, pp. 477-494, 2013.

- [7] J. Zhang, P. Qin, J. Raisch, and R. Wang, "Predictive modeling of human operator cognitive state via sparse and robust support vector machines," *Cognitive Neurodynamics*, vol. 7(5), pp. 395-407, 2013.
- [8] J. Zhang, S. Yang, and R. Wang, "Operator functional state estimation based on EEG-data-driven fuzzy model," *Cognitive Neurodynamics*, vol. 10(5), pp. 375-383, 2016.
- [9] G. Huang, S. Song, J.N.D. Gupta, and C. Wu, "Semi-supervised and unsupervised extreme learning machines," *IEEE Trans. on Cybernetics*, vol. 44(12), pp. 2405-2417, 2014.
- [10] D.K. Alves et al., "Real-time power measurement using the maximal overlap discrete wavelet-packet transform," *IEEE Trans. on Industrial Electronics*, vol. 64(4), pp. 3177-3187, 2017.
- [11] Y. Wang, J. Zhang, and R. Wang, "Mental workload recognition by combining wavelet packet transform and kernel spectral regression techniques," in *Proc. of 13th IFAC Symp. on Analysis, Design, and Evaluation of Human-Machine Systems (HMS2016)*, Kyoto, Japan, Aug. 30-Sep. 02, 2016; *IFAC-PapersOnLine*, vol. 49(19), pp. 561-566, 2016.
- [12] M. Okamoto et al., "Three-dimensional probabilistic anatomical cranio-cerebral correlation via the international 10-20 system oriented for transcranial functional brain mapping," *Neuroimage*, vol. 21(1), pp. 99-111, 2004.
- [13] Z. Yin and J. Zhang, "Operator functional state classification using least-square support vector machine based recursive feature elimination technique," *Computer Methods and Programs in Biomedicine*, vol. 113(1), pp. 101-115, 2014.
- [14] Z. Yin and 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," *Computer Methods and Programs in Biomedicine*, vol. 115(3), pp. 119-134, 2014.