COMPUTER METHODS AND PROGRAMS IN BIOMEDICINE 114 (2014) 88–101

journal homepage: www.intl.elsevierhealth.com/journals/cmpb

Segmentation and detection of breast cancer in mammograms combining wavelet analysis and genetic algorithm

Danilo Cesar Pereira[a,1], *Rodrigo Pereira Ramos*[b,*], *Marcelo Zanchetta do Nascimento*[c,1]

[a] *Centro de Matemática, Computação e Cognição, Universidade Federal do ABC, Rua Santa Adélia, 166, 09210170 Santo André, SP, Brazil*
[b] *Colegiado de Engenharia Elétrica, Universidade Federal do Vale do São Francisco, Avenida Antônio Carlos Magalhães, 510, 48902300 Juazeiro, BA, Brazil*
[c] *Faculty of Computation (FACOM), Federal University of Uberlândia (UFU), Av. João Naves de Ávila, 2121, 38400902 Uberlândia, MG, Brazil*

ARTICLE INFO

Article history:
Received 6 February 2013
Received in revised form
9 January 2014
Accepted 14 January 2014

Keywords:
Wavelet transform
Genetic algorithm
Mammographic images
Image segmentation
Computer-aided systems

ABSTRACT

In Brazil, the National Cancer Institute (INCA) reports more than 50,000 new cases of the disease, with risk of 51 cases per 100,000 women. Radiographic images obtained from mammography equipments are one of the most frequently used techniques for helping in early diagnosis. Due to factors related to cost and professional experience, in the last two decades computer systems to support detection (Computer-Aided Detection – CADe) and diagnosis (Computer-Aided Diagnosis – CADx) have been developed in order to assist experts in detection of abnormalities in their initial stages. Despite the large number of researches on CADe and CADx systems, there is still a need for improved computerized methods. Nowadays, there is a growing concern with the sensitivity and reliability of abnormalities diagnosis in both views of breast mammographic images, namely cranio-caudal (CC) and medio-lateral oblique (MLO). This paper presents a set of computational tools to aid segmentation and detection of mammograms that contained mass or masses in CC and MLO views. An artifact removal algorithm is first implemented followed by an image denoising and gray-level enhancement method based on wavelet transform and Wiener filter. Finally, a method for detection and segmentation of masses using multiple thresholding, wavelet transform and genetic algorithm is employed in mammograms which were randomly selected from the Digital Database for Screening Mammography (DDSM). The developed computer method was quantitatively evaluated using the area overlap metric (AOM). The mean ± standard deviation value of AOM for the proposed method was 79.2 ± 8%. The experiments demonstrate that the proposed method has a strong potential to be used as the basis for mammogram mass segmentation in CC and MLO views. Another important aspect is that the method overcomes the limitation of analyzing only CC and MLO views.

* *Corresponding author*. Tel.: +55 74 2102 7630; fax: +55 74 2102 7630.
E-mail addresses: danilo.pereira@ufabc.edu.br (D.C. Pereira), rodrigo.ramos@univasf.edu.br, godoga@gmail.com (R.P. Ramos), marcelo.zanchetta@gmail.com, nascimento@facom.ufu.br (M.Z. do Nascimento).
[1] Tel.: +55 34 3239 4571; fax: +55 34 3239 4571.

http://dx.doi.org/10.1016/j.cmpb.2014.01.014

1. Introduction

Over the past few years, the cancer has been one of the most responsible for the high number of deaths, and could become one of the main responsible for most deaths in the next decades. According to the World Health Organization, the number of deaths due to cancer, which was just 13% in 2008, is currently having a significant increase and one estimates that this number could reach approximately 12 million until 2030 [1].

Breast cancer is the second-most common and leading cause of cancer death among women. In Brazil, the National Cancer Institute (INCA) reports more than 50,000 new cases of the disease, with risk of 51 cases per 100,000 women [2]. Since it still has unknown causes, a large technology investment and a large number of human resources have to be available for early detection in order to reduce the mortality rate of the patients.

The screen-film mammography is an important tool used by radiologists to detect cancer at an early stage. In this examination, four images are obtained, two corresponding to the right breast and two to the left breast of the projections cranio-caudal (CC) and medio-lateral oblique (MLO). The use of CC and MLO images improves visualization of breast tissue and increases the chances of detecting the presence of non-palpable breast cancer. During the examination, the radiologist combines information from these two views to increase the chances of determining a priori regions with abnormalities defined as true positive (TP) and reduce the number of regions without abnormalities, i.e., reduce false positive (FP) regions [3]. However, screen-film mammography is a repetitive task, making radiologists prone to oversight errors. As a result, radiologists fail to detect from 10% up to 30% of malignant lesion on mammograms [4].

Recently, Computer-Aided Detection (CADe) and Computer-Aided Diagnosis (CADx) have been applied to mammographic images to assist radiologists on lesions analysis such as microcalcification, mass and architectural distortions [5]. CADe schemes automatically detect and segment suspicious lesions in mammograms, i.e., perform a localization task. CADx systems extend the computer analysis to characterize suspicious regions or estimate the probability of malignancy of a lesion, and are focused on the classification task. For a survey of the state of the art of CADx systems, the reader is referred to a recent review article [6]. The present work is focused on an automatic segmentation system of suspicious lesions for mammographic images.

In spite of the relevant research contributions for the area, CADe systems exhibit different performances on detecting and/or segmenting mass lesions or microcalcification clusters. Specifically for microcalcifications, recent papers show some progress for automatic detection methods, reaching sensitivity rates of about 98%. However, mass detection is a more complex task, because the mass is frequently indistinguishable from adjacent tissues. Mass detection poses a difficult challenge because masses are often: (a) very pronounced in size, shape and density; (b) poor in image contrast; (c) highly connected to the surrounding parenchymal tissue density, particularly for speculated lesions and (d) surrounded by no uniform tissue background with similar characteristics [7]. That makes progress be considerably slow for reliable detection of masses.

To improve diagnostic accuracies, the strategy of double reading has been used to increase the sensitivity level. However, the double reading performed by two radiologists increases operating costs and it is not always available. Comparing mammographic images from the same patient is a common practice for diagnosis purposes. Usually, images from both views (CC and MLO) of the same breast are compared. Some researches show that this approach can improve detection and diagnosis performances in contrast to employing a single mammographic view [8–11]. This methodology reduces patient recalls for a second inspection. Frequently, CADe commercial systems operate independently on each mammographic view. Nevertheless, there are situations when the system detects abnormalities in only one of the views. Radiologists believe that there is an inconsistency when a given lesion is not found, still being perceptible on both views, leading to a system reliability reduction and indicating that those results can be ignored. Several authors have been conducting researches to build CADe systems that can analyze two views on breast cancer detection [10,11].

The present work shows a set of computer-aided tools developed for segmentation of masses using two views, namely CC and MLO mammographic projections. To eliminate built-in patient informations, an algorithm is firstly applied to remove artifacts present at the mammograms background. To reduce the degradation/noise occurring in the process of image acquisition, a pre-processing technique based on Wiener filtering in the wavelet domain was applied to each image. After that, a computational process combining genetic algorithm (GA) and wavelet transform was employed to determine an appropriate number of threshold levels as well as the values of these levels required to mass segmentation. The GA and wavelet combination was used to decrease computational times for threshold selection. The threshold levels were then used in a multimodal thresholding procedure for detection and segmentation of masses obtained from CC and MLO views. The proposed computational tools were evaluated on a set of selected cases from the public Digital Database for Screening Mammography (DDSM) [12].

This article is organized as follows: in Section 2, it is described the algorithm employed on artifact elimination, breast image enhancement and multimodal segmentation of images obtained from both views. Section 3 presents experimental results obtained with the developed algorithm. Section 4 shows result discussions and conclusions of the work.

2. Materials and methods

This section describes the procedures employed for mammogram masses segmentation considering both CC and MLO views. Fig. 1 presents the block diagram of the applied algorithms.

2.1. *Artifacts elimination*

The performance of segmentation algorithms can be influenced by several factors such as artifacts and personal patient