Neurocomputing 128 (2014) 175–184

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Breast tumor detection in digital mammography based on extreme learning machine

Zhiqiong Wang [a,b], Ge Yu [b], Yan Kang [a,*], Yingjie Zhao [c], Qixun Qu [a]

[a] Sino-Dutch Biomedical and Information Engineering School, Northeastern University, Shenyang 110819, Liaoning Province, China
[b] College of Information Science and Engineering, Northeastern University, Shenyang 110819, Liaoning Province, China
[c] Medical Imaging Department, Tumor Hospital of Liaoning Province, Shenyang 110042, Liaoning Province, China

ARTICLE INFO

Article history:
Received 11 September 2012
Received in revised form
17 May 2013
Accepted 20 May 2013
Available online 31 October 2013

Keywords:
Extreme learning machine
Breast tumor detection
Mammography
Image segmentation
Feature extraction

ABSTRACT

Breast tumor detection in digital mammography is one of the most important methods of breast cancer prevention. Computer-aided diagnosis (CAD) based on extreme learning machine (ELM) has significant meanings for breast tumor detection as it has good generalization abilities and a high learning efficiency. In this paper, a breast tumor detection algorithm in digital mammography based on ELM is proposed. First, a median filter is used for noise reduction, and contrast enhancement of the digital mammography in data preprocessing. Next, methods of wavelet modulus maxima transform, morphological operation and region growth are used for the breast tumor edge segmentation. Then, five textural features and five morphological features are extracted. Finally, an ELM classifier is used to detect the breast tumor. Comparing breast tumor detection based on Support Vector Machines (SVM), with breast tumor detection based on ELM, not only does ELM have a better classification accuracy than SVM, but it also has a greatly improved training speed.

1. Introduction

Breast cancer has the priority position among cancers for women and it seriously affects the fitness of women. A breast tumor examination can help to detect the tumor in the early stages, which is meaningful for the cure. The medical image examination in digital mammography is the most effective method of the breast cancer detection. However, image examination largely depends on experience, and tumors can be easily overlooked if radiologists are tired while reading. Computer-aided diagnosis (CAD) on breast tumor detection will provide a confirmation for the radiologists in detecting the suspicious regions in images and also improve accuracy and efficiency [1]. A CAD of breast tumor detection based on machine learning turns out to be an efficient method for tumor diagnosis.

Due to excellent generalization performance, rapid training speed and little human intervention, extreme learning machine (ELM) [2–8] is attracting increased attention from more and more researchers [9]. ELM and its variants have been extensively used in many fields, such as text classification, image recognition, mobile object management and bioinformatics [10–14].

ELM and similar algorithms have been widely used in CAD, and have obtained decent accuracy and efficiency [15–18]. Savitha et al. [15] presents a Circular Complex-valued ELM classifier, which attained wonderful accuracy on smaller sample of breast tumor images. Malar et al. [16] studied micro calcification detection in mammograms based on textural features by using ELM, and found that ELM with wavelet texture features could achieve the best classification accuracy. Cordeiro et al. [17] researched locating regions of a tumor by ELM in tumor detection, and found that ELM has a faster learning time with a higher training and testing accuracy than MLP network. Gomathi and Thangaraj [18] researched a computer-aided lung cancer nodules detection system, which is able to detect the false positive nodules accurately using CT images by ELM classifier. But systematic research focused on breast tumor preprocessing, segmentation, feature model building and tumor detection with ELM in a large scale data set has not been found in the existing literature.

In this paper, we present a novel breast tumor detection method based on ELM. After doing the noise reduction and the enhancement in digital mammography, we segment the edge of the tumor. The features of the tumor are extracted, and input into ELM, as well as the judgment of radiologists which acts as the standard indicative of the classification, then the training of ELM is finished. As a result, we can detect the existence of a tumor in the images and assist radiologists in examining images. The contributions of this paper can be summarized as follows.

(1) The wavelet transformation of the local modulus maxima algorithm is applied in digital mammography. Improvement

* Corresponding author. Tel.: +86 24 83676663.
E-mail address: kangyan@bmie.neu.edu.cn (Y. Kang).

0925-2312/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.neucom.2013.05.053

of the accuracy for tumor's segmentation can be achieved by highlighting the tumor region in the images.

(2) Building a two-dimensional vector model to represent the features extracted from the tumors. The vector model contains five geometrical features and five textural features.
(3) Detecting breast tumors in digital mammography based on ELM. At first, training the ELM using the feature vector as the input and the judgment of radiologists as the output. And then detecting the tumor with imported test data .
(4) 482 digital mammography of the Chinese women are the data that was used in our experiment. The experimental results show that ELM has advantages in training speed and detection accuracy over SVM.

The remainder of this paper is organized as follows. Section 2 briefly introduces ELM. Section 3 introduces the details of breast tumor detection based on ELM, which includes four steps. In Section 4, we analyze the performance of breast tumor detection based on ELM. And the conclusions are presented in Section 5.

2. Extreme learning machine

ELM [3,4] is a single hidden-layer feed-forward neural networks (SLFNs) learning algorithm and can be extended to the generalized SLFNs where the hidden layer need not be neuron like [5,6]. It first randomly assigns the input weights and hidden layer biases, and then analytically determines the output weights of SLFNs. ELM can achieve better generalization performance than other conventional learning algorithms at an extremely fast learning speed. Besides, ELM is also less sensitive to user-specified parameters, and can be deployed faster and more conveniently.

For N arbitrary distinct samples $(\mathbf{x}_j, \mathbf{t}_j)$, where $\mathbf{x}_j = [x_{j1}, x_{j2}, \ldots, x_{jn}]^T \in \mathbb{R}^n$ and $\mathbf{t}_j = [t_{j1}, t_{j2}, \ldots, t_{jm}]^T \in \mathbb{R}^m$, the standard SLFNs with L hidden nodes and activation function $g(x)$ are mathematically modeled as

$$\sum_{i=1}^{L} \boldsymbol{\beta}_i g_i(\mathbf{x}_j) = \sum_{i=1}^{L} \boldsymbol{\beta}_i g(\mathbf{w}_i \cdot \mathbf{x}_j + b_i) = \mathbf{o}_j \quad (j = 1, 2, \ldots, N) \tag{1}$$

where $\mathbf{w}_i = [w_{i1}, w_{i2}, \ldots, w_{in}]^T$ is the weight vector connecting the ith hidden node and the input nodes, $\boldsymbol{\beta}_i = [\beta_{i1}, \beta_{i2}, \ldots, \beta_{im}]^T$ is the weight vector connecting the ith hidden node and the output nodes, b_i is the threshold of the ith hidden node, and $\mathbf{o}_j = [o_{j1}, o_{j2}, \ldots, o_{jm}]^T$ is the jth output vector of the SLFNs.

The standard SLFNs with L hidden nodes and activation function $g(x)$ can approximate these N samples with zero error. It means $\sum_{j=1}^{L} ||\mathbf{o}_j - \mathbf{t}_j|| = 0$ and there exist $\boldsymbol{\beta}_i$, $\mathbf{w}_i$ and b_i such that

$$\sum_{i=1}^{L} \boldsymbol{\beta}_i g(\mathbf{w}_i \cdot \mathbf{x}_j + b_i) = \mathbf{t}_j \quad (j = 1, 2, \ldots, N) \tag{2}$$

The equation above can be expressed compactly as

$$\mathbf{H}\boldsymbol{\beta} = \mathbf{T} \tag{3}$$

where

$$\mathbf{H}(\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_L, b_1, b_2, \ldots, b_L, \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_L) = \begin{bmatrix} g(\mathbf{w}_1 \cdot \mathbf{x}_1 + b_1) & g(\mathbf{w}_2 \cdot \mathbf{x}_1 + b_2) & \ldots & g(\mathbf{w}_L \cdot \mathbf{x}_1 + b_L) \\ g(\mathbf{w}_1 \cdot \mathbf{x}_2 + b_1) & g(\mathbf{w}_2 \cdot \mathbf{x}_2 + b_2) & \ldots & g(\mathbf{w}_L \cdot \mathbf{x}_2 + b_L) \\ \vdots & \vdots & \vdots & \vdots \\ g(\mathbf{w}_1 \cdot \mathbf{x}_N + b_1) & g(\mathbf{w}_2 \cdot \mathbf{x}_N + b_2) & \ldots & g(\mathbf{w}_L \cdot \mathbf{x}_N + b_L) \end{bmatrix}_{N \times L} \tag{4}$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_{11} & \beta_{12} & \ldots & \beta_{1m} \\ \beta_{21} & \beta_{22} & \ldots & \beta_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ \beta_{L1} & \beta_{L2} & \ldots & \beta_{Lm} \end{bmatrix}_{L \times m} \quad \text{and} \quad \mathbf{T} = \begin{bmatrix} t_{11} & t_{12} & \ldots & t_{1m} \\ t_{21} & t_{22} & \ldots & t_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ t_{N1} & t_{N2} & \ldots & t_{Nm} \end{bmatrix}_{N \times m} \tag{5}$$

$\mathbf{H}$ is called the hidden layer output matrix of the neural network and the ith column of $\mathbf{H}$ is the ith hidden node output with respect to inputs $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N$. The smallest norm least-squares solution of the above linear system is

$$\hat{\boldsymbol{\beta}} = \mathbf{H}^{\dagger}\mathbf{T} \tag{6}$$

where $\mathbf{H}^{\dagger}$ is the Moore–Penrose generalized inverse of matrix $\mathbf{H}$. Then the output function of ELM can be modeled as

$$f(\mathbf{x}) = \mathbf{h}(\mathbf{x})\boldsymbol{\beta} = \mathbf{h}(\mathbf{x})\mathbf{H}^{\dagger}\mathbf{T} \tag{7}$$

There are three important input parameters in the ELM training. They are training set $\mathcal{N} = \{(\mathbf{x}_j, \mathbf{t}_j) | \mathbf{x}_j = [x_{j1}, x_{j2}, \ldots, x_{jn}]^T \in \mathbb{R}^n, \mathbf{t}_j = [t_{j1}, t_{j2}, \ldots, t_{jm}]^T \in \mathbb{R}^m, j = 1, 2, \ldots, N$, the hidden node output function $g(\mathbf{w}_i, b_i, \mathbf{x}_j)$ and the hidden node number L. Only after the related parameters are set properly, ELM can start its training process. First, ELM randomly generates L pairs of hidden node parameters $(\mathbf{w}_i, b_i)$. Then, according to the input and randomly generated parameters, it calculates the hidden layer output matrix $\mathbf{H}$ by using Eq. (4). Finally, utilizing Eq. (6), it calculates the corresponding output weight vector $\boldsymbol{\beta}$. After completing the above training process, the classification result of the new data set can be predicted by ELM according to Eq. (7).

3. Breast tumor detection based on ELM

3.1. Framework overview

As shown in Fig. 1, there are 4 steps of tumor detection in digital mammography: image preprocessing, edge segmentation, feature extraction and breast tumor detection based on ELM. Image preprocessing includes two steps, noise reduction and enhancement of image. After that, the succeeding procedure to

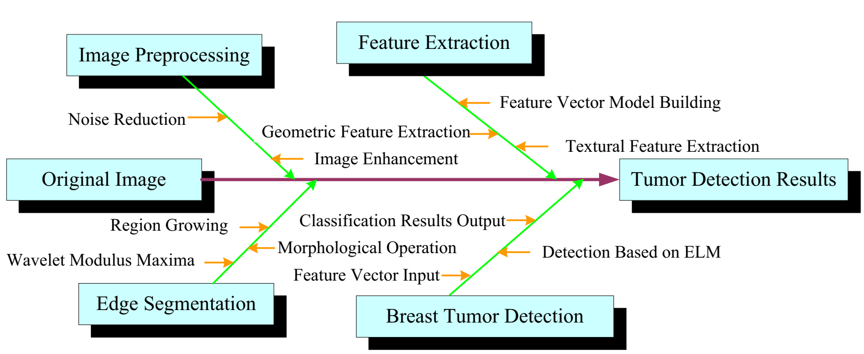

Fig. 1. The framework overview of breast tumor detection.