Expert Systems with Applications 42 (2015) 6990–6999

Contents lists available at ScienceDirect

Expert Systems with Applications

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

Local energy-based shape histogram feature extraction technique for breast cancer diagnosis

Summrina Kanwal Wajid *, Amir Hussain [1]

Division of Computing Science & Maths, School of Natural Sciences, University of Stirling, Stirling, Scotland, UK

ARTICLE INFO

Article history:
Available online 1 May 2015

Keywords:
Computer-aided decision support system (CADSS)
Local energy-based shape histogram (LESH)
Support vector machine (SVM)
Local energy model
Receiver operating characteristic (ROC) curve

ABSTRACT

This paper proposes a novel local energy-based shape histogram (LESH) as the feature set for recognition of abnormalities in mammograms. It investigates the implication of this technique on mammogram datasets of the Mammographic Image Analysis Society and INbreast. In the evaluation, regions of interest were extracted from the mammograms, their LESH features calculated, and fed to support vector machine (SVM) classifiers. In addition, the impact of selecting a subset of LESH features on classification performance was also observed and benchmarked against a state-of-the-art wavelet based feature extraction method. The proposed method achieved a higher classification accuracy of 99.00 ± 0.50, as well as an A_z value of 0.9900 ± 0.0050 with multiple SVM kernels, where a linear kernel performed with 100% accuracy for distinguishing between the abnormalities (masses vs. microcalcifications). Hence, the general capability of the proposed method was established, in which it not only distinguishes between malignant and benign cases for any type of abnormality but also among different types of abnormalities. It is therefore concluded that LESH features are an excellent choice for extracting significant clinical information from mammogram images with significant potential for application to 3-D MRI images.

1. Introduction

Breast cancer is a fatal disease that originates in breast tissue. It primarily affects women; however, men can also develop it. Breast cancer caused more deaths than any other cancer in women in the US in 2011, when it was the second-most diagnosed cancer after skin cancer. In the UK, breast cancer accounts for 30% of all female cancers; almost one in nine women is estimated to develop it. In the EU, a woman is diagnosed with breast cancer every 2.5 min (American Cancer Society, 2011). Although the risk of breast cancer is lowest in less-developed countries, it is increasing every year (Hollander, 2002; Jemal et al., 2011; Shulman, Willett, Sievers, & Knaul, 2010).

The fatality rate of breast cancer can be reduced by early diagnosis and treatment. Computer-aided decision support systems (CADSSs), combined with mammography, help doctors diagnose breast cancer at an early stage. CADSSs function in four key steps: preprocessing, feature extraction, feature selection, and classification. The preprocessing step includes breast image segmentation and filtering, which is followed by image normalization to improve image quality and reduce noise. In feature extraction, images of lesions are extracted from enhanced images by using various techniques. In feature selection, an additional set of features is selected. In classification, the selected feature set is classified to separate false signals from true ones.

The essence of an accurate diagnosis exists in the selection of suitable features that can differentiate between normal and abnormal cases. The literature reports many techniques for feature extraction. Yu, Li, and Huang (2006) used a combination of model-based and statistical texture features to detect microcalcifications. This approach first detects the region of interest (ROI)—in this case, the area that contains the microcalcifications—by using wavelets and thresholds. It then extracts texture features from the ROI by using Markov random fields, fractal models, and statistical features. The performance of this approach was evaluated by using the area under the free-response receiver operating characteristic curve (FROC). A true positive rate of 94% was achieved with a 1.0 false positive per image rate.

Multi-resolution methods greatly interest researchers in image processing, analysis, biology, and other fields. Eltoukhy, Faye, and Samir (2010) applied curvelet transforms to mammograms. They

* Corresponding author at: Division of Computing Science & Maths, School of Natural Sciences, University of Stirling, Stirling FK9 4LA, Scotland, UK. Tel.: +44 (0)1786 - 476437.

E-mail addresses: skw1@cs.stir.ac.uk (S.K. Wajid), ahu@cs.stir.ac.uk (A. Hussain).

[1] Visiting Professor, Anhui University, Hefei, China.

http://dx.doi.org/10.1016/j.eswa.2015.04.057
0957-4174/

used the largest coefficients of the curvelet as feature vectors. The method resulted in a 98.59% classification accuracy when using SVM classifiers. Lladó, Oliver, Freixenet, Martí, and Martí (2009) used extended local binary pattern (LBP) histogram features. These features were extracted from regions that showed masses on the mammogram. The method resulted in an A_z value of 0.94 ± 0.02. Wang, Shi, and Heng (2009) classified malignant and benign masses by using SVM for features selected from the ROI. The feature set included curvilinear, texture, Gabor, and multi-resolution features. Performance was measured by using ROC and reached 0.97 with a maximum accuracy of 91.4%. Karahaliou et al. (2008) found that the texture properties of the area surrounding the microcalcification were significant when detecting malignancies in mammograms. This approach considered features such as grey-level first-order statistics, grey-level co-occurrence matrices, and Laws' texture energy measures, which were extracted from the surrounding tissue of regions of interest (ST-ROI). A redundant discrete wavelet transform (RWT) was then applied to the image. The wavelet coefficient first-order statistics and wavelet coefficient co-occurrence matrices for ST-ROI were then used as features. This combination of features resulted in an A_z performance of 0.9989. Delogu, Fantacci, Kasae, and Retico (2007) extracted 16 different features, including mass perimeter, circularity, mean of the normalized radial length, and others, from segmented masses and studied their various combinations. These features, which were further selected by using feature discriminating power and linear correlation interplay techniques, were fed to a multi-layer perceptron neural network for classification. It achieved 97.8% accuracy. Rashed, Ismail, and Zaki (2007) experimented with different wavelet transforms to analyze their ability to discriminate among different classes of abnormalities, such as clusters, speculated lesions, circumscribed masses, and ill-defined lesions. (Mousa, Munib, & Moussa, 2005) used wavelet (Daubechies) coefficients as feature vectors. In this approach, the horizontal, vertical, and diagonal detailed coefficients were first extracted from the image decomposition via Daubechies wavelets. The coefficients were then normalized and the energy for each feature vector was calculated. This output was used for classification purposes. The features were reduced depending on the level of energy required.

Verma, McLeod, and Klevansky (2010) used density and shape features, such as mass margins, abnormality assessment rank, patient age, and subtlety values, as features and then applied soft clustering by using k-mean to separate each class further. After clustering within malignant and benign cases, a neural network was applied with two additional layers other than the conventional three layer neural network architecture. One of these additional layers was a fixed weight layer, while the other one was a cluster layer, which prescribes the natural tendency of the output towards a malignant or benign class. The algorithm achieved a maximum accuracy of 97.5% on the test set. Diaz-Huerta et al. (2014) applied contrast enhancement with extended maxima threshold for image enhancement; next, they extracted special, texture, and spectral domain features from it. They further applied a support vector machine classifier to distinguish between malignant and benign microcalcifications. The result reached an overall sensitivity of 85.9%. (Rouhi et al., 2015) applied an artificial neural network (ANN)- and cellular neural network (CNN)-based segmentation to mammograms. Furthermore, they extracted texture and shape features from the segmentation area and fed it to the artificial neural network for classification between malignant and benign masses. The result reached an accuracy of 96.87% at maximum. Soltanian-Zadeh et al. (2004) experimented with texture, including multiwavelet, wavelet, Haralick, and shape features, to classify between malignant and benign microcalcifications. Features were further selected by using a genetic algorithm. The result was a maximum 0.89 ROC value. do Nascimento et al. (2013) considered wavelets (Daubechies-8 (db8), Symlet 8 (sym8), and bi-orthogonal 3.7 (bior3.7)) as features and applied a polynomial classifier to discriminate abnormal cases from normal ones. They achieved an A_z value of 0.98 ± 0.03 for the classification performance on a DDSM mammogram dataset.

Most of the above mentioned algorithms focus on one type of abnormality (either mass or microcalcification) and try to diagnose malignancy, whereas this paper proposes a method that can efficiently deal with different types of abnormalities simultaneously while discriminating between malignant and benign cases with a higher accuracy rate. The algorithm also produces good results in discriminating between different types of abnormalities, as can been seen in the results section.

The main objective of this research is to evaluate the implication of a novel LESH feature extraction technique for diagnosing malignancies in mammograms; this algorithm generates more accurate results than the other mentioned algorithms above. LESH has been successfully applied in pattern matching applications with promising results. Zakir, Zafar, and Edirisinghe (2011) applied LESH to automatically detect and recognize different road signs. Sarfraz and Hellwich (2008a, 2008b, 2009) experimented with LESH in face recognition systems for different face and head poses. LESH works by calculating a histogram of the local energy pattern within the image. A histogram is a simple technique that forms the basis of many spatial domain image processing techniques (Gonzalez & Woods, 2002). It provides useful image statistics that can be used to further analyze and process the image as described later in this paper.

The remainder of this paper is organized as follows. Section 2 introduces the proposed LESH scheme. Section 3 describes the experimental work; the results and discussion are presented in Section 4. Conclusions and future work are presented in Section 5.

2. Local energy-based shape histogram feature extraction

LESH works by converting an image into a combination of local energies along different orientations. Morrone and Owens (1987) suggested that features extracted at the points of maximum phase congruency can be helpful in image analysis. The type of phase and amplitude of local maxima of the energy function determines the type, sign, and contrast of a feature. This framework for calculating the phase congruency in two-dimensional images while using a high-pass filter to obtain image features at different scales is given as (Kovesi, 1999):

$$PC(z) = \frac{\sum_n \mathrm{W}(z)\lfloor A_n(z)\Delta\Phi_n(z) - T\rfloor}{\sum_n A_n + \epsilon} \tag{1}$$

where

$$\Delta\Phi_n(z) = \cos\left(\phi_n(z) - \bar{\phi}_n(z)\right) - \left|\sin(\phi_n(z) - \bar{\phi}_n(z))\right| \tag{2}$$

and T is the noise cancelation factor, while $W(z)$ is the weighting of the frequency spread. A_n and ϕ_n represent the amplitude and phase angle, respectively, of local complex value Fourier components at location $z = (x, y)$ in the image of size n. ϵ is a constant value incorporated to avoid division by zero (Kovesi, 1999).

Here, A_n and ϕ_n are calculated by using the logarithmic Gabor wavelets filter (Morrone & Owens, 1987). It detects low-level features that are invariant to image illumination, contrast, and image magnification (Kovesi, 2000). These one-dimensional symmetric/anti-symmetric filters are transformed into a two-dimensional form by applying a Gaussian spreading function across the direction that is perpendicular to its orientation. The image is convolved with a bank of Gabor kernels at each of the