Computers in Biology and Medicine 43 (2013) 765–774

Contents lists available at SciVerse ScienceDirect

Computers in Biology and Medicine

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

Mammographical mass detection and classification using Local Seed Region Growing–Spherical Wavelet Transform (LSRG–SWT) hybrid scheme

Pelin Görgel [a,*], Ahmet Sertbas [a], Osman N. Ucan [b]

[a] Department of Computer Engineering, Faculty of Engineering, Istanbul University (IU), Istanbul, Turkey
[b] Department of Electrical Engineering, Faculty of Engineering, Istanbul Aydin University (IAU), Istanbul, Turkey

ARTICLE INFO

Article history:
Received 13 October 2010
Accepted 19 March 2013

Keywords:
Breast cancer
Image enhancement
Tumor classification
Spherical Wavelet Transform (SWT)
Homomorphic filtering
Local Seed Region Growing (LSRG)
Support Vector Machines (SVM)

ABSTRACT

The purpose of this study is to implement accurate methods of detection and classification of benign and malignant breast masses in mammograms. Our new proposed method, which can be used as a diagnostic tool, is denoted Local Seed Region Growing–Spherical Wavelet Transform (LSRG–SWT), and consists of four steps. The first step is homomorphic filtering for enhancement, and the second is detection of the region of interests (ROIs) using a Local Seed Region Growing (LSRG) algorithm, which we developed. The third step incoporates Spherical Wavelet Transform (SWT) and feature extraction. Finally the fourth step is classification, which consists of two sequential components: the 1st classification distinguishes the ROIs as either mass or non-mass and the 2nd classification distinguishes the masses as either benign or malignant using a Support Vector Machine (SVM). The mammograms used in this study were acquired from the hospital of Istanbul University (I.U.) in Turkey and the Mammographic Image Analysis Society (MIAS). The results demonstrate that the proposed scheme LSRG–SWT achieves 96% and 93.59% accuracy in mass/non-mass classification (1st component) and benign/malignant classification (2nd component) respectively when using the I.U. database with k-fold cross validation. The system achieves 94% and 91.67% accuracy in mass/non-mass classification and benign/malignant classification respectively when using the I.U. database as a training set and the MIAS database as a test set with external validation.

1. Introduction

Among various cancers, breast cancer places at the top in women, both in the developed and the developing countries. There is a parallel increase in the incidence of this disease with life expectancy and urbanization [1,2]. Previously, the most effective way to be able to survive breast cancer is detecting it in an early phase. The significance of mammography is to reduce deaths from breast cancer by early detection of masses. Although this technology has been developing, it remains difficult in some cases to interpret a dense mammogram, including some suspicious region of interest (ROIs). Whether the radiologist is not experienced enough or the contrast is inadequate, unnecessary biopsy tests are performed against the possibility of breast cancer. As biopsy tests are expensive and invasive, computer aided methods, which help to detect true positive masses (TPs) and eliminate false positives (FPs), have to be developed. Such methods have recently achieved adequate performance in assisting radiologists to make a malignant/benign decision by providing a "second eye" for breast cancer diagnosis.

As wavelets present an efficient decomposition in signals and images, several wavelet-based studies have been developed related to mammographical mass detection and classification in recent years [1–4]. However there are less studies about spherical wavelet and curvelet transforms because these methods are new in the literature. Karahaliou et al. [3] investigate clusters of microcalcifications with their texture properties. Three level multi-resolution decomposition is implemented using Laws' exture energy measures, first order statistics and cooccurrence matrices features to extract ROIs from the surrounding tissue. Their system, which uses a probabilistic neural network, produces 86% accuracy rate in classifying the masses as normal, benign or malignant. In the study of Angelini et al. [4] the system classifies the ROIs as either mass or non-mass. A pixel-based, Discrete Wavelet Transform-based (DWT) and Overcomplete Wavelet Transform-based (OWT) image representations are applied to an SVM system subsequently. The best results are obtained by DWT and OWT representations. Hwang et al. [5] extract mammographic image texture features using a Haar wavelet transform.

* Corresponding author. Tel.: +90 212 473 7070.
E-mail addresses: paras@istanbul.edu.tr (P. Görgel), asertbas@istanbul.edu.tr (A. Sertbas), uosman@aydin.edu.tr (O.N. Ucan).

0010-4825/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.compbiomed.2013.03.008

They use neural networks, statistical discriminant analysis and SVM for classification and their system achieves 88% accuracy.

Curvelets represent the discontinuities through edges or curves in objects or images efficiently. Some studies performing curvelet transforms in image processing are as follows. Ali et al. [6] implement a curvelet transform approach to computed tomography (CT) images. Their system achieves satisfying results for the fusion of magnetic resonance. In the study of Binh and Thanh [7], a curvelet transform-based method is developed for object detection in speckled images. The constructed segmentation method presents a sparse expansion for typical smooth-contoured images.

In recent years Buciu and Gacsadi [8] present a study, in which the mammograms are filtered with Gabor wavelets, and directional features at different orientation and frequencies are extracted. Principal component analysis (PCA) is implemented to reduce the high dimension of filtered and unfiltered data and an SVM is used to classify the data. They achieve 97.56% sensitivity and 78.26% specificity. Tahmasbi et al. [9] present a study aiming to reduce the false negative rate by using Zernike moments as shape descriptors and margin characteristics. Two groups of the moments are extracted from the pre-processed mammograms. The moments that are the most effective ones are chosen and a backpropagation multilayer perceptron is used for classification, which performs at a 92.8% accuracy rate.

This paper presents a computer-aided diagnosis system including mammographic image enhancement, segmentation and diagnosis stages via filtering, mass detection and classification. SWT, which fits the geometric structure of spherical breast masses, helps to optimize a multiresolution transform prior to feature extraction. This study uses two different databases to indicate the superiority of SWT over DWT and the last scale coefficients over all coefficients. The new proposed method in this paper is based on a four-stage algorithm: enhancement with homomorphic filtering; segmentation with Local Seed Region Growing (LSRG); feature extraction with Spherical Wavelet Transform (SWT) and finally classification the ROIs and masses with SVM. The proposed system that we have called LSRG–SWT can be helpful to extract specific characteristics from raw data and provide true interpretation to radiologists.

The remainder of this paper is organized as follows. Section 2 gives a brief introduction to homomorphic filtering, Wavelet Transform, LSRG–SWT and SVM methods. Section 3 discusses the experimental work while Section 4 presents the results and Section 5 includes the conclusion.

2. Methodology

In this study the diagnosis task begins with contrast enhancement as seen in Fig. 1. First, we enhance the images by using homomorphic filtering and in this way local contrast is improved. Next, the suspicious regions such as masses are extracted using the proposed LSRG algorithm proposed by adding some local rules and descriptions to a Seed Region Growing algorithm. The detected ROIs are not always true positive masses, some of them are non-mass breast tissue and relatively brighter than the surrounding tissue. To prevent the increment of false positives and improve true positive detection, a Spherical Wavelet Transform is implemented prior to feature extraction. Each detected ROI is passed from a five-level SWT as the optimum results are achieved with a two-level DWT having six coefficients (approximation$_2$ (a_2), horizantal$_2$ (h_2), vertical$_2$ (v_2), diagonal$_2$ (d_2), approximation$_1$ (a_1) and the mean of (h_1, v_1, d_1)). To generate six coefficients ($w_1, w_2, w_3, w_4, w_5, c_5$) in SWT as well, the decomposition should contain five levels. Moreover, according to some previous studies in the literature [10] a five-level SWT performs better performance.

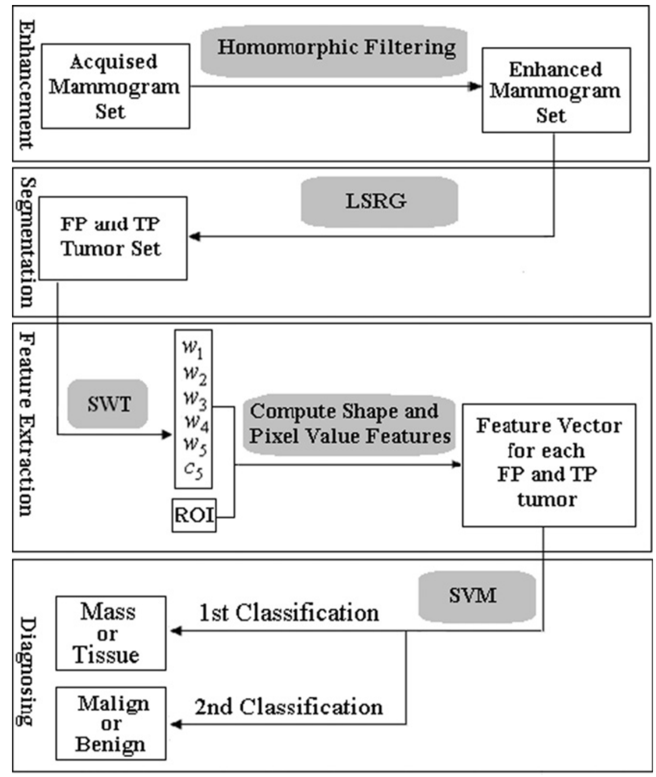

Fig. 1. The flow chart of LSRG–SWT method.

Each ROI is represented both with its own and SWT coefficients' shape and gray level-based feature matrices. 1st classification determines whether the ROI is mass or non-mass and the 2nd classification determines whether the mass is benign or malignant, which provides the breast cancer diagnosis. The software is developed with MATLAB Version 7.6 and the feature matrices are given to the SVM using WEKA 3.7.1.

2.1. Enhancement using homomorphic filtering

For correct segmentation and diagnosis mammogram contrast enhancement is implemented using homomorphic filtering,, which provides a good deal of control over the components of illumination and reflectance. This control requires the specification of a filter function $H(u,v)$ that affects the low and high frequency components of Fourier transform differently. An image $f(x,y)$ can be expressed as the product of illumination $i(x,y)$ and reflectance $r(x,y)$ components [11]:

$$f(x,y) = i(x,y)r(x,y) \tag{1}$$

and we define:

$$z(x,y) = \ln f(x,y) = \ln i(x,y) + \ln r(x,y), \tag{2}$$

$$Z(u,v) = F_i(u,v) + F_r(u,v) \tag{3}$$

where $Z(u,v)$, $F_i(u,v)$ and $F_r(u,v)$ demonstrate the Fourier transforms of $z(x,y)$, $\ln i(x,y)$ and $\ln r(x,y)$ respectively. If it is processed by means of a $H(u,v)$ filter function, $S(u,v)$ is yielded:

$$S(u,v) = H(u,v)Z(u,v) = H(u,v)F_i(u,v) + H(u,v)Fr(u,v) \tag{4}$$

so $s(x,y)$ is the inverse Fourier transform of $S(u,v)$ and can be expressed in the form:

$$s(x,y) = i'(x,y) + r'(x,y) \tag{5}$$

finally, the desired enhanced image is obtained as seen in Eq. (6).

$$g(x,y) = e^{s(x,y)} = e^{i'(x,y)} \times e^{r'(x,y)} = i_E(x,y)r_E(x,y) \tag{6}$$