Received July 12, 2020, accepted July 22, 2020, date of publication July 27, 2020, date of current version August 13, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3012093

Computer-Aided Diagnosis Based on Extreme Learning Machine: A Review

ZHIQIONG WANG[1], YIQI LUO[1], JUNCHANG XIN[2,3], (Member, IEEE), HAO ZHANG[4], LUXUAN QU[1], ZHONGYANG WANG[1], YUDONG YAO[5], (Fellow, IEEE), WANCHENG ZHU[6], AND XINGWEI WANG[7]

[1]College of Medicine and Biological Information Engineering, Northeastern University, Shenyang 110169, China
[2]School of Computer Science and Engineering, Northeastern University, Shenyang 110169, China
[3]Key Laboratory of Big Data Management and Analytics, Northeastern University, Shenyang 110169, China
[4]Department of Breast Surgery, Liaoning Cancer Hospital and Institute, Cancer Hospital of China Medical University, Shenyang 110042, China
[5]Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ 07450, USA
[6]Center for Rock Instability and Seismicity Research, School of Resources and Civil Engineering, Northeastern University, Shenyang 110819, China
[7]College of Software, Northeastern University, Shenyang 110169, China

Corresponding authors: Junchang Xin (xinjunchang@mail.neu.edu.cn) and Hao Zhang (haozhang840514@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61472069 and Grant 61402089, in part by the China Postdoctoral Science Foundation under Grant 2019T120216 and Grant 2018M641705, in part by the Fundamental Research Funds for the Central Universities under Grant N2019007, Grant N180101028, Grant N180408019, and Grant N2024005-2, in part by the China Electronics Technology Group Corporation (CETC) Joint Fund, and in part by the Fund of Acoustics Science and Technology Laboratory.

ABSTRACT Computer-Aided Diagnosis (CAD) can improve the accuracy of diagnosis effectively, reduce the rate of misdiagnosis, and provide the support for the valid decision. In clinical applications, high requirements are often imposed on the execution speed and accuracy of CAD systems. The classifier is regarded as the core of the CAD system, that is, the performance of the classifier will have a decisive influence on the operating affection of the CAD system. Extreme Learning Machine (ELM) is a fast learning algorithm using Single Hidden Layer Feedforward Neural Network (SLFN) structure. With its advantages in training speed, generalization performance and accuracy, ELM has draw attention in many research fields, including the development of CAD system. The applications of ELM in CAD are reviewed in this research. First, the mathematical model of ELM and framework of CAD system are briefly introduced. Then, the application of ELM in CAD is reviewed in detail, including the feature modeling method combined with ELM in CAD and the specific application of ELM. Finally, we summarized the current research status of CAD systems based on ELM, and the future work is prospected.

INDEX TERMS Computer-aided diagnosis, extreme learning machine, machine learning, review.

I. INTRODUCTION

Since the last century, with the continuous efforts of researchers in various fields, our knowledge of human anatomy and physiology has grown significantly. Meanwhile, human examination tools based on imaging technologies such as X-ray, ultrasound, and Magnetic Resonance Imaging (MRI) have also made great progress [1], [2]. Nevertheless, in terms of clinical diagnose, the complexity of medical diagnosis process, the diversity of diseases, and the increasing number of medical data significantly greatly increase the workload and difficulty of doctors, resulting in the possibility of misdiagnosis due to fatigue or empiricism. Short after the arrival of the computer age, biomedical researchers began to explore the possibility of using computers to research and solve biological and medical problems [3]. In 1963, Lodwick *et al.* [4]proposed the method of digitizing X-ray films. This provides a practical foundation for the use of computers to extract multidimensional information from medical data to assist doctors in diagnosis. CAD refers to the combination of imaging, medical image processing technology and other possible physiological and biochemical methods with computer analysis and calculation, which is used to assist in the detection of lesions or the classification of benign and malignant diseases [5].Through the objective judgment provided by CAD, it plays an active role in improving the efficiency of doctors, the accuracy of diagnosis, reducing the rate of misdiagnose and so on. In order to further meet

The associate editor coordinating the review of this manuscript and approving it for publication was Inês Domingues.

the requirements of clinical diagnosis, to achieve an efficient and accurate CAD system is still a ultimate goal of many researchers.

At present, many CAD systems are oriented to medical image. The main components of these CAD systems include preprocessing, segmentation, feature modeling, and classification (detection/diagnosis). Among these components, classification is often regarded as the core of CAD system. It refers to a data mining process that assigns labels or classes to different groups. Selecting an appropriate machine learning method to establish a classifier responsible for detecting or distinguishing different types of lesions is a key component of CAD system development [6]. On the other hand, ELM is a new type of fast learning algorithm which has attracted much attention in recent years. Compared with other classifiers, it has significant advantages in training speed and accuracy. As a promising algorithm, it is widely used in various related researches [7], including the design and implementation of CAD system.

ELM is a training algorithm for SLFNs proposed by Huang *et al.* [8] of Nanyang Technological University, Singapore. In the past 10 years of its research and development, ELM has attracted the attention of a large number of researchers, and related improved algorithms have also been proposed to deal with some specific problems: Online Sequential Extreme Learning Machine (OS-ELM) [9] can learn from a growing data set. Furthermore, the Convex Incremental Extreme Learning Machine (CI-ELM), which can solve the problem of new node training in the incremental model, was proposed in [10]. Wang *et al.* [11] implemented Effective Extreme Learning Machine (EELM), which adjusts the weights and biases of the input layer before calculating the output layer weights so that the output conditions of the hidden layer satisfy the column full rank condition. The improved EELM algorithm can reduce training time, improve network robustness and classification accuracy. Cao *et al.* [12] proposed Voting based Extreme Learning Machine (V-ELM) to avoid the instability of classification results caused by randomly generated hidden layer weights and biases. In order to solve the classification problem of non-equilibrium data, Cao *et al.* [12] proposed Weighted Extreme Learning Machine (WELM). This algorithm can be directly used for multi-classification problems, and it can also be extended to cost-sensitive learning. Liu *et al.* [13] proposed the Multiple Kernel Extreme Learning Machine (MK-ELM) as a general-purpose learning framework that can be used to solve the selection and optimization of ELM kernel functions.

Existing CAD systems based on ELM or improved algorithms have achieved good performance. The purpose of this study is to systematically organize and review these existing research results, to provide reference for researchers in the field of ELM algorithm and CAD system design. The structure of the rest of this paper is as follows: ELM algorithm and CAD framework respectively are introduced in Section II. Section III sorts out the feature modeling methods used in ELM based CAD systems. Section IV summarizes the specific application of ELM and its improved algorithm in CAD. Section V demonstrates the effectiveness of ELM in CAD. Section VI provides an outlook for the future development of ELM. In Section VII, the use of open data sets in references and systematic evaluation methods are sorted out. Finally, Section VIII summarizes this research.

II. BACKGROUND

In this section, we will briefly introduce the classic ELM algorithm and the basic structure of the CAD system.

A. CLASSIC ELM

The network structure of ELM [8] is shown in Figure 1. In simple terms, the network structure of ELM model is the same as that of SLFN, except that in the training stage, it is no longer the gradient based algorithm (backward propagation) in the traditional neural network, but the random weight and deviation of input layer are used, and the output layer weight is calculated by the generalized inverse matrix theory. The training of ELM is completed after the weights and deviations of all network nodes are obtained. Therefore, when the test data comes, the output layer weights just obtained can be used to calculate the network output to complete the prediction of data. The specific principle of ELM is as follows:

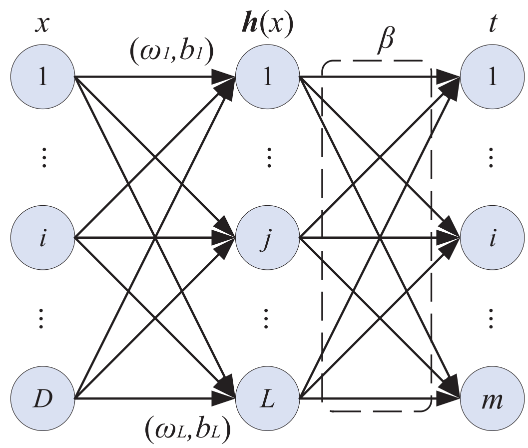

FIGURE 1. Network structure of ELM.

For any N different samples (x_j, t_j), where $x_j = [x_{j1}, x_{j2}, \cdots, x_{jn}]^T \in R^D$ and $t_j = [t_{j1}, t_{j2}, \cdots, t_{jm}]^T \in R^m$. x_j represents the jth data example, t_j represents the label corresponding to the jth data example, and the set (x_j, t_j) refers to all training data. Obviously, in Figure 1, the input of the neural network from left to right is the training sample set x, and there is a hidden layer in the middle. From the input layer to the hidden layer, there is a full connection. Note that the output of the hidden layer is $H(x)$, and the calculation formula of the output $H(x)$ of the hidden layer is as follows:

$$H(x) = [h_1(x), \cdots, h_L(x)] \tag{1}$$

The output of the hidden layer is obtained by multiplying the input by the corresponding weight plus the deviation, and then summing the results of all nodes of a nonlinear function. $H(x) = [h_1(x), \cdots, h_L(x)]$ is the ELM nonlinear mapping (hidden layer output matrix), and $h_i(x)$ is the output of the *ith* hidden layer node. The output function of the hidden layer node is not unique. Different output functions can be used for different hidden layer neurons. Generally, in practical application, $h_i(x)$ is expressed as follows:

$$h_i(x) = g(x) \tag{2}$$

where $g(x)$is the activation function, which is a nonlinear piecewise continuous function satisfying the general approximation ability theorem of ELM. The standard SLFNs mathematical model with L hidden nodes and activation function $g(x)$ is modeled as:

$$\sum_{i=1}^{L} \beta_i g_i(x_j) = \sum_{i=1}^{L} \beta_i g_i(w_i \cdot x_j + b_i) = o_j \tag{3}$$

where $w_i = [w_{i1}, w_{i2}, \cdots, w_{in}]^T$ is the weight vector connecting the i_{th} hidden node with the input node. $\beta_i = [\beta_{i1}, \beta_{i2}, \cdots, \beta_{im}]^T$ is the weight vector connecting the i_{th} hidden node with the output node. b_i is the threshold of i_{th} hidden node. $o_j = [o_{j1}, o_{j2}, \cdots, o_{jm}]^T$ is the j_{th} output vector of SLFNs.

Standard SLFNs with L hidden nodes and activation function $g(x)$ can approximate N samples with zero error. It means that $\sum_{j=1}^{L} ||o_j - t_j|| = 0$ and there are β_i, w_i and b_i as follows:

$$\sum_{i=1}^{L} \beta_i g(w_i \cdot x_j + b_i) = t_j \quad (j = 1, 2, \cdots, N) \tag{4}$$

The above equation can be succinctly expressed as:

$$H\beta = T \tag{5}$$

where (6) and (7), as shown at the bottom of the page, H is called the hidden layer output matrix of the neural network and the i_{th} column of H is the i_{th} hidden node output with respect to inputs $x_1, x_2, \cdots, x_N$. The smallest norm least-squares solution of the above linear system is:

$$\widehat{\beta} = H^{\dagger} T \tag{8}$$

where $H^{\dagger}$ is the Moore-Penrose generalized inverse of matrix H. The the output function of ELM can be modeled as follows:

$$f(x) = h(x)\beta = h(x)H^{\dagger}T \tag{9}$$

B. WORKFLOW OF CAD SYSTEM

The CAD system based on the medical image can be divided into two categories: one is the Computer-Aided Detection (CADe) system which detects and locates anomalies on medical images; the other is the Computer-Aided Diagnosis (CADx) system which detects anomalies on medical images and helps doctors determine the types of anomalies and malignant levels. The specific links of CAD systems for different diseases and application areas are slightly different, but the main structures are almost similar. The general processing framework for CADe and CADx systems is shown in Figure 2.

CAD systems generally include the following modules:

1) **Image acquisition**: Image acquisition refers to the way the system acquires medical images. Generally, there are three ways. The first is to acquire images from self-built image libraries, which are generally built using medical images obtained from partner hospitals [14]. The second is to acquire images through the system attached to the image generation equipment, such as Picture Archiving and Communication Systems (PACS) [15]. The third is to obtain data directly from the imaging system in real-time [16].
2) **Preprocessing**: The preprocessing process refers to correcting the distortion caused by media attenuation, noise, or motion artifacts, normalizing the original image [17], and enhancing the display quality of the

$$\begin{aligned} &H(w_1, w_2, \cdots, w_L, b_1, b_2, \cdots, b_L, x_1, x_2, \cdots, x_L) \\ &= \begin{bmatrix} g(w_1 \cdot x_1 + b_1) & g(w_2 \cdot x_1 + b_2) & \cdots & g(w_L \cdot x_1 + b_L) \\ g(w_1 \cdot x_2 + b_1) & g(w_2 \cdot x_2 + b_2) & \cdots & g(w_L \cdot x_2 + b_L) \\ \vdots & \vdots & \vdots & \vdots \\ g(w_1 \cdot x_N + b_1) & g(w_2 \cdot x_N + b_2) & \cdots & g(w_L \cdot x_N + b_L) \end{bmatrix}_{N \times L} \end{aligned} \tag{6}$$

$$\beta = \begin{bmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1m} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ \beta_{L1} & \beta_{L2} & \cdots & \beta_{Lm} \end{bmatrix}_{L \times m}$$

$$T = \begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1m} \\ t_{21} & t_{22} & \cdots & t_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ t_{N1} & t_{N2} & \cdots & t_{Nm} \end{bmatrix}_{N \times m} \tag{7}$$

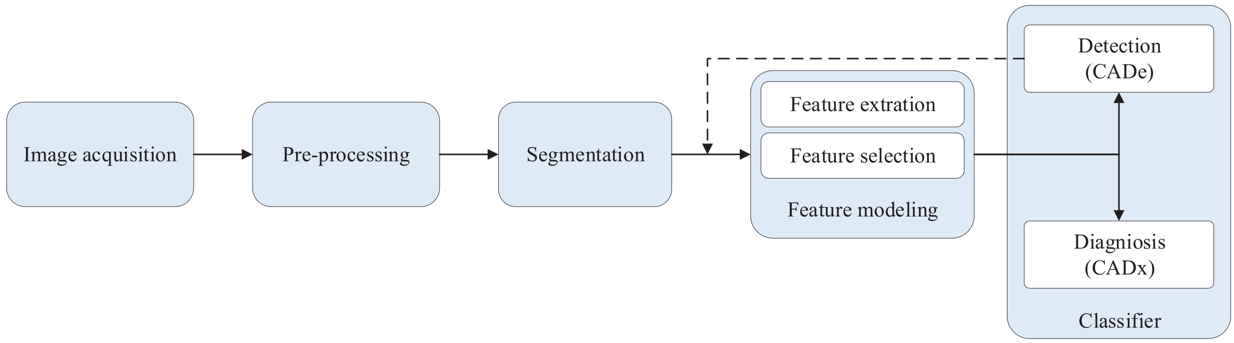

FIGURE 2. Workflow of computer aided detection/diagnosis systems.

image by denoising and increasing the contrast [18] for subsequent processing.

3) **Segmentation**: In order to reduce the interference of peripheral tissues or background on the detection of the region of interest, and reduce the amount of calculation, some CAD systems also need to perform a one-step image segmentation operation after preprocessing to separate the area to be studied from the background or surrounding tissues. Image segmentation is the basis for this step of the CAD system [19]. Most image description and recognition techniques rely heavily on the results of segmentation [20].
4) **Feature modeling**: Feature extraction is performed in the feature modeling component, and if necessary, feature selection or dimensionality reduction is performed. Feature extraction uses algorithms to calculate various feature values of the Region Of Interest (ROI), such as color features, texture features, shape features, and spatial relationship features. When the feature dimension is large, in order to ensure the performance of the system, it is necessary to make the optimal selection of the features, and only the features that have a large effect on the classification result which is the feature selection. The purpose is to reduce redundant features and reduce feature dimensions to improve computational efficiency and maximize classification accuracy.
5) **Detection/Diagnosis**: Detection refers to labeling and locating abnormal symptoms. Diagnosis refers to the classification of benign and malignant lesions. These two functions are the core of CADe and CADx respectively, and both rely on classifiers for implementation.

III. FEATURE MODELING

The high-dimensional irrelevance of medical image features, the heterogeneity of feature subsets, and the uneven distribution of sample categories have been obstacles to improving the accuracy of detection and diagnosis of various diseases. Therefore, feature modeling can be regarded as a key part of detecting lesions or identifying benign and malignant lesions. The general feature modeling part mainly includes feature extraction and feature selection. The method of feature extraction and selection will directly affect the performance of classifier.

A. FEATURE EXTRACTION

The features of the image can reveal the basic attributes of the image. The CAD system can extract feature values on the ROI region from the perspectives of morphology, texture, shape, color, and so on. Then the algorithm model is trained according to the sample database marked by experienced doctors, and these feature values are used to distinguish different classes of similar objects. The feature extraction method used in the ELM-based CAD system is organized in Table 1.

It can be found that image texture and shape features have been widely used in ELM-based CAD systems. For these two common feature types, some experiments have carried out intuitive experimental comparisons on feature extraction methods, which provides a reference for the selection of feature extraction methods in CAD systems based on ELM. For shape features, [28] combines three shape feature extraction methods, Scale Invariant Feature Transform (SIFT), Harris corner detection and Zernike Moments, with Deep Neural Network (DNN) and ELM respectively for brain tumor classification based on MRI. The experimental results show that the combination of Zernike moment and ELM is the best. Reference [38] discusses the availability of multiple feature extraction methods and classifiers when using thermal images for breast disease. When ELM is used as a classifier, the Haralick moment and Zernike moment are combined to obtain the best result, and this result is superior to other combinations. This indicates that both texture and shape informations are related to the identification of breast lesions by thermography. For texture features, three texture-based feature extraction methods, wavelet feature, Gray Level Spatial Dependence Matrix (GLSDM) and Gabor filter-based techniques, are compared in [31], [32]. Wavelet-based tissue texture analysis combined with ELM or CC-ELM for microcalcification detection in digitized mammograms can achieve better classification performance.

In order to fully express the image features or imitate the doctor's diagnostic process to obtain better classification performance, many studies have adopted mixed feature

TABLE 1. Summary of feature extraction methods.

Feature types	Methods	References
Geometric features	-	[21]–[24]
	Roundness	[21], [22], [24]
	Entropy of Standardized Radius	[21], [22], [24]
	Variance of Standardized Radius	[21], [22], [24]
	Ratio of Area	[21], [22], [24]
	Roughness	[21], [22], [24]
	Circularity	[24]
	Length-width ratio	[24]
	Squareness	[24]
Shape feature	-	[25]–[29]
	Harris Corner Detection	[28]
	Zernike Moments	[25], [27], [28]
	SIFT	[28]
	Histogram of Oriented Gridients (HOG)	[29]
Textural features	-	[21]–[27], [30]–[47]
	Grey-Level Co-occurrence Matrix (GLCM)	[22], [24], [26], [30], [31], [35], [38], [44], [47], [48]
	3D GLCM	[42]–[44]
	GLSDM	[31], [32]
	Haralick	[27]
	Gabor	[31], [35], [38], [46]
	LBP	[44]
	Wavelet features	[31], [32], [34], [49]
	SURF	[35]
	Discrete Tchebichef Transform (DT-T)	[36]
	Gray Run Length	[42]
	3D-GRLM	[42]
	Run-Length Matrix (RLM)	[43]
CNN	-	[49], [50]
	MFC-CNN	[51]
	3D-CNN	[52]
Color features	-	[45]
Gray level features	-	[33]
Density feature	-	[24]
EIS features	-	[53]
Spectroscopic features	-	[54]

TABLE 2. Evaluation indicators of diagnosis in [24].

Category	Classfier	Accuracy	Sensitivity	Specificity	TP Ratio	TN Ratio	AUC
GT	BP	0.73	0.713	0.718	0.784	0.603	0.71
	SVM	0.812	0.8	0.793	0.877	0.694	0.798
	ELM	0.833	0.821	0.81	0.903	0.719	0.824
GD	BP	0.744	0.761	0.741	0.802	0.644	0.738
	SVM	0.827	0.848	0.819	0.894	0.729	0.818
	ELM	0.851	0.866	0.842	0.917	0.753	0.848
TD	BP	0.753	0.771	0.734	0.818	0.725	0.767
	SVM	0.838	0.859	0.814	0.901	0.812	0.849
	ELM	0.864	0.882	0.835	0.925	0.83	0.862
GTD	BP	0.789	0.814	0.776	0.84	0.747	0.798
	SVM	0.871	0.903	0.847	0.929	0.838	0.871
	ELM	**0.895**	**0.926**	**0.873**	**0.948**	**0.846**	**0.881**

models or extracted features from a new perspective. In [48], the eigenvector model was established by mathematical methods, and the geometric and texture feature sets were combined for breast cancer diagnosis on digital mammography. On the basis of this feature model, [22] proposed a fused feature model that blends features of single views with comparative features of double views to simulate the process of doctor's film reading. In [24], the feature model and classifier are validated respectively in breast mass detection, and local fusion features with sub-region density are established. This method combines geometric features, texture features, and density features to build a local feature model for breast cancer detection and diagnosis. When the classifier is ELM, this feature model is superior to the geometry features + texture features (GT) model, geometry features + density features (GD) model and the nature features + density features (TD) model. The specific experimental results are shown in Table 2. The GTD model mentioned in this paper has obvious advantages in large-scale diagnosis. In [25], wavelet transform and Zernike moments were used to extract the texture features

TABLE 3. Summary of feature selection methods.

Feature selection method	References
Heuristic search	[30]
Principal component analysis (PCA)	[32], [36]
GAS	[22]–[24], [34], [42], [47], [49], [62], [63]
Impact value selection	[22], [23]
SFS	[22], [23]
CBF	[26]
Glowworm Swarm Optimization(GSO)	[35]
Linear discriminant analysis (LDA)	[36]
Rough set (RS)	[64]
Coefficient approach	[65]
Information theoretic criterion	[53]
Differential evolution (DE) algorithm	[43]
Choose the feature with the largest variance	[44]
SVM+ELM	[33]
ReliefF method	[61]

and shape features on the brain magnetic resonance images (MRI), respectively, for the benign and malignant identification of brain tumors. Although wavelet-based feature extraction methods are widely used in CAD systems, most methods are limited to expressing the correlation within each wavelet scale, while ignoring the correlation between wavelet scales. Reference [34] proposed a Hidden Markov Tree model of Dual-Tree Complex Wavelet Transform (DTCWT-HMT), and combined DTCWT-HMT based features with DTCWT based features for a microcalcification diagnosis system. This method effectively simulates the statistical distribution of wavelet coefficients and better reflects the correlation between wavelet coefficients. In [44], for feature extraction of brain MRI, first use LBP method to extract local feature information, and then use GLCM method to extract global features. In this way, the local and global features are used to fully describe the brain image.

Although various excellent algorithms appear in the field of medical image diagnosis, feature extraction and model selection established by traditional methods are always difficult to generalize robustly. Convolutional Neural Network (CNN) is one of the representative network structures in deep learning technology, and its application has solved this problem well. CNN is based on artificial neural networks. For different classification tasks, the backpropagation algorithm is used to automatically strengthen or reduce the weight of corresponding features. The filters are used to automatically extract features in the convolution layer to achieve image feature extraction [55]. The images can be directly used as the input of CNN, thus avoiding the complex feature extraction design and data reconstruction process in traditional algorithms. Some studies have proved that using CNN combined with ELM or its improved method achieve a better performance in the diagnosis of malignant diseases based on medical image data. Reference [56] uses CNN for feature extraction in brain tumor diagnosis systems. In [50], CNN and Discrete Wavelet Transform-Singular Value Decomposition (DWT-SVD) are combined to propose a new perceptual hash function. This method effectively reduces the execution time of CNN structure and the space occupied by image data on the hard disk, and achieves good classification performance in ELM. The accuracy rate of benign and malignant classification of liver tumors in CT images is 97.3%.

B. FEATURE SELECTION

Whether the feature set contains irrelevant or redundant information directly affects the performance of the classifier. Feature selection refers to the process of selecting a subset of features that make the classification results most significant [57]. Its purpose is to reduce redundant features and reduce feature dimensions to improve computational efficiency and maximize classification accuracy [58]. When the size of the feature set is too large, an appropriate feature selection method is the key to giving full play to the performance of the ELM classifier. The feature selection methods used in conjunction with ELM in related studies are summarized in table 3.

There are two classical algorithms for feature dimensionality reduction: Linear Discriminant Analysis (LDA) [59] and Principal Component Analysis (PCA) [60]. LDA performs a new projection on the feature values. After projection, the distances of data points of different properties are greater, and the distances of data points of the same nature are more compact. PCA maps high-dimensional features to low-dimensional space from the perspective of covariance and expects that the variance of the data is the largest in the projected dimension. In order to obtain better classification performance and simplify classification tasks, PCA and LDA are fused in [36]. Experiments have proved that compared with PAC, PCA + LDA can do the same with relatively few features. Feature selection does not do projection or mapping, but only selects some features from all features. Common algorithms include ReliefF Method [61], Sequential Forward Selection (SFS) [22], [23] and Genetic Algorithm Selection (GAS) [22]–[24], [34], [42], [47], [49], [62], [63]. In the research process of references [22] and [23], three popular feature selection algorithms, GAS, impact value selection and SFS, are compared and tested. From the results, GAS is the algorithm with the best obvious effect on ELM classifier performance optimization.

In addition, there are some feature selection methods and conclusions that have certain reference significance for future work. Reference [30] analyzes the number of features and neurons needed to achieve the best classification performance when using ELM to classify breast benign and malignant tumors, through heuristic search. It provides a reference for the determination of parameters in the future research. In [33], ELM and Support Vector Machine (SVM) are combined for feature eliminations. The mean of accuracy obtained by SVM and ELM classifier is used as the score of each feature, and the features that have little influence on SVM and ELM are deleted recursively. This method can select the most suitable feature combination for the two classifiers. In [26], the Correlation Based Feature (CBF) selection method is used for feature selection, which is faster than other methods in essence. Reference [64] used Rough Set (RS) to reduce the attribute set in the database, and proposed the RS-ELM model. RS theory provides an effective tool for studying the analysis and reasoning of inaccurate data, mining the relationships between data, and discovering potential knowledge.

IV. APPLICATION OF ELM IN CAD

In addition to detection and diagnosis, ELM has also been applied in cancer prognosis prediction [65], tumor segmentation [48] and feature selection [33], which shows that ELM has great application space and development potential in CAD system. Table 7 lists the application of ELM and its improved algorithms in CAD.

In order to further improve the efficiency and performance of the system, some improved algorithms based on Classic ELM are applied to the CAD system. The performance of ELM depends on the input weights and the bias. In order to avoid high computational complexity and fall into local optimal solutions, it is especially important to set the appropriate parameter values. In [31], krill herd algorithm is used to optimize the weight. As a low-level animal foraging algorithm, krill herd algorithm has few function parameters and strong group tendency, which provides a good structural framework for solving the optimization problem. In [65], the parameters of ELM were optimized by BAT algorithm. The biomimetic model, BATELM, was used to predict the recurrence and recurrence time of breast cancer, which provided an important reference for cancer risk prediction. However, the above two methods of optimizing parameters have shortcomings of slow convergence speed and poor globality. Compared with the former two optimization algorithms, Particle Swarm Optimization (PSO) has faster convergence speed and higher accuracy. PSO-ELM model and Improved PSO-ELM (IPSO-ELM) model start from the best individual obtained in the learning stage, have good generalization ability, and have been applied in the fields of brain tumor diagnosis [42], [54] and abnormal detection of ceramic cancer cells [56]. A breast cancer diagnosis system based on Multi-Layer ELM (ML-ELM) is proposed in [66]. The Area Under the Receiver Operating Characteristic (AUROC) curves is used as the performance index to analyze the performance of ELM, and the system is optimized by weight attenuation method according to the analysis results. In [36], a parameter less based global optimization algorithm called Improved Gray Wolf Optimization-based ELM (IGWO-ELM) is proposed. The algorithm uses GWO to adaptively calculate the optimization value of ELM hidden node parameters, and then uses Moore Penrose inverse to analyze. In the experiment of benign and malignant diagnosis of breast tumor, the highest accuracy rate is 100%.

For dealing with real value classification, in [67], the Circular Complex Valued ELM (CC-ELM) is proposed. This method shows obvious performance advantages in the case of highly unbalanced data sets. In this study, CC-ELM is used in mammography classification. Compared with traditional ELM classifier, the performance of classification is improved by nearly 9%. In [32], CC-ELM was tested on DDSM database, which was used to diagnose microcalcifications on mammogram images. Good results were also obtained, with an accuracy of 96.2%.

For the improvement of operation efficiency, in [29], aiming at the problem of brain tumor cell recognition, the H-ELM based on Histogram Orientation Gradient (HOG) is implemented. On this basis, Parallel H-ELM (PH-ELM) is proposed, which is accelerated by GPU, further improving the performance of H-ELM in high-dimensional and large image data set computing complexity. In the experiment, compared with ELM and H-ELM, the speed of PH-ELM is increased by 7 times and 3 times respectively.

For the multi-classification problem, it has been proved in [68] that ELM can be directly used to solve regression and multi-classification problems, which provides a theoretical basis for the use of ELM in the diagnosis of multiple types of tumors. In [28] and [46], Kernel based ELM (KELM) and Regulated ELM (RELM) are applied to the classification of various brain tumors. Among them, KELM has a strong ability to solve multi-class recognition problems, and the RELM not only avoids the number of iterations and local minimums, but also has better generalization, robustness, and controllability [69].

The ensemble method is a meta-algorithm which combines several machine learning technologies into a prediction model to achieve the effect of reducing variance and boosting or improving prediction. With reference to this idea, the ensemble learning of multiple-view 3D-CNNs model for micro-nodules identification is now available in [52]. In this model, 5 3D-CNN components are integrated using ELM, and the final classification results are generated. In particular, for the integration method, ELM has better performance than majority voting, averaging, operators, and autoencoders. The difference in the number of hidden neurons in the ELM will lead to different results in a particular classification task, that is, each ELM classifier with a different structure can provide different classification information. According to the idea of ensemble learning, combining these classifiers can improve the efficiency and accuracy of the overall system [70]. In [53], ELM classifiers with different structures are used for classification, and then their amount classification results are combined using SVM to classify breast tissue.

Learning sample scarcity is a problem that must be faced in all machine learning research. The size and quality of the data set have a significant impact on the performance of the system. In some studies, solutions have been proposed for data shortages. Huang *et al.* [41] used one-class ELM to provide a preliminary detection scheme for liver tumors on CT images. This method can still detect most tumors even when the training set is incomplete. Reference [39] further improves the method in [41], and proposes Random Feature Subspace Ensemble based ELM (RFSE-ELM). In this method, KELM is selected as the basic classifier, and then the classification results of the basic classifier set are fused using the majority voting method. Without training data, the one-class RFSE-ELM can also detect liver tumors. The previous algorithm is improved again in [40], and Data and Feature Mixed Ensemble based ELM (DFEN-ELM) was proposed. This method combines multiple weak classifiers to implement a strong classifier, which overcomes the problem of overfitting while maintaining the advantage of ELM in training speed. This study implements ELM based on one and two types of integrations, and uses the new training data to improve system performance. Sequential kernel learning was further used in the experiments to achieve fast retraining and iteratively enhance image segmentation performance. In [47],Semi Supervised ELM (SS-ELM) was used to achieve the auxiliary diagnosis of pulmonary nodules. This method can input both labeled feature sets and unlabeled feature sets. It has faster learning speed and higher test accuracy, and it also has better generalization performance. Furthermore, in Document [24], Unsupervised ELM (US-ELM) is used to cluster the density features on the sub-regions to realize the detection of breast tumors. Then use ELM to complete the benign and malignant diagnosis of breast tumors.

V. PERFORMANCE OF ELM IN CAD

ELM is a very simple and fast neural network learning algorithm. In the past decade, the theory and application of ELM has been widely studied. From the point of view of learning efficiency, ELM has the advantages of less training parameters, fast learning speed and strong generalization ability. In a large number of experiments on the standard UCI data set, it is shown that ELM has faster training speed and better generalization performance than Back-Propagation (BP) [71] algorithm and SVM method [72]. In order to further explain the performance of ELM in CAD, we will discuss the performance of ELM in specific application by analyzing the experimental results in related research. The comparison experiments between ELM and other classifiers are widely used in related researches. In Table 4, we give a brief example of some experimental results. It can be seen that ELM generally performs well in the research.

ELM not only has remarkable performance when it is used alone, but also can be combined with other algorithms to get better system performance. Combined with the current popular CNN network, ELM and its improved algorithm also have good performance in the experiment [49], [50]. Table 5 shows the experimental results in [49]. It can be seen that the addition of KELM improves the overall accuracy of nearly 10% compared with CNN network, and is significantly better than other classifiers in this study. It is worth noting that ELM can also be used in combination with 3D CNN and improve the overall efficiency, which provides an effective solution for the realization of 3D medical data oriented CAD [52]. In [53], the combination of ELM and SVM is used to classify breast tissue, and the effect of this method is better than that of SVM alone or ELM alone. Similarly, in [26], the combination of RF and ELM is much better than the effect of using RF alone.

To sum up, from many related research results, it can be seen that using ELM or its improved algorithm as classifier in CAD can often obtain satisfactory results, and has advantages in training and testing time. In addition, ELM can be used with many other algorithms, and can improve the overall performance of the system. Therefore, ELM is suitable for solving the classification problems in CAD, and has a broad application prospect and high research value in CAD-related fields.

TABLE 4. Comparison of different classifiers.

References	Classifier	Acc	Sn	Sp
[30]	SVM	0.82	0.86	0.78
	VSVM	0.84	0.93	0.74
	RFDC	0.9	0.9	0.98
	KNN	0.84	0.91	1
	ELM	**0.91**	**0.9**	**0.98**
[73]	SVM	0.968	-	-
	ELM	**1**	**-**	**-**
	Naive Bayes	0.959	-	-
[49]	SVM	0.864	0.882	0.863
	ELM	**0.924**	**0.913**	**0.921**
[47]	SVM	0.942	0.951	0.922
	ELM	**0.95**	**0.964**	**0.928**
	PNN	0.833	0.894	0.857
	MLP	0.859	0.861	0.883

TABLE 5. The performance of E-CNN method and compare with other classifiers in [49].

Features	Classifier	Accuracy
Feature selected using CNN	**KELM**	**0.937**
	MLP	0.888
	Stacking	0.869
	XGBoost	0.873
	SVM	0.875
	RBF	0.868
CNN	Fully-connected layer	0.811

VI. PROBLEMS AND POSSIBLE RESEARCH DIRECTIONS

It can be seen from the above overview that more and more CAD systems based on ELM are implemented, and the superiority of ELM algorithm can be shown by many experiments. It can not only ensure high accuracy and short training time of CAD, but also be widely used in many forms of medical data. Although remarkable experimental results have been obtained in most studies, a large part of them are tested in small samples or under specific circumstances, and the CAD system based on ELM still faces some challenges in practical application. In the clinical situation, the experimental subjects will become ordinary random cases, but the content structure of the actual unprocessed medical related data set is not standardized and there may be incomplete information, these factors may have a significant impact on the diagnosis results. To sum up, the CAD system based on ELM still cannot fully meet the needs of the practical clinical application, and there is room for improvement. According to the development trend, we think that further research can be carried out from the following aspects:

1) Because of the randomness of hidden layer parameters, the generalization performance of ELM trained by different initial parameters is different, which affects its stability and robustness. In order to avoid the instability of classification results caused by uncertain parameters, some researches are devoted to improving the ELM hidden layer node structure or optimizing the network parameter selection method [36], [65]. Although the existing research has proposed effective selection methods for ELM parameters, but the performance of these methods will be affected by the parameters in the introduced algorithm, resulting in the hidden danger of reducing the performance of the whole model caused by the parameters has not been fundamentally eliminated. The future research can take the realization of nonparametric method as the goal, mainly study the hybrid model of adaptive optimization parameter algorithm and ELM, in order to reduce or even completely avoid the negative impact of unreasonable parameter selection on system performance.
2) With the rapid popularization of Electronic Medical Record (EMR) systems in medical institutions, a large amount of important medical-related information is stored in the medical information system in electronic form. These data records important information in clinical medicine, such as examination results, diagnostic information, medications, etc., and their data types can be roughly divided into three types: text, numbers, and images. Researching a certain type of data alone cannot completely inherit the doctor's experience, so a complete auxiliary diagnosis system must combine these three types of data as the research object. Therefore, in addition to medical imaging, future research can realize multi-modal data analysis based on pathology, electronic medical records and other data to assist clinical department diagnosis and treatment plan recommendation. And further form the whole-process intelligent assistance system from screening to tumor grading and staging, and then to treatment plan recommendation.
3) Medical image classification based on visual semantics has always been a challenging research field. For medical images, there are not only many kinds of images, but also many variables (such as illumination change, dislocation, deformation, etc.) in each kind of image, which may affect the accuracy of classification. In this respect, deep learning shows good performance in image feature learning. The existing research has proved that the combination of deep learning method and ELM can effectively avoid the negative impact of segmentation error and human subjectivity brought in by hand-designed feature extraction model on the final classification and improve the accuracy of CAD diagnosis. Moreover, ELM can also have a positive impact on the efficiency of the system [49], [51]. Furthermore, in order to comprehensively analyze the information contained in the 3D medical data to obtain more accurate diagnosis results, deep learning algorithms for 3D data, such as 3D CNN, have also begun to be used in the diagnosis of diseases. 3D CNN can be used in conjunction with ELM, and can also achieve excellent results [52]. At present, the main problem of deep learning related research is that due to the complexity of the calculation process, it is often accompanied by high storage space requirements and calculation complexity while obtaining excellent accuracy. Using GPU to train neural network has become the standard of deep learning algorithm. However, in the actual deployment, the traditional general computing platform, including GPU and CPU, can not meet the comprehensive needs of power consumption and performance (or energy efficiency ratio) in the actual model deployment in most cases, which also makes some hardware with higher comprehensive energy efficiency ratio, such as Field Programmable Gate Array (FPGA) and Application Specific Integrated Circuit (ASIC) chips for AI applications gradually get attention. In order to meet the computing power and energy efficiency requirements of the deep neural network model in different application scenarios, on one hand, we could use the inherent redundancy of the existing deep neural network model to cut and optimize the model from the algorithm level without losing the accuracy of the model. On the other hand, we could also design a high-energy hardware architecture to optimize the calculation mode of the deep neural network model, which is used to accelerate the calculation process of the model. Of course, we could also combine these two aspects to design and optimize the algorithm and hardware together.
4) The lack of training data set is one of the difficulties for machine learning algorithm. For medical data, unlabeled data is easy to collect, but these data cannot be directly used for ELM training. Traditional ELM classifier can only use labeled data for training. However, it is very difficult to obtain the complete labeled

TABLE 6. Add caption.

Dataset	Data categories	Sample size	Reference
DDSM	Mammograms	2620 cases	[32]–[34], [36]
Mini-MIAS	Mammograms	322 cases, 207 are normal, 63 are benig, 52 are malignant	[33], [34], [36] [26], [30], [31] [48]
LIDC	Pulmonary CT, CR, DX	1010 cases	[47], [52], [62]
WBCD	Non-image data	11 Variables, 699 Observations	[64], [66], [73], [76]
WPBC	Non-image data	35 attributes and 198 instances	[65]

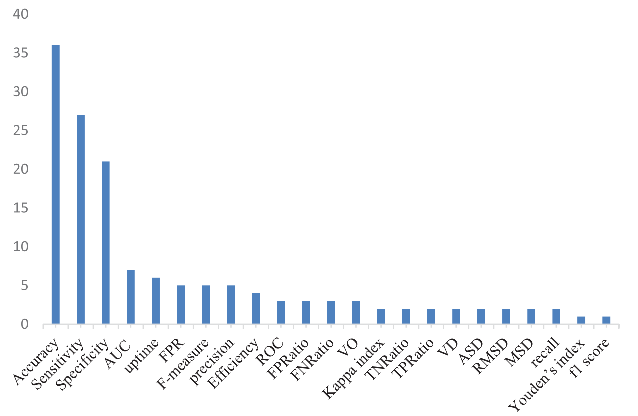

FIGURE 3. Distribution of evaluation metrics used in references.

medical sample data, which requires the efforts of experienced annotators. For the utilization of unlabeled data, semi-supervised learning and unsupervised learning can provide solutions to this problem. In some researches, semi-supervised learning or unsupervised learning has been combined with ELM to solve the problem of insufficient labeled samples by making full use of unlabeled data [24], [47]. However, the current research is not deep enough. The specific improvements of algorithms specifically for medical diagnosis have not been discussed in depth, and there is still room for development. To solve the problem of insufficient label data, transfer learning can transfer the model suitable for large label data set to small data set. By using the existing knowledge, it can solve the problem of insufficient label data in the target domain, which can broaden the application scope of existing data and improve the utilization of effective resources. For rare diseases, transfer learning may be a good way to solve the problem of data set shortage. ELM is more and more popular in the field of transfer learning because of its simplicity, training speed and ease of use in online sequential learning [74]. But the research of this kind of algorithm in disease diagnosis is very few. It is a very promising research direction to apply the ELM based transfer learning algorithms to the diagnosis of diseases.

VII. DATA SETS AND EVALUATION METHODS

After the training of the classifier, the performance of the classifier will be evaluated with test data. The use of public medical sample data sets for testing is the basis of effective, objective and fair evaluation of the performance of various CAD systems [75]. The use of the private data sets hinders the analysis and comparison of different algorithms and makes them invalid. In the current related researches, the commonly used public data sets are Digital Database for Screening Mammography (DDSM), the mini-MIAS database of mammograms, Lung Image Database Consortium (LIDC), Wisconsin Breast Cancer Database (WBCD), and database on Wisconsin Prognostic Breast Cancer (WPBC). We summarize the research using public data sets in Table 6.

The frequency of the use of the evaluation metrics in the related research is calculated in Figure 3. It can be seen that the first three evaluation metrics with the highest frequency of use are Accuracy (Acc), Sensitivity (Sn), and Specificity (Sp). Acc is the rate at which true positive and true negative individuals in a subject are correctly identified. Sn, also known as True Positive Rate (TPR), is the ratio of correctly identified positives in abnormal areas, and is a measure of the true positive recognition performance of a system. Sp, also known as True Negative Rate (TNR), is the ratio of correctly identified negatives in the normal category. It measures how well a system can correctly identify negative individuals. Equations of Acc, Sn, and Sp are given as follows:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \tag{10}$$

$$Sn = \frac{TP}{TP + FN} \tag{11}$$

$$Sp = \frac{TN}{TN + FP} \tag{12}$$

where TP, TN, FP, and FN are the detection and diagnosis results that four CAD systems may output: (1) TP means that the diagnosis is positive, and the true value of the object is also positive. (2) TN means that the diagnosis is negative and the true value of the subject is also negative. (3) FP means that the diagnosis is positive, but the true value of the subject is negative. (4) FN means that the diagnosis is negative, but the true value of the subject is positive.

In addition, Receiver Operating Characteristic (ROC) curve is based on statistical decision theory, which is widely used in the evaluation of CAD system. The TPR of the system is expressed by the ROC curve as a function of $1 - Sn$, and the overall performance of the CAD system is measured by the area under the ROC curve (AUC). The closer AUC is to 1, the better the performance of the system. When the AUC is 1, the system is perfect, it can correctly classify all samples. But in general, when the true positive rate of the system increases, the corresponding false positive rate also increases, so the AUC of the system will not reach 1.

TABLE 7. The application of ELM and its improved algorithm in CAD.

Location	Data type	Scope	Improvement	Reference	Data set	Results
Breast	Mammography	Segment tumor regions	-	[48]	MIAS	$Kappaindex = 0.49$
						$Acc = 0.85$
		Feature dimensionality reduction	Combine with SVM	[33]	MIAS DDSM	$Acc = 0.9672$
						$Sn = 0.9629$
						$Sp = 0.9432$
						$AUC = 0.9659$
		Breast mass detection	-	[48]	Private 482 images, 246 have tumors	$Acc = 0.83$
						$Sn = 0.86$
						$FPRatio = 0.82$
				[22]	Private 222 pairs of mammograms, 112 have tumors	$Acc = 0.87$
						$Sn = 0.96$
						$Particularity = 0.90$
						$TPRatio = 0.89$
						$TNRatio = 0.84$
				[23]	Private 490 images, 246 have tumors	$Acc = 0.89$
						$Sn = 0.87$
						$Sp = 0.87$
			US-ELM	[24]	Private 480 mammograms, 246 have tumors,116 are benig,130 are malignant	$Acc = 0.911$
						$Sn = 0.933$
						$Sp = 0.901$
						$TPRatio = 0.952$
						$TNRatio = 0.869$
						$AUC = 0.938$
			IGWO-ELM	[36]	MIAS	$Acc = 1$
					DDSM	$Acc = 0.995$
		Classification of benign and malignant masses	-	[30]	MIAS	$Efficiency = 0.91$
						$Sn = 0.9$
						$Sp = 0.98$
				[33]	MIAS DDSM	$Acc = 0.9672$
						$Sn = 0.9629$
						$Sp = 0.9432$
						$AUC = 0.9659$
				[34]	Nijmegen dataset	$AUC = 0.9856$
					MIAS	$AUC = 0.9941$
					DDSM	$AUC = 0.9168$
				[24]	Private 480 images, 246 have tumors, 116 are benig, 130 are malignant	$Acc = 0.911$
						$Sn = 0.933$
						$Sp = 0.901$
						$TPRatio = 0.952$
						$TNRatio = 0.869$
						$AUC = 0.938$
			CC-ELM	[24]	Private	$Efficiency = 1$
						$Traintime = 0.047$
				[32]	Private 400 ROIs, 200 are benign and 200 are malignant	$TPR = 0.962$
						$FPR = 0.038$
						$Precision = 0.962$
						$Acc = 0.962$
			IGWO-ELM	[36]	MIAS	$Acc = 1$
					DDSM	$Acc = 0.985$
		Classification of tumors and tissues	RF-ELM	[26]	MIAS	$Acc = 0.98$
						$Sn = 0.89$
						$Sp = 0.91$

TABLE 7. *(Continued.)* **The application of ELM and its improved algorithm in CAD.**

Location	Data type	Scope	Improvement	Reference	Data set	Results
Breast	Mammography	Microcalcifications detection	-	[31]	MIAS	$TPR = 0.98$
						$FPR = 0.05$
						$F-measure = 0.96$
						$Precision = 0.95$
						$AUC = 0.98$
						Training efficiency=1
						Testing efficiency=0.94
		Diagnosis of microcalcification	ELM-FOA	[35]	Private 184 images	$Acc = 0.96$
					MIAS	$Acc = 0.98$
	EIS	Breast tissue classification	Use SVM to organize multiple ELMs	[53]	The UCI Machine Learning Repository	Classification of carcinoma tissue against other tissues: $Acc = 0.9775$
						Classification of all six breast tissues: $Acc = 0.8895$
	Thermography	Ductal carcinoma in situ detection	-	[37]	Private 40 thermograms	$Sn = 0.93$
						$Sp = 0.925$
						$Acc = 0.928$
		Classification of cysts and lesions	-	[27]	Private 219 cyst, 371 benign lesions, 235 malignant lesions	$Acc = 0.7006$
						$Kappaindex = 0.6566$
	Non-image data	Classification of benign and malignant masses	-	[64]	WBCD	$Acc = 1$
						$Sn = 1$
						$Sp = 1$
				[76]	The Breast Cancer Wisconsin dataset 699 cases	$Acc = 0.964$
						$Sn = 0.948$
						$Sp = 0.974$
				[73]	The Breast Cancer Wisconsin dataset 699 cases	$Acc = 0.9899$
			ML-ELM	[66]	The Breast Cancer Wisconsin dataset 699 cases	Training time(s)$\leq 10^{-4}$
						Testing time(s)$\leq 10^{-4}$
						$Acc = 0.93$
		Predict cancer recurrence and the time of recurrence	BATELM	[65]	WBCP 35 attributes and 198 instances	$TrainingAcc = 0.94$
						$TestingAcc = 0.93$
						Training time(s)=1.49
						Testing time(s)=0.98
Liver	US	Risk stratification of fatty liver disease	-	[38]	Private 63 patients, 36 abnormal , 27 normal	$Sn = 0.913$
						$Sp = 0.921$
						$Acc = 0.924$
						$AUC = 0.92$
	CT	Liver tumor detection and segmentation	KELM	[41]	Private 20 tumors	$VO = 0.6715$
						$VD = 0.1416$
						$ASD = 2.27mm$
						$RMSD = 2.47mm$
						$MSD = 8.46mm$
			One-class RFSE-ELM	[39]	Private 20 tumors	$VO = 0.6882$
						$VD = 0.1412$
						$ASD = 1.65mm$
						$RMSD = 2.11mm$
						$MSD = 7.14mm$

TABLE 7. ***(Continued.)*** **The application of ELM and its improved algorithm in CAD.**

Location	Data type	Scope	Improvement	Reference	Data set	Results
Liver	CT	Liver tumor detection and segmentation	Two-class RFSE-ELM	[39]	Private 20 tumors	$VO = 0.7475$ $VD = 0.1189$ $ASD = 1.03mm$ $RMSD = 1.28mm$ $MSD = 4.77mm$
			DFEN-ELM	[40]	Private	$VO = 0.7526$
		Classification of benign and malignant masses	PH-C-ELM	[50]	Private 200 CT images	$Sn = 0.964$ $Sp = 0.982$ $Acc = 0.973$ Youden's index= 0.946
	Cancerous tissue pathological image	Epatocellular carcinoma (HCC) nuclei grading	Hybrid CNN-ELM model	[51]	Private 127 liver pathology images	$Acc = 0.967$ $Sn = 0.995$ $Sp = 0.975$ $Precision = 0.998$ $F1score = 0.996$
Brain	MRI	Classification of benign and malignant brain tumors	-	[25]	Private	$Acc = 0.72$ $Sn = 0.79$ $Sp = 0.67$
			ELM-LRF	[77]	Private 16 patients' data	$Acc = 0.9718$ $Sn = 0.968$ $Sp = 0.9712$
		Brain tissue and pathological tumor classification	-	[28]	Private 1000 images, 5 classes and each class has 200 images	$Sn = 0.51$ $Sp = 0.16$ $Acc = 0.99$ $Errorrate = 0.01$ $F-measure = 0.66$
			IPSO-ELM	[42]	Private	$Acc = 0.98$
			Hybrid Krill herd- ELM	[44]	Private 400 sample images	$Sn = 0.985$ $Sp = 0.979$ $Acc = 0.989$
			KELM	[49]	Private	$Acc = 0.9368$
			ELM-RGSO	[43]	Private	$Acc = 0.9368$
			RELM	[46]	Private	$Acc = 0.9423$
	MRS and MR	Brain tissues and tumors classification	IPSO-ELM	[54]	Private 35 clinical routine cases	$Sn = 0.9801$ $Sp = 0.95$ $Acc = 0.9915$
	Microscopy images of cells	Brain tumor cell recognition	PH-ELM	[29]	Private 13560 tumor-images	$Acc = 0.9$
Lung	CT	Classification of benign and malignant pulmonary nodule	SS-ELM	[47]	LIDC-IDRI 1439 pulmonary nodule images, 454 benign, 613 malignant, 372 for training data	$Sn = 0.9656$ $Sp = 0.9503$ $TrainingAcc = 0.9757$ $TestingAcc = 0.9591$ $AUC = 0.961$ $FPRatio = 0.9635$ $FNRatio = 0.9538$
		Micro-nodules detection	-	[52]	LIDC-IDRI 19081 samples	$Acc = 0.9113$ $Sn = 0.7349$
	Non-image data	Multi classification of lung cancer	Fuzzy ELM	[63]	The UCI Machine Learning Repository 32 samples	$Acc = 0.9885$
Cervix uteri	FLIM	Early diagnosis	-	[78]	Private	$Sn = 0.946$ $Sp = 0.843$

TABLE 7. *(Continued.)* The application of ELM and its improved algorithm in CAD.

Location	Data type	Scope	Improvement	Reference	Data set	Results
Cervix uteri	Microscopy images of cells	Abnormality detection of cells	Fast PSO-ELM	[56]	Private 50 images	$Acc = 0.9476$
Kidney	Tissue microarrays	Tumor detection	-	[45]	Private 90 tissue ROIs	$Acc = 0.9173$
Thyroid	US	Classification of benign and malignant tumors	-	[61]	Private 114 benign nodules and 89 malignant nodules	$Acc = 0.8772$
						$AUC = 0.8672$
						$Sp = 0.9455$
						$Sn = 0.7889$

VIII. CONCLUSION

CAD has always been a research hotspot in the field of medical information processing. The establishment of a powerful, high-performance CAD system can better help doctors find and diagnose diseases, especially malignant diseases, improve the survival rate of patients, and improve the quality of life of patients. Many research results can prove that ELM can be applied to the construction of CAD, and the research in this field has important medical and social value.

This paper mainly discusses the feature extraction method, feature selection method, the application of ELM and its improved algorithm in CAD, the performance of ELM and its future development prospect. It can be seen that ELM algorithm not only has short processing time, but also has good generalization performance. The application prospect of ELM in CAD system is broad, and there is still room for development and improvement, which is worthy of further study.

REFERENCES

[1] H. N. Wagner and P. S. Conti, "Advances in medical imaging for cancer diagnosis and treatment," *Cancer*, vol. 67, no. S4, pp. 1121–1128, Feb. 1991.
[2] G. Li, D. Citrin, K. Camphausen, B. Mueller, C. Burman, B. Mychalczak, R. W. Miller, and Y. Song, "Advances in 4D medical imaging and 4D radiation therapy," *Technol. Cancer Res. Treatment*, vol. 7, no. 1, pp. 67–81, Feb. 2008.
[3] R. S. Ledley, "Digital electronic computers in biomedical science," *Science*, vol. 130, no. 3384, pp. 1225–1234, 1959.
[4] G. S. Lodwick, T. E. Keats, and J. P. Dorst, "The coding of roentgen images for computer analysis as applied to lung cancer," *Radiology*, vol. 81, no. 2, pp. 185–200, Aug. 1963.
[5] J. Yanase and E. Triantaphyllou, "A systematic survey of computer-aided diagnosis in medicine: Past and present developments," *Expert Syst. Appl.*, vol. 138, Dec. 2019, Art. no. 112821.
[6] R. Fusco, M. Sansone, S. Filice, G. Carone, D. M. Amato, C. Sansone, and A. Petrillo, "Pattern recognition approaches for breast cancer DCE-MRI classification: A systematic review," *J. Med. Biol. Eng.*, vol. 36, no. 4, pp. 449–459, Aug. 2016.
[7] M. Eshtay, H. Faris, and N. Obeid, "Metaheuristic-based extreme learning machines: A review of design formulations and applications," *Int. J. Mach. Learn. Cybern.*, vol. 10, no. 6, pp. 1543–1561, Jun. 2019.
[8] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: A new learning scheme of feedforward neural networks," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, vol. 2, Jul. 2004, pp. 985–990.
[9] G.-B. Huang, N.-Y. Liang, H.-J. Rong, P. Saratchandran, and N. Sundararajan, "On-line sequential extreme learning machine.," *Comput. Intell.*, vol. 2005, pp. 232–237, Jul. 2005.
[10] G.-B. Huang and L. Chen, "Convex incremental extreme learning machine," *Neurocomputing*, vol. 70, nos. 16–18, pp. 3056–3062, Oct. 2007.
[11] Y. Wang, F. Cao, and Y. Yuan, "A study on effectiveness of extreme learning machine," *Neurocomputing*, vol. 74, no. 16, pp. 2483–2490, Sep. 2011.
[12] J. Cao, Z. Lin, G.-B. Huang, and N. Liu, "Voting based extreme learning machine," *Inf. Sci.*, vol. 185, no. 1, pp. 66–77, Feb. 2012.
[13] X. Liu, L. Wang, G.-B. Huang, J. Zhang, and J. Yin, "Multiple kernel extreme learning machine," *Neurocomputing*, vol. 149, pp. 253–264, Feb. 2015.
[14] M. Firmino, A. H. Morais, R. M. Mendoça, M. R. Dantas, H. R. Hekis, and R. Valentim, "Computer-aided detection system for lung cancer in computed tomography scans: Review and future prospects," *Biomed. Eng. OnLine*, vol. 13, no. 1, p. 41, 2014.
[15] L. Bogoni, J. P. Ko, J. Alpert, V. Anand, J. Fantauzzi, C. H. Florin, C. W. Koo, D. Mason, W. Rom, M. Shiau, M. Salganicoff, and D. P. Naidich, "Impact of a computer-aided detection (CAD) system integrated into a picture archiving and communication system (PACS) on reader sensitivity and efficiency for the detection of lung nodules in thoracic CT exams," *J. Digit. Imag.*, vol. 25, no. 6, pp. 771–781, Dec. 2012.
[16] Y. Kominami, S. Yoshida, S. Tanaka, Y. Sanomura, T. Hirakawa, B. Raytchev, T. Tamaki, T. Koide, K. Kaneda, and K. Chayama, "Computer-aided diagnosis of colorectal polyp histology by using a real-time image recognition system and narrow-band imaging magnifying colonoscopy," *Gastrointestinal Endoscopy*, vol. 83, no. 3, pp. 643–649, Mar. 2016.
[17] H. Fujita, X. Zhang, S. Kido, T. Hara, X. Zhou, Y. Hatanaka, and R. Xu, "An introduction and survey of computer-aided detection/diagnosis (CAD)," in *Proc. IEEE Int. Conf. Future Comput., Control Commun.*, Sep. 2010, pp. 200–205.
[18] A. K. Dhara, S. Mukhopadhyay, and N. Khandelwal, "Computer-aided detection and analysis of pulmonary nodule from ct images: A survey," *IETE Tech. Rev.*, vol. 29, no. 4, pp. 265–275, 2012.
[19] K. Elakkia and P. Narendran, "Survey of medical image segmentation using removal of Gaussian noise in medical image," *Int. J. Eng. Sci.*, vol. 6, no. 6, pp. 7593–7595, 2016.
[20] C. Jacobs, E. M. van Rikxoort, E. T. Scholten, P. A. de Jong, M. Prokop, C. Schaefer-Prokop, and B. van Ginneken, "Solid, part-solid, or non-solid: Classification of pulmonary nodules in low-dose chest computed tomography by a computer-aided diagnosis system," *Investigative Radiol.*, vol. 50, no. 3, pp. 168–173, Mar. 2015.
[21] Z. Wang, G. Yu, Y. Kang, Y. Zhao, and Q. Qu, "Breast tumor detection in digital mammography based on extreme learning machine," *Neurocomputing*, vol. 128, pp. 175–184, Mar. 2014.
[22] Z. Wang, Q. Qu, G. Yu, and Y. Kang, "Breast tumor detection in double views mammography based on extreme learning machine," *Neural Comput. Appl.*, vol. 27, no. 1, pp. 227–240, Jan. 2016.
[23] Z. Wang, Y. Kang, G. Yu, and Y. Zhao, "Breast tumor detection algorithm based on feature selection elm," *J. Northeast Univ. (Nat. Sci.)*, vol. 34, no. 6, pp. 792–796, 2013.
[24] Z. Wang, Y. Huang, M. Li, H. Zhang, C. Li, J. Xin, and W. Qian, "Breast mass detection and diagnosis using fused features with density," *J. X-Ray Sci. Technol.*, vol. 27, no. 2, pp. 321–342, 2019.
[25] M. Sornam, M. S. Kavitha, and R. Shalini, "Segmentation and classification of brain tumor using wavelet and Zernike based features on MRI," in *Proc. IEEE Int. Conf. Adv. Comput. Appl. (ICACA)*, Oct. 2016, pp. 166–169.

[26] R. D. Ghongade and D. G. Wakde, "Detection and classification of breast cancer from digital mammograms using RF and RF-ELM algorithm," in *Proc. 1st Int. Conf. Electron., Mater. Eng. Nano-Technol. (IEMENTech)*, Apr. 2017, pp. 1–6.

[27] M. A. D. Santana, J. M. S. Pereira, F. L. D. Silva, N. M. D. Lima, F. N. D. Sousa, G. M. S. D. Arruda, R. D. C. F. D. Lima, W. W. A. D. Silva, and W. P. D. Santos, "Breast cancer diagnosis based on mammary thermography and extreme learning machines," *Res. Biomed. Eng.*, vol. 34, no. 1, pp. 45–53, Mar. 2018.

[28] A. A. Pandian and R. Balasubramanian, "Analysis on shape image retrieval using DNN and ELM classifiers for MRI brain tumor images," *Int. J. Inf. Eng. Electron. Bus.*, vol. 8, no. 4, pp. 63–72, Jul. 2016.

[29] W. Phusomsai, C. So-In, C. Phaudphut, C. Thammasakorn, and W. Punjaruk, "Brain tumor cell recognition schemes using image processing with parallel ELM classifications on GPU," in *Proc. 13th Int. Joint Conf. Comput. Sci. Softw. Eng. (JCSSE)*, Jul. 2016, pp. 1–6.

[30] G. Vani, R. Savitha, and N. Sundararajan, "Classification of abnormalities in digitized mammograms using extreme learning machine," in *Proc. 11th Int. Conf. Control Autom. Robot. Vis.*, Dec. 2010, pp. 2114–2117.

[31] E. Malar, A. Kandaswamy, D. Chakravarthy, and A. G. Dharan, "A novel approach for detection and classification of mammographic microcalcifications using wavelet analysis and extreme learning machine," *Comput. Biol. Med.*, vol. 42, no. 9, pp. 898–905, Sep. 2012.

[32] M. Elangeeran, S. Ramasamy, and K. Arumugam, "A novel method for benign and malignant characterization of mammographic microcalcifications employing waveatom features and circular complex valued—Extreme learning machine," in *Proc. IEEE 9th Int. Conf. Intell. Sensors, Sensor Netw. Inf. Process. (ISSNIP)*, Apr. 2014, pp. 1–6.

[33] W. Xie, Y. Li, and Y. Ma, "Breast mass classification in digital mammography based on extreme learning machine," *Neurocomputing*, vol. 173, pp. 930–941, Jan. 2016.

[34] K. Hu, W. Yang, and X. Gao, "Microcalcification diagnosis in digital mammography using extreme learning machine based on hidden Markov tree model of dual-tree complex wavelet transform," *Expert Syst. Appl.*, vol. 86, pp. 135–144, Nov. 2017.

[35] J. G. Melekoodappattu and P. S. Subbian, "A hybridized ELM for automatic micro calcification detection in mammogram images based on multi-scale features," *J. Med. Syst.*, vol. 43, no. 7, p. 183, Jul. 2019.

[36] F. Mohanty, S. Rup, B. Dash, B. Majhi, and M. N. S. Swamy, "A computer-aided diagnosis system using tchebichef features and improved grey wolf optimized extreme learning machine," *Int. J. Speech Technol.*, vol. 49, no. 3, pp. 983–1001, Mar. 2019.

[37] S.-H. Wang, K. Muhammad, P. Phillips, Z. Dong, and Y.-D. Zhang, "Ductal carcinoma *in situ* detection in breast thermography by extreme learning machine and combination of statistical measure and fractal dimension," *J. Ambient Intell. Humanized Comput.*, vol. 2017, pp. 1–11, Nov. 2017.

[38] V. Kuppili, M. Biswas, A. Sreekumar, H. S. Suri, L. Saba, D. R. Edla, R. T. Marinhoe, J. M. Sanches, and J. S. Suri, "Extreme learning machine framework for risk stratification of fatty liver disease using ultrasound tissue characterization," *J. Med. Syst.*, vol. 41, no. 10, p. 152, Oct. 2017.

[39] W. Huang, Y. Yang, Z. Lin, G.-B. Huang, J. Zhou, Y. Duan, and W. Xiong, "Random feature subspace ensemble based extreme learning machine for liver tumor detection and segmentation," in *Proc. 36th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2014, pp. 4675–4678.

[40] W. Zhu, W. Huang, Z. Lin, Y. Yang, S. Huang, and J. Zhou, "Data and feature mixed ensemble based extreme learning machine for medical object detection and segmentation," *Multimedia Tools Appl.*, vol. 75, no. 5, pp. 2815–2837, Mar. 2016.

[41] W. Huang, N. Li, Z. Lin, G.-B. Huang, W. Zong, J. Zhou, and Y. Duan, "Liver tumor detection and segmentation using kernel-based extreme learning machine," in *Proc. 35th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2013, pp. 3662–3665.

[42] B. Arunadevi and S. N. Deepa, "Brain tumor tissue categorization in 3D magnetic resonance images using improved PSO for extreme learning machine," *Prog. Electromagn. Res. B*, vol. 49, pp. 31–54, Feb. 2013.

[43] K. Kothavari, B. Arunadevi, and S. N. Deepa, "A hybrid DE-RGSO-ELM for brain tumor tissue categorization in 3D magnetic resonance images," *Math. Problems Eng.*, vol. 2014, pp. 1–14, Jan. 2014.

[44] J. Preethi, "A bio inspired hybrid krill herd-extreme learning machine network based on LBP and GLCM for brain cancer tissue taxonomy," in *Proc. 3rd Int. Conf. Comput. Intell. Appl. (ICCIA)*, Jul. 2018, pp. 140–144.

[45] X. Yang, S. Y. Yeo, S. T. Wong, G. Lee, Y. Su, J. M. Hong, A. Choo, and S. Chen, "Color-texture based extreme learning machines for tissue tumor classification," *Proc. SPIE*, vol. 9791, p. 97910W, Mar. 2016.

[46] A. Gumaei, M. M. Hassan, M. R. Hassan, A. Alelaiwi, and G. Fortino, "A hybrid feature extraction method with regularized extreme learning machine for brain tumor classification," *IEEE Access*, vol. 7, pp. 36266–36273, 2019.

[47] Z. Wang, J. Xin, P. Sun, Z. Lin, Y. Yao, and X. Gao, "Improved lung nodule diagnosis accuracy using lung CT images with uncertain class," *Comput. Methods Programs Biomed.*, vol. 162, pp. 197–209, Aug. 2018.

[48] F. R. Cordeiro, S. M. Lima, A. G. Silva-Filho, and W. Santos, "Segmentation of mammography by applying extreme learning machine in tumor detection," in *Proc. Int. Conf. Intell. Data Eng. Automated Learn.*, 2012, pp. 92–100.

[49] A. Pashaei, H. Sajedi, and N. Jazayeri, "Brain tumor classification via convolutional neural network and extreme learning machines," in *Proc. 8th Int. Conf. Comput. Knowl. Eng. (ICCKE)*, Oct. 2018, pp. 314–319.

[50] A. Doğantekin, F. Özyurt, E. Avcı, and M. Koç, "A novel approach for liver image classification: PH-C-ELM," *Measurement*, vol. 137, pp. 332–338, Apr. 2019.

[51] S. Li, H. Jiang, and W. Pang, "Joint multiple fully connected convolutional neural network with extreme learning machine for hepatocellular carcinoma nuclei grading," *Comput. Biol. Med.*, vol. 84, pp. 156–167, May 2017.

[52] P. Monkam, S. Qi, M. Xu, H. Li, F. Han, Y. Teng, and W. Qian, "Ensemble learning of multiple-view 3D-CNNs model for micro-nodules identification in CT images," *IEEE Access*, vol. 7, pp. 5564–5576, 2019.

[53] M. R. Daliri, "Combining extreme learning machines using support vector machines for breast tissue classification," *Comput. Methods Biomech. Biomed. Eng.*, vol. 18, no. 2, pp. 185–191, Jan. 2015.

[54] D. S. Nachimuthu and A. Baladhandapani, "Multidimensional texture characterization: On analysis for brain tumor tissues using MRS and MRI," *J. Digit. Imag.*, vol. 27, no. 4, pp. 496–506, Aug. 2014.

[55] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in *Proc. Eur. Conf. Comput. Vis.* Zürich, Switzerland: Springer, 2014, pp. 818–833.

[56] P. Sukumar and R. K. Gnanamurthy, "Segmentation and abnormality detection of cervical cancer cells using fast elm with particle swarm optimization," *Genetika*, vol. 47, no. 3, pp. 863–876, 2015.

[57] G. D. Tourassi, E. D. Frederick, M. K. Markey, and C. E. Floyd, "Application of the mutual information criterion for feature selection in computer-aided diagnosis," *Med. Phys.*, vol. 28, no. 12, pp. 2394–2402, Dec. 2001.

[58] B. K. Elfarra and I. S. Abuhaiba, "New feature extraction method for mammogram computer aided diagnosis," *Int. J. Signal Process., Image Process. Pattern Recognit.*, vol. 6, no. 1, pp. 13–36, 2013.

[59] C. Bauckhage and J. K. Tsotsos, "Separable linear discriminant classification," in *Proc. Joint Pattern Recognit. Symp.* Vienna, Austria: Vienna Univ. Technol., 2005, pp. 318–325.

[60] A. Daffertshofer, C. J. C. Lamoth, O. G. Meijer, and P. J. Beek, "PCA in studying coordination and variability: A tutorial," *Clin. Biomech.*, vol. 19, no. 4, pp. 415–428, May 2004.

[61] J. Xia, H. Chen, Q. Li, M. Zhou, L. Chen, Z. Cai, Y. Fang, and H. Zhou, "Ultrasound-based differentiation of malignant and benign thyroid nodules: An extreme learning machine approach," *Comput. Methods Programs Biomed.*, vol. 147, pp. 37–49, Aug. 2017.

[62] X. Jin, F. Zhu, L. Li, and Q. Xia, "Unsupervised representation learning for pulmonary nodule classification," in *Proc. 10th Int. Symp. Comput. Intell. Design (ISCID)*, vol. 1, Dec. 2017, pp. 362–365.

[63] M. R. Daliri, "A hybrid automatic system for the diagnosis of lung cancer based on genetic algorithm and fuzzy extreme learning machines," *J. Med. Syst.*, vol. 36, no. 2, pp. 1001–1005, Apr. 2012.

[64] Y. Kaya, "A new intelligent classifier for breast cancer diagnosis based on a rough set and extreme learning machine: RS+ELM," *Turkish J. Electr. Eng. Comput. Sci.*, vol. 21, no. 1, pp. 2079–2091, 2013.

[65] Doreswamy and M. U. Salma, "BAT-ELM: A bio inspired model for prediction of breast cancer data," in *Proc. Int. Conf. Appl. Theor. Comput. Commun. Technol. (iCATccT)*, Oct. 2015, pp. 501–506.

[66] A. Malik and J. Iqbal, "Extreme learning machine based approach for diagnosis and analysis of breast cancer," *J. Chin. Inst. Eng.*, vol. 39, no. 1, pp. 74–78, Jan. 2016.

[67] R. Savitha, S. Suresh, and N. Sundararajan, "Fast learning circular complex-valued extreme learning machine (CC-ELM) for real-valued classification problems," *Inf. Sci.*, vol. 187, pp. 277–290, Mar. 2012.

[68] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 42, no. 2, pp. 513–529, Apr. 2012.

[69] W.-Y. Deng, Q.-H. Zheng, L. Chen, and X.-B. Xu, "Research on extreme learning of neural networks," *Chin. J. Comput.*, vol. 33, no. 2, pp. 279–287, Apr. 2010.

[70] J. Kittler, M. Hatef, R. Duin, and J. Matas, ''On combining classifiers,'' *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 3, pp. 225–239, Mar. 1998.
[71] S. Ding, W. Jia, C. Su, L. Zhang, and L. Liu, ''Research of neural network algorithm based on factor analysis and cluster analysis,'' *Neural Comput. Appl.*, vol. 20, no. 2, pp. 297–302, Mar. 2011.
[72] S. Ding, H. Zhao, Y. Zhang, X. Xu, and R. Nie, ''Extreme learning machine: Algorithm, theory and applications,'' *Artif. Intell. Rev.*, vol. 44, no. 1, pp. 103–115, Jun. 2015.
[73] A. Toprak, ''Extreme learning machine (ELM)-based classification of benign and malignant cells in breast cancer,'' *Med. Sci. Monitor Int. Med. J. Exp. Clin. Res.*, vol. 24, p. 6537, Jul. 2018.
[74] S. M. Salaken, A. Khosravi, T. Nguyen, and S. Nahavandi, ''Extreme learning machine based transfer learning algorithms: A survey,'' *Neurocomputing*, vol. 267, pp. 516–524, Dec. 2017.
[75] G. Han, X. Liu, F. Han, I. N. T. Santika, Y. Zhao, X. Zhao, and C. Zhou, ''The LISS—A public database of common imaging signs of lung diseases for computer-aided detection and diagnosis research and medical education,'' *IEEE Trans. Biomed. Eng.*, vol. 62, no. 2, pp. 648–656, Feb. 2015.
[76] C. Prasetyo, A. Kardiana, and R. Yuliwulandari, ''Breast cancer diagnosis using artificial neural networks with extreme learning techniques,'' *Int. J. Adv. Res. Artif. Intell.*, vol. 3, no. 7, pp. 10–14, 2014.
[77] A. Ari and D. Hanbay, ''Deep learning based brain tumor classification and detection system,'' *TURKISH J. Electr. Eng. Comput. Sci.*, vol. 26, no. 5, pp. 2275–2286, Sep. 2018.
[78] J. Gu, C. Y. Fu, B. K. Ng, L. B. Liu, S. K. Lim-Tan, and C. G. L. Lee, ''Enhancement of early cervical cancer diagnosis with epithelial layer analysis of fluorescence lifetime images,'' *PLoS ONE*, vol. 10, no. 5, May 2015, Art. no. e0125706.

ZHIQIONG WANG received the M.Sc. and Ph.D. degrees in computer science and technology from Northeastern University, China, in 2008 and 2014, respectively. She visited the National University of Singapore, in 2010, and The Chinese University of Hong Kong, in 2013, as an Academic Visitor. She is currently an Associate Professor with the College of Medicine and Biological Information Engineering, Northeastern University. She has published more than 60 articles. Her main research interests include computer aided diagnosis, medicine information, big health data analysis, cloud computing, and machine learning.

YIQI LUO received the bachelor's degree in information security from Northeastern University, in 2017, where she is currently pursuing the degree with the College of Medicine and Biological Information Engineering. Her main research interests include machine learning, computer-aided diagnosis, and medical informatics.

JUNCHANG XIN (Member, IEEE) received the B.Sc., M.Sc., and Ph.D. degrees in computer science and technology from Northeastern University, China, in 2002, 2005, and 2008, respectively. He visited the National University of Singapore, as a Postdoctoral Visitor, from April 2010 to April 2011. He is currently a Professor with the School of Computer Science and Engineering, Northeastern University. He has published more than 60 research articles. His research interests include big data, uncertain data, and bioinformatics.

HAO ZHANG received the bachelor's, master's, and Ph.D. degrees in medicine from China Medical University, China, in 2008, 2010, and 2018, respectively. He is currently an Assistant Director Physician with the Department of Breast Surgery, Liaoning Cancer Hospital and Institute, Cancer Hospital of China Medical University, China. He has published more than 10 research articles, including PNAS. His research interests include combined treatment, mechanism of recurrence and metastasis, and bioinformatics of breast cancer.

LUXUAN QU received the bachelor's degree in automatic from Northeast Dianli University, in 2010, and the master's degree in biomedical engineering from Northeastern University, in 2014, where she is currently pursuing the Ph.D. degree with the College of Medicine and Biological Information Engineering. Her main research interests include gene regulatory networks, machine learning, and cloud computing.

ZHONGYANG WANG received the B.Sc. and M.Sc. degrees in biomedical engineering from Northeastern University, in 2012 and 2015, respectively, where he is currently pursuing the Ph.D. degree with the College of Medicine and Biological Information Engineering. His main research interests include machine learning, digital image processing, bioinformatics, brain functional network research, data mining, and data analysis.

YUDONG YAO (Fellow, IEEE) received the B.Eng. and M.Eng. degrees in electrical engineering from the Nanjing University of Posts and Telecommunications, Nanjing, China, in 1982 and 1985, respectively, and the Ph.D. degree in electrical engineering from Southeast University, Nanjing, in 1988. From 1989 to 1990, he was with Carleton University, Ottawa, ON, Canada, as a Research Associate in mobile radio communications. From 1990 to 1994, he was with Spar Aerospace Ltd., Montreal, QC, Canada, where he was involved in research on satellite communications. From 1994 to 2000, he was with Qualcomm Inc., San Diego, CA, USA, where he has participated in research and development in wireless code-division multiple-access (CDMA) systems. He has been with the Stevens Institute of Technology, Hoboken, NJ, USA, since 2000. He is currently a Professor and the Director of the Department of Electrical and Computer Engineering. He is also a Professor with the Medicine and Biological Information Engineering, Northeastern University, and the Director of the Stevens' Wireless Information Systems Engineering Laboratory (WISELAB). He holds one Chinese patent and 12 U.S. patents. His research interests include wireless communications and networks, spread spectrum and CDMA, antenna arrays and beamforming, cognitive and software defined radio (CSDR), and digital signal processing for wireless systems. He served as an Associate Editor for the IEEE COMMUNICATIONS LETTERS and the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. He also served as an Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.

WANCHENG ZHU received the Ph.D. degree in mining engineering from Northeastern University, China, in 2001. He was invited to the Department of Civil and Structural Engineering, Hong Kong University of Technology, for four times, for cooperative research. He was a Postdoctoral Researcher with The University of Western Australia, from 2004 to 2006. From 2006 to 2007, he was supported by the Humboldt Foundation of Germany and a Humboldt Scholar with the University of Bohong, Germany. He is currently a Professor with the College of Resources and Civil Engineering, Northeastern University. He has published more than 90 articles. His main research interests include damage and fracture of rock and concrete materials, rock fracture process under dynamic load and induced mine dynamic disaster, and multi physical field coupling effect in rock mechanics.

XINGWEI WANG received the M.Sc. and Ph.D. degrees in computer application from the Computer Department, Northeastern University, China, in 1992 and 1998, respectively. From 1996 to 1999, he was a Deputy Director with the Computer Department, School of Information, Northeastern University. From 1998 to 1999, he was with the Department of Information Engineering, Nagoya University, Japan, for inter-school communication. In 1999, he was an Assistant Dean with the College of Information, Northeastern University. He was also a Senior Visiting Scholar with the Department of Computing, Hong Kong University of Technology, from August to October 2001, and in May 2002. He has been a Principal of CERNET Northeast Network Center, since July 2002. He is currently a Professor with the College of Software, Northeastern University. He has published more than 200 articles and three academic works. His main research interests include next generation internet, self-organizing networks, IP/DWDM optical internet, mobile wireless internet, network information security, and grid computing. He is a member of the CERNET Expert Committee of China Education and Scientific Research Computer Network.