

EEG-based tonic cold pain assessment using extreme learning machine

Mingxin Yu^a, Mingli Dong^a, Jing Han^b, Yingzi Lin^{c,d,*}, Lianqing Zhu^{a,*}, Xiaoying Tang^d, Guangkai Sun^a, Yanlin He^a and Yikang Guo^c

^a*Key Laboratory of the Ministry of Education for Optoelectronic Measurement Technology and Instrument, Beijing Information Science and Technology University, Beijing 100015, China*

^b*Emergency and Critical Care Center, Beijing Anzhen Hospital, Capital Medical University, Beijing 100029, China*

^c*School of Life Sciences, Beijing Institute of Technology, Beijing 100086, China*

^d*Intelligent Human-Machine Systems Lab, College of Engineering, Northeastern University, Boston, MA 02115, USA*

Abstract. The purpose of this study is to present a novel method which can objectively identify the subjective perception of tonic pain. To achieve this goal, scalp EEG data are recorded from 16 subjects under the cold stimuli condition. The proposed method is capable of classifying four classes of tonic pain states, which include No pain, Minor Pain, Moderate Pain, and Severe Pain. Due to multi-class problem of our research an extended Common Spatial Pattern (ECSP) method is first proposed for accurately extracting features of tonic pain from captured EEG data. Then, a single-hidden-layer feedforward network is used as a classifier for pain identification. With the aid of extreme learning machine (ELM) algorithm, the classifier is trained here. The advantages of ELM-based classifier can obtain an optimal and generalized solution for multi-class tonic cold pain. Experimental results demonstrate that the proposed method discriminates the tonic pain successfully. Additionally, to show the superiority for the ELM-based classifier, compared results with the well-known support vector machine (SVM) method show the ELM-based classifier outperform than the SVM-based classifier. These findings may pay the way for providing a direct and objective measure of the subjective perception of tonic pain.

Keywords: Common spatial pattern (CSP), electroencephalogram (EEG), extreme learning machine (ELM), tonic cold pain

1. Introduction

Pain is a complex human experience and a symptom or sign of numerous medical conditions. Therefore, pain is very difficult to gauge and quantify [1–3]. To help physicians and caregivers guide treatment accurately, the intensity and pattern of pain are necessary to be assessed. So far, the most common measures for pain intensity are visual analog scales, numerical rating scales and verbal rating scales [4]. However, those conventional methods have intrinsic limitations. On the one hand, human intervention by physicians or caregivers is that patients need to be required to perform the rating form. On the other

*Corresponding authors: Yingzi Lin, Intelligent Human-Machine Systems Lab, College of Engineering, Northeastern University, 360 Huntington Ave., Boston, MA 02148, USA. Tel.: +1 6173738610; Fax: +1 6173732921; E-mail: yilin@coe.neu.edu. Lianqing Zhu, Key Laboratory of the Ministry of Education for Optoelectronic Measurement Technology and Instrument, Beijing Information Science and Technology University, 6 Hongxia Road, Chaoyang District, Beijing 100015, China. Tel.: +86 13901168502; Fax: +86 10 68918820; E-mail: lianqingbistu@sina.com.

hand, those measures cannot achieve time-dependent dynamic changes in pain. In addition, according to [5], patient-reported pain is not always reliable, valid, and even meaningless, e.g. for demented patients and newborn patients. Furthermore, some patients who are addicted to drugs lie their pain intensity on purpose which is to get more medicines. Consequently, objective measures of pain have been expected for clinical pain assessment and pain management.

Significant efforts have been made to identify reliable and valid indicators of pain. Various medical imaging techniques are used to investigate the human pain, including single-photon emission tomography [6], positron emission tomography [7], and functional magnetic resonance imaging [8]. However, those methods are more expensive and used inconvenience for patients and physicians. Therefore, more and more experts are looking for other methods to achieve this purpose. The most comprehensive method is to adopt the facial expression identification for pain intensity analysis [9,10]. However, the challenge of this method is how to identify characteristics of whether patients are in real pain, since some patients could have abilities to pretend themselves in pain condition as an actor. Fortunately, pain could be expressed through multiple modalities and efforts are needed to extend technologies beyond facial expression to other methods. A prime example is the “Emotion and Pain Projection (EmoPain)” initiated by Engineering and Physical Sciences Research Council (EPSRC) [11]. Its purpose is to design and develop an intelligent system that will enable ubiquitous monitoring and assessment of patients’ pain-related mood and movements in the clinical environment. The project involves multidisciplinary collaboration of computer vision and machine learning scientists, clinical psychologists, physiotherapists, experts in pain management and rehabilitation.

Electroencephalogram (EEG) plays an important role in neurophysiological studies, since it is an inexpensive and non-invasive measure of brain activity. Therefore, it has gained much interest from the scientific community in searching for brain measures to characterize the perception of pain. Especially, related dynamic changes in the brain activity are captured appropriately, efficient analysis of EEG data may effectively reflect dynamic changes of pain perception due to the high temporal resolution of EEG.

This study attempts to research and assess the tonic pain under the innocuous cold stimuli condition, which is built with iced water simulation experiment. In this experiment, we recruited twenty subjects and collected EEG data from each subject under the tonic cold pain stimuli condition. In order to identify tonic cold pain, this paper proposes a new method which is capable of classifying four classes of pain states. It consists of two modules: feature extraction and classification. Because this study investigated the multi-classification task, we propose an extended Common Spatial Pattern (ECSP) method which can extract the features of four types of pain states. Then, a single-hidden-layer feedforward network classifier is utilized as a pain classifier. The classifier is trained by extreme learning machine (ELM) algorithm. To show the superiority for the proposed pain classification, compared results with the well-known support vector machine method are reported. Experimental results demonstrate that the method can be regarded as an objective cortical measure associated with the subjective perception of tonic cold pain and a promising candidate for the development of an expert system for clinical pain application.

The rest of this paper is organized as follows. Some related works to pain classification are given in Section 2. The methods used in our research are described in Section 3. Section 4 presents the process of the experimental design and the dataset establishment. Experimental results are reported in Section 5. Finally, in Section 6, conclusions are given.

2. Related works

In early studies, many researchers took advantage of EEG to investigate the association and relationship between EEG data and the perception of pain [12–14].

Steriade et al. [15] and Silva [16] showed that a decrease of alpha frequency band and an increase of beta frequency band were performed by an experimental pain stimulus. Similar research results were reported by other researchers [17–21]. With the analysis of power spectral changes, Dowan et al. [22] successfully identified features that tonic pain-related cortical activity under the cold condition. Nir et al. [23] provided a research work focused on the association between subjective perception of tonic pain and peak frequency of alpha band during stimulation and at rest. Then, Nir et al. [24] carried out another research for the prediction of subjective pain perception on alpha-1 frequency band (8–10 Hz) and found that it was able to be served as a direct, objective and experimentally stable measure of subjective perception of tonic pain. Shao et al. [25] investigated electrocortical responses in five frequency bands, i.e., 1–4 Hz, 4–8 Hz, 8–12 Hz, 12–18 Hz, and 18–30 Hz. Then, the pain states were discriminated using frequency-domain analysis. Gram et al. [26] performed a research to investigate the reliability of tonic cold pain scores and EEG spectral indices. The results showed that the EEG assessed during tonic cold pain was a valid experimental pain model.

In above research works, pain-related EEG signals were only analyzed in frequency domain, which focused mainly on the power spectrum peaks and coherence values at specific EEG frequency bands. Other researchers considered to utilize machine learning methods to identify pain states from EEG data. Hajileontiadis [27] have successfully attempted to discriminate relax from tonic cold pain state using the wavelet higher order spectral as a feature extraction method. Kumar et al. [28] designed an method based on fuzzy model for pain estimation and observed that EEG parameters ‘Hjorth Activity’ and ‘Spectral Entropy’ could reflect the level of pain experienced by the surgical patient during the postoperative period. Similar research provided by [29] was to develop an adaptive neuro-fuzzy interface system based on Support Vector Machine (SVM) to perform pain classification.

Except for those methods, Becker et al. [30] and Granocsky et al. [31] made efforts in investigating the pain intensity and NRS using evoked potentials (EP). The EP component was related to phasic and short stimuli, since it could be measured as a distinct waveform up to several hundreds of milliseconds after the stimulus [32]. However, [24] reported that phasic and short stimuli was not able to well adequately simulate natural and clinical painful experiences. Compared with phasic pain modes, therefore, it is commonly believed that tonic cold pain could better simulate clinic pain [23]. Additionally, it is well known that human brain can perceive distinct cold (or hot) somatic stimulation and discriminate its intensity with different affective responses [33].

3. The proposed method

Figure 1 illustrates the framework of our proposed tonic cold pain recognition method. This method consists of two stages, i.e., training stage and testing stage. In each stage, it contains three successive modules: 1) preprocessing module, 2) feature extraction module, and 3) ELM module. All training and testing EEG data come from the experimental data (see Section 4.4). The preprocessing module aims to remove noise signals. In the feature extraction module, an extended CSP (ECSP) algorithm is proposed for extracting the pain features from EEG data. The ELM module is taken as a pain classifier in this method.

Specifically, in the training stage, the ELM algorithm is used to estimate output weights of the SFNN with all training EEG data in a batch learning mode. The training inputs include a feature matrix $\mathbf{X}$ and a class label vector $\mathbf{Y}$. Each row of the feature matrix is composed of the feature vector from the preprocessed training EEG data with ECSP. The class label is represented to a pain state corresponding to a feature of training EEG data. Then, $\mathbf{X}$ and $\mathbf{Y}$ are fed to the ELM module to train the SFNN to

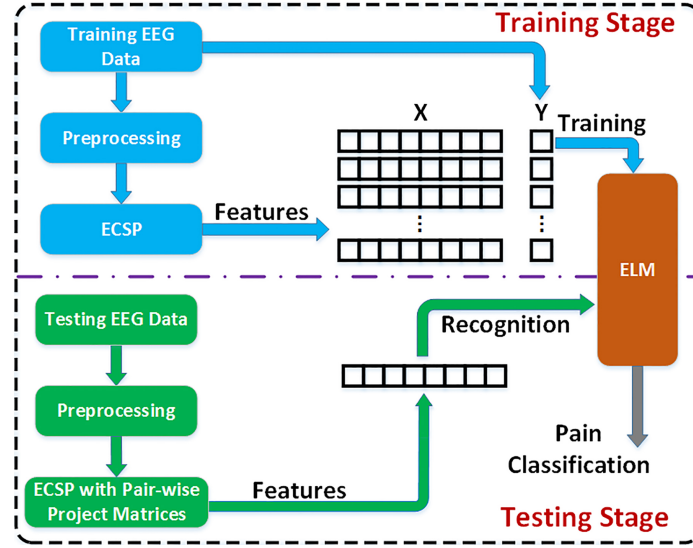


Fig. 1. Framework of the proposed procedure for pain assessment.

classify four classes of pain states. In this paper, we use two types of operations in the hidden nodes, i.e., dot product and kernel operation. The dot product operation represents for the ELM and kernel operation for kernel ELM.

In the testing stage, the same pre-treatments of EEG data are carried out for noise signals. Then the ECSP with trained pair-wise project matrices extracts features of received EEG data. The features pass through the trained ELM module to output pain class labels.

3.1. Preprocessing

Severe contamination of EEG activity by strong muscle, eye movements, eye blinks, and heart noise is serious problem for EEG interpretation and analysis. Three steps are used to preprocess EEG raw data. Firstly, EEG signals contaminated with strong muscle artifacts are manually rejected by visual inspection. Secondly, using an independent component analysis (ICA) with scalp topography, eye blinks, eye movements, and heart artifacts are removed from EEG data. Lastly, EEG data are filtered to five frequency bands, i.e., delta (1–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz), and gamma (30–49 Hz), with a band-pass Butterworth filter. All 32 channels have been filtered and used for further processing.

3.2. Extended common spatial pattern

After preprocessing the acquired EEG data, next step is to extract the features from EEG data. In this paper, we investigate the classification of multi-class pain of EEG signals. For this purpose, an extended CSP algorithm is proposed for pain features extraction.

3.2.1. Common spatial pattern

The method of CSP is to decompose raw EEG signals into spatial patterns, which are extracted from two classes of single trial EEG [46]. Obtained patterns can maximize the difference between two classes. It is first applied to EEG for detection of abnormalities [34].

The EEG data of a single trail is represented as an $N \times M$ matrix $\mathbf{X}$, where N is the number of channels (i.e., recording electrodes) and M is the number of samples per channel. For two classes of pain EEG data the matrix can be represented as $\mathbf{X}_1$ and $\mathbf{X}_2$. Therefore, the normalized spatial covariance of two classes of pain EEG signals can be calculated as:

$$\mathbf{C}_1 = \frac{\mathbf{X}_1 \mathbf{X}_1^T}{\text{trace}(\mathbf{X}_1 \mathbf{X}_1^T)} \quad (1)$$

$$\mathbf{C}_2 = \frac{\mathbf{X}_2 \mathbf{X}_2^T}{\text{trace}(\mathbf{X}_2 \mathbf{X}_2^T)} \quad (2)$$

where $\mathbf{X}_1^T$ and $\mathbf{X}_2^T$ denote the transpose of $\mathbf{X}_1$ and $\mathbf{X}_2$, trace is the sum of all diagonal entries of a matrix. For two classes of pain, the average normalized covariance $\overline{\mathbf{C}}_d, \in [1, 2]$ can be obtained by averaging over all trails of each group. The combined average normalized covariance is given as:

$$\mathbf{C} = \overline{\mathbf{C}}_1 + \overline{\mathbf{C}}_2 = \mathbf{U}_0 \mathbf{\Lambda} \mathbf{U}_0^T \quad (3)$$

where $\mathbf{U}_0$ is the matrix of eigenvectors and $\mathbf{\Lambda}$ is the corresponding diagonal matrix of eigenvalues.

The whitening transformation is expressed as:

$$\mathbf{P} = \mathbf{\Lambda}^{-1/2} \mathbf{U}_0^T \quad (4)$$

Then, using the whitening transformation $\overline{\mathbf{C}}_1$ and $\overline{\mathbf{C}}_2$ are transformed as:

$$\mathbf{S}_1 = \mathbf{P} \overline{\mathbf{C}}_1 \mathbf{P}^T \quad (5)$$

$$\mathbf{S}_2 = \mathbf{P} \overline{\mathbf{C}}_2 \mathbf{P}^T \quad (6)$$

where $\mathbf{S}_1$ and $\mathbf{S}_2$ share common eigenvectors [44], that is

$$\text{if } \mathbf{S}_1 = \mathbf{U} \mathbf{\Lambda}_1 \mathbf{U}^T \text{ then } \mathbf{S}_2 = \mathbf{U} \mathbf{\Lambda}_2 \mathbf{U}^T \text{ and } \mathbf{\Lambda}_1 + \mathbf{\Lambda}_2 = \mathbf{I} \quad (7)$$

where $\mathbf{I}$ is the identity matrix. According to Eq. (7), the eigenvector with the largest eigenvalue corresponding to one class of pain states and the eigenvector with the smallest eigenvalue corresponding to another class of pain states. Therefore, the whitening transformation of EEG onto eigenvectors of largest eigenvalues provides feature vectors which are optimal for discriminating two classes of pain EEG signal matrices. The projection matrix can be computed as:

$$\mathbf{W} = \mathbf{U}^T \mathbf{P} \quad (8)$$

With the projection matrix (2), the single trail EEG $\mathbf{X}$ can be transformed into uncorrelated components:

$$\mathbf{Z} = \mathbf{W} \mathbf{X} \quad (9)$$

According to the formulation, the original EEG signals is reconstructed by multiplying the inverse of $\mathbf{W}$ with $\mathbf{Z}$, as follows:

$$\mathbf{X} = \mathbf{W}^{-1} \mathbf{Z} \quad (10)$$

The columns of $\mathbf{W}^{-1}$ are the common spatial patterns and it can be considered as EEG source distribution vectors. The spatial filtered signal $\mathbf{Z}$ can maximize the differences in the variance of two classes of pain EEG measurements.

Using above formulas, the features used for pain classification are obtained by decomposing EEG data. A $2m \times N$ matrix $\mathbf{Z}$ is obtained by decomposing the inverse of $\mathbf{X}_i (i = 1, 2)$ using the projection

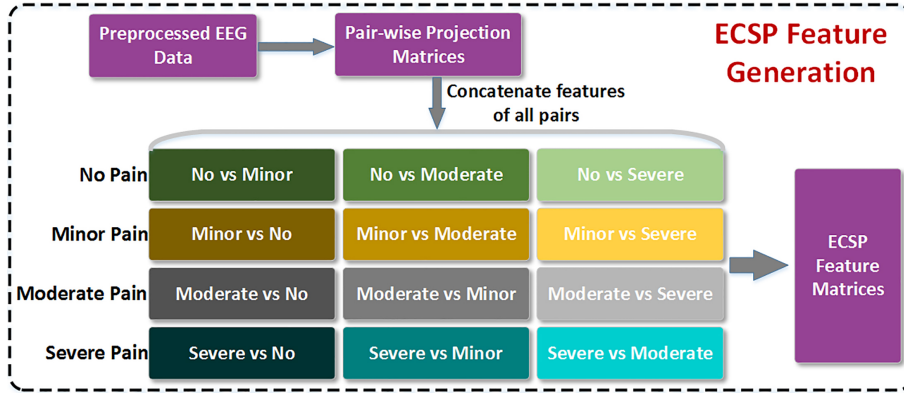


Fig. 2. The proposed ECSP scheme.

matrix $\mathbf{W}$. The variance of each row of $\mathbf{Z}$, i.e., $\text{var}(\mathbf{Z}_p)$, ($p = 1, 2, \dots, 2m$), is seen as features used for pain classification. The normalization of features can be computed as:

$$S_{i,p} = \log \left(\frac{\text{var}(\mathbf{Z}_{i,p})}{\sum_{i=1}^{2m} \text{var}(\mathbf{Z}_{i,p})} \right), (i = 1, 2) \quad (11)$$

Therefore, the features of two classes of pain are obtained by:

$$\begin{cases} \mathbf{x}_1 = [S_{1,1}, S_{1,2}, \dots, S_{1,2m}]^T \\ \mathbf{x}_2 = [S_{2,1}, S_{2,2}, \dots, S_{2,2m}]^T \end{cases} \quad (12)$$

where $\mathbf{X}_1$ and $\mathbf{X}_2$ represent features of two classes of pain states.

3.2.2. ECSP

The CSP intends to find the directions in the pattern space to optimally distinguish the two classes. In our research, however, this mission is a four-classification problem. For this purpose, we propose an extended CSP algorithm. Firstly, three CSP algorithms are used to extract three pair-wise projection matrices with each one representing a pair of pain states. Then, the three pair-wise projection matrices are concatenated. The detailed information is shown as follows.

Using Eq. (8) the pair-wise projection matrix $\mathbf{Q}_w(i, j)$ is calculated, where i and j are a pair of classes, $i \neq j$, $i, j = 1, 2, \dots, n_l$, n_l denotes the number of classes. Here, n_l is four, which denotes four classes of pain states.

The preprocessed EEG data corresponding to four classes of pain states is projected into pair-wise projection matrices to obtain feature vectors in term of Eqs (11) and (12), respectively. Then feature vectors of all pairs are concatenated to final ECSP feature matrices, which are used to train the ELM classifier. The illustration of the proposed ECSP scheme is shown in Fig. 2.

Algorithm 1 gives the generation process of ECSP features matrix based on pair-wise projection matrices.

3.3. Extreme learning machine

Extreme learning machine (ELM) proposed by [35] is used to train a SFNN as shown in Fig. 3. The method well overcome innate slow learning ability of traditional optimization techniques. It has been successfully applied many applications such as motor-imagery EEG classification for brain-computer

Algorithm 1: Establishment of ECSP features matrix**Inputs:** $\mathbf{Q}_w(i, j)$, $\mathbf{S}(i, j)$, $i, j = 1, 2, 3, 4$ **Output:** $\mathbf{F}_e$ $\mathbf{F}_e = \text{null};$ $l = 4;$ For $i = 1$ to n_l For $j = 1$ to n_l If $i \neq j$ Then $\mathbf{F}_c(i, j) = \mathbf{Q}_w(i, j)\mathbf{S}(i, j);$ $\mathbf{F}_e = [\mathbf{F}_e; \mathbf{F}_c(i, j)];$

End If

End For

End For

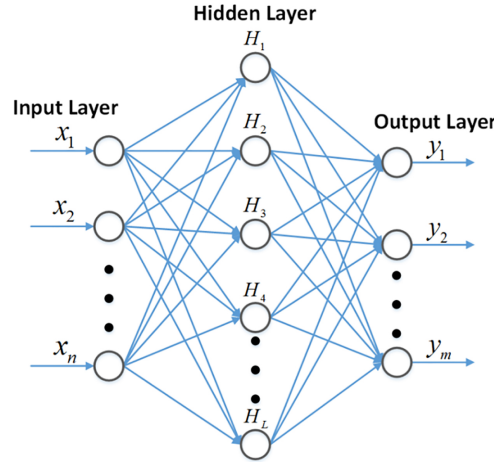
Return $\mathbf{F}_e;$ 

Fig. 3. A diagram of a single-hidden-layer feedforward neural network (SFNN).

interfaces [45], handwritten digit classification [40], human action recognition [41], traffic sign recognition [42], and ship detection [43]. In this paper, we consider known and unknown feature mapping functions applied into ELM, i.e., ELM-based classifier and kernel ELM-based classifier.

3.3.1. ELM-based classifier

The ELM-based classifier consists of three layers: input layer, hidden layer and output layer. The input layer is the feature vector $\mathbf{x}$ of one class of pain states. According to Section 3.2, the dimension of $\mathbf{x}$ is denoted as $2m$. As for the hidden layer, we assume that it has L hidden nodes. The output of a hidden node with index i can be expressed as follows:

$$g(\mathbf{a}_i, b_i, \mathbf{x}) = g(\mathbf{a}_i \cdot \mathbf{x} + b_i) \quad (13)$$

where $i = 1, 2, \dots, L$, $\mathbf{a}_i$ is the input weight vector between the hidden node and input node, b_i is the bias of this node, g is the activation function.

According to [36–38], almost all nonlinear piecewise continuous functions can be used as the activation function. In this paper, the sigmoid function is chosen as the activation function, it is expressed as:

$$g(\mathbf{a}_i, b_i, \mathbf{x}) = \frac{1}{1 + \exp(-(\mathbf{a}_i \cdot \mathbf{x} + b_i))} \quad (14)$$

Through the activation function, a $2m$ -dimensional space is transformed into an L -dimensional space, feature mappings can be denoted as:

$$\mathbf{h}(\mathbf{x}) = [g(\mathbf{a}_1, b_1, \mathbf{x}), \dots, g(\mathbf{a}_L, b_L, \mathbf{x})] \quad (15)$$

where $\{(\mathbf{a}_i, b_i)\}_{i=1}^L$ are randomly generated with any continuous probability distribution.

In the output layer, the number of neurons is denoted as M . M is the number of classified pain states. The output weight $\beta_{i,j}$ connects the hidden layer with the output layer, where $j = 1, 2, \dots, M$. Then, the output of the neuron is given by:

$$f_j(\mathbf{x}) = \sum_{i=1}^L \beta_{i,j} \times g(\mathbf{a}_i, b_i, \mathbf{x}) \quad (16)$$

According to Eq. (16), for the input sample $\mathbf{x}$, its output vector of the output layer can be expressed as:

$$\mathbf{f}(\mathbf{x}) = [f_1(x), \dots, f_M(x)] = \mathbf{h}(\mathbf{x})\boldsymbol{\beta} \quad (17)$$

where

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_L \end{bmatrix} = \begin{bmatrix} \beta_{1,1} & \dots & \beta_{1,M} \\ \vdots & \vdots & \vdots \\ \beta_{L,1} & \dots & \beta_{L,M} \end{bmatrix}$$

Above the description it can be learned that the structure of ELM by an input feature vector from the training pain EEG data. Then, the train process of ELM-based classifier is given as follows.

The proposed pain classification method belongs to the supervised learning problem. We have a training set with N samples, $\{\mathbf{X}, \mathbf{T}\} = \{\mathbf{x}_k, \mathbf{t}_k\}_{k=1}^N$. Here, $\mathbf{x}_k$ is a feature vector, $\mathbf{x}_k \in R^{2m}$, $\mathbf{t}_k$ is a M -dimensional binary vector (class label), $\mathbf{t}_k = [t_{k,1}, \dots, t_{k,M}]$, where $2m$ and M are the dimension of input and output, respectively. For the label vector, each entry is equal to one class which corresponds to the pain class that $\mathbf{x}_k$ belongs to. The real label vector for pain classes can be denoted as a matrix $\mathbf{T} = [\mathbf{t}_1, \dots, \mathbf{t}_N]^T$.

The weights and biases, i.e., $\{(\mathbf{a}_i, b_i)\}_{i=1}^L$, are randomly generated with any continuous probability distribution. Therefore, only parameters $\boldsymbol{\beta}$ in ELM needed to be trained. Here, the actual output vector is denoted as $\mathbf{y}_k$. The Eq. (17) can be transformed into the following linear representation:

$$\mathbf{H}\boldsymbol{\beta} = \mathbf{Y} \quad (18)$$

where

$$\boldsymbol{\beta} = \begin{bmatrix} \mathbf{h}(\mathbf{x}_1) \\ \vdots \\ \mathbf{h}(\mathbf{x}_N) \end{bmatrix} = \begin{bmatrix} g(\mathbf{a}_1, b_1, \mathbf{x}_1) & \dots & g(\mathbf{a}_L, b_L, \mathbf{x}_1) \\ \vdots & \vdots & \vdots \\ g(\mathbf{a}_1, b_1, \mathbf{x}_N) & \dots & g(\mathbf{a}_L, b_L, \mathbf{x}_N) \end{bmatrix}$$

and

$$\mathbf{Y} = \begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_N \end{bmatrix} = \begin{bmatrix} y_{1,1} & \dots & y_{1,M} \\ \vdots & \vdots & \vdots \\ y_{N,1} & \dots & y_{N,M} \end{bmatrix}$$

Based on the theory in [39], the ELM aims to achieve not only the smallest error but also the smallest norm of output weights. Therefore, its purpose is to minimize the training error $\|\mathbf{T} - \mathbf{Y}\|^2 =$

$\|\mathbf{T} - \mathbf{H}\beta\|^2$ and $\|\beta\|$. With the constrained-optimization algorithm, the training process can be formulated as:

$$\begin{aligned} \text{Minimize: } \phi(\beta, \xi) &= \frac{1}{2}\|\beta\|^2 + C\frac{1}{2}\|\xi\|^2 \\ \text{Subject to: } \mathbf{H}\beta &= \mathbf{T} - \xi \end{aligned} \quad (19)$$

where $\xi = [\xi_1, \dots, \xi_N]$ is the training error vector, C is a regularization factor.

In our experimental data, the number of training samples are not huge, so β can be obtained as:

$$\beta = \mathbf{H}^T \left(\frac{\mathbf{I}}{C} + \mathbf{H}\mathbf{H}^T \right)^{-1} \mathbf{T} \quad (20)$$

Based on the KKT theorem, the optimization problem of the ELM can be solved [39]. According to the above equations, we found that just two parameters need to be tuned, i.e., the number of hidden neurons L and the regulation factor C . As a whole, the training process of the ELM-based classifier can be summarized in Algorithm 2.

Algorithm 2: Training process of the ELM-based classifier

Input: Given a training set $\{(\mathbf{x}_k, \mathbf{t}_k) | \mathbf{x}_k \in \mathbf{R}^{2m}, \mathbf{t}_k \in \mathbf{R}^M, k = 1, \dots, N\}$, activation function g , and the number of hidden neurons L .

Output: Output weight matrix β

- 1: Randomly generate hidden neuron parameters $(\mathbf{a}_i, b_i)$, where $i = 1, \dots, L$;
 - 2: Calculate the matrix $\mathbf{H}$ in the hidden layer;
 - 3: Calculate the output weight β by Eq. (20).
-

3.3.2. Kernel ELM-based classifier

The kernel ELM-based classifier is also considered in this paper. If the feature mapping function $\mathbf{h}(\mathbf{x})$ is unknown, the Mercer's condition can be used into ELM, i.e., kernel technique. The kernel matrix for ELM is defined as follows [34]:

$$K(\mathbf{x}_i, \mathbf{x}_j) = h(\mathbf{x}_i) \cdot h(\mathbf{x}_j) \quad (21)$$

Then, based on Eq. (20), the output function of the ELM classifier can be represented compactly as:

$$\begin{aligned} \mathbf{f}(\mathbf{x}) &= \mathbf{h}(\mathbf{x})\beta = \mathbf{h}(\mathbf{x})\mathbf{H}^T \left(\frac{\mathbf{I}}{C} + \mathbf{H}\mathbf{H}^T \right)^{-1} \mathbf{T} \\ &= \begin{bmatrix} K(\mathbf{x}, \mathbf{x}_1) \\ \vdots \\ K(\mathbf{x}, \mathbf{x}_{N_k}) \end{bmatrix} \left(\frac{\mathbf{I}}{C} + \Omega_{\text{ELM}} \right)^{-1} \mathbf{T} \end{aligned} \quad (22)$$

where

$$\Omega_{\text{ELM}} = \mathbf{H}\mathbf{H}^T = \begin{bmatrix} K(x_1, x_1) & \cdots & K(x_1, x_{N_k}) \\ \vdots & \ddots & \vdots \\ K(x_{N_k}, x_1) & \cdots & K(x_{N_k}, x_{N_k}) \end{bmatrix}$$

and N_k is the number of training EEG data for the kernel ELM-based classifier. The Gaussian function is used as the kernel K which can be expressed as:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp \left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{\sigma^2} \right) \quad (23)$$

where σ denotes the standard deviation of Gaussian function.

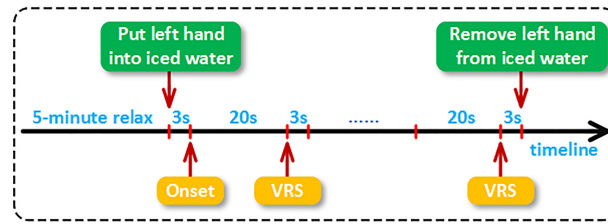


Fig. 4. Illustration of the experimental procedure.

4. Experimental procedure and data

This section presents the experimental apparatus and procedure for collecting and storing experimental data. Then the consistency verification of pain intensity ratings across the subjects is given. Lastly, the obtained data are used to train and test the proposed pain assessment method.

4.1. Experimental setup

The EEG data of subjects are recorded using the Enobio 32-electrode EEG wireless recording system (Neuroelectronics Inc., Barcelona, Spain). The captured data are streamed through the standard Bluetooth ISM band, operating distance range is 10 meters. The electrodes are inserted and arranged in the cap (Neuroelectrical cap) according to the international 10-10 system. The recorded scalp sites include P7, P4, Cz, Pz, P3, P8, O1, O2, T8, F8, C4, F4, Fp2, Fz, C3, F3, Fp1, T7, F7, Oz, PO4, FC6, FC2, AF4, CP6, CP2, CP1, CP5, FC1, FC5, AF3, and PO3 positions. The scalp electrodes are referenced to a pair of electrodes which are connected to CMS and DRL channels. The system is digitized at a sampling rate of 500 Hz.

4.2. Experimental procedure

Twenty subjects, nine female and eleven males, aged from 23–42 years, take part in the experiment. All are right-handed, not on any medications, and without any history of neurological and psychiatric disease. The whole experimental process is completed within five days. To be specific, in the first day, the participated subjects need to be familiar with the experimental procedure. From second day to fourth day, the tonic cold pain stimulation experiment is carried out once in each day. During these experiments, EEG data does not need to be captured. The purpose is to determine the tolerance time of hand submerged in iced water and further validate the pain level consistency on all subjects in three experiments. The last day, the EEG data are captured from each subject. The detailed process of the experiment is described as follows.

Each subject with eyes open is comfortably seated in an upright chair with a distance of 1 meter from a computer screen. The diagram of experimental process in a trail is shown in Fig. 4. Firstly, the subject is given 5 minutes to relax prior to the initiation of each trail. Then the subjects are asked to put his/her left-hand into a barrel with iced water mixture ($1^{\circ}\text{C} \pm 0.5^{\circ}\text{C}$) for tonic cold stimuli. The first 3 seconds are spent on keeping hands stable and still in iced water. Its purpose is easy to remove the muscular artifact due to hand movements when preprocessing EEG data. After 3 seconds, it is thought as the onset of tonic cold stimuli experiment. At the same time, the subject are required to remain as still as possible and to focus their sight on the displaying green dot in order to minimize muscular and occipital artifacts.

After each green dot, 3-second break is given to each subject and a number will be displayed on

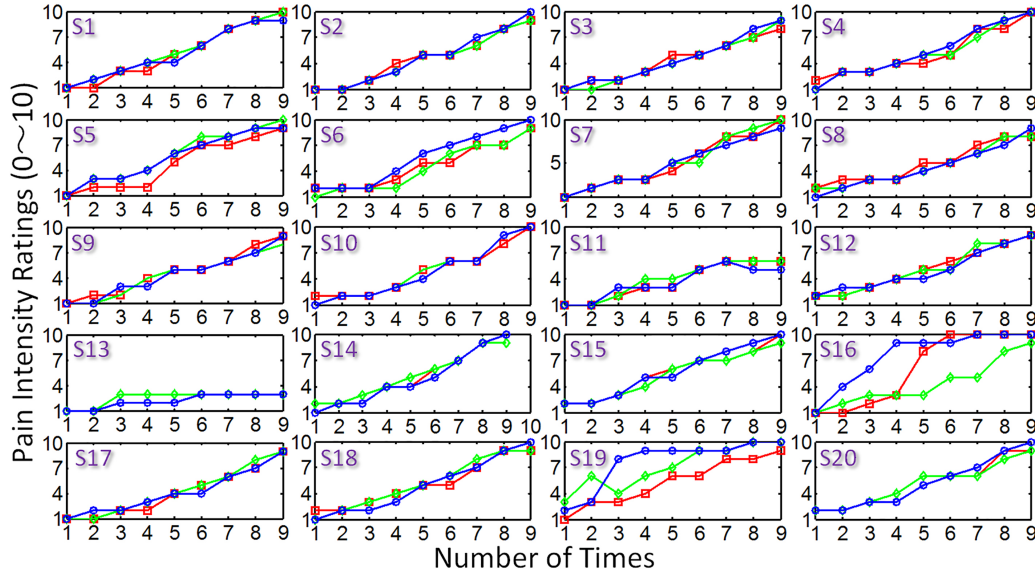


Fig. 5. Pain intensity ratings across 20 subjects in three experimental phases from second day to fourth day. (red line: 2nd day; blue line: 3rd day; green line: 4th day).

the screen. Then the subject is required to rate the perceived pain intensity on 11-point (0–10) verbal rating scales (VRSs, 0: no pain, 1: barely noticeable pain, 5: mild pain and 10: worst pain ever). The experimental process is repeated until the subject feel the pain became unbearable.

Some subjects were from Northeastern University, Boston, USA, and the experimental procedure was approved by the Northeastern University Institutional Review Board (IRB). Left subjects were volunteers and gave written informed consent prior to the beginning of the experiments.

4.3. Consistency verification of pain intensity ratings

Through experiments in three days, the longest tolerance time with a hand submerged into iced water is 210 seconds. In other words, it covers nine verbal rating scores for pain perception. However, each subject probably rates different perceived pain intensity on same point-in-time due to nature of subjectivity and uncertainty. The results are shown in Fig. 5.

In this paper, the proposed method is capable of classifying four classes of pain states, which include No Pain (0), Minor Pain (1–3), Moderate Pain (4–6), and Severe Pain (7–10), respectively. Hence, to ensure the reliability and availability of pain EEG data, the consistency of perceived pain scores on self-reported VRS needs to be validated with two criteria. One is that nine pain scores for each subject are required to fall into the range of four classes of pain states averagely, i.e., each pain contains at least two pain scores. Another is that two adjacent subjective scores do not exceed 2. Figure 5 shows that subjects 11, 13, 16, and 19 do not meet the criteria. Other 16 subjects with subjective pain intensity ratings have a good consistency and continue to complete the tonic cold stimulation experiment.

4.4. Experimental data

According to the experimental procedure described last subsection, selected 16 subjects participate the experiment in last day. Figure 6 shows subjective pain scores of 16 subjects. Each pain state has at

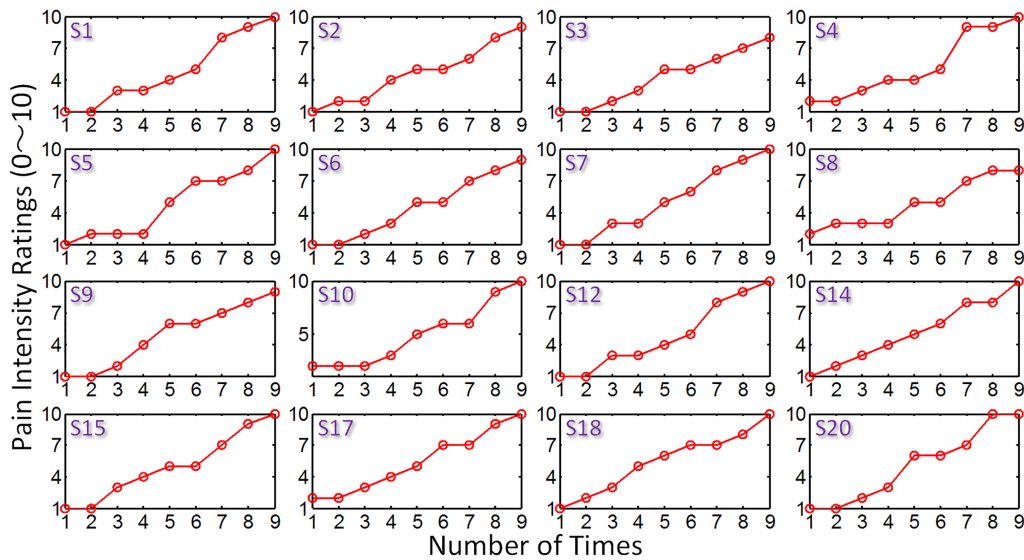


Fig. 6. Pain intensity ratings across 16 subjects in formal experimental phase.

least two 20 seconds EEG data. For no pain state, the EEG data are extracted from the relax phase prior to the initiation of each trail.

Then, EEG data of pain are partitioned for establishing pain dataset used for training and testing proposed classification method of pain states. From Fig. 6, we determine to extract two 20 seconds EEG data for each pain state. In our work, the size of one sample is set as 0.5 second EEG data. Hence 80 samples (Sampling Frequency: 500 Hz) are created for each pain state. The data preprocessing and proposed model were completed in MATLAB 2015.

4.5. *N*-fold stratified cross-validation and classification process

To assess the generalization capability and choose optimal parameters of the proposed classifier, the *N*-fold cross-validation method is used in our research. Firstly, the EEG dataset corresponding to each pain state is divided into *N* subsets. Then, the *N*-1 subsets are used as the training sets and the remaining one subset is used as the testing set. The training data window is moved with one subset interval. The training process is performed iteratively for *N* times. Lastly, the classification accuracies of *N* folds are averaged, the optimal parameters are obtained from the best average validation results.

More formally, considering to the lower mean square and bias, we decide to perform ten rounds of 10-fold cross-validation on the all dataset. This process is also called the 10-fold stratified cross-validation. Firstly, the dataset corresponding to each pain state is divided into 10 subsets with equal data size. In each round, then, nine out of ten subsets are utilized for the 10-fold cross-validation procedure, and the left one is used for testing. This process is repeated 10 times until all EEG signals are included. Its aim is to enable a different subset used as the testing data in each time. The classification performance across all ten rounds of testing for each subject is averaged for the final experimental results.

5. Results

To test the performance of tonic cold pain identification, this section gives the experimental results

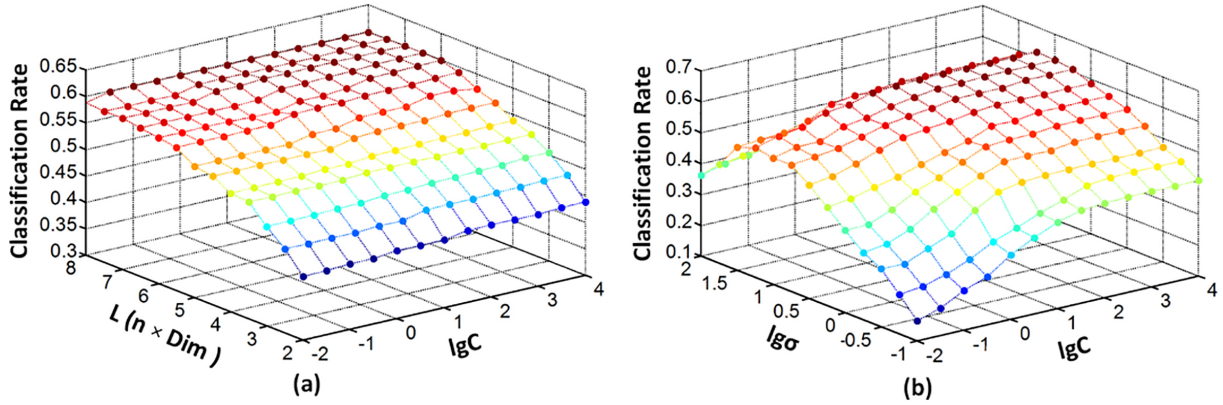


Fig. 7. Performance of (a) the ELM-based classifier and (b) the kernel ELM-based classifier to different combinations of (L , $\lg C$) and ($\lg \sigma$, $\lg C$) in theta band of EEG data from Subject 10.

obtained by our proposed method on EEG database. The procedure of parameters selection is first presented for the ELM model. Then, classification and performance evaluation results of the proposed method are reported. Lastly, to show the superiority of the ELM model, we present the comparison result with support vector machine on identifying the tonic cold pain.

5.1. Parameters selection

Because this paper investigates pain classification accuracy in five frequency bands of EEG data, i.e., delta (1–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz), and gamma (30–49 Hz). Taking this into consideration, the procedure of parameters selection needs to be performed in each frequency band. For a more concise description of this procedure, this section will give results of parameters selection in alpha and theta bands as two examples.

In this study, the grid search approach is utilized to select the optimal parameters for the ELM model. There are four parameters in the ELM model: 1) The ELM-based classifier has two parameters including regularization factor C and number of hidden neurons L ; 2) The kernel ELM-based classifier consists of Gaussian kernel spread σ and regularization factor C . The grid search approach seeks to compute these parameters during the 10-fold stratified cross-validation procedure.

For the ELM-based classifier, the value of C is represented in the form of $\lg C$, which is varied between -2 and 4 at intervals of 0.5 . The number of neurons in the hidden layer is varied between $2 \times \text{Dim}$ and $8 \times \text{Dim}$ at intervals of $0.5 \times \text{Dim}$. The Dim denotes the number of input nodes. Therefore, for each training, we try $13 \times 13 = 169$ combinations of parameters (L , $\lg C$) for the ELM-based classifier. Average classification rate across all 10 training is calculated, then the best optimal parameters are chosen in terms of best average validation results.

Figures 7a and 8a show average classification rate curves with different combinations of L and $\lg C$ in alpha and theta bands from one subject's EEG data. As we can see, the classification rate curves have tiny fluctuation as parameter $\lg C$ varied. However, classification rate curves goes up obviously with the increase of L , and this trend is stable when L is up to $5.5 \times \text{Dim}$. Therefore, it can be generally concluded that the classification rate is not very sensitive to the number of hidden neurons in the ELM-based classifier.

For the kernel ELM-based classifier, the Gaussian kernel spread σ is also presented in the form of $\lg \sigma$, which is varied between -1 and 2 at intervals of 0.25 . The form of the value of C is same as the procedure of the optimization selection in ELM-based classifier.

Table 1
Optimal parameters of the ELM-based classifier

Parameter	Delta	Theta	Alpha	Beta	Gamma
L	$4.5 \times \text{Dim}$	$4.5 \times \text{Dim}$	$6 \times \text{Dim}$	$6.5 \times \text{Dim}$	$5.5 \times \text{Dim}$
C	$10^{1.5}$	10^1	10^2	10^2	$10^{1.5}$

Table 2
Optimal parameters of the kernel ELM-based classifier

Parameter	Delta	Theta	Alpha	Beta	Gamma
C	$10^{2.5}$	10^2	$10^{3.5}$	$10^{3.5}$	10^2
σ	$10^{0.5}$	$10^{0.5}$	$10^{0.75}$	10^1	$10^{1.25}$

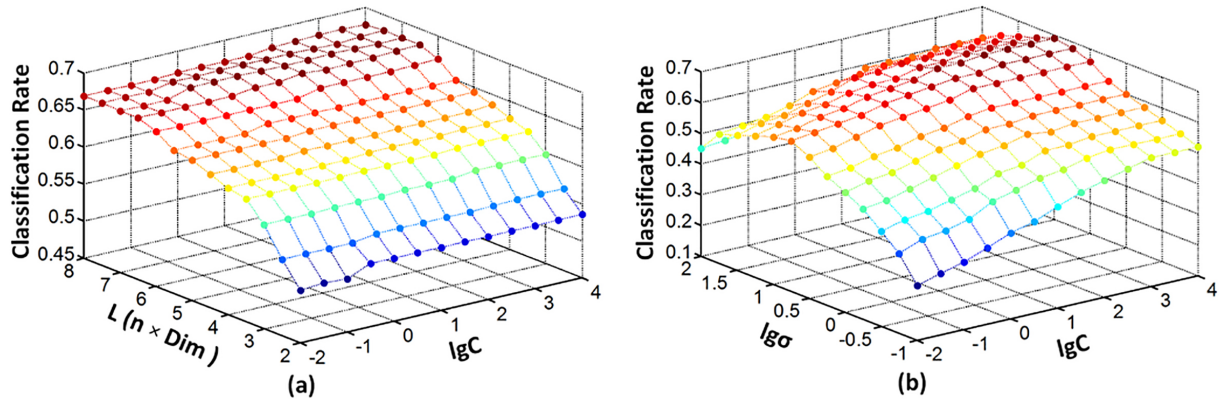


Fig. 8. Performance of (a) the ELM-based classifier and (b) the kernel ELM-based classifier to different combinations of (L , $\lg C$) and ($\lg \sigma$, $\lg C$) in alpha band of EEG data from Subject 10.

Figures 7b and 8b show average classification rate curves with different combinations of $\lg \sigma$ and $\lg C$. It can be seen that the classification rate goes up with the increase of $\lg \sigma$ and $\lg C$. From Fig. 7b we observe that there is a large flat area near the optimal value on the landscape, which means the classification rate is insensitive to combination of $\lg \sigma$ and $\lg C$.

Therefore, based on this procedure, we can obtain the optimal parameter combinations of the ELM-based and kernel ELM-based classifiers in each frequency band, as shown in Tables 1 and 2.

5.2. Classification results

Once optimal parameters are chosen, those are used to train the model over all the training data. In order to obtain lower mean square and bias, we performed ten rounds of 10-fold cross-validation on the all dataset. The procedure is described in subsection 4.5. Tables 3 and 4 show the classification results and average classification results which are evaluated across all 16 subjects with the ELM-based and kernel ELM-based classifier in five frequency bands of EEG data.

In Table 3, each column illustrates the classification performance comparison in five bands. As highlighted by red in this table, it can be seen that the ELM-based classifier achieves best classification results appeared in alpha and beta bands. Because of individual differences, we observe that the highest classification accuracy is 71.67% in beta band of subject 9 and the lowest classification accuracy is just 57.50% in beta band of subject 18. Across all 16 subjects, the average classification accuracy 66.77% obtained in alpha band is little higher than 66.17% in beta band.

Table 3

Classification accuracy obtained by the ELM-based classifier for 16 subjects in five bands of EEG data

Subjects	Delta	Theta	Alpha	Beta	Gamma
S1	52.06 ± 2.28%	62.50 ± 1.47%	70.00 ± 1.14%	70.41 ± 0.93%	57.50 ± 1.86%
S2	57.08 ± 1.14%	60.42 ± 0.00%	70.41 ± 0.93%	71.25 ± 0.93%	58.75 ± 2.72%
S3	53.75 ± 1.75%	56.25 ± 0.00%	67.08 ± 1.74%	65.42 ± 1.14%	54.58 ± 2.28%
S4	56.67 ± 2.28%	55.83 ± 1.74%	60.84 ± 0.93%	60.42 ± 0.00%	57.08 ± 1.14%
S5	50.00 ± 0.00%	50.83 ± 1.86%	66.67 ± 0.00%	65.00 ± 0.93%	53.33 ± 1.87%
S6	52.08 ± 1.47%	50.84 ± 3.16%	67.50 ± 1.14%	68.33 ± 2.28%	53.33 ± 1.14%
S7	42.50 ± 1.14%	47.50 ± 0.93%	65.41 ± 1.86%	67.08 ± 2.72%	52.92 ± 1.14%
S8	54.58 ± 1.74%	56.25 ± 4.66%	70.41 ± 0.93%	70.83 ± 1.47%	57.08 ± 1.86%
S9	61.67 ± 6.00%	63.75 ± 1.14%	69.58 ± 1.14%	71.67 ± 1.14%	63.75 ± 2.80%
S10	58.33 ± 1.47%	56.25 ± 2.08%	64.16 ± 0.93%	63.75 ± 1.14%	58.33 ± 1.47%
S12	55.00 ± 1.14%	56.25 ± 1.47%	66.67 ± 0.00%	65.00 ± 0.93%	57.50 ± 1.14%
S14	55.42 ± 1.14%	59.58 ± 1.14%	65.42 ± 1.14%	65.00 ± 1.75%	56.25 ± 2.08%
S15	54.58 ± 1.74%	57.50 ± 1.14%	69.17 ± 0.93%	67.50 ± 1.14%	55.42 ± 1.14%
S17	53.33 ± 1.14%	56.25 ± 0.00%	67.54 ± 1.11%	65.00 ± 1.75%	55.42 ± 1.14%
S18	54.17 ± 0.00%	52.50 ± 2.28%	61.25 ± 1.14%	57.50 ± 1.14%	59.17 ± 1.14%
S20	55.42 ± 1.14%	56.67 ± 0.93%	66.25 ± 0.93%	64.58 ± 0.00%	57.50 ± 1.86%
Average	54.17 ± 4.13%	56.20 ± 4.28%	66.77 ± 2.92%	66.17 ± 3.88%	56.74 ± 2.71%

Table 4

Classification accuracy obtained by the kernel ELM-based classifier for 16 subjects in five bands of EEG data

Subjects	Delta	Theta	Alpha	Beta	Gamma
S1	61.25 ± 2.28%	63.75 ± 1.14%	72.92 ± 2.09%	72.08 ± 1.14%	60.00 ± 0.93%
S2	57.50 ± 1.86%	60.42 ± 1.47%	72.92 ± 1.47%	73.75 ± 1.14%	61.67 ± 1.14%
S3	56.67 ± 0.93%	57.91 ± 0.93%	69.17 ± 1.74%	66.67 ± 1.47%	58.33 ± 0.00%
S4	57.50 ± 2.37%	57.08 ± 2.79%	62.50 ± 1.47%	61.25 ± 1.14%	60.00 ± 2.28%
S5	51.25 ± 1.14%	53.75 ± 0.93%	70.00 ± 1.14%	67.09 ± 0.93%	52.50 ± 1.75%
S6	52.92 ± 1.14%	55.83 ± 2.28%	67.50 ± 3.16%	67.50 ± 4.79%	55.42 ± 1.14%
S7	44.17 ± 2.28%	50.42 ± 1.74%	68.75 ± 2.55%	68.75 ± 2.95%	49.59 ± 2.72%
S8	55.83 ± 1.74%	57.50 ± 4.57%	72.08 ± 3.16%	72.92 ± 0.00%	56.67 ± 2.28%
S9	65.83 ± 1.14%	64.16 ± 0.93%	72.08 ± 1.14%	73.34 ± 0.93%	67.92 ± 1.14%
S10	59.59 ± 1.86%	57.92 ± 2.28%	66.67 ± 2.09%	65.42 ± 1.14%	60.83 ± 2.72%
S12	57.08 ± 1.14%	58.33 ± 0.00%	67.92 ± 1.14%	67.50 ± 1.14%	58.75 ± 0.93%
S14	58.33 ± 1.47%	61.67 ± 1.14%	68.40 ± 0.85%	65.42 ± 1.14%	61.25 ± 1.14%
S15	56.25 ± 2.08%	59.17 ± 1.14%	71.25 ± 0.93%	69.58 ± 1.14%	64.17 ± 1.74%
S17	55.00 ± 1.14%	57.50 ± 1.14%	69.58 ± 1.11%	68.75 ± 0.00%	59.58 ± 3.16%
S18	55.42 ± 1.14%	54.17 ± 1.47%	62.92 ± 2.28%	60.84 ± 0.93%	59.58 ± 1.14%
S20	56.67 ± 0.93%	57.08 ± 1.14%	67.92 ± 1.14%	67.92 ± 1.14%	60.00 ± 1.75%
Average	56.32 ± 4.62%	57.92 ± 3.55%	68.90 ± 3.12%	68.05 ± 3.82%	59.14 ± 4.28%

Table 4 lists the classification accuracies obtained by the kernel ELM-based classifier in five bands of EEG data. It is also noted that best classification results appeared in alpha and beta bands. Compared with results obtained by the ELM-based classifier, the kernel ELM-based classifier achieves higher classification accuracies highlighted by red in five bands for each subject. For average classification results, it can be seen that the kernel ELM-based classifier achieved 68.90% of classification accuracy which is higher than the ELM-based classifier in alpha band.

5.3. Performance evaluation

In this subsection, two performance measures account for evaluating outcomes of our proposed method, being the ‘Sensitivity’ and ‘Specificity’ metrics. The process is performed in alpha and beta bands which obtain higher classification accuracy showed in last subsection. Here, four classes of pain

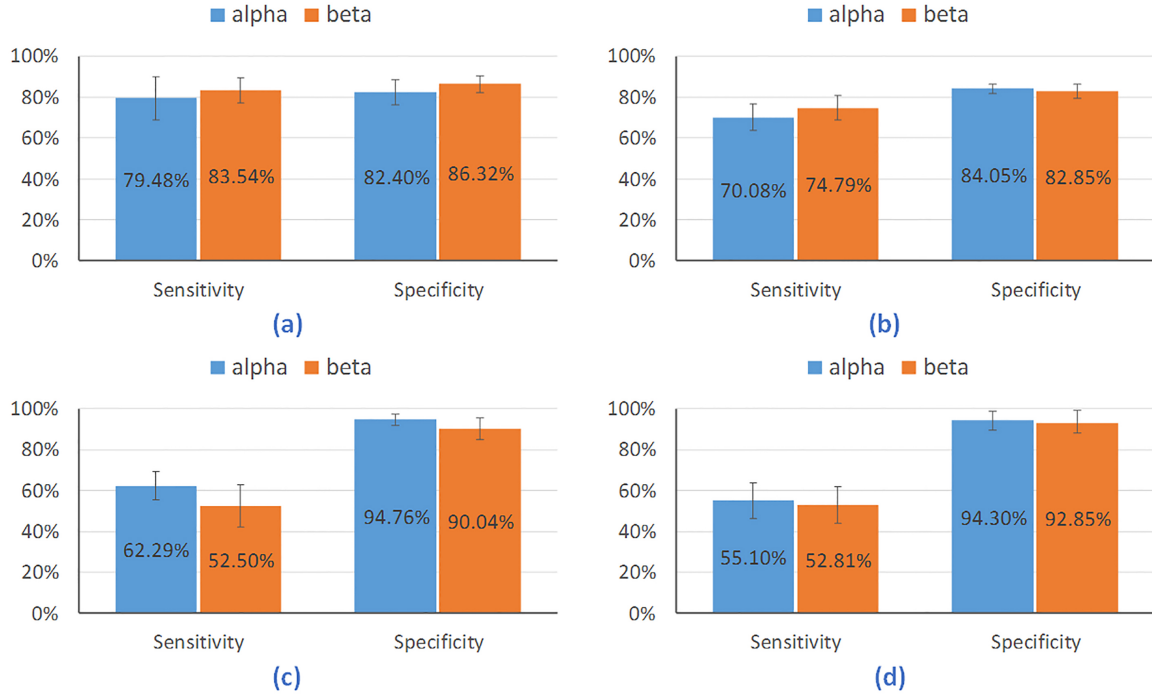


Fig. 9. Average sensitivity and specificity obtained by ELM-based classifier in (a) P1, (b) P2, (c) P3, and (d) P4 classification.

states, which include No pain, Minor Pain, Moderate Pain, and Severe Pain, are shortly expressed as P1, P2, P3, and P4, respectively.

We first present the results of performance evaluation for pain classification using the ELM-based classifier. Figure 9 shows sensitivity and specificity metrics in each pain class. It can be seen that the method achieves higher sensitivity metrics 79.48% and 83.54% in P1 classification. For the sensitivity of P4 classification, the results are lower than ones of P1 classification obtained 55.10% and 52.81%. However, in P4 classification, the specificity metrics are up to 94.30% and 92.85%. Similarity, we observe that the sensitivity metrics of P3 classification are 62.29% and 52.50%. Otherwise, it obtains 94.76% and 90.04% in specificity metrics.

Based on above results, it is concluded that the method used for correctly identifying P1 and P2 classes is better than identifying P3 and P4 classes. However, the ability of the method used for correctly identifying P1 and P2 classes is weaker than identifying P3 and P4 classes.

The results of performance evaluation for the kernel ELM-based classifier are given in Fig. 10. For the alpha band, it can be seen that sensitivity values of P1, P2, and P4 classification are 72.19%, 73.54%, and 72.54%, respectively. However, the sensitivity of P3 classification is 57.19%, its performance is weaker than other pain classes. For the beta band, the sensitivity values of P1 and P2 classification are better than ones of P3 and P4 classification. We observe clearly that the specificity value is over 85% in each pain class. Therefore, it can be easily concluded that the kernel ELM-based classifier obtains better performance on correctly identifying each pain class.

5.4. Performance comparison with support vector machine

To test the performance of the ELM-based and kernel ELM-based classifier, this subsection takes the

Table 5
Performance comparison between ELM and SVM

Methods	Alpha	Beta	Average training time (s)	Average testing time (s)
SVM	$65.95 \pm 3.67\%$	$65.65 \pm 3.77\%$	13.856	0.823
Kernel SVM	$68.72 \pm 3.45\%$	$67.57 \pm 3.99\%$	15.432	0.912
ELM	$66.77 \pm 2.92\%$	$66.17 \pm 3.88\%$	0.239	0.078
Kernel ELM	$68.90 \pm 3.12\%$	$68.05 \pm 3.82\%$	0.198	0.065

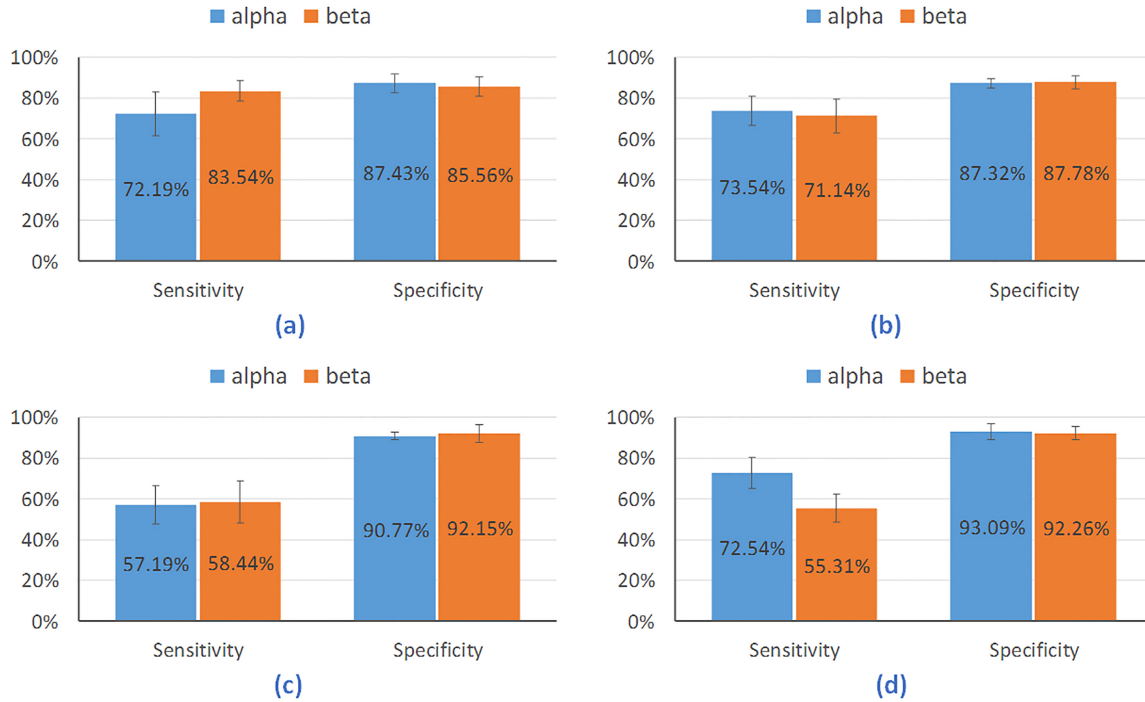


Fig. 10. Average sensitivity and specificity obtained by kernel ELM-based classifier in (a) P1, (b) P2, (c) P3, and (d) P4 classification.

SVM-based and kernel SVM-based model as two competing classifiers. The classification accuracies in alpha and beta bands, average training time, and average testing time are used as evaluation measures. All of classifiers are evaluated using the same database, same data preprocessing, same training process, same testing process, and same computation platform. In this paper, SVM and kernel SVM algorithms are implemented with LIBSVM toolkit.

From Table 5, it can be seen that the kernel ELM-based classifier achieves a slightly higher classification accuracies (alpha: $68.90 \pm 3.12\%$, beta: $68.05 \pm 3.82\%$) in comparison with those achieved by the kernel SVM-based classifier (alpha: $68.72 \pm 3.45\%$, beta: $67.57 \pm 3.99\%$). Similar results obtained by this experiment shows that the ELM-based classifier outperforms the SVM-based classifier in alpha and beta bands. As for average training time, the kernel ELM-based classifier is around 77 times faster than the kernel SVM-based classifier. In addition, average testing time is only 0.065 seconds for the kernel ELM-based classifier. Compared with the kernel SVM-based classifier, it is about 14 times faster than average testing time achieved by the kernel SVM-based classifier. For average testing time of the ELM-based classifier, it achieves a slightly shorter time than the kernel ELM-based classifier, and far faster time than the SVM-based classifier. Totally, the computation time of training and testing for the kernel

SVM-based and SVM-based classifier is far shorter than that for the kernel ELM-based and ELM-based classifier.

Based on the results obtained in this experiment, we found that the computation speeds of the kernel ELM-based and ELM-based classifier were much faster than the kernel SVM-based and SVM-based classifier during the model training. The reason is that input weights and hidden node biases randomly are chosen and output weights analytically are computed by the ELM model. Because the processes of complicated iterative operation are avoided well, the training speed of the ELM model is very fast. However, in the SVM model, the training procedure contains a convex quadratic programming problem. Hence, the computation time is quadratic to the number of training instances.

6. Conclusions

A new framework used for objectively identifying the subjective perception of tonic cold pain has been proposed in this paper. In particular, this method presents an extended Common Spatial Pattern (ECSP) algorithm as a feature extractor, which is capable of dealing with multi-class pain identification, to extract features of tonic pain EEG data. Then, the ELM algorithm is utilized for training the tonic pain classifier, which is a single-hidden-layer neural network. To verify this purpose of the study, 20 subjects are recruited and a procedure of cold stimuli simulation experiment is designed and performed.

The proposed method, in this paper, is capable of classifying four classes of pain states, which include No pain, Minor Pain, Moderate Pain, and Severe Pain. The promising experimental results reveal the potentiality of objectively identifying the subjective perception of tonic pain. Compared with the well-known SVM method, the proposed method costs much less computation time during the training and testing. Therefore, it is very promising for real-time applications. Overall, the proposed approach contributes with an alternative way to the endeavor toward object quantification of the subjective characterization of tonic pain. Future work will be focused on enhancing the recognition accuracy of tonic cold pain. All programs run on a laptop with 2.4 GHz i7-4700MQ CPU and 8 GB DDR3 RAM.

Acknowledgments

This work was supported by Program for Changjiang Scholars and Innovative Research Team in University [grant number IRT_16R07], National Natural Science Foundation of China [grant number 51705024], U.S. National Science Foundation (NSF) [grant number 1838796]. We thank all the participants who have participated in this work.

References

- [1] R. Melzack, The McGill Pain Questionnaire: major properties and scoring methods, *Pain* **1** (1975), 277–299. doi: 10.1016/0304-3959(75)90044-5.
- [2] R.L. Daut, C.S. Cleeland and R.C. Flanery, Development of the Wisconsin Brief Pain Questionnaire to assess pain in cancer and other diseases, *Pain* **17** (1983), 197–210. doi: 10.1016/0304-3959(83)90143-4.
- [3] T.D. Walsh, Practical problems in pain measurements, *Pain* **19** (1984), 96–98.
- [4] J.S. Shieh, C.Y. Dai, Y.R. Wen and W.S. Sun, A novel fuzzy pain demand index derived from patient-controlled analgesia for postoperative pain, *IEEE Transactions on Biomedical Engineering* **54** (2007), 2123–2132. doi: 10.1109/TBME.2007.896584.
- [5] S.M.G. Zwakhalen, J.P.H. Hamers, H.H. Abu-Saad and P.F.B. Martjin, Pain in elderly people with severe dementia: a systematic review of behavioral pain assessment tools, *BMC Geriatrics* **6** (2006). doi: 10.1186/1471-2318-6-3.

- [6] D.P. Vittorio, F. Stefano, S. Umberto, P. Patrizia, C. Giorgio and L.L. Gian, A cerebral blood flow study on tonic pain activation in man, *Pain* **56** (1994), 167–173. doi: 10.1016/0304-3959(94)90091-4.
- [7] P. Petrovic, K.M. Petersson, P.H. Ghatan, S. Stone-Elander and M. Ingvar, Pain-related cerebral activation is altered by a distracting cognitive task, *Pain* **85** (2000), 19–30. doi: 10.1016/S0304-3959(99)00232-8.
- [8] M. Ingvar, Pain and functional imaging, *Philosophical Transactions of the Royal Society B* **354** (1999), 1347–1358. doi: 10.1098/rstb.1999.0483.
- [9] K.D. Craig, K.M. Prkachin and R.V.E. Grunau, The facial expression of pain, 2nd ed., In D.C. Turk and R. M., eds, *Handbook of pain assessment*. Guilford, New York, 1992.
- [10] K.M. Prkachin and P.E. Solomon, The structure, reliability and validity of pain expression: evidence from patients with shoulder pain, *Pain* **139** (2008), 267–274. doi: 10.1016/j.pain.2008.04.010.
- [11] The Emotion & Pain Project, [Online]. Available: <http://www.emo-pain.ac.uk>.
- [12] P.F. Chang, L.A. Nielsen, T.G. Thomas and A.C.N. Chen, Psychophysical and EEG responses to repeated experimental muscle pain in humans: pain intensity encodes EEG activity, *Brain Research Bulletin* **59** (2003), 533–543. doi: 10.1016/S0361-9230(02)00950-4.
- [13] A.C.N. Chen, New perspectives in EEG/MEG brain mapping and PET/fMRI neuroimaging of human pain, *International Journal of Psychophysiology* **42** (2001), 147–159. doi: 10.1016/S0167-8760(01)00163-5.
- [14] T. Nomura, S. Fukudo, H. Matsuoka and M. Hongo, Abnormal electroencephalogram in irritable bowel syndrome, *Scandinavian Journal of Gastroenterology* **34** (1999), 478–484. doi: 10.1080/003655299750026209.
- [15] M. Steriade, P. Gloor, R.R. Linás, F.H.L.D. Silva and M.M. Mesulam, Basic mechanisms of cerebral rhythmic activities, *Electroencephalography and Clinical Neurophysiology* **76** (1990), 481–508. doi: 10.1016/0013-4694(90)90001-Z.
- [16] F.L.D. Silva, Neural mechanisms underlying brain waves: from neural membranes to networks, *Electroencephalography and Clinical Neurophysiology* **79** (1991), 81–93. doi: 10.1016/0013-4694(91)90044-5.
- [17] M. Backonja, E.W. Howland, J. Wang, J.L. Smith, M. Salinsky and C.S. Cleeland, Tonic changes in alpha power during immersion of the hand in cold water, *Electroencephalography and Clinical Neurophysiology* **79** (1991), 192–203. doi: 10.1016/0013-4694(91)90137-S.
- [18] P.F. Chang, L.A. Nielsen, T.G. Nielsen, P. Svensson and A.C.N. Chen, Topographic effects of tonic cutaneous nociceptive stimulation on human electroencephalograph, *Neuroscience Letters* **305** (2001), 49–52. doi: 10.1016/S0304-3940(01)01802-X.
- [19] A.C.N. Chen, S.F. Dworkin, J. Huang and J. Gehrig, Topographic brain measures of human pain and pain responsively, *Pain* **37** (1989), 129–141. doi: 10.1016/0304-3959(89)90125-5.
- [20] A.C.N. Chen, S.D. Dworkin, J. Huang and J. Gehrig, Human pain responsivity in a tonic pain model: psychological determinants, *Pain* **37** (1989), 143–160. doi: 10.1016/0304-3959(89)90126-7.
- [21] S. Ferracuti, S. Seri, D. Mattia and G. Cruccu, Quantitative EEG modifications during the Cold Water Pressor Test: hemispheric and hand differences, *International Journal of Psychophysiology* **17** (1994), 261–268. doi: 10.1016/0167-8760(94)90068-X.
- [22] R. Dowman, D. Rissacher and S. Schuckers, EEG indices of tonic pain-related activity in the somatosensory cortices, *Clinical Neurophysiology* **119** (2008), 1201–1212. doi: 10.1016/j.clinph.2008.01.019.
- [23] R.R. Nir, A. Sinai, E. Raz, E. Sprecher and D. Yarnitsky, Pain assessment by continuous EEG: association between subjective perception of tonic pain and peak frequency of alpha oscillations during stimulation and at rest, *Brain Research* **1344** (2010), 77–86. doi: 10.1016/j.brainres.2010.05.004.
- [24] R.R. Nir, A. Sinai, R. Moont, E. Harari and D. Yarnitsky, Tonic pain and continuous EEG: Prediction of subjective pain perception by alpha-1 power during stimulation and at test, *Clinical Neurophysiology* **123** (2012), 605–612. doi: 10.1016/j.clinph.2011.08.006.
- [25] S.Y. Shao, K.Q. Shen, K. Yu, E.P.V. Wilder-Smith and X.P. Li, Frequency-domain EEG source analysis for acute tonic pain perception, *Clinical Neurophysiology* **123** (2012), 2042–2049. doi: 10.1016/j.clinph.2012.02.084.
- [26] M. Gram, C. Graversen, S.S. Olesen and A.M. Drewes, Dynamic spectral indices of the electroencephalogram provide new insights into tonic pain, *Clinical Neurophysiology* **126** (2015), 763–771. doi: 10.1016/j.clinph.2014.07.027.
- [27] L.J. Hadjileontiadis, EEG-based tonic cold pain characterization using wavelet higher-order spectral features, *IEEE Transactions on Biomedical Engineering* **62** (2015), 1981–1991. doi: 10.1109/TBME.2015.2409133.
- [28] S. Kumar, A. Kumar, A. Trikha and S. Anand, Electroencephalogram based quantitative estimation of pain for balanced anesthesia, *Measurement* **59** (Jan. 2015), 296–301. doi: 10.1016/j.measurement.2014.09.021.
- [29] M. Vatankhah, V. Asadpour and R. Fazel-Rezai, Perceptual pain classification using ANFIS adapted RBF kernel support vector machine for therapeutic usage, *Applied Soft Computing* **13** (2013), 2537–2546. doi: 10.1016/j.asoc.2012.11.032.
- [30] D.E. Becker, D.W. Haley, V.M. Ureña and C.D. Yingling, Pain measurement with evoked potentials: combination of subjective ratings, randomized intensities, and long interstimulus intervals produces a P300-like confound, *Pain* **84** (2000), 37–47. doi: 10.1016/S0304-3959(99)00182-7.
- [31] Y. Granovsky, M. Granot, R.R. Nir and D. Yarnitsky, Objective correlate of subjective pain perception by contact heat-evoked potentials, *The Journal of Pain* (2008), 53–63. doi: 10.1016/j.jpain.2007.08.010.

- [32] J. Lorenz and L. Gracia-Larrea, Contribution of attentional and cognitive factors to laser evoked brain potentials, *Neurophysiologie Clinique/Clinical Neurophysiology* **33** (2003), 293–301. doi: 10.1016/j.neucli.2003.10.004.
- [33] P.F. Chang, L. Arendt-Nielsen and A.C.N. Chen, Comparative cerebral responses to non-painful warm vs. cold stimuli in man: EEG power spectra and coherence, *International Journal of Psychophysiology* **55** (2005), 73–83. doi: 10.1016/j.ijpsycho.2004.06.006.
- [34] B. Nasihatkon, R. Boostani and M.Z. Jahromi, An efficient hybrid linear and kernel CSP approach for EEG feature extraction, *Neurocomputing* **73** (2009), 432–437. doi: 10.1016/j.neucom.2009.07.012.
- [35] G.B. Huang, Q.Y. Zhu and C.K. Siew, Extreme learning machine: theory and applications, *Neurocomputing* **70** (2006), 489–501. doi: 10.1016/j.neucom.2005.12.126.
- [36] G.B. Huang and L. Chen, Convex incremental extreme learning machine, *Neurocomputing* **70** (2007), 3056–3062. doi: 10.1016/j.neucom.2007.02.009.
- [37] G.B. Huang and L. Chen, Enhanced random search based incremental extreme learning machine, *Neurocomputing* **71** (2008), 3460–3468. doi: 10.1016/j.neucom.2007.10.008.
- [38] G.B. Huang, L. Chen and C.K. Siew, Universal approximation using incremental constructive feedforward networks with random hidden nodes, *IEEE Transactions on Neural Networks* **17** (2006), 879–892. doi: 10.1109/TNN.2006.875977.
- [39] G.B. Huang, H.M. Zhou, X.J. Ding and R. Zhang, Extreme learning machine for regression and multiclass classification, *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics* **42** (2011), 513–529. doi: 10.1109/TSMCB.2011.2168604.
- [40] M.D. McDonnell, M.D. Tissera, T. Vladusich, A.V. Schaik and J. Tapson, Fast, simple and accurate handwritten digit classification by training shallow neural network classifiers with the ‘extreme learning machine’ algorithm, *PLOS One* **10** (2015), e0134254. doi: 10.1371/journal.pone.0.134254.
- [41] R. Minhas, A. Baradarani, S. Seifzadeh and Q.M.J. Wu, Human action recognition using extreme learning machine based on visual vocabularies, *Neurocomputing* **73** (2010), 1906–1917. doi: 10.1016/j.neucom.2010.01.020.
- [42] Z.Y. Huang, Y.L. Yu, J. Gu and H.P. Liu, An efficient method for traffic sign recognition based on extreme learning machine, *IEEE Transactions on Cybernetics* **47** (2017), 920–933. doi: 10.1109/TCYB.2016.2533424.
- [43] J. Tang, C. Deng, G.B. Huang and B. Zhao, Compressed-domain ship detection on spaceborne optical image using deep neural network and extreme learning machine, *IEEE Transactions on Geoscience and Remote Sensing* **53** (2015), 1174–1185. doi: 10.1109/TGRS.2014.2335751.
- [44] X.P. Bai, X.Z. Wang, S.H. Zheng and M.X. Yu, The offline feature extraction of four-class motor imagery EEG based on ICA and Wavelet-CSP, in: *The 33rd Chinese Control Conference*, Nanjing, China, 2014, pp. 1934–19. doi: 10.1109/ChiCC.2014.6896188.
- [45] Y. Zhang, Y. Wang, G.X. Guo, J. Jin, B. Wang, X.Y. Wang and A. Cichocki, Multi-kernel extreme learning machine for EEG classification in Brain-computer interfaces, *Expert Systems with Applications* **96** (2018), 302–310. doi: 10.1016/j.eswa.2017.12.015.
- [46] J.X. Chen, D.M.J. and Y.N. Zhang, An Extended Common Spatial Pattern Framework for EEG-Based Emotion Classification, in: *2018 International Conference on Brain Inspired Cognitive Systems*, Xi’an, China, 2018, pp. 282–292.

Copyright of Intelligent Data Analysis is the property of IOS Press and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.