IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 11, 2018, accepted January 1, 2019, date of publication January 16, 2019, date of current version August 14, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2892795

Breast Cancer Detection Using Extreme Learning Machine Based on Feature Fusion With CNN Deep Features

ZHIQIONG WANG[1,2], MO LI[3], HUAXIA WANG[4], HANYU JIANG[4], YUDONG YAO[4], (Fellow, IEEE), HAO ZHANG[5], AND JUNCHANG XIN[3]

[1]Sino-Dutch Biomedical and Information Engineering School, Northeastern University, Shenyang 110169, China
[2]Neusoft Research of Intelligent Healthcare Technology, Co. Ltd., Shenyang 110179, China
[3]Key Laboratory of Big Data Management and Analytics (Liaoning), School of Computer Science and Engineering, Northeastern University, Shenyang 110169, China
[4]Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ 07030, USA
[5]Department of Breast Surgery, Shengjing Hospital of China Medical University, Shenyang 110004, China

Corresponding author: Junchang Xin (xinjunchang@mail.neu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61472069, Grant 61402089, and Grant U1401256, in part by the China Postdoctoral Science Foundation under Grant 2018M641705, in part by the Fundamental Research Funds for the Central Universities under Grant N161602003, Grant N161904001, and Grant N160601001, and in part by the Open Program of Neusoft Research of Intelligent Healthcare Technology, Co. Ltd. under Grant NRIHTOP1802.

ABSTRACT A computer-aided diagnosis (CAD) system based on mammograms enables early breast cancer detection, diagnosis, and treatment. However, the accuracy of the existing CAD systems remains unsatisfactory. This paper explores a breast CAD method based on feature fusion with convolutional neural network (CNN) deep features. First, we propose a mass detection method based on CNN deep features and unsupervised extreme learning machine (ELM) clustering. Second, we build a feature set fusing deep features, morphological features, texture features, and density features. Third, an ELM classifier is developed using the fused feature set to classify benign and malignant breast masses. Extensive experiments demonstrate the accuracy and efficiency of our proposed mass detection and breast cancer classification method.

INDEX TERMS Mass detection, computer-aided diagnosis, deep learning, fusion feature, extreme learning machine.

I. INTRODUCTION

Breast cancer is a serious threat to women's life and health, and the morbidity and mortality of breast cancer are ranked first and second out of all female diseases [1]. Early detection of lumps can effectively reduce the mortality rate of breast cancer [2]. The mammogram is widely used in early screening of breast cancer due to its relatively low expense and high sensitivity to minor lesions [3]. In the actual diagnosis process, however, the accuracy can be negatively affected by many factors, such as radiologist fatigue and distraction, the complexity of the breast structure, and the subtle characteristics of the early-stage disease [4], [5]. The computer-aided diagnosis (CAD) for breast cancer can help address this issue.

The classical CAD for breast cancer contains three steps: (a) finding the Region of Interest (ROI) in the preprocessed mammogram, and hence locating the region of the tumor. (b) then, extracting features of the tumor based on expert knowledge, such as shape, texture, and density, to manually generate feature vectors. (c) finally, diagnosing benign and malignant tumors by classifying these feature vectors [6], [7].

Although the classical diagnosis method has been commonly used, its accuracy still needs to be improved [8]. The quality of the handcrafted feature set directly affects the diagnostic accuracy, and hence an experienced doctor plays a very important role in the process of manual feature extraction. The commonly used features, including morphology, texture, density and other characteristics are manual set, which are obtained based on doctors' experience, that is, subjective features. In recent years, deep learning methods, such as the convolutional neural network (CNN), that can extract hierarchical features from image data without the manual selection, which is also called objective features, have been successfully applied with a great improvement on accuracies

in many applications, such as image recognition, speech recognition, and natural language processing [9], [10]. There are some shortcomings in either subjective or objective features. Subjective features ignore the essential attributes of images, while objective features ignore artificial experience. Therefore, the subjective and objective features are fused so that these features can reflect the essential properties of the image as well as the artificial experience. Meantime, Extreme Learning Machine (ELM) has better classification effect on multi-dimensional features than other classifiers including SVM, decision tree, etc., based on our previous research. Thus, we use ELM to classify the extracted breast mass features. Therefore, in this paper, we propose a novel diagnosis method that merges several deep features. The main contributions are shown as follows.

- In the detection phase, we propose a method that utilizes CNN and US-ELM for feature extraction and clustering, respectively. First, a mammogram will be segmented into several sub-regions. Then, CNN is used to extract features based on each sub-region, followed by utilizing US-ELM to cluster features of sub-regions, which eventually locate the region of breast tumor.
- In the phase of feature integration, we design an 8-layer CNN architecture and obtain 20 deep features. In addition, we integrate extra 5 shape features, 5 texture features and 7 density features of the tumor with those deep features to form a fusion deep feature set.
- In the diagnosis phase, we use the fusion deep feature set of each mammogram as an input of ELM for classification. The output directly indicates whether the patient has either a benign or a malignant breast tumor.
- Finally, the experimental results demonstrate that our proposed methods, the sub-regional US-ELM clustering and the ELM classification with fusion deep feature sets, achieve the best performance in the diagnosis of breast cancer. The experimental dataset contains 400 female mammograms.

II. RELATED WORK

The research efforts related to breast cancer CAD mainly focus on the detection and diagnosis of breast tumor. This section briefly summarizes existing works related to these two aspects. In the aspect of breast tumor detection, Sun *et al.* [11] proposed a mass detection method, where an adaptive fuzzy C-means algorithm for segmentation is employed on each mammogram of the same breast. A supervised artificial neural network is used as a classifier to judge whether the segmented area is a tumor. Saidin *et al.* [12] employed pixels as an alternative feature and used a region growing method to segment breast tumor in the mammogram. Xu *et al.* [13] proposed an improved watershed algorithm. They first make a coarse segmentation of breast tumor, followed by the image edge detection via combining regions that has similar gray-scale mean values. Hu *et al.* [14] proposed a novel algorithm to detect suspicious masses in the mammogram, in which they utilized an adaptive global and local thresholding segmentation method on the original mammogram. Yap *et al.* [15] used three different deep learning methods to detect lession in breast ultrasound images based on a Patch-based LeNet, a U-Net, and a transfer learning approach with a pretrained FCN-AlexNet, respectively.

In the aspect of differentiating benign and malignant breast tumors, Kahn, Jr., *et al.* [16] created a Bayesian network that utilizes 2 physical features and 15 manually marked probabilistic characteristics to conduct the computer-aided diagnosis for breast cancer. Wang *et al.* [17] utilized ELM to classify features of breast tumors and compare results with SVM classifier. Qiu *et al.* [18] applied CNN to the risk prediction of breast cancer by training CNN with a large amount of time series data. Sun *et al.* [19] also used deep neural networks to predict the risk of breast cancer in the near term based on 420 time series records of mammography. Jiao *et al.* [20] proposed a deep feature based framework for breast masse classification, in which CNN and a decision tree process are utilized. Arevalo *et al.* [21] used CNN to abstract representations of breast tumor and then classified the tumor as either benign or malignant. Carneiro *et al.* [22] proposed an automated mammogram analysis method based on deep learning to estimate the risk of patients of developing breast cancer. Kumar *et al.* [23] presented an image retrieval system using Zernike moments (ZMs) for extracting features since the features can affect the effectiveness and efficiency of a breast CAD system. Aličković and Subasi [24] proposed a breast CAD method, in which genetic algorithms are used for extraction of informative and significant features, and the rotation forest is used to make a decision for two different categories of subjects with or without breast cancer.

III. METHODS

In this paper, we consider the following five steps in breast cancer detection: breast image preprocessing, mass detection, feature extraction, training data generation, and classifier training. In the breast image preprocessing, denoising and enhancing contrast processes on the original mammogram have been utilized to increase the contrast between the masses and the surrounding tissues. The mass detection is then performed to localize the ROI. After that, features including deep features, morphological features, texture features and density features, are extracted from the ROI. During the training process, the classifiers have been trained with every image from the breast image dataset using their extracted features and corresponding labels. Thus, the mammogram underdiagnosis can be identified using the well-trained classifiers. Fig. 1 presents the flowchart of the entire diagnosis process.

A. BREAST IMAGE PREPROCESSING

There are several preprocessing methods in [25]–[29]. The adaptive mean filter algorithm [25] is selected to eliminate noise on the original mammograms in order to avoid the impact of noise on subsequent auxiliary diagnosis. The main idea is to use a fixed size window sliding in the line direction

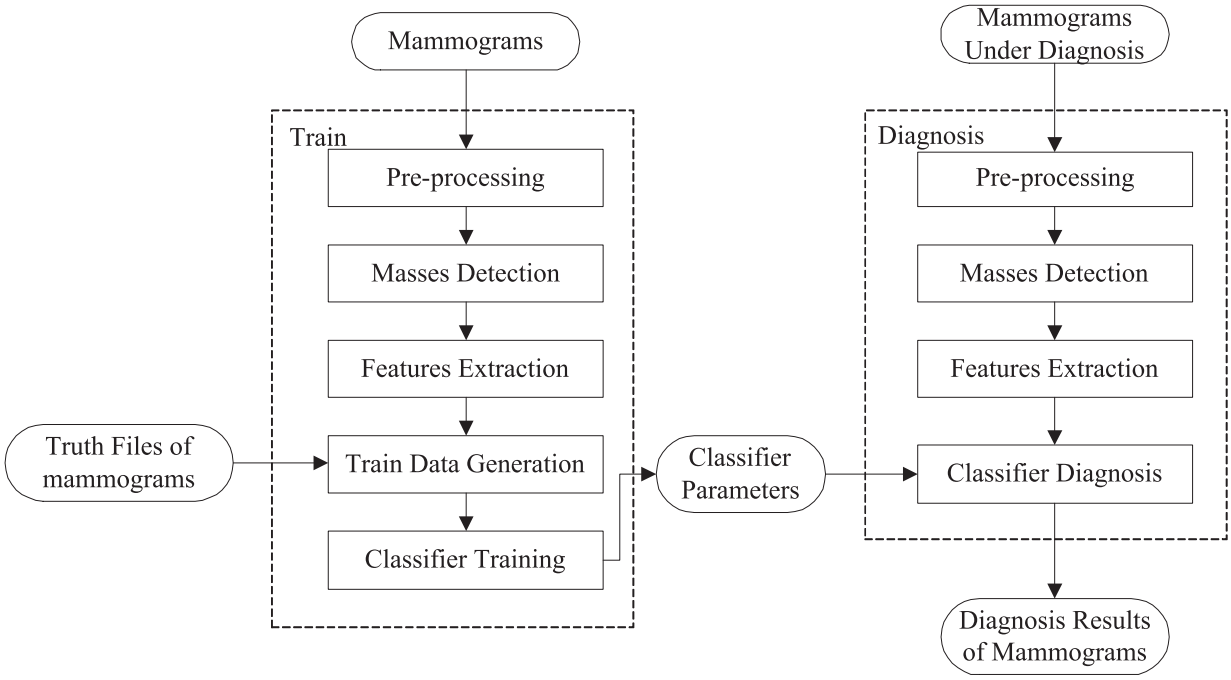


FIGURE 1. Flowchart of mass detection process.

of the image, calculating the mean, variance, and spatial correlation values of each sliding window to determine whether the window contains noise. If the noise is detected, replace the pixel values of the selected window with the mean value.

In this paper, a contrast enhancement algorithm [30] has been used to increase the contrast between the suspected masses and surrounding tissues. The main idea is to transform the histogram of the original image into uniformly distributed. After this process, the gray scale of the image is enlarged, thus the contrast has been enhanced and the image details become more clear.

Fig. 2 shows the mammogram before and after preprocessing. Fig. 2(a) is the original image, Fig. 2(b) and 2(c) are the images after denosing and enhancing contrast processes, respectively. By comparing these three images, we can see that the boundaries of the mammary and the background area in the original image (Fig. 2(a)) are often ambiguous and irregular. After preprocessing, the contrast between the mammary region and the background area has been significantly enhanced, thus reducing considerable computational burden in image post-processing.

B. MASS DETECTION

The purpose of mass detection is to extract the mass region from the normal tissues. The more precise the mass segmentation, the more accurate the extracted features. In this paper, we propose a mass detection method based on sub-domain CNN deep features and US-ELM clustering. The processing flowchart is shown in Fig. 3. The first step is to extract the ROI from the images after preprocessing. The ROI is then divided into several non-overlapping sub-regions using a sliding window. After that, we determine whether all sub-regions have been successfully traversed. If yes, extract the deep features of the sub-regions; otherwise, clustering the deep features of the sub-regions, obtain the mass area boundary and complete the mass detection process.

1) EXTRACT ROI

In mammography, there are a large number of 0 gray value areas, which have no impact on the breast CAD. In order to improve the mammary image processing efficiency and ensure the accuracy of follow-up diagnosis, it is necessary to separate the mammary area from the whole mammogram as ROI.

In this paper, an adaptive mass region detection algorithm has been utilized to extract the breast mass region. Specifically, in a mammogram, all rows are scanned sequentially to find the first nonzero pixel (with abscissa denoted as x_s) and the last nonzero pixel (with abscissa denoted as x_d), and all columns are then scanned sequentially to find the first nonzero pixel (with ordinate denoted as y_s) and the last nonzero pixel (with ordinate denoted as y_d). Algorithm 1 presents the details of this algorithm, the size of the mammography I is $m \times n$.

2) PARTITION SUB-REGION

In this section, a method to divide the ROI into several non-overlapping sub-regions is proposed. The searching area to determine the masses from the ROI is fixed in a rectangular area $[x_s, y_s, x_d, y_d]$, where the length of the search-

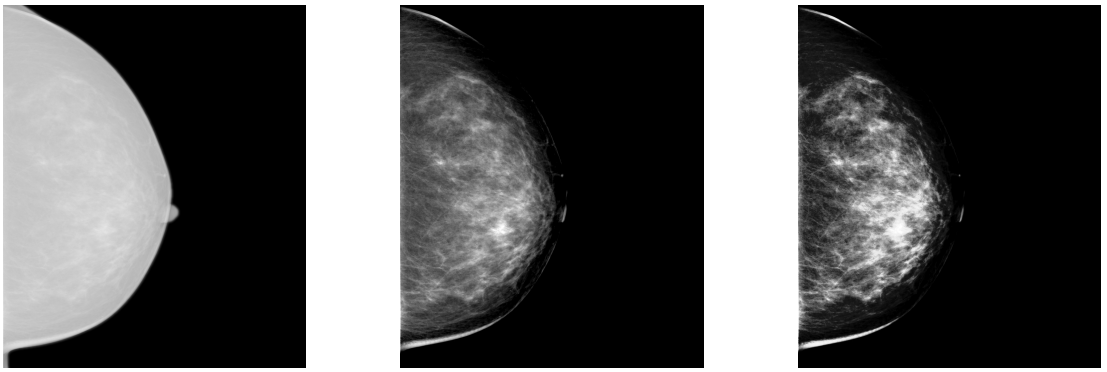


FIGURE 2. Images after denoising and enhancement. (a) Initial mammogram. (b) Denoising after (a). (c) Enhancement after (b).

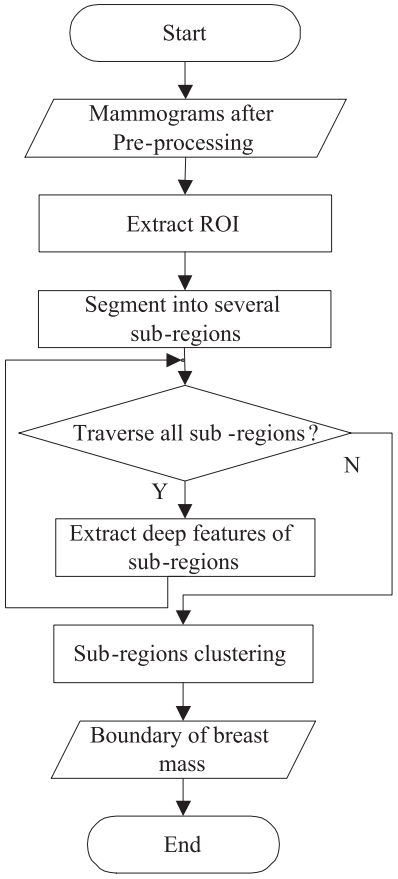


FIGURE 3. Flowchart of mass detection process.

Algorithm 1 Self-Adaptive Mass Region Detection Algorithm

1 Input: Mammography I
2 Output: Mass area M
3 for $i = 1$ to L **do**
4 find the first and the last nonzero pixel x_s, x_d.
5 end for
6 for $i = 1$ to n **do**
7 find the first and the last nonzero pixel y_s, y_d.
8 end for
9 Cut off a rectangle M by the coordinates (x_s, x_d) and (y_s, y_d).
10 Return M.

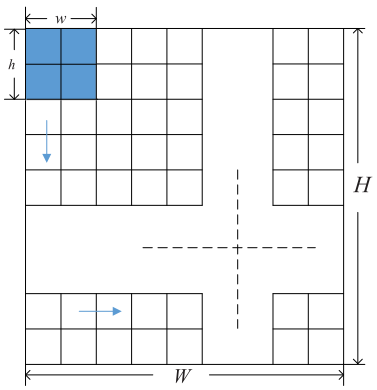


FIGURE 4. Using sliding window to divide the ROI.

ing rectangular is $W = x_d - x_s$ and width is $H = y_d - y_s$. The rectangular searching area is segmented using a sliding window with length w and width h ($W \geq w$, $H \geq h$). The segmentation procedure can be performed as follows.

First, we generate a rectangular searching area (as shown in Fig. 4). In the rectangular searching area ($W \times H$), the sliding window ($w \times h$) is moved with a certain step size, traversing the searching area without crossing the ROI boundary. Thus, the ROI is divided into several equal size ($w \times h$), non-overlapping sub-regions and such sub-regions will serve as the basis for subsequent feature extraction. In this paper, the size of the sliding window is fixed as 48×48 and the searching step size is equal to 48. Finally, the ROI has been divided into N non-overlapping sub-regions $(s_1, s_2, \cdots, s_N)$.

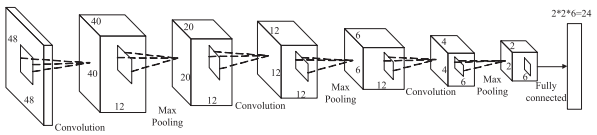


FIGURE 5. CNN architecture for mass detection.

3) EXTRACT DEEP FEATURES USING CNN

In this paper, CNN is used to extract deep features from the ROI sub-regions. Fig. 5 presents a 7-layer CNN architecture, which contains 3 convolution layers, 3 max-pooling layers, and one fully-connected layer. The input of CNN is a 48×48 dimension sub-region image captured from previous steps. The first convolution layer filters the $48 \times 48 \times 3$ input images with 12 kernels of size $9 \times 9 \times 3$ and obtains the output with size $40 \times 40 \times 12$.

$$\mathrm{Conv}^k(i,j) = \sum_{u,v} W^{k,l}(u,v) \cdot \mathrm{input}^i(i-u, j-v) + b^{k,l} \tag{1}$$

where $W^{k,l}$ represents the k^{th} kernel and $b^{k,l}$ denotes the bias of k^{th} layer. The activation value is constrained in the range $[-1, 1]$ using tanh as the activation function.

$$\mathrm{Output}^k(i,j) = \tanh\left(\mathrm{Conv}^k(i,j)\right) \tag{2}$$

The output of the first convolution layer is connected with a max-pooling layer. Then the second and third convolution/max-pooling layer are connected to one another until we have the output with size $2 \times 2 \times 6$. The fully-connected layer has $2 \times 2 \times 6 = 24$ neurons which are the features for the following clustering analysis.

4) CLUSTERING DEEP FEATURES USING US-ELM

In this paper, we use the US-ELM algorithm to cluster deep features extracted from the previous CNN architecture. The cluster number is set to 2 and sub-region features are clustered into two categories: A. suspicious mass areas; B. non-suspicious mass areas.

When the amount of training data is small, the effect of the model obtained by supervised learning cannot satisfy the demand. Therefore, semi-supervised learning is used to enhance the effect, meanwhile, it can also perform some clustering tasks [31], [32]. US-ELM algorithm is one of the semi-supervised learning algorithms and it can find out the internal structure relationships that exist among the unlabeled dataset [33]. The details of this algorithm are shown in Algorithm 2. The input of the algorithm is the deep feature matrix $\mathbf{X}$, and the output is the feature clustering results. Specifically, the Laplacian operator $\mathbf{L}$ is first constructed from the training set $\mathbf{X}$, then a hidden layer neuron output matrix is randomly generated. If the hidden neuron number is smaller than the input neuron number, we use the equation $\min_{\beta \in \mathbb{R}^{n_h \times n_o}} \| \beta \|^2 + \lambda Tr\left(\beta^T \mathbf{H}^T L \mathbf{H} \beta\right)$ to calculate output weights, where β represents the weights between hidden layer and output layer. Otherwise, we use the equation $\left(I_0 + l\mathbf{H}^T L \mathbf{H}\right) v = \gamma \mathbf{H}^T \mathbf{H} v$ to calculate output weights. After that, we calculate the embedding matrix and use k-means algorithm clustering N points into K categories.

Algorithm 2 US-ELM Algorithm

Input: Deep feature matrix $\mathbf{X} \in \mathbb{R}^{N \times n_0}$
Output: Embedding matrix $\mathbf{E} \in \mathbb{R}^{N \times n_0}$
Output: Clustering index vector $\mathbf{y} \in \mathbb{R}^{N \times 1}$
Construct the Laplacian operator $\mathbf{L}$ from the training set $\mathbf{X}$
Randomly generate hidden layer neuron output matrix $\mathbf{H} \in \mathbb{R}^{N \times n_h}$
if $n_h \leq N$ **then**
    Use equation $\min_{\beta \in \mathbb{R}^{n_h \times n_o}} \| \beta \|^2 + \lambda Tr\left(\beta^T \mathbf{H}^T L \mathbf{H} \beta\right)$ to calculate output weights
else
    Use equation $\left(I_0 + l\mathbf{H}^T L \mathbf{H}\right) v = \gamma \mathbf{H}^T \mathbf{H} v$ to calculate output weights
end if
Calculate the embedding matrix: $\mathbf{E} = \mathbf{H}\beta$
Use k-means algorithm clustering N points into K categories
Denote $\mathbf{y}$ as the dimension vector for all point classification indexes
return $\mathbf{y}$

C. CLASSIFY BENIGN AND MALIGNANT MASSES BASED ON FEATURES FUSED WITH CNN FEATURES

In this subsection, a diagnosis method using ELM classifier with fusion deep features is proposed. The main idea is to extract the deep features using CNN, and also extract the morphological, texture and density characteristics from the breast mass area. Then use the ELM classifier to classify the fusion features and obtain the benign and malignant diagnosis results.

1) FEATURE MODELING

In clinical, breast masses are common signs of early breast disease, according to the pathological characteristics of the masses, where the masses are divided into two categories: malignant and benign. On one hand, CNN extracts the deep features of the masses, which can represent the essential properties of the masses. On the other hand, based on the doctors' experience, malignant masses in mammography are often have the following characteristics: the irregular shape with a burr-like edge, unsmooth surface with hard nodules, and the densities are significantly different with surrounding tissues. While the benign masses often have regular shape with the clear edge, smooth surface, rarely accompanied with small nodules, and the densities are uniformly distributed. The types of the extracted features used in this paper are listed in Table 1.

The fusion features can be modeled as

$$F = [F_1, F_2, F_3, F_4] \tag{3}$$

TABLE 1. Types of extracted features.

Types	Features
Deep features	20 deep features extracted from CNN
Morphological features	Roundness, normalized radius entropy, normalized radius variance, acreage ratio, roughness
Texture features	Inverse difference moment, entropy, energy, correlation and contrast coefficient
Density features	Density mean, density variance, density skew, density peak gray density variance, gray density skew, gray density peak

where F_1 denotes deep features, F_2 denotes morphological features, F_3 represents texture features, and F_4 represents density features.

a: DEEP FEATURES

CNN has great advantages in feature extraction due to its convolution-pooling operations. It can extract the essential image characteristics without human participation [34]. A 10-layer CNN architecture used in feature extraction is shown in Fig. 6. The CNN input is the suspicious mass area. After a series of convolution/max-pooling layers, the last fully-connected layer has $2 \times 2 \times 5 = 20$ neurons which are the deep features denoted as $F_1 = [c_1, c_2, \cdots, c_{20}]$.

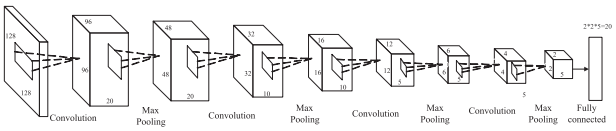


FIGURE 6. CNN architecture for mass diagnosis.

b: MORPHOLOGICAL FEATURES

According to experienced doctors, malignant breast masses often have irregular shapes and blur boundaries with surrounding tissues [35]. Morphological features are important indicators to distinguish between benign and malignant masses. In this paper, we have extracted mass roundness, normalized radius entropy, normalized radius variance, acreage ratio, and roughness as the morphological features and the corresponding model can be expressed as $F_2 = [g_1, g_2, g_3, g_4, g_5]$. Table 2 illustrates the detailed equations to calculate each morphological feature.

c: TEXTURE FEATURES

Texture features are effective parameters that reflect the benign and malignant characteristics of the breast masses, and contribute to the early diagnosis of breast cancer. The gray-level co-occurrence matrix which was first introduced by Haralick is a classic gray-scale texture feature. The gray-level co-occurrence matrix describes the gray distribution over all image pixels, which is based on the second-order joint condition probability [36]. In this paper, we have extracted inverse moment, entropy, energy, correlation and contrast coefficient as the texture features and the corresponding model can be expressed as $F_3 = [t_1, t_2, t_3, t_4, t_5]$. Table 3 illustrates the equations to calculate each texture feature.

d: DENSITY FEATURES

Recent studies have shown that the density features of breast masses have a significant correlation with mass benign and malignant characteristics. Using density features to predict the breast cancer is also a common method [37]. In this paper, we have extracted seven density features and the corresponding model can be expressed as $F_4 = [d_1, d_2, d_3, d_4, d_5, d_6, d_7]$. Table 4 presents the equations to calculate each density feature.

2) CLASSIFIER

The ELM algorithm is a single hidden layer feed-forward neural network proposed by Huang *et al.* [38], which has a good generalization performance and fast learning speed, and insensitive to manual parameters setup. In this paper, we use ELM as the classifier to obtain breast cancer benign and malignant diagnosis results.

The entire ELM algorithm consists of training and testing processes, and detailed algorithm training steps are shown in Algorithm 3. First, randomly generate weights $\mathbf{w}_i$ and bias b_i of the hidden layer. Then calculate the single hidden layer output matrix $\mathbf{H}$ based on parameters $\mathbf{w}_i$, b_i, and fusion features F. After that, obtain the output weight vector β based on the training data labels.

Algorithm 3 ELM Training Algorithm

1 Input: $N = \{(\mathbf{x}_j, \mathbf{t}_j) \,|\mathbf{x}_j \in \mathbb{R}^n, \mathbf{t}_j \in \mathbb{R}^m, j = 1, 2, \cdots, N\}$
L: number of hidden neurons; N: labeled dataset size
2 Output: Three parameters of ELM: $\mathbf{w}$, b, β
3 for $i = 1$ to L **do**
4 randomly generate weights $\mathbf{w}_i$ and bias b_i of the hidden layer
5 end for
6 Calculate the single hidden layer output matrix $\mathbf{H}$
7 Calculate the output weight vector $\beta = \mathbf{H}^T\mathbf{T}$
8 Return $\mathbf{w}$, b, β

Algorithm 4 shows the testing steps of ELM algorithm. First, calculate the single hidden layer output matrix $\mathbf{H}$ based on parameters $\mathbf{w}_i, b_i, \beta$. Then extract the fusion features of the testing image as the algorithm input, obtaining breast cancer diagnosis results R.

Algorithm 4 ELM Testing Algorithm

1 Input: $F, N, L, \mathbf{w}, b, \beta$
2 Output: R: diagnosis results
3 Calculate the single hidden layer output matrix $\mathbf{H}$
4 Calculate the diagnosis results $R = f(\mathbf{x})$
5 Return R

TABLE 2. Morphological features.

Features	Calculation Equations	Parameter Details
Roundness	$g_1 = \frac{p^2}{A}$	A: Masses acreage; P: Masses edge circumference
Normalized radius entropy	$g_2 = -\sum_{k=1}^{100} p_k \left(\log\left(p_k\right)\right)$	p_k: The probability that the normalized edge point falls in the k^{th} interval
Normalized radius variance	$g_3 = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left(d_i - d_{\text{avg}}\right)^2}$	N: The number of all edge points; d_i: Normalized radius of the i^{th} edge point.; d_{avg}: Edge points average standardized radius
Acreage ratio	$g_4 = \frac{1}{d_{\text{avg}} N}\sum_{i=1}^{N}\left(d\left(i\right) - d_{\text{avg}}\right)$	
Roughness	$g_5 = \frac{1}{N}\sum_{i=1}^{N}\left\|d_i - d_{i+1}\right\|$	

TABLE 3. Texture features.

Features	Calculation Equations	Parameter Details
Inverse moment	$t_1 = \sum \frac{P(i,j)}{1+(i-j)^2}$	$P\left(i,j\right)$: $(i,j)^{\text{th}}$ element of gray covariance matrix
Entropy	$t_2 = \sum P\left(i,j\right)\cdot\left(-\ln P\left(i,j\right)\right)$	
Energy	$t_3 = \sum P\left(i,j\right)^2$	
Correlation coefficient	$t_4 = \sum \frac{P(i,j)\cdot(i-\mu_z)\cdot\left(j-\mu_y\right)}{\delta_x \delta_y}$	$\mu_x, \mu_y, \delta_x, \delta_y$ denote the mean value and variance of P
Contrast coefficient	$t_5 = \left(i-j\right)^2 \cdot P\left(i,j\right)$	

TABLE 4. Density features.

Features	Calculation Equations	Parameter Details
Histogram mean	$d_1 = \sum_{i=0}^{M-1} z_i p\left(z_i\right)$	M: Gray level of the pixel; $p\left(z_i\right)$: The proportion of pixel with gray level equals to i
Histogram variance	$d_2 = \sum_{i=0}^{M-1}\left(z_i - n\right)^2 p\left(z_i\right)$	n: The gray scale of the pixel
Histogram skew	$d_3 = \frac{1}{d_1^{3/2}}\sum_{i=0}^{M-1}\left(z_i - n\right)^3 p\left(z_i\right)$	
Histogram peak	$d_4 = \frac{1}{d_1^2}\sum_{i=0}^{M-1}\left(z_i - n\right)^4 p\left(z_i\right)$	
Histogram gray density variance	$d_5 = \sum_{i=0}^{M-1}\left(z_i - p_n\right)^3 p\left(z_i\right)$	p_n: The gray scale density of the pixel
Histogram gray density skew	$d_6 = \frac{1}{d_1^{3/2}}\sum_{i=0}^{M-1}\left(z_i - p_n\right)^3 p\left(z_i\right)$	
Histogram gray density peak	$d_7 = \frac{1}{d_1^2}\sum_{i=0}^{M-1}\left(z_i - p_n\right)^4 p\left(z_i\right)$	

IV. EXPERIMENT

In this section, the effectiveness of the mass detection method based on the CNN and US-ELM algorithms, and breast cancer diagnostic method based on the fusion features are investigated. First, the parameter setup and experiment procedures are introduced. Then, the evaluation metrics of each experiment are described. Finally, the experimental results of each experiment are listed and analyzed.

A. EXPERIMENT DATA

In this paper, there are 400 mammograms in the image dataset, which contains 200 malignant mass images and 200 benign mass images. These images are generated using the Senographe 2000D all-digital mammography camera from 32 to 74 years old female patients. The location of the masses of all images have been marked by the professional doctor, and the diagnosis results are also confirmed by the pathologist.

B. EXPERIMENT PROCEDURES AND PARAMETERS

In the following experiments, the effectiveness of the mass detection method based on the CNN and US-ELM algorithms, and breast cancer diagnostic method based on the fusion features are investigated based on the above experiment dataset. The experimental procedures and parameters used in the verification process are as follows.

1) MASS DETECTION EXPERIMENT

In the mass detection process, the feature extraction algorithms used in the experiment are CNN, DBN, and SAE. The clustering algorithms are US-ELM and k-means. The names of the experiments are simplified, and the combination of the feature extraction method and the clustering method used in each experiment is shown in Table 5.

TABLE 5. Mass detection experiment procedures.

Experiment Name	CNN	DBN	SAE	US-ELM	k-means
CNN-US	√			√	
CNN-K	√				√
DBN-US		√		√	
DBN-K		√			√
SAE-US			√	√	
SAE-K			√		√

The marker-controlled watershed algorithm (MCWA) [13] and adaptive thresholding algorithm (ATA) [14] have been used for comparison and to further verify the accuracy of the proposed method.

In the diagnosis of breast masses, the CNN architecture parameters are shown in Fig. 5. Other parameters include the activation function f and the learning rate η. In our experiment, the learning rate is equal to 0.003 and the activation function is tanh. When using DBN and SAE for feature extraction, we use the same activation function and learning rate. When using ELM for clustering, the parameters involved are the activation function and the number of hidden layer neurons. In the experiment, the selected activation function f is ''sigmoid'', and the number of hidden layer neurons is set to $L = 1000$. Since the density feature is binary clustered based on US-ELM, the k value is set to 2. When clustering is performed using k-means, the parameter k involved in the experiment is the same as the US-ELM algorithm, which is equal to 2.

2) BREAST CANCER DIAGNOSIS EXPERIMENT

In the process of breast-assisted diagnosis, the deep, morphological, texture and density features of suspected masses were extracted and used for fusion feature modeling. In the experiment, the deep feature model is mainly divided into single deep feature, double features and multi-features. Single deep feature models (SF) contain the features extracted from CNN, DBN, and SAE algorithms only. Double features (DF) models contain deep features and one of the morphological, texture and density features, i.e., ''deep feature + morphological feature (CNN-G)''. Multi-features models (MF) contain deep features and two or more other features, i.e., ''deep feature + morphological feature + texture feature (CNN-GTD)''. To further evaluate our proposed method, we also select the state-of-art algorithm mentioned in [24] as the baseline, which is can be simplified as ''GARF''. Detailed feature combinations and selections are shown in Table 6.

In the diagnosis of breast masses, the CNN architecture parameters are shown in Fig. 6. Other parameters include the activation function f and the learning rate η. In our experiment, the learning rate is equal to 0.003 and the activation function is tanh. When using DBN and SAE for feature extraction, we use the same activation function and learning rate. When using ELM for clustering, the parameters involved are the activation function and the number of hidden layer neurons. In the experiment, the selected activation function f is ''sigmoid'', and the number of hidden layer neurons is set to $L = 1000$. When considering the SVM classification technique, the parameters involved are kernel function R, penalty coefficient c, and kernel function parameter g. RBF is selected as kernel function R, $c = 0.5$, and $g = 0.0206$.

C. EVALUATION METRICS

In the above experimental scheme, the mass detection method based on subdomain CNN deep feature through US-ELM clustering and the mass diagnosis method based on deep feature fusion are investigated respectively. The quantitative evaluation metrics of the experimental results are described as follows.

1) DETECTION METRICS

In this paper, Misclassified Error (ME), Area Overlap Metric (AOM), Area Over-segmentation Metric (AVM), Area Under-segmentation Metric (AUM), and Comprehensive Measure (CM) are used to evaluate the accuracy of mass segmentation in the step of breast mass detection. The detailed calculation formula of each evaluation metrics is shown in Table 7. A smaller value of ME, AVM and AUM, and larger value of AOM and CM, correspond to a better segmentation result. However, in practical applications, AVM and AUM often can not be optimal at the same time. ME and CM as comprehensive measurement parameters are often more important in the measurement process.

2) DIAGNOSIS METRICS

Accuracy, sensitivity, specificity, TPRatio, TNRatio, and area under ROC curve (AUC) are used to compare and analyze the results of masses. The evaluation formula of the evaluation metrics is shown in Table 8, and the explanation of the evaluation metrics parameters in Table 9 are presented in Table 13. Among the above six metrics, the higher the value of accuracy, sensitivity, specificity, TP Ratio, TN Ratio and AUC, the more accurate the diagnosis will be. The k-fold cross validation method [39] has been used to make the evaluation metrics more general. In this paper, the above evaluation indices are derived from the 5-folded cross validation method.

D. EXPERIMENT RESULTS AND ANALYSIS

In this section, we verify that the mass detection method based on subdomain CNN deep feature through US-ELM clustering and the mass diagnosis method based on deep feature fusion are superior to other detection and diagnosis methods considering the above evaluation metrics.

TABLE 6. Breast cancer diagnosis experiment feature selection.

Feature Types	Feature Models	DBN	SAE	CNN	Morphological Feature (G)	Texture Feature (T)	Density Feature (D)
Single feature	CNN			√			
	DBN	√					
	SAE		√				
Double feature	CNN-G			√	√		
	CNN-T			√		√	
	CNN-D			√			√
Multiple feature	CNN-GT			√	√	√	
	CNN-GD			√	√	√	
	CNN-TD			√		√	√
	CNN-GTD			√	√	√	√

TABLE 7. Detection indicators.

Indicators	Calculation Equations	Parameter Details
ME	$\mathrm{ME}=\frac{\mathrm{Area}\{S_A\cup S_B\}-\mathrm{Area}\{S_A\cap S_B\}}{\mathrm{Area}\{S_B\}}$	S_A: masses area segemented from algorithm
AOM	$\mathrm{AOM}=\frac{\mathrm{Area}\{S_A\}\cap\mathrm{Area}\{S_B\}}{\mathrm{Area}\{S_A\}\cup\mathrm{Area}\{S_B\}}$	S_B: masses area from standard
AUM	$\mathrm{AUM}=\frac{\mathrm{Area}\{S_B\}-\mathrm{Area}\{S_A\}}{\mathrm{Area}\{S_B\}}$	
AVM	$\mathrm{AVM}=\frac{\mathrm{Area}\{S_A\}-\mathrm{Area}\{S_B\}}{\mathrm{Area}\{S_A\}}$	
comprehensive measure	$\mathrm{CM}=\alpha\mathrm{AOM}+\beta\left(1-\mathrm{AUM}\right)+\lambda\left(1-\mathrm{AVM}\right)$	α,β,λ are weights

TABLE 8. Evaluation indices.

Indices	Calculation Equations
Accuracy	(TP + TN) / (TP + TN + FP + FN)
Sensitivity	(TP) / (TP + FN)
Specificity	(TN) / (TN + FP)
TP Ratio	(TP) / (TP + FP)
TN Ratio	(TN) / (TN + FN)
AUC	$\mathrm{AUC}=\sum_{i=1}^{N}\left(y_i+y_{i-1}\right)\times\left(x_{i-1}-x_i\right)/2$

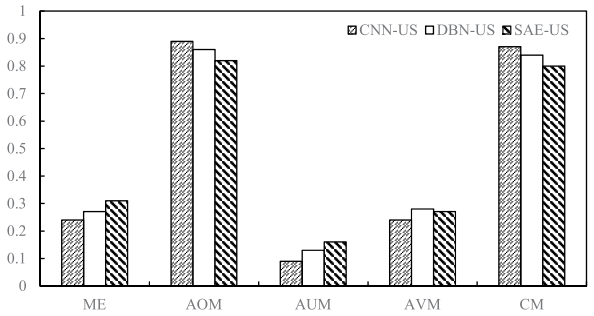


FIGURE 7. Collation map of evaluation indices of experiments.

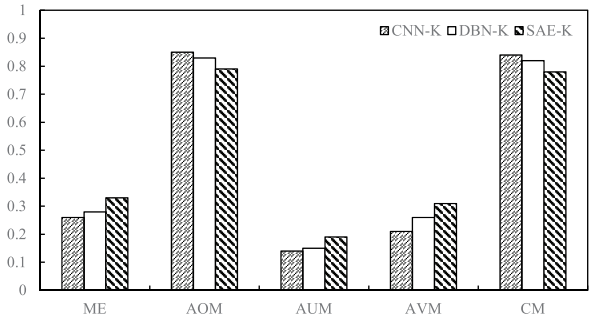


FIGURE 8. Collation map of evaluation indices of experiments.

The experimental results and corresponding analysis are presented as follows.

1) DETECTION RESULTS ANALYSIS

Based on Table 5, ME, AOM, AVM, AUM and CM are compared using three feature models and two clustering algorithms. The comparison results are shown in Fig. 7 and 8. From those figures, we can see that, the mass detection method based on CNN sub-region deep feature clustering gives the lowest ME regardless of which feature model is used, and the mass detection method based on US-ELM clustering gives the lowest ME and the best detection result. In summary, the mass detection method based on sub-domain CNN deep feature through US-ELM clustering can achieve the best mass detection effect.

Fig. 9 shows the performance difference among the proposed method and the MCWA, ATA algorithms on the five metrics (ME, AOM, AUM, AVM and CM). From Fig. 9 we can see that, although the proposed method in the AVM evaluation is slightly worse than other algorithms, the ME and CM evaluation performance have significant advantages. Therefore, the proposed method based on sub-domain CNN deep feature througth US-ELM clustering is better than other segmentation methods.

2) MASS DETECTION BASED ON FUSION FEATURE

According to the experimental scheme described in Section III.C.2, the accuracy, sensitivity, specificity,

TABLE 9. Parameters meaning in Table 12.

Parameter	Parameter Meaning
TP	True Positive: the true value is vicious and the predicted value is also vicious
TN	True Negative: the true value is benign and the predictive value is also benign
FN	False Negative: the true value is vicious and the predicted value is benign
FP	False Positive: he true value is benign and the predicted value is vicious
x_i	The abscissa of the i^{th} point in the ROC curve
y_i	The ordinate of the i^{th} point in the ROC curve

TABLE 10. Breast cancer diagnosis experiment feature selection.

Classifier	Feature Types	Feature Models	Accuracy	Sensitivity	Specificity	Benign Accuracy	Malignant Accuracy	AUC
ELM	SF	CNN	76.25	76.38	76.12	76.50	76.00	0.828
		DBN	73.25	70.85	76.27	79.00	67.50	0.804
		SAE	70.50	67.69	72.99	76.50	63.50	0.738
	DF	CNN-G	79.50	77.31	82.07	83.50	75.50	0.861
		CNN-T	80.75	80.00	81.53	82.00	79.50	0.862
		CNN-D	78.50	77.67	79.38	80.00	77.00	0.844
	MF	CNN-GT	84.50	83.17	85.94	86.50	82.50	0.881
		CNN-GD	82.50	81.86	83.16	83.50	81.50	0.879
		CNN-TD	83.25	83.08	83.42	83.50	83.00	0.889
		CNN-GTD	86.50	85.10	88.02	88.50	84.50	0.923
SVM	SF	CNN	74.50	73.33	75.79	77.00	72.00	0.808
		DBN	71.75	70.23	73.51	75.50	68.00	0.782
		SAE	68.25	66.52	70.39	73.50	63.00	0.732
	DF	CNN-G	77.25	74.01	81.50	84.00	70.50	0.836
		CNN-T	78.25	74.67	83.04	85.50	71.00	0.859
		CNN-D	75.50	71.98	80.36	83.50	67.50	0.825
	MF	CNN-GT	81.75	79.81	83.96	85.00	78.50	0.869
		CNN-GD	80.00	76.79	84.09	86.00	74.00	0.865
		CNN-TD	80.75	79.43	82.20	83.00	78.50	0.849
		CNN-GTD	83.75	81.40	86.49	87.50	80.00	0.911
	Baseline	GARF	80.75	80.47	81.05	82.43	78.97	0.800

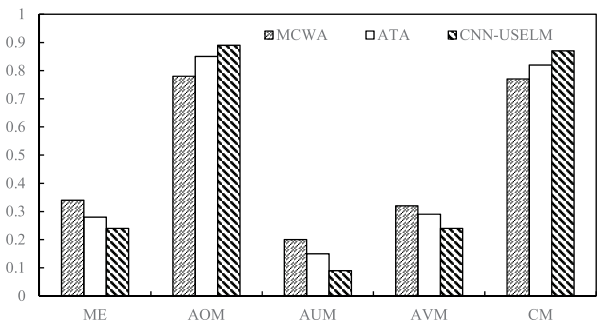


FIGURE 9. Collation map of evaluation indices of experiments.

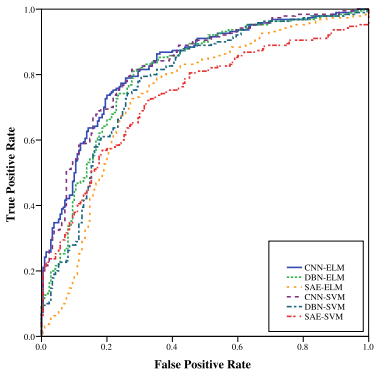


FIGURE 10. ROC Curve of single feature with ELM and SVM classifier.

TP Ratio, TN Ratio and AUC of the benign and malignant breast mass classification are analyzed based on three types of deep feature models (10 feature sets) and two classifiers. We also show the results of the method mentioned in [24] using our datasets. The evaluation results are shown in Table 10, and the ROC curve are shown in Fig. 10-12.

When the classifier is ELM, the accuracy, sensitivity and specificity of the CNN feature model are the best in the single deep feature model, which shows that the CNN model chosen in this paper is the most suitable method. In the double feature model, the CNN deep feature model combined with texture feature has the best performance in the diagnosis accuracy, sensitivity and specificity.

When the classifier is SVM, the accuracy, sensitivity and specificity of the CNN feature model are the best in the single deep feature model, which shows that the CNN model chosen in this paper is the most suitable method. In the double feature

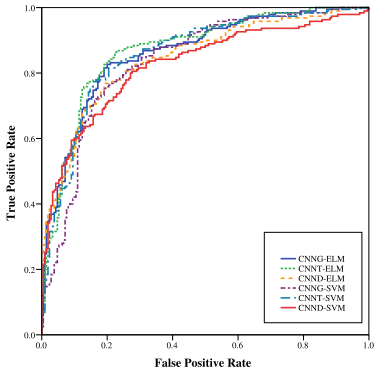


FIGURE 11. ROC Curve of double feature with ELM and SVM classifier.

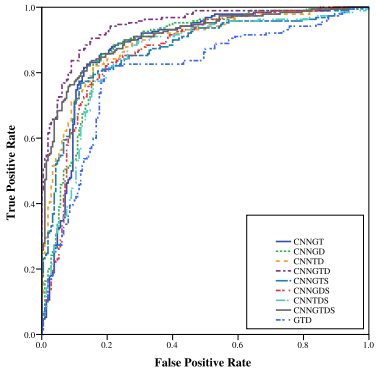


FIGURE 12. ROC Curve of multiple feature with ELM and SVM classifier.

model, the CNN deep feature model combined with texture feature has the best performance in the diagnosis accuracy, sensitivity and specificity.

According to the deep fusion model, comparing the analysis of each evaluation metrics obtained by using ELM and SVM classifier for benign and malignant tumor classification, it can be seen that ELM classifier gives better diagnostic accuracy, sensitivity, and specificity. Meantime, it is obvious that the mass diagnosis method based on deep fusion feature are better than GARF [24] in the diagnosis accuracy, sensitivity and specificity. Thus it is desirable to combine the deep features with ELM in the diagnosis of breast cancer.

V. CONCLUSION

This paper proposes a breast CAD method based on fusion deep features. Its main idea is to apply deep features extracted from CNN to the two stages of mass detection and mass diagnosis. In the stage of mass detection, a method based on sub-domain CNN deep features and US-ELM clustering is developed. In the stage of mass diagnosis, an ELM classifier is utilized to classify the benign and malignant breast masses using a fused feature set, fusing deep features, morphological features, texture features, and density features. In the process of breast CAD, the choice of features is the key in determining the accuracy of diagnosis. In previous studies, either traditional subjective features or objective features are used, in which traditional subjective features include morphology, texture, density, etc., and objective features include features extracted from CNN or DBN. These features are flawed to some extent. In this paper we combine subjective and objective features, taking the doctor's experience and the essential attributes of the mammogram into account at the same time. After extracting the features, the classifier is used to classify the benign and malignant of the breast mass. In this paper, ELM, which has a better effect on multi-dimensional feature classification, is selected as the classifier. Through the experiments using breast CAD of 400 cases of female mammograms in the northeastern China, it demonstrates that, in mass detection and mass diagnosis, our proposed methods outperform other existing methods.

REFERENCES

[1] R. L. Siegel *et al.*, "Colorectal cancer statistics," *CA, Cancer J. Clinicians*, vol. 64, no. 2, pp. 104–117, Mar. 2014
[2] J. B. Harford, "Breast-cancer early detection in low-income and middle-income countries: Do what you can versus one size fits all," *Lancet Oncol.*, vol. 12, no. 3, pp. 306–312, Mar. 2011.
[3] C. Lerman *et al.*, "Mammography adherence and psychological distress among women at risk for breast cancer," *J. Nat. Cancer Inst.*, vol. 85, no. 13, pp. 1074–1080, Jul. 1993.
[4] P. T. Huynh, A. M. Jarolimek, and S. Daye, "The false-negative mammogram," *Radiographics*, vol. 18, no. 5, pp. 1137–1154, Sep. 1998.
[5] M. G. Ertosun and D. L. Rubin, "Probabilistic visual search for masses within mammography images using deep learning," in *Proc. IEEE Int. Conf. BioInform. Biomed. (BIBM)*, Nov. 2015, pp. 1310–1315.
[6] S. D. Tzikopoulos, M. E. Mavroforakis, H. V. Georgiou, N. Dimitropoulos, and S. Theodoridis, "A fully automated scheme for mammographic segmentation and classification based on breast density and asymmetry," *Comput. Methods Programs Biomed.*, vol. 102, no. 1, pp. 47–63, 2011.
[7] D. C. Pereira, R. P. Ramos, and M. Z. do Nascimento, "Segmentation and detection of breast cancer in mammograms combining wavelet analysis and genetic algorithm," *Comput. Methods Programs Biomed.*, vol. 114, no. 1, pp. 88–101, Apr. 2014.
[8] S. A. Taghanaki, J. Kawahara, B. Miles, and G. Hamarneh, "Pareto-optimal multi-objective dimensionality reduction deep auto-encoder for mammography classification," *Comput. Methods Programs Biomed.*, vol. 145, pp. 85–93, Jul. 2017.
[9] X.-W. Chen and X. Lin, "Big data deep learning: Challenges and perspectives," *IEEE Access*, vol. 2, pp. 514–525, 2014.
[10] K. Ganesan, U. R. Acharya, C. K. Chua, L. C. Min, K. T. Abraham, and K. H. Ng, "Computer-aided breast cancer detection using mammograms: A review," *IEEE Rev. Biomed. Eng.*, vol. 6, pp. 77–98, 2012.
[11] X. Sun, W. Qian, and D. Song, "Ipsilateral-mammogram computer-aided detection of breast cancer," *Comput. Med. Imag. Graph.*, vol. 28, no. 3, pp. 151–158, Apr. 2004.
[12] N. Saidin, U. K. Ngah, H. A. M. Sakim, N. S. Ding, M. K. Hoe, and I. L. Shuaib, "Density based breast segmentation for mammograms using graph cut and seed based region growing techniques," in *Proc. 22nd Int. Conf. Comput. Res. Develop.*, 2010.
[13] S. Xu, H. Liu, and E. Song, "Marker-controlled watershed for lesion segmentation in mammograms," *J. Digit. Imag.*, vol. 24, no. 5, pp. 754–763, Oct. 2011.
[14] K. Hu, X. Gao, and F. Li, "Detection of suspicious lesions by adaptive thresholding based on multiresolution analysis in mammograms," *IEEE Trans. Instrum. Meas.*, vol. 60, no. 2, pp. 462–472, Feb. 2011.
[15] M. H. Yap *et al.*, "Automated breast ultrasound lesions detection using convolutional neural networks," *IEEE J. Biomed. Health Inform.*, vol. 22, no. 4, pp. 1218–1226, Jul. 2017.
[16] K. C. Khan, Jr, L. M. Roberts, K. A. Shaffer, and P. Haddawy, "Construction of a Bayesian network for mammographic diagnosis of breast cancer," *Comput. Biol. Med.*, vol. 27, no. 1, pp. 19–29, Jan. 1997.
[17] 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, no. 5, pp. 175–184, Mar. 2014.

[18] Y. Qiu *et al.*, "An initial investigation on developing a new method to predict short-term breast cancer risk based on deep learning technology," in *Proc. SPIE*, vol. 9785, p. 978521, Mar. 2016.
[19] W. Sun, T. L. Tseng, B. Zheng, and W. Qian, "A preliminary study on breast cancer risk analysis using deep neural network," in *Proc. Int. Workshop Breast Imag.* Malmö, Sweden: Springer, Jun. 2016, pp. 385–391.
[20] Z. Jiao, X. Gao, Y. Wang, and J. Li, "A deep feature based framework for breast masses classification," *Neurocomputing*, vol. 197, pp. 221–231, Jul. 2016.
[21] J. Arevalo, F. A. González, R. Ramos-Pollán, J. L. Oliveira, and M. A. G. Lopez, "Representation learning for mammography mass lesion classification with convolutional neural networks," *Comput. Methods Program. Biomed.*, vol. 127, pp. 248–257, Apr. 2016.
[22] G. Carneiro, J. Nascimento, and A. P. Bradley, "Automated analysis of unregistered multi-view mammograms with deep learning," *IEEE Trans. Med. Imag.*, vol. 36, no. 11, pp. 2355–2365, Nov. 2017.
[23] Y. Kumar, A. Aggarwal, S. Tiwari, and K. Singh, "An efficient and robust approach for biomedical image retrieval using Zernike moments," *Biomed. Signal Process. Control*, vol. 39, pp. 459–473, Jan. 2018.
[24] E. Aličković and A. Subasi, "Breast cancer diagnosis using GA feature selection and rotation forest," *Neural Comput. Appl.*, vol. 28, no. 4, pp. 753–763, Apr. 2017.
[25] H. D. Cheng, J. Shan, W. Ju, Y. Guo, and L. Zhang, "Automated breast cancer detection and classification using ultrasound images: A survey," *Pattern Recognit.*, vol. 43, no. 1, pp. 299–317, Oct. 2010.
[26] Y. Pathak, K. V. Arya, and S. Tiwari, "Low-dose ct image reconstruction using gain intervention-based dictionary learning," *Modern Phys. Lett. B*, vol. 32, no. 14, p. 1850148, May 2018.
[27] S. Tiwari, "A variational framework for low-dose sinogram restoration," *Int. J. Biomed. Eng. Technol.*, vol. 24, no. 4, pp. 356–367, 2017.
[28] Z. Gao *et al.*, "Motion tracking of the carotid artery wall from ultrasound image sequences: A nonlinear state-space approach," *IEEE Trans. Med. Imag.*, vol. 37, no. 1, pp. 273–283, Jan. 2018.
[29] Z. Gao *et al.*, "Robust estimation of carotid artery wall motion using the elasticity-based state-space approach," *Med. Image Anal.*, vol. 37, pp. 1–21, Apr. 2017.
[30] H. Ibrahim and N. S. P. Kong, "Brightness preserving dynamic histogram equalization for image contrast enhancement," *IEEE Trans. Consum. Electron.*, vol. 53, no. 4, pp. 1752–1758, Nov. 2007.
[31] Q. Miao, R. Liu, P. Zhao, Y. Li, and E. Sun, "A semi-supervised image classification model based on improved ensemble projection algorithm," *IEEE Access*, vol. 6, pp. 1372–1379, 2018.
[32] H. Gan, Z. Li, Y. Fan, and Z. Luo, "Dual learning-based safe semi-supervised learning," *IEEE Access*, vol. 6, pp. 2615–2621, 2017.
[33] G. Huang, S. Song, J. N. D. Gupta, and C. Wu, "Semi-supervised and unsupervised extreme learning machines," *IEEE Trans. Cybern.*, vol. 44, no. 12, pp. 2405–2417, Dec. 2014.
[34] R. K. Samala, H.-P. Chan, L. Hadjiiski, M. A. Helvie, J. Wei, and K. Cha, "Mass detection in digital breast tomosynthesis: Deep convolutional neural network with transfer learning from mammography," *Med. Phys.*, vol. 43, no. 12, pp. 6654–6666, Dec. 2016.
[35] J. Tang, R. M. Rangayyan, J. Xu, I. E. Naqa, and Y. Yang, "Computer-aided detection and diagnosis of breast cancer with mammography: Recent advances," *IEEE Trans. Inf. Technol. Biomed.*, vol. 13, no. 2, pp. 236–251, Mar. 2009.
[36] G. D. Tourassi, B. Harrawood, S. Singh, J. Y. Lo, and C. E. Floyd, "Evaluation of information-theoretic similarity measures for content-based retrieval and detection of masses in mammograms," *Med. Phys.*, vol. 34, no. 1, pp. 140–150, Jan. 2007.
[37] L. Liu, J. Wang, and K. He, "Breast density classification using histogram moments of multiple resolution mammograms," in *Proc. Int. Conf. Biomed. Eng. Informat.*, Oct. 2010, pp. 146–149.
[38] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, nos. 1–3, pp. 489–501, 2006.
[39] Q. Dai, "A competitive ensemble pruning approach based on cross-validation technique," *Knowl.-Based Syst.*, vol. 37, no. 2, pp. 394–414, Jan. 2013.

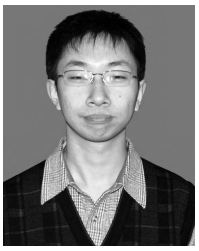
ZHIQIONG WANG received the M.Sc. and Ph.D. degrees in computer science and technology from Northeastern University, China, in 2008 and 2014, respectively. She visited the National University of Singapore, in 2010, and The Chinese University of Hong Kong, in 2013, as an Academic Visitor. She is currently an Associate Professor with the Sino-Dutch Biomedical and Information Engineering School, Northeastern University. She has published more than 50 papers. Her main research interests include biomedical, biological data processing, cloud computing, and machine learning.

MO LI received the B.E. degree from the College of Information and Computer Engineering, Northeast Forestry University, in 2014, and the M.E. degree from the Sino-Dutch Biomedical and Information Engineering School, Northeastern University, in 2017. She is currently pursuing the Ph.D. degree with the School of Computer Science and Engineering, Northeastern University. Her main research interests include bioinformatics, machine learning, and big data management.

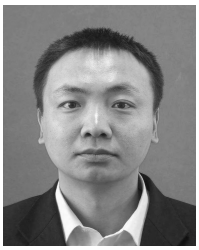
HUAXIA WANG received the B.Eng. degree in information engineering from Southeast University, Nanjing, China, in 2012. He is currently pursuing the Ph.D. degree with the Electrical and Computer Engineering Department, Stevens Institute of Technology, NJ, USA. In 2016, he was a Research Intern with the Mathematics of Networks and Systems Research Department, Nokia Bell Labs, Murray Hill, NJ, USA. His current research interests include wireless communications, cognitive radio networks, and machine learning.

HANYU JIANG received the B.S. degree in control science and engineering from the Harbin Institute of Technology, Harbin, China, in 2012, and the M.Eng. degree in computer engineering from the Stevens Institute of Technology, Hoboken, NJ, USA, in 2014, where he is currently pursuing the Ph.D. degree. He has been a Research Co-Op with Nokia Bell Labs, Murray Hill, NJ, USA, since 2017. His current research interests include heterogeneous and parallel computing, multi-/many-core processor architecture, bioinformatics, and artificial intelligence.

YUDONG YAO (S'88–M'88–SM'94–F'11) 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 and 1990, he was a Research Associate with Carleton University, Ottawa, Canada, focusing on mobile radio communications. From 1990 to 1994, he was with Spar Aerospace Ltd., Montreal, 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 participated in the research and development of wireless code-division multiple-access (CDMA) systems. Since 2000, he has been with the Stevens Institute of Technology, Hoboken, NJ, USA, and is currently a Professor and the Department Director of electrical and computer engineering. He is also a Professor with the Sino-Dutch Biomedical and Information Engineering School, Northeastern University, and the Director of the Stevens' Wireless Information Systems Engineering Laboratory. 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, and digital signal processing for wireless systems. He was an Associate Editor of IEEE COMMUNICATIONS LETTERS and the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, and an Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.



HAO ZHANG received the B.Sc., M.Sc., and Ph.D degrees in medicine from China Medical University, Shenyang, China, in 2008, 2010, and 2018, respectively. He is currently an Assistant Director Physician with the Department of Breast Surgery, Shengjing Hospital of China Medical University, Shenyang, China. He has published more than ten research papers as the First Author, including PNAS. His research interests include the combined treatment, the mechanism of recurrence and metastasis, and bioinformatics of breast cancer.



JUNCHANG XIN 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. From 2010 to 2011, he visited the National University of Singapore as a Postdoctoral Visitor. He is currently a Professor with the School of Computer Science and Engineering, Northeastern University. He has published more than 60 research papers. His research interests include big data, uncertain data, and bioinformatics.