Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 93 (2016) 389 – 395

Procedia Computer Science

6th International Conference On Advances In Computing & Communications, ICACC 2016, 6-8 September 2016, Cochin, India

A study of different texture features based on local operator for benign-malignant mass classification

Rinku Rabidas[a,*], Abhishek Midya[b], Jayasree Chakraborty[c], Wasim Arif[a]

[a]*Department of Electronics and Communication Engineering,National Institute of Technology Silchar, Assam, 788010, India*
[b]*Department of Electronics and Instrumentation Engineering,National Institute of Technology Silchar, Assam, 788010, India*
[c]*Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, NY, 10065, USA*

Abstract

In this paper, a comparative analysis of different texture features based on local operator has been produced for the determination of mammographic masses as benign or malignant. Local Binary Pattern (LBP), LBP Variance (LBPV), and Completed LBP (CLBP) descriptors are extracted to evaluate their potential for mass classification in a Computer-Aided Diagnosis (CAD) system. An A_z value of **0.97 ± 0.02** and an accuracy of **92.25 ± 0.01%** have been achieved, while experimenting on 200 mass cases from the DDSM database, by selecting the optimal set of features employing stepwise logistic regression method, followed by classification via Fisher Linear Discriminant Analysis (FLDA) using 10-fold cross validation.

Peer-review under responsibility of the Organizing Committee of ICACC 2016

Keywords: Breast Cancer; Mammography; Mass Classification; Local Binary Pattern(LBP); LBP Variance(LBPV); Completed LBP(CLBP)

1. Introduction

Among different types of cancers, breast cancer, though rare in male, frequent in women especially aged above 40 years has been ranked top in estimated number of new cases and deaths in 2015[1]. The only way to reduce this mortality rate is to detect and diagnose breast cancer at its earlier stages. Mammography, an X-ray imaging technique, has been considered as the best effective technique for the detection of abnormalities present in the breast. Anomalies like calcification, architectural distortion, bilateral asymmetry, and masses are the common peculiarities visible in mammogram, but among these, detection of masses are the most difficult due to their subtle nature i.e. variations in shape, size, and margin. The growing number of new cases and the difficulty involves in the detection of masses, have made examination of mammograms a challenging task for the radiologists. Hence, in order to assist them, Computer-Aided Diagnosis (CAD) system has been developed as a second evaluator.

Generally the categorization of a mammographic mass as benign and malignant is done on the basis of shape, margin, and density. A mass having round/oval shape, well defined margin, and low density can be categorized

* Corresponding author. Tel.: +91-9085601375
E-mail address: rabidas.rinku@gmail.com

Peer-review under responsibility of the Organizing Committee of ICACC 2016
doi:10.1016/j.procs.2016.07.225

into benign case otherwise malignant case. As a result, several feature extraction methods based on shape, margin, and texture have been proposed for the classification of mammographic masses in the state-of-the-art. In general, textural information has been widely exploited against shape-[2,3,4] and margin[5,6,7]-based features because the later approaches require accurate segmentation. In case of texture analysis, Haralick's features, extracted from Gray Level Co-occurrence Matrix (GLCM)[8], Angle Co-occurrence Matrices (ACM)[9], Optical Density Co-occurrence Matrix (ODCM)[10] have been extensively utilized. Sahiner et al.[11] has measured Haralick's features using GLCM matrix obtained from the Rubber-band straightening transform(RBST) images and achieved an A_z value of 0.94 using 160 mammograms. Similar analysis has been produced by Chakraborty et al.[12] where Haralick's features are extracted from ACM matrices and an A_z value of 0.77 has been reported using 433 Region of interests (ROIs). Multi-resolution analysis of the oriented patterns has been evaluated by Midya and Chakraborty using Haralick's features and 0.86 has been observed as the highest A_z value using 433 ROIs[13]. Tai et al.[10] has also computed Haralick's descriptors from ODCM matrix and obtained an A_z value of 0.98 for 398 mammograms. Frequency domain analysis has also been performed where coefficients of Wavelet transform have been used as feature vector[14] to achieve an accuracy of 98.6% using 140 cases. Similar analysis has been reported by Eltoukhy et al.[15] where an accuracy of 97.3% is achieved with curvelet transform. A few works have advocated the efficiency of combinational methods to get improved performance. Laroussi et al. has clubbed Zernike moments and Local Binary Pattern (LBP) features to achieve an A_z value of 0.96 using 160 mammograms[16].

In this correspondence, the authors have analyzed the discriminative capability of different texture feature extraction methods, based on the local operators, in a CAD system. Local binary pattern (LBP), LBP Variance (LBPV) and Completed LBP (CLBP) features with their variants have been extensively examined with an aim to achieve more precision in the classification of mammographic masses as benign or malignant.

The reminder of the paper is as follows: An overview of the DDSM database, utilized to carry out several experiments, has been provided in section 2. Section 3 presents a brief introduction of a CAD system and feature extraction methods investigated in the present work. The performance analysis of different features for benign-malignant mass classification has been discussed in section 4 and finally, the paper is concluded in section 5.

2. Image Database

To observe the performance of LBP, LBPV, and CLBP features for the classification of masses, several experiments have been conducted on the DDSM database which is an open source database provided by the University of South Florida[17]. The mammographic images of the DDSM database are of resolution of 42μm/pixel, 43.5μm/pixel, and 50μm/pixel and the boundary points inscribing the abnormalities present in the breast are provided with the database. In this present study, 200 randomly selected mass cases, 100 cases each of benign masses and malignant masses have been considered and the sample mass cases are shown in Fig. 1.

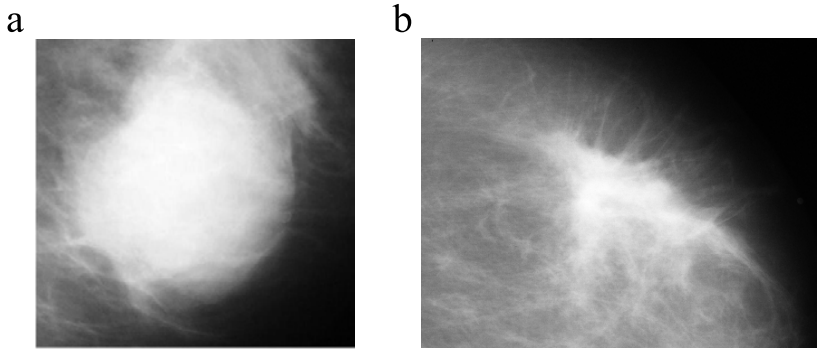

Fig. 1: Selected ROIs containing (a) Benign mass and (b) Malignant mass from the original images of the DDSM database.