

## Review Article

# Extreme Learning Machines on High Dimensional and Large Data Applications: A Survey

Jiuwen Cao<sup>1</sup> and Zhiping Lin<sup>2</sup>

<sup>1</sup>Key Lab for IOT and Information Fusion Technology of Zhejiang, Hangzhou Dianzi University, Zhejiang 310018, China

<sup>2</sup>School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 643798

Correspondence should be addressed to Jiuwen Cao; [jwcao@hdu.edu.cn](mailto:jwcao@hdu.edu.cn)

Received 12 November 2014; Accepted 21 March 2015

Academic Editor: Roman Lewandowski

Copyright © 2015 J. Cao and Z. Lin. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Extreme learning machine (ELM) has been developed for single hidden layer feedforward neural networks (SLFNs). In ELM algorithm, the connections between the input layer and the hidden neurons are randomly assigned and remain unchanged during the learning process. The output connections are then tuned via minimizing the cost function through a linear system. The computational burden of ELM has been significantly reduced as the only cost is solving a linear system. The low computational complexity attracted a great deal of attention from the research community, especially for high dimensional and large data applications. This paper provides an up-to-date survey on the recent developments of ELM and its applications in high dimensional and large data. Comprehensive reviews on image processing, video processing, medical signal processing, and other popular large data applications with ELM are presented in the paper.

## 1. Introduction

It is well known that the back-propagation (BP) based algorithms [1–3] played dominant roles in training feedforward neural networks (FNNs) in the past several decades. Although many efforts have been paid to enhance the BP algorithm, challenging issues such as local minima, time-costing in learning, and manual parameter setups still remain in the training phase and are not well addressed in the literature. These drawbacks may limit its applications in high dimensional and large data. Other than BP based neural network, support vector machine has been comprehensively investigated and implemented for model regression and classification applications [4–8]. Comparing to artificial neural networks, the relatively high generalization performance of SVM attracted increasing attention from researchers and engineers in the past two decades. However, both the BP based neural networks and the SVM are likely to face a relatively long learning time issue for complex data and suboptimal generalization performance. In the current big data and complex system era especially [9], data are explosive

with the rapid development of the internet, computer, and electronic equipments.

The timely proposed extreme learning machine (ELM) has shown its efficiency in training feedforward neural networks and overcoming the limitations faced by the BP algorithm and its variants [10–15]. The essences of ELM lie in two aspects, that is, random neurons and the tuning-free strategy. The learning phase of ELM generally includes two steps, namely, (1) constructing the hidden layer output matrix with random hidden neurons and (2) finding the output connections. Thanks to using random hidden neuron parameters which remain unchanged during the learning phase, ELM enjoys a very low computational complexity. The computational burden has been greatly reduced as the only cost is solving a linear system. At the same time, numerous applications have shown that ELM can provide a comparable or better generalization performance than the popular support vector machine (SVM) [4, 6] and the BP method in most cases [11, 16–18].

Theoretical support and analyses of ELM have been comprehensively studied in the literature. Such contributions

include the universal approximation capability [12], the classification capability [19], the unified learning platform and its comparisons to SVM [16], the feasibility assessments on its generalization performance, and pros and cons [20, 21]. Fruitful results covering the algorithm developments [14, 17, 18, 22–39] and applications [39–103] based on ELM have been achieved in the past several years. For the detailed summary on ELM variants and the insight into ELM on randomness and learning, we recommend the readers to refer to the interesting work [104, 105]. As the main concerns in [104, 105] are on the survey of theory and algorithm developments for ELM and its possible connections with human brain biological learning mechanisms, this paper aims to provide a detailed review on high dimensional and large data applications with ELM and its variants.

Benefiting from the low computation complex and the reasonable performance, ELM has shown powerful capability in handling large data, such as image processing and computer vision. In this paper, we aim to provide an up-to-date survey on applications in high dimensional and large data using the ELM method. A review on basic ELM and its recent development is first given. Then, applications using ELM and its variants on large data are reviewed, including image processing, video processing, and medical signal processing. Conclusions and future perspectives are also provided in the paper.

## 2. ELM Algorithms

*2.1. Basic ELM Algorithm.* Given a training data  $\mathcal{A} = \{(\mathbf{x}_i, \mathbf{t}_i)\}_{i=1}^N$ , the output function of the single hidden layer feedforward neural network (SLFN) with  $L$  hidden neurons can be expressed as

$$f(\mathbf{x}_i) = \sum_{j=1}^L \beta_j h_j(\mathbf{a}_j, b_j, \mathbf{x}_i) = \mathbf{h}(\mathbf{x}_i) \boldsymbol{\beta}, \quad i = 1, \dots, N, \quad (1)$$

where  $\boldsymbol{\beta} = [\beta_1, \dots, \beta_L]^T$  is the output weight matrix,  $\mathbf{h}(\mathbf{x}_i) = [h_1(\mathbf{a}_1, b_1, \mathbf{x}_i), \dots, h_L(\mathbf{a}_L, b_L, \mathbf{x}_i)]$  is the network output corresponding to the training sample  $\mathbf{x}_i$ ,  $h_j(\cdot)$  is a nonlinear piecewise continuous function, and  $\mathbf{a}_j \in \mathcal{R}^d$  and  $b_j \in \mathcal{R}$  ( $j = 1, 2, \dots, L$ ) are parameters of the  $j$ th hidden node, respectively. The network training is to find suitable network parameters to minimize the error function  $\|\mathbf{H}\boldsymbol{\beta} - \mathbf{T}\|_2$ , where

$$\mathbf{H} = \begin{bmatrix} \mathbf{h}(\mathbf{x}_1) \\ \vdots \\ \mathbf{h}(\mathbf{x}_N) \end{bmatrix}, \quad \mathbf{T} = \begin{bmatrix} \mathbf{t}_1^T \\ \vdots \\ \mathbf{t}_N^T \end{bmatrix} \quad (2)$$

are the hidden layer output matrix and the target output, respectively. The illustration of a SLFN with  $L$  hidden neurons is depicted in Figure 1.

Other than updating the network parameters iteratively as done in conventional gradient descent algorithms, ELM employs random hidden node parameters and the tuning-free training strategy for feedforward neural networks. The learning is then transferred to solving a linear system which

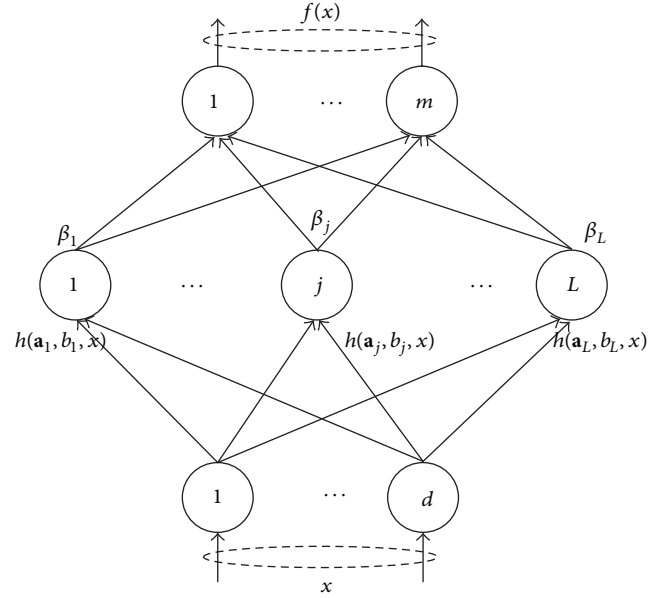


FIGURE 1: Single hidden layer feedforward neural network (SLFN).

has been well suited via the minimal norm least square approach [10, 11]. As shown in the universal approximation capability theorems [12], ELM is flexible with hidden activation functions. Almost any nonlinear piecewise continuous functions and their linear combinations work well in ELM algorithms [11, 104]. Thanks to these advantages, ELM has shown superiority of the fast learning speed and reasonable generalization performance over SVM and its variants [10, 11, 16].

*2.2. Unified ELM.* A unified learning platform of ELM which minimizes the training errors and the norm of the output weight matrix has been presented in [16]. A regulation coefficient is introduced to provide a trade-off between minimizing the training errors and the norm of output weights. The equality optimization constraints based ELM is expressed as

$$\begin{aligned} \min \quad & \frac{1}{2} \|\boldsymbol{\beta}\|_2^2 + \frac{C}{2} \sum_{i=1}^N \|\boldsymbol{\xi}_i\|^2, \\ \text{s.t.} \quad & \mathbf{h}(\mathbf{x}_i) \boldsymbol{\beta} = \mathbf{t}_i^T - \boldsymbol{\xi}_i^T, \quad i = 1, \dots, N, \end{aligned} \quad (3)$$

where  $\mathbf{t}_i$  and  $\boldsymbol{\xi}_i$  are the target output and the estimation error corresponding to the training sample  $\mathbf{x}_i$ , respectively, and  $C$  is the regulation coefficient. According to the Karush-Kuhn-Tucker (KKT) theorem [106], the network output weights are derived by solving the following dual optimization problem:

$$\begin{aligned} \mathbf{L}_{\text{DELM}} := \quad & \frac{1}{2} \|\boldsymbol{\beta}\|_2^2 + \frac{C}{2} \sum_{i=1}^N \|\boldsymbol{\xi}_i\|^2 \\ & - \sum_{i=1}^N (\mathbf{h}(\mathbf{x}_i) \boldsymbol{\beta} - \mathbf{t}_i^T + \boldsymbol{\xi}_i^T) \boldsymbol{\alpha}_i, \end{aligned} \quad (4)$$

where  $\alpha_i = [\alpha_{i,1}, \dots, \alpha_{i,L}]^T$ . The optimal network output weight matrix  $\beta$  can be derived by finding the derivatives of  $L_{\text{DELM}}$  with respect to  $\beta$ ,  $\alpha_i$ , and  $\xi_i$  and setting them to 0. The nonkernel case and the kernel case solutions of  $\beta$  have been presented in [16], respectively, that is,

(1) nonkernel case:

$$\beta = \left( \frac{\mathbf{I}}{C} + \mathbf{H}^T \mathbf{H} \right)^{-1} \mathbf{H}^T \mathbf{T}; \quad (5)$$

(2) kernel case:

$$\begin{aligned} \beta &= \mathbf{H}^T \left( \frac{\mathbf{I}}{C} + \mathbf{H} \mathbf{H}^T \right)^{-1} \mathbf{T} \\ &= \mathbf{H}^T \left( \frac{\mathbf{I}}{C} + \mathbf{K}_{\text{ELM}} \right)^{-1} \mathbf{T}. \end{aligned} \quad (6)$$

The detailed calculation of (5) and (6) can be referred to in [16]. Here,  $\mathbf{I}$  denotes the identity matrix and  $\mathbf{K}_{\text{ELM}}$  is the ELM kernel defined in [16] with  $\mathbf{K}_{\text{ELM}} = \{\mathbf{K}(\mathbf{x}_i, \mathbf{x}_j)\} = \{\mathbf{h}(\mathbf{x}_i) \cdot \mathbf{h}(\mathbf{x}_j)\}$ ,  $i, j = 1, \dots, N$ .  $\mathbf{h}(\cdot)$  is named the random feature mapping where all the parameters are randomly generated.

**2.3. ELM Variants.** Besides the basic ELM and the regularized ELM algorithms, a great number of variants based on ELM have been developed in the past several years. These developments can be broadly categorized into the fully complex ELM, the incremental ELM, the online sequential ELM, the ensemble ELM, and the pruning ELM.

In addition to discussions on real-valued neural networks, Huang et al. [15] extended the basic ELM to complex-valued neural networks, which can be employed to efficiently address equalization problems in digital communications. It was shown in [15] that ELM works well with fully complex activation functions in the hidden neurons. The universal approximation capability of ELM with complex continuous discriminatory or complex bounded nonlinear piecewise continuous function has been provided in [15].

The developments of incremental ELM (I-ELM) enriched the ELM family in the incremental learning field [12–14]. Huang et al. [12, 13] utilized the random hidden neuron to incrementally construct the FNNs. It was also shown [12] that I-ELM is flexible in hidden neurons activations. Thus, I-ELM and its variants enhanced I-ELM (EI-ELM) and error minimized ELM (EM-ELM) has shown superiorities over conventional incremental learning techniques as the usage of them is usually limited to certain types of activations functions. For example, the popular resource allocation networks only work for radial basis functions (RBF) [107, 108]. Other than constructing the network incrementally, Zhang et al. [14] have recently developed the dynamic ELM (D-ELM), which can iteratively add or delete hidden neurons according to their contributions to network performance. It has been demonstrated that D-ELM achieves a good generalization performance with a reduced network size.

Based on ELM and the recursive least square algorithm (RLS), Liang et al. [25] presented a novel online sequential

ELM algorithm (OS-ELM) for SLFNs. OS-ELM has well addressed the problem encountered by the conventional batch-mode learning algorithms which generally require that all training samples are available at hand for model constructions. OS-ELM involves two steps for sequential data learning, namely, (1) the model initialization with the existing training data and (2) the sequential learning phase with new collected samples. In the initialization phase, OS-ELM adopts the basic ELM to train the SLFN. In the sequential learning phase, the network output weight matrix is updated via the RLS algorithm by only exploiting the new received data. With this new learning strategy, OS-ELM enjoys the merits of low computational complexity and flexibilities in processing the number of sequential data; that is, OS-ELM is able to handle new captured data in both one-by-one and chunk-by-chunk with fixed or varying chunk sizes cases. Improvements based on OS-ELM include the online sequential fuzzy ELM [26], the ensemble based OS-ELM, the voting based OS-ELM classifier [28], and the ensemble of subset OS-ELM for class imbalance applications [39].

An alternative to the basic ELM with a single SLFN is the ensemble based ELM [17, 18]. To enhance the generalization performance of ELM and overcome the suboptimal solution issue which may be induced by random parameters, the ensemble based ELM employs multiple SLFNs to construct the decision model using a plurality consensus scheme. Various strategies have been utilized for the implementation of ELM ensembles in the past several years [17, 18, 28–30]. Liu and Wang [17] partitioned training data into various subsets with equal number of samples and implemented independent ELMs on each subset by using a training and cross validation approach to improve the generalization performance and avoid overfitting. Lan et al. took the averaging among multiple ELMs for online sequential training to reduce the network deviation. A majority voting based ELM ensemble (V-ELM) approach has been presented in [18] to enhance the classification rate. The method has been then extended to the majority voting online sequential ELM ensemble (VOS-ELM) in [28]. To further ameliorate the classification performance, Lu et al. [29] recently introduced two dissimilarity measures to calculate the similarities among ELMs. This approach can efficiently remove similar and redundant ELM ensembles existing in V-ELM and improve the recognition accuracy. To emphasize the contribution difference of each ELM ensemble to the final classification performance, Liu et al. [30] presented an improved approach named the evolutionary V-ELM (EV-ELM) for signal classification. The contribution difference of each ensemble is realized by a weighting scheme where the differential evolutionary algorithm is involved to automatically update the weights.

Pruning ELM and its variants are another group of learning approaches which have recently received comprehensive attention from the research community [22, 23]. To design a systematic classifier using ELM, firstly generated a network with a large size and then eliminated hidden neurons with low relevance to the class labels by means of the Chi-squared and information gain measures. Miche et al. [22] presented a robust and generic algorithm named the optimally pruned

ELM (OP-ELM) for regression and classification. The leave-one-out strategy is adopted in OP-ELM to select a proper number of hidden neurons where the multiresponse sparse regression algorithm is used to rank the significance of hidden neurons. It is stated in [22] that OP-ELM is more efficient than ELM in handling irrelevant or correlated data. Based on OP-ELM and fuzzy theory, Pouzols and Lendasse [23] developed an enhanced method named the evolving fuzzy OP-ELM (eF-OP-ELM) for the Takagi-Sugeno systems identification. Recently, Mozaffari and Azad presented an improved pruning algorithm based on ELM for the engine cold-start hydrocarbon emission identification. The novel method is referred to the ensemble of regularized OP-ELM using the negative correlation learning selection criterion (OP-ELM-ER-NCL).

Besides the above-mentioned variants on ELM, prominent contributions on algorithms in the past three years mainly include the Bayesian ELM [33] and the sparse Bayesian ELM [35], the bidirectional ELM [32], the parallel ELM (PELM), the hierarchical structure of ELM (HELM), the ELM tree [38], the timeless OS-ELM (TOSELM), the dissimilarity based ensemble of ELM (DE-ELM) [29], and the random projection based ELM (RP-ELM) [34]. These fruitful results have widely enriched the ELM family, not only in the development of methodologies, but also in their contributions to real-world applications.

### 3. ELM in High Dimensional and Large Data Applications

The property of low computational complexity in ELM attracts extensive attention from the research community. The reduced computational burden makes ELM and its variants more feasible in dealing with high dimensional and large data applications than conventional iterative algorithms and SVM. Simulation results presented in [11] showed that, for very large complex applications, ELM not only learns thousands times faster than conventional popular learning algorithm for FNNs but also produces good generalization performance. A recent progress on ELM for deep learning is given in [36]. The representational learning using multilayer ELM for deep networks is developed for big data processing. In this section, we aim to provide an up-to-date review on the high dimensional and large data processing with the ELM based approach. For the convenience of presentation, we conclude the applications in four parts in the following, that is, ELM in image processing, ELM in video applications, ELM in medical applications, and other applications.

*3.1. ELM in Image Processing.* Thanks to the reliable performance and fast learning speed, image recognition and objective detection using ELM have attracted increasing attention in recent years [40–55]. Among these applications, face recognition based on ELM and its developments is one of the hot topics which has been widely discussed by many researchers [40–44]. Zong and Huang studied the multilabel face recognition performance using the ELM classifier. Two

approaches, namely, ELM based on the one-against-all (ELM-OAA) and one-against-one (ELM-OAO) strategies, are used for face recognition. Discussions and comparisons on four benchmark face databases have shown that the ELM based classifier is able to achieve a comparable recognition rate to SVM but wins the convenience in parameter selection. Marques and Graña [40] proposed a novel feature extraction method named the lattice independent component analysis (LICA) for face image representation. The basic ELM and the regularized ELM are then adopted for face classification. The conventional feature extractions including principal component analysis (PCA), independent component analysis (ICA), linear discriminant analysis (LDA), and two state-of-the-art algorithms SVM and random forest are introduced for comparison. It is illustrated via experimental results in [40] that combining LICA with ELM obtains the best recognition performance. Choi et al. [41] proposed an incremental face recognition algorithm to address the real-time retraining problem for simultaneous recognitions in social network services such as Twitter and Facebook. The reduced Gabor features learned through the binarization of a Gabor filter by considering orientations in different grid positions are employed for face image representation. The OS-ELM is used to perform the sequential learning on each subregion after dividing the face image into equally sized local patches in [41]. Wang et al. [42] presented a discriminant tensor subspace analysis based face representation approach. The SLFN is then trained as classifier using the extracted features with the basic ELM algorithm. He et al. introduced a novel and fast face recognition by combining the sparse coding and ELM. The OP-ELM is utilized to learn the common feature hypothesis directly from the randomly collected universal images. To speed up the recognition process, ELM is embedded in finding the sparse representation coefficients. To enhance the recognition accuracy, Zhao et al. developed the ensemble of polyharmonic ELM (EP-ELM) for human face classification. The facial features are first obtained with the fast discrete curvelet transform (FDCT) and then the 2-dimensional principal component analysis (2DPCA) is used for feature dimensionality reduction. The EP-ELM classifier which exploits the performance of multiple P-ELM networks is implemented for face recognition in. In addition, Uçar [43] investigated the performance of color face image recognition with local features extracted by the steerable pyramid transform (SPT) and the basic ELM. Uçar et al. [44] analyzed the facial expression recognition via OS-ELM with the RBF networks. In [44], the curvelet transform coefficients on local cells of the image are first calculated and then the spherical clustering (SC) method is performed on the feature set to determine the optimal RBF network parameters in OS-ELM, where the feature set is derived by finding the entropy, the standard deviation, and the mean of curvelet coefficients of each region.

Besides applications in face recognition, much concern has been paid to object detection and image classification with ELM [45–51]. Lu et al. [45] applied the basic ELM to the palmprint recognition. Combining with the PCA and locality preserving projection based dimensionality reduction approach, ELM achieves a higher recognition rate on

the palmprint image classification than the BP based FNNs. In the meantime, the cost time of ELM on data learning and image recognition is far less than one second in dealing with hundreds of images. Minhas et al. implemented the original ELM to image object recognitions. The global and local pieces of information extracted via the two-dimensional PCA and Ferns style approach [109] are first employed for image representation. Then, parallel ELMs are learned as the classifier for object recognition. It is shown that ELM classifier is able to achieve a high recognition performance on several standard databases with more than ten thousand features. Moreover, the ELM classifier can recognize an image within a second compared to some related modern techniques, which generally take around 2-3 seconds per image. Cao et al. developed an effective ELM (EELM) for image classification. Incorporating the curvelet transform for image decomposition, the discriminative locality alignment for dimensionality reduction, extreme  $k$ -means method for feature set generating, and the proposed recognition framework with the EELM achieves a higher classification rate than the basic ELM. Man et al. [46] investigated the handwritten digit image recognition performance with an optimal weight learning machine based on ELM. The input weights in ELM are optimized in [46] by referring to the idea of model reference control in control engineering. Luo and Zhang [47] presented a hybrid method for handwritten image and face classification based on ELM and the sparse representation classifier (ELM-SRC). Combining the fast learning speed merit of ELM and the relative high accuracy advantage of SRC, the new classifier ELM-SRC is shown to have a faster training speed than SRC and a higher recognition rate than ELM. Zhu et al. [48] applied ELM for vehicle detection in a driving simulation platform. Comparing with SVM and BP, ELM has shown a comparable performance in virtual road segmentation and vehicle recognition with a fast learning and testing speed, which makes it suitable in real-time implementation. Rong et al. [49] adopted ELM to identify the images containing different types of aircrafts. Multiple SLFNs training with ELM are used as the classifier. The final category is decided using a weighted sum strategy on all the classification outputs obtained by each modular SLFN. To speed up the image recognition, Zhang et al. applied ELM into the standard ICA. The combination has well addressed the extremely high computational complexity issue encountered by ICA for real-time image recognition. Recently, Cao et al. [50] studied the performance of mobile landmark recognition with ELM. Due to the fast response requirement by mobile terminal users, an efficient classifier is highly desired. In experimental results show that SVM achieves a slightly higher recognition rate than ELM but with the price of long training and testing times. In [50], a robust and efficient online landmark learning approach has been presented. Incorporating the spatial pyramid kernel bag-of-words (SPK-BoW) image representation, ELM, and the RLS algorithm, the proposed online algorithm is efficient in processing new collected landmarks without performing retraining using the old data. To speed up the recognition phase of landmark images, Cao et al. [51] investigated the discriminative feature extraction approach for dimension

reduction. Measuring the representative capability and the discriminative capability of visual words based on relative entropy, the significance value of visual words is calculated and the PageRank approach is then introduced to filter out nondiscriminative visual words and thus reduce the feature dimension. To enhance the classification performance, the ensemble based ELM is adopted in [51] for the landmark recognition system. Samat et al. [54] and Bazi et al. [55] implemented ELM for high dimensional hyperspectral remote sensing image classification. In [54], two improved ensemble based ELM algorithms, namely, the bagging-based ELM and the AdaBoost-based ELM, are proposed to enhance the classification rate. Alternative to the ensemble based ELM, Bazi et al. [55] implemented an automatic-solution-based differential evolution to optimize the network parameters in ELM for hyperspectral image recognition.

Other than applications in face recognition and object detection, ELM based approaches have been also used for the image quality assessment [52], segmentation, superresolution [53], and so forth. Suresh et al. [52] developed two improved classifiers to measure the visual quality of JPEG-coded images based on ELM. To enhance the image quality assessment, the  $k$ -fold selection scheme and the real-coded genetic algorithm are utilized to optimize the input weights and the bias values of hidden neurons. Pan et al. utilized ELM classifier as a visual neuron system to detect the target object in leukocyte images. Efficient samples are automatically found by using the proposed ELM based segmentation approach. An and Bhanu [53] applied ELM to learn a neural network model for image superresolution (SR). To construct such a SR model, features extracted from low resolution images are considered as input signals for the neural network while high-frequency components from the corresponding high resolution image are set to be the target output.

*3.2. ELM in Video Applications.* In pace with the rapid development of computer vision and machine learning, intelligent video signal processing is another hot topic in high dimensional and large data applications with ELM. Most of the concern has been paid to human action recognition and tracking [56–65]. Minhas et al. [56, 57] studied human action recognition performance with ELM. In [56], novel hybrid features consisting of the spatial-temporal and the local static information are presented for human action recognition. The 3D dual-tree complex wavelet transform and the affine scale invariant feature transform (SIFT) descriptor are first employed to extract the hybrid features. The dimension reduction is then conducted through the bidirectional two-dimensional PCA approach in both the row- and columnwise directions, respectively. Visual vocabularies based on these two kinds of features are built for video action representation and the SLFN is finally trained with ELM as the classifier for human action recognition. In [57], an incremental learning framework for human action classification based on ELM and snippets has been developed. The contour of the object in video frame is used to initialize the classifier. The articulated target object is first tracked in a window in the video subsequent frame and then the human body contour approximated

with small rectangular boxes inside the tracking window is utilized to represent the target. Pyramids of histogram of oriented gradient (PHOG) features calculated from the tracking window and rectangular boxes are concatenated as a vector for human action representations. Such PHOG features are fed to a SLFN where the OS-ELM is employed for constructing an incremental classifier to avoid retraining. Iosifidis and his group contributed fruitful achievements to action recognitions with ELM [58–61]. In [58], human action recognition is performed with a multiple layer FNN where the number of layers and the number of hidden neurons in each layer are adaptively adjusted in the learning phase. The developed recognition scheme is flexible to the usage of human action representation approaches. In [58], Iosifidis et al. used the dynamic based action representation approach [110] to represent actions to verify the proposed algorithm. A dynamic recognition scheme with multiple classification levels is developed. In each level, the regularized ELM is performed as the learning algorithm for a SLFN and the most similar to the test action instance labeled vectors acted as input signals. In [59], a novel technique named the minimum class variance ELM (MCVELM) is presented for human action recognition. The action description, representation, and classification are performed with the spatiotemporal local shape and motion information, the fuzzy vector quantization, and the MCVELM classifier, respectively. A plenty of experiments on several benchmark databases covering the simple human actions, the sports single-view containing complex human action, the multiview database, and the facial expression recognition from videos are conducted to show the effectiveness of the proposed recognition framework. In [60], semisupervised action classifications with ELM are researched. The discriminative subspace learning and the ELM are combined for training a SLFN. An iterative optimization scheme is incorporated into the discriminative ELM for semisupervised multiview human action recognition. The minimum variance ELM (MVELM) is given in [61] to detect human action. The bag-of-words (BoW) approach is used to represent human actions and the MVELM is developed to minimize the output weights norm as well as the dispersion of the training data in the projection space. Budiman and Fanany [62] studied the 3D human motion pose-based classification with ELM. The  $k$ -means algorithm is used to cluster the motion features and the performance on badminton sport action and traditional dance are utilized as the databases in the experiments. Deng et al. [63] employed ELM to the cross-person activity classification for developing mobile based human-centric pervasive applications. An improved sequential learning method is designed to accelerate the data processing speed and enhance the recognition rate, in which a new transfer learning reduced kernel ELM (TransRKELM) is introduced for classifier initialization and the updating phase with new inputs is performed with the online sequential TransRKELM. Oh et al. [64] presented a novel signature recognition system based on extracting hand gestures with ELM. The system is realized using the Microsoft Kinect for data collecting. Four feature groups, that is, the hand position in horizontal and vertical directions and the hand movement in the horizontal and vertical directions, are extracted for

model learning and the total error rate minimization of ELM (TERELM) is adopted for classifier training. Yu et al. [65] proposed an online gesture recognition system using an adaptive and iterative online sequential ELM. The experiments are conducted on recognizing the writing gestures of 10 Arabic digits 0~9 and 26 letters a~z collected by the Microsoft Kinect.

Besides human action recognition, the merits of ELM have been explored in many other video applications including hand motion classification [66], visual tracking, video based semantic concept detection [67], and video watermarking [68]. Shi et al. [66] investigated the hand motion recognition using surface electromyography (SEMG) signals. The cumulative residual entropy (CREn) which measures the uncertainty in the SEMG signals is extracted as features for hand motion representation. The basic ELM is involved for classification. After comparing the performance to SVM, it is suggested that the proposed CREn-ELM based is applicable for real-time control of the SEMG-based multifunctional prosthesis as well as the SEMG-based hand motion recognition. Liu et al. applied ELM for visual tracking. A multitask ELM is proposed where a colearning semisupervised ELM is developed to address the rare labeled samples problem in tracking. The  $\ell_{2,1}$  norm penalty is imposed to achieve the joint sparse coding for discovering intrinsic relationship between the two ELMs with different cues and pruning redundant nodes in ELM. In the visual tracking, the model updating is realized through online sequential learning ELM. Lu et al. [67] exploited superiorities of using ELM for the video based semantic concept detection. The ELM based multimodality classifier combination framework including three ELM classifiers tested with color, edge, and texture features, a robust probability-based fusion method of predictions from each classifier, and the incorporation of contextual correlation among concepts into the classifier is presented to enhance the recognition performance. Agarwal et al. [68] adopted the ELM trained SLFN to realize a fast and robust video watermarking. The subband coefficients of video frames with the discrete wavelet transform are used as training data for ELM and a binary watermark in video frames is embedded in the output of SLFN. Experimental results verify that such framework is able to achieve high visual quality in the resultant videos.

*3.3. ELM in Medical Applications.* Nowadays, intelligent medical signal processing has gradually become an important auxiliary tool for clinical diagnosis. Accurate predictions and fast processing speed are two essential assessments in evaluating the intelligent medical processing systems. Benefiting from the easy implementation for real-time diagnosing and the relatively convincing performance, adopting ELM for medical signal processing has attracted increasing attention from the research community in the past several years. To the best of our capability, we have collected dozens of articles discussing ELM on various medical applications in this paper, including cardiac arrhythmia classification [69], gene cancer identification [70, 71], mammographic microcalcifications

detection [72], epileptic diagnosis [73–75], liver parenchyma and tumor detection [76–78], EEG vigilance [79], magnetic resonance images (MRI) data processing [80], gene selection [81], protein sequence applications [82–85], hypoglycemia prediction [86], and Parkinson classification [87].

Kim et al. [69] analyzed the cardiac arrhythmia classification performance with the ELM trained SLFN. Seven ECG-type beats consisting of one normal rhythm and six arrhythmias from the MIT-BIH arrhythmia database are used for classifier verification. In [69], three features, the R peak amplitude, instance RR interval, and R peak morphology data, are used as input signals and the PCA is adopted for feature dimension reduction. Zhang et al. [70] utilized ELM for microarray gene expression cancer diagnosis. Three benchmark microarray databases, that is, the GCM dataset, the lung dataset, and the lymphoma dataset, are involved in the experiments and the performance is compared to SVM to show the advantages of using ELM. Saraswathi et al. [71] developed a novel classification scheme for multiclass cancer classification. The integer-coded genetic algorithm is first implemented to select optimal sets of genes for recognition, and, then, the particle swarm optimization combining with ELM is introduced for classifier construction. Comparisons made on many state-of-the-art approaches show that the new method enjoys a high recognition rate in cancer classification. Malar et al. [72] applied ELM to detect and classify the microcalcifications in digitized mammograms. In the poor contrast of the mammogram image, the microcalcifications existing in the dense breast tissue are first exploited and represented using the wavelet transform features. Then, ELM is employed to learn these feature to build a SLFN detector. With an obvious reduction on the training time, it is shown that ELM wins a higher recognition rate than Bayes net classifier, naive Bayes classifier, and SVM. Yuan et al. [73] and Song et al. [74, 75] studied the performance of electroencephalogram (EEG) epileptic recognition with ELM. In [73], nonlinear dynamic features containing the approximate entropy, the Hurst exponent, and the scaling exponent are extracted to represent the interictal and ictal EEGs. These features are fed to a SLFN learned by ELM to form an intelligent classifier. In [74], the automatic epileptic detection is achieved through using an optimized sample entropy for feature extractions from EEG signals and the basic ELM as the classifier for recognition. An alternative realization of automatic epileptic recognition using multiresolution features, the basic ELM, and the genetic algorithm (GA) has been presented in [75]. First of all, the multiresolution feature extraction is obtained by decomposing the original EEG signal into several frequency bands through wavelet transform and exploiting complexity based features on all frequency bands. Then, representative subsets of features are selected by the efficient GA algorithm and learned by the basic ELM algorithm.

Huang et al. [76–78] studied the liver segmentation and liver tumor detection from 3D computed tomography (CT) images with the basic ELM, the kernel ELM, and the ensemble ELM, respectively. In [76], the liver segmentation is treated as a recognition problem where the 3D CT image is detected either containing the region of liver or a nonliver region.

Texture features including the mean, variance, and sum-and-difference histograms are extracted and used as the input signals to establish the classifier with the basic ELM algorithm. The liver segmentation and tumor classification with the kernel based ELM are considered in [77]. Real CT data collected from 7 patients are used to test the effectiveness of the adopted approach. An enhanced liver tumor detection approach with the ensemble based ELM is recently presented in [78]. Shi and Lu [79] investigated the vigilance estimation for human machine interaction systems using the EEG signal and ELM. Three ELM based vigilance estimators, that is, the basic ELM, the modified ELM with  $\ell_1$  norm, and  $\ell_2$  norm penalties, are presented for comparisons. Termenon et al. [80] identified the cocaine dependent patient through the structural magnetic resonance images (sMRI) using ELM. After selecting the most relevant watershed regions in the brain sMRI, representative features extracted from these regions are calculated for the classifier establishment. The sMRI intensity value of the selected regions and its corresponding mean and median values are used as the features in [80]. The basic ELM is used to recognize cocaine dependent patients and the classification performance is compared with SVM, OP-ELM, and the nearest neighborhood (NN) on real collected brain MRI. Han et al. [81] introduced a novel method for gene selection from microarray data based on ELM. The newly predictive gene selection strategy includes three steps: (1) the gene-to-class sensitivity (GCS) is calculated using a SLFN trained with the basic ELM algorithm; (2) the  $k$ -means clustering method is adopted to group genes according to their GCS values while representative genes are selected and redundant genes with low GCS values are filtered out; (3) for reminders genes, a binary particle swarm optimization (BPSO) is developed to conduct further selection. The fast learning speed and good generalization performance of ELM are well exploited for gene selection applications. Protein sequence classification, interaction, and analysis with ELM are other popular topics in medical applications [82–85]. Wang and Huang [82] applied the basic ELM for protein sequence classification. The experiments conducted on the Protein Information Resource (PIR) (<http://pir.georgetown.edu/>) database have shown that ELM with both sigmoid and RBF activation functions learns thousands times faster than SVM and BP. You et al. [83] and Wang et al. [84] studied the ELM based method for protein-protein interactions (PPIs). In general, identifying PPIs is time-consuming and expensive in experiments and applications. The emergent development of ELM well meets the requirements in PPIs applications. In [83], the autocovariance quantifying interactions between amino acids are first used to transform numerical protein sequences into uniform matrices for feature representations in PPIs database. In [84], the protein-protein interface predictions on both multichain sets and single-chain sets are investigated via the ELM algorithm. Three complex datasets including more than 19 thousand records are implemented to verify the effectiveness of ELM on both the model training time and the generalization performance. Savojardo et al. [85] developed a machine learning tool based on ELM for transmembrane beta barrel proteins detection. Mo et al. [86]

adopted the basic ELM and the regularized ELM to predict hypoglycemia in blood glucose for the purpose of diabetes management. The prediction accuracy on three prediction horizons with different time durations is compared between the basic ELM and the regularized ELM. Sachnev and Kim [87] studied the Parkinson disease (PD) classification using ELM. The benchmark ParkPD database consisting of more than 22 thousand genes' expressions collected from normal and PD patients is tested with the classifier. Before building the classifier, a binary-coded genetic algorithm is presented to select representative genes which can discriminate PD patients from normal patients as the input signal for the SLFN.

*3.4. Other Applications.* In addition to image, video, and medical applications, ELM has also been widely researched and implemented on other high dimensional and large data applications [39, 88–103]. These achievements covered time series prediction and forecasting [88–93], terrain reconstruction and navigation [94, 95], power loss analysis [96], company internationalization search [97], XML document classification and text categorization [98, 99], cloud computing [100], activity recognition for miniwearable devices [101], imbalance data processing [39, 102, 103], and so forth. Such fruitful results enlarged the application fields of ELM.

Intelligent time series prediction and forecasting play a vital role in industrial production, financial data analysis, and human life. In general, intelligent approaches with a good accuracy and fast processing speed are highly desired. Tian and Mao [88] implemented ELM for molten steel's temperature prediction in ladle furnace (LF) applications. The AdaBoost.RT algorithm is combined with the ensemble ELMs to improve the model approximation and temperature prediction accuracy. Experiments with real data obtained from a 300-ton LF in Baoshan Iron and Steel Co., China, are conducted to show the effectiveness of using ELM for prediction. Zhang et al. [89] and Chen et al. [90] exploited electricity applications with SLFN. In [89], the short-term electricity load forecasting of Australian national electricity market is realized through the ELM algorithm. Multiple ensembles of SLFNs trained with ELM are utilized and the median value is considered as the final output for electricity load prediction. The electricity price prediction is considered using ELM and the bootstrapping method in [90] where the usage of ELM aims to accelerate the forecasting speed and the bootstrapping method addresses uncertainty estimations to enhance the electricity price intervals forecast accuracy. Wind power forecasting is an important part in the wind power generation system. Utilizing ELM and the bootstrapping algorithm, Wan et al. [91] developed an intelligent wind power prediction system. Three different bootstrapping strategies, namely, the pairs bootstrap, the standard residuals bootstrap, and the wild bootstrap, are employed to estimate model uncertainty intervals. Financial data prediction and analysis with ELM also attract attention from researchers in the community. Li et al. [92] implemented ELM for stock price movement predictions. Analyses have been conducted on the stock tick prices using the tick prices in H-share market of year 2001. Yu et al. [93] considered the bankruptcy prediction problem

with ELM. The bankruptcy prediction problem is transferred to a binary classification application while the leave-one-out-incremental ELM (LOO-IELM) is introduced as the learning algorithm for the classifier. Ensemble based model is also used in [93] to enhance the prediction accuracy.

Terrain reconstruction and navigation with ELM based approaches for aiding unmanned aerial vehicles (UAVs) have been studied in [94, 95]. Yeu et al. [94] realized the multiresolution terrain reconstruction through learning the stored digital elevation information with ELM. Experiments show that, to achieve the same MSE, ELM requires a lower memory for terrain reconstruction in computer than the linear interpolation approach and SVM. Kan et al. [95] applied ELM as a navigation assistant for UAVs to deal with the problem of missing the global positioning system (GPS) information. Besides series prediction and terrain reconstruction, Nizar et al. [96] studied the electricity nontechnical losses (NTL) problem with ELM. Accurate identifying of NTLs could help to provide the preventative and corrective ways to reduce the losses with suitable inventions and implementations. The knowledge from individual electricity customer behavior is shown to be important in strategies designs and decisions making for electricity service providers. With this objective, the basic ELM and the OS-ELM methods are utilized in [96] to analyze customers' behavior for detecting the NTLs. Landa-Torres et al. [97] explored a new application area of ELM in evaluating the success of internationalization of a company. To build such an evaluation model, a grouping-based harmony search method combining with the ELM ensembles is developed to improve the accuracy. Experiments on real data from several Spanish exporting manufacturers are conducted and analyzed to verify the proposed hybrid learning scheme. Zhao et al. [98] and Zheng et al. [99] employed ELM for multiclass XML document classification and text categorization, respectively. In [98], the distribution based structured vector model is introduced for XML documents representations and the probability based voting ELM is conducted for classification. In [99], the text categorization is performed using a regularization ELM. The latent semantic analysis is adopted for dimensionality reduction for text representations. Lin et al. [100] first outsourced ELM in cloud computing for addressing highly complex structure of data. The randomly and independently assigned hidden neurons make ELM suitable for being outsourced to cloud computing. A partitioned ELM is realized as a secure and practical outsourcing mechanism for large-scale data processing in [100]. One of the most recent literatures on ELM applications is the activities detection using miniwearable devices [101]. To meet the low-computational-complexity, lightweight, and high-accuracy requirements on recognition model, a novel ELM based algorithm named the bias constrained-optimization-based ELM (b-COELM) is presented for the miniwearable devices activities identification. The performance of the proposed technique is tested via the real experiments of human motion tracking where the real data collected by a portable device of human motion tracker named XSens MTx are analyzed.

Imbalance data processing with ELM algorithms became popular in the past two years [39, 102, 103]. The imbalance

data classification problems for both binary and multiclass databases are dealt with by a novel weighted ELM in [102]. The weight information for each class is generated from the number of class samples belonging to the same category. The revised ELM which incorporates the class weights into ELM is then adopted to improve the recognition performance. Mirza et al. [103] extended the basic weighted ELM to a weighted online sequential ELM for class imbalance data processing (WOS-ELM). An improved tuning total error rate (TER) approach is first introduced to assign weights for each class. Mirza et al. claim that a high  $G_{\text{mean}}$  value, which measures the classification performance of class imbalance processing [103], can be guaranteed through using the new technique. The weighted least square solutions are then introduced for network output weight calculation in the initial phase, the weight updating in the sequential learning phase. Fresh extension to online sequential learning for class imbalance and concept drifts with the ensemble approach is done in [39] recently, where a novel ensemble of subset online sequential ELM (EOS-ELM) is developed. Combining the subset learning with OS-ELM and the ensembles, EOS-ELM shows efficiency and superiority in handling class imbalance problems in both stationary and nonstationary environments.

#### 4. Conclusions

This paper presented an up-to-date review on the recent developments of ELM algorithms and its applications for high dimensional and large data processing. The survey covered applications in a wide field including image and video signal processing and medical data processing. As presented in the above section, the fast data learning speed and easy implementation characteristics of ELM boosted its applications in various fields. Applications on intelligent high dimensional and large data processing benefited a lot from ELM and its variants due to the significantly reduced computational complexity brought by the randomness in network parameters and the tuning-free learning strategy. In many applications, ELM has well addressed the out-of-memory, time-costing, premature performance encountered by conventional gradient based learning approaches and SVM. Although a great number of achievements on high dimensional data applications have been presented in the past several years, the following three issues are worth considering.

- (1) Tuning-free is one of the most important contributions to ELM. However, various approaches and applications have applied the iterative updating processing into the original ELM to produce good generalization performance, such as the usage of genetic algorithms, the boosting approaches, the pruning methods, and the evolutionary ensembles. Although the model regression accuracy and data classification performance are more or less improved by introducing such strategies, there is no doubt that the computational complexity is also increased. Thus, how to balance the

performance and the processing time is an open issue, especially for applications in high dimensional data.

- (2) How to choose the optimal number of hidden neurons for a certain application is not well addressed yet. In most of the existing works, few efforts have been paid to discussions on the selection of hidden neurons and almost all of them are chosen manually in a tentative way. Although some researchers claimed that the performance of ELM and its variants tend to be stable and acceptable when a large number of hidden neurons are used, redundancy and high computation burden would also occur.
- (3) Designing real-time processing systems and devices for applications with ELM is highly desired. Although plentiful achievements have been reported in the past 10 years or so, most of them are still conducted in the laboratory via computer simulations. Real-world devices for different applications are always facing various challenges, which are more obvious for large data applications.

#### Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

#### Acknowledgments

This work was supported by the Natural Science Foundation of Zhejiang province, China, under Grant no. LY15F030017 and in part by the National Natural Science Major Foundation of Research Instrumentation of China under Grant no. 61427808.

#### References

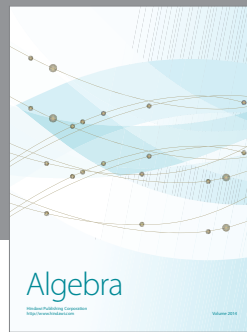
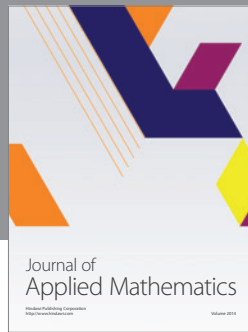
- [1] D. W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *SIAM Journal on Applied Mathematics*, vol. 11, pp. 431–441, 1963.
- [2] P. J. Werbos, *The Roots of Backpropagation: From Ordered Derivatives to Neural Networks and Political Forecasting*. John Wiley & Sons, New York, NY, USA, 1994.
- [3] S. Haykin, *Neural Networks, A Comprehensive Foundation*, Pearson Education Press, New York, NY, USA, 2001.
- [4] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [5] P. Lingras and C. Butz, "Rough set based 1-v-1 and 1-v-r approaches to support vector machine multi-classification," *Information Sciences*, vol. 177, no. 18, pp. 3782–3798, 2007.
- [6] C.-W. Hsu and C.-J. Lin, "A comparison of methods for multiclass support vector machines," *IEEE Transactions on Neural Networks*, vol. 13, no. 2, pp. 415–425, 2002.
- [7] A. Benchabane, A. Bennia, F. Charif, and A. Taleb-Ahmed, "Multi-dimensional Capon spectral estimation using discrete Zhang neural networks," *Multidimensional Systems and Signal Processing*, vol. 24, no. 3, pp. 583–598, 2013.

- [8] S. Liu, S. Hu, Y. Xiao, and Y. An, "Bayesian Shearlet shrinkage for SAR image de-noising via sparse representation," *Multidimensional Systems and Signal Processing*, vol. 25, no. 4, pp. 683–701, 2014.
- [9] V. Mayer-Schonberger and K. Cukier, *Big Data: A Revolution That Will Transform How We Live, Work and Think*, Houghton Mifflin Harcourt Publishing Company, 2013.
- [10] G.-B. Huang and C.-K. Siew, "Extreme learning machine: RBF network case," in *Proceedings of the 8th International Conference on Control, Automation, Robotics and Vision (ICARCV '04)*, vol. 2, pp. 1029–1036, Kunming, China, December 2004.
- [11] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: theory and applications," *Neurocomputing*, vol. 70, no. 1–3, pp. 489–501, 2006.
- [12] 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*, vol. 17, no. 4, pp. 879–892, 2006.
- [13] G. Feng, G.-B. Huang, Q. Lin, and R. Gay, "Error minimized extreme learning machine with growth of hidden nodes and incremental learning," *IEEE Transactions on Neural Networks*, vol. 20, no. 8, pp. 1352–1357, 2009.
- [14] R. Zhang, Y. Lan, G.-B. Huang, Z.-B. Xu, and Y. C. Soh, "Dynamic extreme learning machine and its approximation capability," *IEEE Transactions on Cybernetics*, vol. 43, no. 6, pp. 2054–2065, 2013.
- [15] M.-B. Li, G.-B. Huang, P. Saratchandran, and N. Sundararajan, "Fully complex extreme learning machine," *Neurocomputing*, vol. 68, no. 1–4, pp. 306–314, 2005.
- [16] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 42, no. 2, pp. 513–529, 2012.
- [17] N. Liu and H. Wang, "Ensemble based extreme learning machine," *IEEE Signal Processing Letters*, vol. 17, no. 8, pp. 754–757, 2010.
- [18] J. Cao, Z. Lin, G.-B. Huang, and N. Liu, "Voting based extreme learning machine," *Information Sciences*, vol. 185, pp. 66–77, 2012.
- [19] G.-B. Huang, Q.-Y. Zhu, K.-Z. Mao, C.-K. Siew, P. Saratchandran, and N. Sundararajan, "Can threshold networks be trained directly?" *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 53, no. 3, pp. 187–191, 2006.
- [20] X. Liu, S. B. Lin, J. Fang, and Z. B. Xu, "Is extreme learning machine feasible? a theoretical assessment (part I)," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 1, pp. 7–20, 2015.
- [21] S. B. Lin, X. Liu, J. Fang, and Z. B. Xu, "Is extreme learning machine feasible? a theoretical assessment (part II)," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 1, pp. 21–34, 2015.
- [22] Y. Miche, A. Sorjamaa, P. Bas, O. Simula, C. Jutten, and A. Lendasse, "OP-ELM: optimally pruned extreme learning machine," *IEEE Transactions on Neural Networks*, vol. 21, no. 1, pp. 158–162, 2010.
- [23] F. Pouzols and A. Lendasse, "Evolving fuzzy optimally pruned extreme learning machine: a comparative analysis," in *Proceedings of the IEEE International Conference on Fuzzy Systems*, Barcelona, Spain, July 2010.
- [24] Q.-Y. Zhu, A. K. Qin, P. N. Suganthan, and G.-B. Huang, "Evolutionary extreme learning machine," *Pattern Recognition*, vol. 38, no. 10, pp. 1759–1763, 2005.
- [25] N.-Y. Liang, G.-B. Huang, P. Saratchandran, and N. Sundararajan, "A fast and accurate online sequential learning algorithm for feedforward networks," *IEEE Transactions on Neural Networks*, vol. 17, no. 6, pp. 1411–1423, 2006.
- [26] H.-J. Rong, G.-B. Huang, P. Saratchandran, and N. Sundararajan, "On-line sequential fuzzy extreme learning machine for function approximation and classification problems," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 39, no. 4, pp. 1067–1072, 2009.
- [27] J. Cao, Z. Lin, and G.-B. Huang, "Self-adaptive evolutionary extreme learning machine," *Neural Processing Letters*, vol. 36, no. 3, pp. 285–305, 2012.
- [28] J. Cao, Z. Lin, and G.-B. Huang, "Voting base online sequential extreme learning machine for multi-class classification," in *Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS '13)*, pp. 2327–2330, Beijing, China, May 2013.
- [29] H.-J. Lu, C.-L. An, E.-H. Zheng, and Y. Lu, "Dissimilarity based ensemble of extreme learning machine for gene expression data classification," *Neurocomputing*, vol. 128, pp. 22–30, 2014.
- [30] N. Liu, J. Cao, Z. Lin, P. P. Pek, Z. X. Koh, and M. E. Ong, "Evolutionary voting-based extreme learning machine," *Mathematical Problems in Engineering*, vol. 2014, Article ID 808292, 7 pages, 2014.
- [31] E. Cambria and G.-B. Huang, "Extreme learning machines [trends & controversies]," *IEEE Intelligent Systems*, vol. 28, no. 6, pp. 30–31, 2013.
- [32] Y. Yang, Y. Wang, and X. Yuan, "Bidirectional extreme learning machine for regression problem and its learning effectiveness," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 23, no. 9, pp. 1498–1505, 2012.
- [33] E. Soria-Olivas, J. Gómez-Sanchis, I. H. Jarman et al., "BELM: Bayesian extreme learning machine," *IEEE Transactions on Neural Networks*, vol. 22, no. 3, pp. 505–509, 2011.
- [34] P. Gastaldo, R. Zunino, E. Cambria, and S. Decherchi, "Combining ELMs with random projections," *IEEE Intelligent Systems*, vol. 28, no. 6, pp. 46–48, 2013.
- [35] J. Luo, C.-M. Vong, and P.-K. Wong, "Sparse bayesian extreme learning machine for multi-classification," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 4, pp. 836–843, 2014.
- [36] L. Kasun, H. Zhou, G.-B. Huang, and C. M. Vong, "Representational learning with ELMs for big data," *IEEE Intelligent Systems*, vol. 28, no. 6, pp. 31–34, 2013.
- [37] J. Chen, G. Zheng, and H. Chen, "ELM-based multiple classifier systems," in *Proceedings of the 10th IEEE International Conference on Control and Automation*, Hangzhou, China, June 2013.
- [38] R. Wang, Y.-L. He, C.-Y. Chow, F.-F. Ou, and J. Zhang, "Learning ELM-tree from big data based on uncertainty reduction," *Fuzzy Sets and Systems*, vol. 258, pp. 79–100, 2015.
- [39] B. Mirza, Z. Lin, and N. Liu, "Ensemble of subset online sequential extreme learning machine for class imbalance and concept drift," *Neurocomputing*, vol. 149, pp. 316–329, 2015.
- [40] I. Marques and M. Graña, "Face recognition with lattice independent component analysis and extreme learning machines," *Soft Computing*, vol. 16, no. 9, pp. 1525–1537, 2012.
- [41] K. Choi, K.-A. Toh, and H. Byun, "Incremental face recognition for large-scale social network services," *Pattern Recognition*, vol. 45, no. 8, pp. 2868–2883, 2012.
- [42] S.-J. Wang, H.-L. Chen, W.-J. Yan, Y.-H. Chen, and X. Fu, "Face recognition and micro-expression recognition based on discriminant tensor subspace analysis plus extreme learning

- machine,” *Neural Processing Letters*, vol. 39, no. 1, pp. 25–43, 2014.
- [43] A. Uçar, “Color face recognition based on steerable pyramid transform and extreme learning machines,” *The Scientific World Journal*, vol. 2014, Article ID 628494, 15 pages, 2014.
- [44] A. Uçar, Y. Demir, and C. Güzeliş, “A new facial expression recognition based on curvelet transform and online sequential extreme learning machine initialized with spherical clustering,” *Neural Computing and Applications*, 2014.
- [45] J. Lu, Y. Zhao, Y. Xue, and J. Hu, “Palmprint recognition via locality preserving projections and extreme learning machine neural network,” in *Proceedings of the 9th International Conference on Signal Processing (ICSP '08)*, pp. 2096–2099, Beijing, China, October 2008.
- [46] Z. Man, K. Lee, D. Wang, Z. Cao, and S. Khoo, “An optimal weight learning machine for handwritten digit image recognition,” *Signal Processing*, vol. 93, no. 6, pp. 1624–1638, 2013.
- [47] M. Luo and K. Zhang, “A hybrid approach combining extreme learning machine and sparse representation for image classification,” *Engineering Applications of Artificial Intelligence*, vol. 27, pp. 228–235, 2014.
- [48] W. Zhu, J. Miao, J. Hu, and L. Qing, “Vehicle detection in driving simulation using extreme learning machine,” *Neurocomputing*, vol. 128, pp. 160–165, 2014.
- [49] H.-J. Rong, Y.-X. Jia, and G.-S. Zhao, “Aircraft recognition using modular extreme learning machine,” *Neurocomputing*, vol. 128, pp. 166–174, 2014.
- [50] J. Cao, T. Chen, and J. Fan, “Fast online learning algorithm for landmark recognition based on BoW framework,” in *Proceedings of the 9th IEEE Conference on Industrial Electronics and Applications (ICIEA '14)*, pp. 1163–1168, Hangzhou, China, June 2014.
- [51] J. Cao, T. Chen, and J. Fan, “Landmark recognition with compact BoW histogram and ensemble ELM,” *Multimedia Tools and Applications*, 2015.
- [52] S. Suresh, R. Venkatesh Babu, and H. J. Kim, “No-reference image quality assessment using modified extreme learning machine classifier,” *Applied Soft Computing Journal*, vol. 9, no. 2, pp. 541–552, 2009.
- [53] L. An and B. Bhanu, “Image super-resolution by extreme learning machine,” in *Proceedings of the 19th IEEE International Conference on Image Processing (ICIP '12)*, pp. 2209–2212, Orlando, Fla, USA, October 2012.
- [54] A. Samat, P. Du, S. Liu, J. Li, and L. Cheng, “E<sup>2</sup>LMs: ensemble extreme learning machines for hyperspectral image classification,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 4, pp. 1060–1069, 2014.
- [55] Y. Bazi, N. Alajlan, F. Melgani, H. AlHichri, S. Malek, and R. R. Yager, “Differential evolution extreme learning machine for the classification of hyperspectral images,” *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 6, pp. 1066–1070, 2014.
- [56] R. Minhas, A. Baradarani, S. Seifzadeh, and Q. M. Jonathan Wu, “Human action recognition using extreme learning machine based on visual vocabularies,” *Neurocomputing*, vol. 73, no. 10–12, pp. 1906–1917, 2010.
- [57] R. Minhas, A. A. Mohammed, and Q. M. J. Wu, “Incremental learning in human action recognition based on Snippets,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 11, pp. 1529–1541, 2012.
- [58] A. Iosifidis, A. Tefas, and I. Pitas, “Dynamic action recognition based on dynemes and extreme learning machine,” *Pattern Recognition Letters*, vol. 34, no. 15, pp. 1890–1898, 2013.
- [59] A. Iosifidis, A. Tefas, and I. Pitas, “Minimum class variance extreme learning machine for human action recognition,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 23, no. 11, pp. 1968–1979, 2013.
- [60] A. Iosifidis, A. Tefas, and I. Pitas, “Regularized extreme learning machine for multi-view semi-supervised action recognition,” *Neurocomputing*, vol. 145, pp. 250–262, 2014.
- [61] A. Iosifidis, A. Tefas, and I. Pitas, “Minimum variance extreme learning machine for human action recognition,” in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Florence, Italy, May 2014.
- [62] A. Budiman and M. I. Fanany, “Pose-based 3D human motion analysis using extreme learning machine,” in *Proceedings of the IEEE 2nd Global Conference on Consumer Electronics (GCCE '13)*, pp. 3–7, Tokyo, Japan, October 2013.
- [63] W.-Y. Deng, Q.-H. Zheng, and Z.-M. Wang, “Cross-person activity recognition using reduced kernel extreme learning machine,” *Neural Networks*, vol. 53, pp. 1–7, 2014.
- [64] B. Oh, J. Jeon, K.-A. Toh, A. Teoh, and J. Kim, “A system for signature verification based on horizontal and vertical components in hand gestures,” *IEEE Intelligent Systems*, vol. 28, no. 6, pp. 52–55, 2013.
- [65] H. Yu, Y. Chen, J. Liu, and G.-B. Huang, “An adaptive and iterative online sequential ELM-based multi-degree-of-freedom gesture recognition system,” *IEEE Intelligent Systems*, vol. 28, no. 6, pp. 55–59, 2013.
- [66] J. Shi, Y. Cai, J. Zhu, J. Zhong, and F. Wang, “SEMG-based hand motion recognition using cumulative residual entropy and extreme learning machine,” *Medical & Biological Engineering and Computing*, vol. 51, no. 4, pp. 417–427, 2013.
- [67] B. Lu, G. Wang, Y. Yuan, and D. Han, “Semantic concept detection for video based on extreme learning machine,” *Neurocomputing*, vol. 102, pp. 176–183, 2013.
- [68] C. Agarwal, A. Mishra, A. Sharma, and G. Chetty, “A novel scene based robust video watermarking scheme in DWT domain using extreme learning machine,” in *Proceedings of the Extreme Learning Machines: Algorithms and Applications, Adaptation, Learning, and Optimization*, Beijing, China, October 2013.
- [69] J. Kim, H. Shin, Y. Lee, and J. Lee, “Algorithm for classifying arrhythmia using extreme learning machine and principal component analysis,” in *Proceedings of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS '07)*, pp. 3257–3260, Lyon, France, August 2007.
- [70] R. Zhang, G.-B. Huang, N. Sundararajan, and P. Saratchandran, “Multicategory classification using an extreme learning machine for microarray gene expression cancer diagnosis,” *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 4, no. 3, pp. 485–494, 2007.
- [71] S. Saraswathi, S. Sundaram, N. Sundararajan, M. Zimmermann, and M. Nilsen-Hamilton, “ICGA-PSO-ELM approach for accurate multiclass cancer classification resulting in reduced gene sets in which genes encoding secreted proteins are highly represented,” *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 8, no. 2, pp. 452–463, 2011.
- [72] E. Malar, A. Kandaswamy, D. Chakravarthy, and A. Giri Dharan, “A novel approach for detection and classification of mammographic microcalcifications using wavelet analysis and extreme learning machine,” *Computers in Biology and Medicine*, vol. 42, no. 9, pp. 898–905, 2012.

- [73] Q. Yuan, W. Zhou, S. Li, and D. Cai, "Epileptic EEG classification based on extreme learning machine and nonlinear features," *Epilepsy Research*, vol. 96, no. 1-2, pp. 29-38, 2011.
- [74] Y. Song, J. Crowcroft, and J. Zhang, "Automatic epileptic seizure detection in EEGs based on optimized sample entropy and extreme learning machine," *Journal of Neuroscience Methods*, vol. 210, no. 2, pp. 132-146, 2012.
- [75] Y. Song and J. Zhang, "Automatic recognition of epileptic EEG patterns via extreme learning machine and multiresolution feature extraction," *Expert Systems with Applications*, vol. 40, no. 14, pp. 5477-5489, 2013.
- [76] W. Huang, Z. M. Tan, Z. Lin et al., "A semi-automatic approach to the segmentation of liver parenchyma from 3D CT images with extreme learning machine," in *Proceedings of the 34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS '12)*, pp. 3752-3755, San Diego, Calif, USA, September 2012.
- [77] W. Huang, N. Li, Z. Lin et al., "Liver tumor detection and segmentation using kernel-based extreme learning machine," in *Proceedings of the 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '13)*, pp. 3662-3665, Osaka, Japan, July 2013.
- [78] W. Huang, Y. Yang, Z. Lin et al., "Random feature subspace ensemble based Extreme Learning Machine for liver tumor detection and segmentation," in *Proceedings of the 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '14)*, pp. 4675-4678, Chicago, Ill, USA, August 2014.
- [79] L.-C. Shi and B.-L. Lu, "EEG-based vigilance estimation using extreme learning machines," *Neurocomputing*, vol. 102, pp. 135-143, 2013.
- [80] M. Termenon, M. Graña, A. Barrós-Loscertales, and C. Ávila, "Extreme learning machines for feature selection and classification of cocaine dependent patients on structural MRI data," *Neural Processing Letters*, vol. 38, no. 3, pp. 375-387, 2013.
- [81] F. Han, W. Sun, and Q. H. Ling, "A novel strategy for gene selection of microarray data based on gene-to-class sensitivity information," *PLoS ONE*, vol. 9, no. 5, Article ID e97530, 17 pages, 2014.
- [82] D. Wang and G.-B. Huang, "Protein sequence classification using extreme learning machine," in *Proceedings of the International Joint Conference on Neural Networks (IJCNN '05)*, vol. 3, pp. 1406-1411, Montreal, Canada, July-August 2005.
- [83] Z. You, L. Li, Z. Ji, M. Li, and S. Guo, "Prediction of protein-protein interactions from amino acid sequences using extreme learning machine combined with auto covariance descriptor," in *Proceedings of the IEEE Workshop on Memetic Computing (MC '13)*, pp. 80-85, Singapore, April 2013.
- [84] D. D. Wang, R. Wang, and H. Yan, "Fast prediction of protein-protein interaction sites based on extreme learning machines," *Neurocomputing*, vol. 128, pp. 258-266, 2014.
- [85] C. Savojardo, P. Fariselli, and R. Casadio, "BETAWARE: a machine-learning tool to detect and predict transmembrane beta-barrel proteins in prokaryotes," *Bioinformatics*, vol. 29, no. 4, pp. 504-505, 2013.
- [86] X. Mo, Y. Wang, and X. Wu, "Hypoglycemia prediction using extreme learning machine (ELM) and regularized ELM," in *Proceedings of the 25th Chinese Control and Decision Conference (CCDC '13)*, pp. 4405-4409, Guigang, China, May 2013.
- [87] V. Sachnev and H. J. Kim, "Parkinson disease classification based on binary coded genetic algorithm and extreme learning machine," in *Proceedings of the IEEE 9th International Conference on Intelligent Sensors, Sensor Networks and Information Processing*, pp. 1-6, Singapore, April 2014.
- [88] H.-X. Tian and Z.-Z. Mao, "An ensemble ELM based on modified AdaBoost.RT algorithm for predicting the temperature of molten steel in ladle furnace," *IEEE Transactions on Automation Science and Engineering*, vol. 7, no. 1, pp. 73-80, 2010.
- [89] R. Zhang, Z. Y. Dong, Y. Xu, K. Meng, and K. P. Wong, "Short-term load forecasting of Australian national electricity market by an ensemble model of extreme learning machine," *IET Generation, Transmission and Distribution*, vol. 7, no. 4, pp. 391-397, 2013.
- [90] X. Chen, Z. Y. Dong, K. Meng, Y. Xu, K. P. Wong, and H. W. Ngan, "Electricity price forecasting with extreme learning machine and bootstrapping," *IEEE Transactions on Power Systems*, vol. 27, no. 4, pp. 2055-2062, 2012.
- [91] C. Wan, Z. Xu, P. Pinson, Z. Y. Dong, and K. P. Wong, "Probabilistic forecasting of wind power generation using extreme learning machine," *IEEE Transactions on Power Systems*, vol. 29, no. 3, pp. 1033-1044, 2014.
- [92] X. Li, H. Xie, R. Wang et al., "Empirical analysis: stock market prediction via extreme learning machine," *Neural Computing and Applications*, 2014.
- [93] Q. Yu, Y. Miche, E. Séverin, and A. Lendasse, "Bankruptcy prediction using Extreme Learning Machine and financial expertise," *Neurocomputing*, vol. 128, pp. 296-302, 2014.
- [94] C.-W. T. Yeu, M.-H. Lim, G.-B. Huang, A. Agarwal, and Y.-S. Ong, "A new machine learning paradigm for terrain reconstruction," *IEEE Geoscience and Remote Sensing Letters*, vol. 3, no. 3, pp. 382-386, 2006.
- [95] E. M. Kan, M. H. Lim, Y. S. Ong, A. H. Tan, and S. P. Yeo, "Extreme learning machine terrain-based navigation for unmanned aerial vehicles," *Neural Computing and Applications*, vol. 22, no. 3-4, pp. 469-477, 2013.
- [96] A. H. Nizar, Z. Y. Dong, and Y. Wang, "Power utility nontechnical loss analysis with extreme learning machine method," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 946-955, 2008.
- [97] I. Landa-Torres, E. G. Ortiz-García, S. Salcedo-Sanz et al., "Evaluating the internationalization success of companies through a hybrid grouping harmony search-extreme learning machine approach," *IEEE Journal on Selected Topics in Signal Processing*, vol. 6, no. 4, pp. 388-398, 2012.
- [98] X. Zhao, X. Bi, and B. Qiao, "Probability based voting extreme learning machine for multiclass XML documents classification," *World Wide Web*, vol. 17, no. 5, pp. 1217-1231, 2014.
- [99] W. Zheng, Y. Qian, and H. Lu, "Text categorization based on regularization extreme learning machine," *Neural Computing and Applications*, vol. 22, no. 3-4, pp. 447-456, 2013.
- [100] J. Lin, J. Yin, Z. Cai, Q. Liu, K. Li, and V. C. M. Leung, "A secure and practical mechanism for outsourcing ELMs in cloud computing," *IEEE Intelligent Systems*, vol. 28, no. 6, pp. 35-38, 2013.
- [101] L. Hu, Y. Chen, S. Wang, and Z. Chen, "b-COELM: a fast, lightweight and accurate activity recognition model for mini-wearable devices," *Pervasive and Mobile Computing*, vol. 15, pp. 200-214, 2014.
- [102] W. Zong, G.-B. Huang, and Y. Chen, "Weighted extreme learning machine for imbalance learning," *Neurocomputing*, vol. 101, pp. 229-242, 2013.
- [103] B. Mirza, Z. Lin, and K.-A. Toh, "Weighted online sequential extreme learning machine for class imbalance learning," *Neural Processing Letters*, vol. 38, no. 3, pp. 465-486, 2013.

- [104] G. Huang, S. Song, and K. You, "Trends in extreme learning machines: a review," *Neural Networks*, vol. 61, pp. 32–48, 2015.
- [105] G.-B. Huang, D. H. Wang, and Y. Lan, "Extreme learning machines: a survey," *International Journal of Machine Learning and Cybernetics*, vol. 2, no. 2, pp. 107–122, 2011.
- [106] R. Fletcher, *Practical Methods of Optimization. Volume 2: Constrained Optimization*, John Wiley & Sons, Chichester, UK, 1981.
- [107] Y. Lu, N. Sundararajan, and P. Saratchandran, "Performance evaluation of a sequential minimal Radial Basis Function (RBF) neural network learning algorithm," *IEEE Transactions on Neural Networks*, vol. 9, no. 2, pp. 308–318, 1998.
- [108] G.-B. Huang, P. Saratchandran, and N. Sundararajan, "A generalized growing and pruning RBF (GGAP-RBF) neural network for function approximation," *IEEE Transactions on Neural Networks*, vol. 16, no. 1, pp. 57–67, 2005.
- [109] M. Ozuysal, M. Calonder, V. Lepetit, and P. Fua, "Fast keypoint recognition using random ferns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 3, pp. 448–461, 2010.
- [110] N. Gkalelis, A. Tefas, and I. Pitas, "Combining fuzzy vector quantization with linear discriminant analysis for continuous human movement recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 11, pp. 1511–1521, 2008.



# Hindawi

Submit your manuscripts at  
<http://www.hindawi.com>

