Texture Analysis of Mammogram Using Local Binary Pattern Method

Athraa H. Farhan and Mohammed Y. Kamil *

College of Sciences, Mustansiriyah University, Baghdad, Iraq.
* Email: m80y98@uomustansiriyah.edu.iq, ORCID: orcid.org/0000-0001-5709-2549

Abstract. Breast cancer is one of the types of cancer that threatens the lives of women in their 40s. Based on the statistic reports, death-rate can be reduced by early detection of breast cancer. Breast cancer detection in early stages and lessen false positives in radiologist diagnosis can be achieved by combination Computer-Aided Diagnosis (CAD) with mammography. In this work, we offer a feature extraction technique as a method to lessen false-positive in breast mass recognition. Distinguishing explicit breast masses and ordinary tissue is the objective we strive to achieve. The mini-MIAS database of mammograms was used in this paper. LBP is the method that was used to extract features from the ROI. Comprehensive detection of this method can be developed, by taking the ROI inside the ground truth, which is automatically identified in the mini-MIAS and classifier majority voting. Better sensitivity, specificity, and accuracy are observed with a logistic regression classifier.

1. Introduction
Cancer is an irregular growth of cells that divide uncontrollably and can spread to normal body tissues[1, 2]. Breast cancer is a type of cancer that affects breast tissue and leads to the death of many women unless it is detected early [3]. Mammography is a specific medical imaging where the X-ray system uses a low dose to investigate the breast as a method for early detection of a breast mass [4]. Mammography technology in rapid development. However, in some cases, the accuracy of the investigation may be reduced. In the case of dense breast, the mass can be hidden by the normal tissue, so that the result of the examination may be inaccurate. Often times, the specialist suggests biopsies for the possible regions of cancer. This procedure is costly, therefore involved systems should be used to increase accuracy without wasting true (cancerous) positive [5]. These systems combine CAD and Specialist experience to help radiologists to recognize a cancerous or noncancerous mass to avoid mistakes and its efficacy has been proven in many studies[6]. One-sided and two-sided are the classes of CAD systems. one-sided CAD used one imaging view. (CC) and (MLO) is a usual view of mammography. To develop the accuracy of breast cancer detection, two-sided CAD combined between Cranio-Caudal (CC) and Medio-Lateral (MLO)[7]. CAD technique divided into three principal steps: pre-processing, feature extraction and classification [8]. Various feature extraction techniques were suggested in the field of mammogram analysis. such as LBP, Hog, and others [9]. S. J. S. Gardezi, et al. (2015) [10] suggested a method to increase the accuracy of recognizing mass breast in mammograms by fusion of texture features. The proposed approach employs texture features collected from the (CLBP) and (GLCM). The nearest neighbor is used as a classifier. classification accuracy 96.68%,98.9% sensitivity. (FP) fall by 40% and 78% respectively. P. Král and L. Lenc (2016) [11] presented (LBP)as a method for feature extraction. This method utilizes a local binary pattern with thresholding and classifier. This approach is estimated on images collected from MIAS and DDSM databases. The proposed method is efficient and effective because the accuracy obtained from this method is about 84%. M. Abdel-Nasser, et al. (2016) [12] suggested several texture analysis techniques to reduce the FPs
in breast diagnosis. They employed (LBP) as a beneficial method for feature extraction. Sensitivity 82.58 ± 0.0712 and Specificity 97.14 ± 0.0206. N. Ponraj and M. Mercy (2017) [13] proposed a comparison of the (LBP) and (LGP) including their histograms to classification the mammogram. The classification accuracy, Sensitivity, and Specificity that was obtained using LBP is 91% 90% 92%, respectively. R. Touahri, et al. (2019) [14] investigated (LBP) as a method for texture analysis of mammograms and considering effective factor for (CAD) system. The approach method produced best results than the CNN approach with an accuracy of 96.32%. C. E. F. Matos, et al. (2019) [15] proposed techniques (SIFT), (ORB), (SURF) and (LBP) for texture analysis to recognizing benign and malignant masses in a mammogram. The sensitivity achieved by this approach 100%, accuracy 99.65%, and specificity 99.24%.

The mini-MIAS database was used in this paper. Which available on the internet. (Ground truth) are selected automatically upon tumor indicating. We used ROI only the region inside the ground truth. LBP applied to the ROI resulting from mini-MIAS cyst indicating and afterward computational processes such as contrast, correlation, energy, homogeneity. Then, we selected the fit features to choose the features meaningful and consider them as the input to the classifier to identify the abnormality masses in mammograms as benign and malignant to improve the performance of breast mass detection.

2. Methodology

In this paper, we focusing on extract features from ROI. The characteristics of ROI can be defined as Features. The features are classified into texture and morphological features. The texture is one of the important aspects gives us information about the objects in an image. It is a measure of gray levels, on LBP to get a comparative analysis of their knowledge in the breast mass classification as normal or abnormal [16]. Fig. 1 shows a sketch of a CAD system for classification of mammography images.

![Figure 1. Schematic block diagram of a CAD system.](image)

(LBP) Suggested by Ojala et al.[17] for texture feature extraction with low computational complexity and low sensitivity to illumination changes in the images. It is an effective descriptor in several tasks of computer vision and face recognition. Features extraction are very important for the classification of mammographic masses. The original LBP coding the pixels of an image by comparing the central pixel with its neighbors and the result counted as a binary pattern number. in basic LBP, powerful features are captured from a small region so that a circular neighborhood was used instead of a 3×3 window to capture greater details around each pixel [18]. In calculating LBP in a circular neighborhood, The difference is taken between the current pixel and its neighbors. It is defined as

\[ LBP_{PR} = \sum_{p=0}^{P-1} s(g_p - g_c)2^p, \quad s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \]  

(1)

where \( g_p \) is the neighbor value, \( g_c \) is the current pixel value, \( P \) is the number of neighbors and \( R \) is the radius of the circular neighborhood. First, the central pixel matches with neighbor pixel if the current pixel less than the neighbor pixel it is labeled 1 otherwise it is labeled 0. the sign parameter for neighborhood pixel can be computed by employing (1). The process of computing LBP is shown in Fig.2.

The signs of the differences just estimate when calculating the last result of LBP. the magnitude of the differences is neglected completely [16]. the variance value can be used to get the dominant direction in a neighborhood. The dominant direction is the mark in a circular neighborhood which in the difference is greatest. Rotation LBP (RLBP) calculates the binary codes depending on the pixel neighborhood. The final values obtained from RLBP are different since the weights are shifted circularly. the arrangement of the pattern is the same because the binary patterns for both operators are similar. RLBP can be 3+4 transformed into invariant LBP by shifted the weights circularly Depending on where the dominant direction. At compute texture features, LBP uses two transitions one is uniform LBP and the other is RILBP. when the code of LBP contains at most two transitions 0 to 1 or 1 to 0, LBP named uniform LBP. As an instance, the binary code 00011100, 01000000 [19]. \( P(\text{P}−1) + 3 \) is the number of uniform
patterns. While rotationally invariant uniform pattern (RILBP) is produced P+2 distinct output values [20].

3. Results and discussion
In this paper, we used a digitized mammogram collected from the mini_MIAS database which is available on the internet [21]. The mini_MIAS database has 322 images. That is taken from 161 different ladies, inside this: 209 normal, 51 malignant and 62 benign. mammograms are reduced to a 200- micron pixel edge so that every image is 1024x1024 pixels. normal and abnormal images extracted from the mini_MIAS database can be shown in Fig. 3 mammograms are classified by the character of background tissue (fatty glandular, dense-glandular, and fatty) and the class of abnormality (ill-define, normal, asymmetry, architectural distortion, speculated, circumscribed, and calcification) [22].

In mini_MIAS, mass regions are selected automatically upon tumor indicating. These indicating were determined by the specialist according to analysis the area of the tumor and marked the contour points (x,y) for each mass in the images. for texture analysis, we utilized ROIs within the specialist mark. The normal cases of mammograms in mini_MIAS, ROIs were determined manually with various dimensions and from arbitrary regions. Figure 4 shows ground truth extracted by the specialist (the circle surrounding the square ) and ROI (the square region ).
Since masses have different sizes in mini-MIAS images, we created ROIs in three sizes (10 × 10, 20 × 20, and 30 × 30) according to the dimension of the smallest mass in a mammogram. To analyze the texture of ROIs, we applied uniform LBP to all ROIs for all images and calculated the features (energy, homogeneity, contrast, correlation). Table 1 shows a sample for the features that we got from applying LBP on the Asymmetry class of abnormality.

| Table 1. Statistical features of Asymmetry class of abnormality (15 images) for ROI at (10 × 10), (20 × 20) and (30 × 30). |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Cont. | Corr. | Energy | Homo. | Cont. | Corr. | Energy | Homo. | Cont. | Corr. | Energy | Homo. |
| 0.0009 | 1.0268 | 0.0007 | 0.0261 | 0.0004 | 1.0383 | 0.0014 | 0.0369 | 0.0004 | 1.0449 | 0.0018 | 0.0429 |
| 0.0011 | 1.0679 | 0.0040 | 0.0636 | 0.0006 | 1.0476 | 0.0021 | 0.0455 | 0.0005 | 1.0446 | 0.0018 | 0.0427 |
| 0.0014 | 1.0707 | 0.0044 | 0.0660 | 0.0014 | 1.0568 | 0.0029 | 0.0538 | 0.0013 | 1.0594 | 0.0031 | 0.0561 |
| 0.0007 | 1.0523 | 0.0025 | 0.0497 | 0.0009 | 1.0618 | 0.0034 | 0.0582 | 0.0006 | 1.0467 | 0.0020 | 0.0446 |
| 0.0010 | 1.0459 | 0.0019 | 0.0439 | 0.0006 | 1.0316 | 0.0009 | 0.0306 | 0.0005 | 1.0284 | 0.0008 | 0.0276 |
| 0.0008 | 1.1093 | 0.0097 | 0.0985 | 0.0005 | 1.0741 | 0.0048 | 0.0689 | 0.0006 | 1.0663 | 0.0039 | 0.0622 |
| 0.0009 | 1.0909 | 0.0069 | 0.0833 | 0.0007 | 1.0560 | 0.0028 | 0.0530 | 0.0006 | 1.0604 | 0.0032 | 0.0570 |
| 0.0006 | 1.0641 | 0.0036 | 0.0602 | 0.0008 | 1.0640 | 0.0036 | 0.0601 | 0.0006 | 1.0512 | 0.0024 | 0.0487 |
| 0.0011 | 1.0647 | 0.0037 | 0.0607 | 0.0008 | 1.0483 | 0.0021 | 0.0460 | 0.0007 | 1.0577 | 0.0030 | 0.0546 |
| 0.0010 | 1.0442 | 0.0018 | 0.0423 | 0.0010 | 1.0357 | 0.0012 | 0.0345 | 0.0008 | 1.0455 | 0.0019 | 0.0435 |
| 0.0009 | 1.0955 | 0.0076 | 0.0872 | 0.0009 | 1.0633 | 0.0035 | 0.0596 | 0.0010 | 1.0583 | 0.0030 | 0.0551 |
| 0.0043 | 1.0117 | 0.0001 | 0.0115 | 0.0066 | 1.0103 | 0.0001 | 0.0101 | 0.0060 | 1.0177 | 0.0003 | 0.0174 |
| 0.0012 | 1.0766 | 0.0051 | 0.0711 | 0.0016 | 1.0580 | 0.0030 | 0.0548 | 0.0009 | 1.0506 | 0.0023 | 0.0481 |
| 0.0011 | 1.0903 | 0.0069 | 0.0828 | 0.0006 | 1.0371 | 0.0013 | 0.0358 | 0.0007 | 1.0481 | 0.0021 | 0.0459 |
| 0.0013 | 1.0362 | 0.0012 | 0.0349 | 0.0009 | 1.0287 | 0.0008 | 0.0279 | 0.0010 | 1.0529 | 0.0025 | 0.0502 |

The features extracted from ROIs were subject to classification. After applying more than one classifier, we found that the best results of texture analysis were obtained with a logistic regression classifier. The suggested method can be evaluated by using 5-fold cross-validation where the database is distributed arbitrarily through five groups and compute accuracy to every group. The mean accuracy for all five groups is the last accuracy of the method. Table 2. shows the results obtained when using a logistic regression classifier. LBP produces the best classification accuracy and Sensitivity with size ROI (30×30), 85.5%, 73% respectively in the case of AS images and best Specificity, 95% in the case of MI images.
Table 2. accuracy, Sensitivity, and specificity of breast mass detection with LBP and logistic regression classifier

| ROI     | type | AR  | AS  | CA  | CI  | MI  | SP  |
|---------|------|-----|-----|-----|-----|-----|-----|
| 10 x 10 |      | 70.7| 70.4| 78.2| 62.7| 75  | 65.5|
| 20 x 20 |      | 78  | 75  | 72  | 65  | 77.3| 68  |
| 30 x 30 |      | 83.1| 85.5| 82.1| 75  | 79.2| 78  |

| ROI     | type | AR  | AS  | CA  | CI  | MI  | SP  |
|---------|------|-----|-----|-----|-----|-----|-----|
| 10 x 10 |      | 21  | 15  | 44  | 5   | 15  | 16  |
| 20 x 20 |      | 63  | 31  | 19  | 10  | 17  | 18  |
| 30 x 30 |      | 68  | 73  | 56  | 50  | 31  | 58  |

| ROI     | type | AR  | AS  | CA  | CI  | MI  | SP  |
|---------|------|-----|-----|-----|-----|-----|-----|
| 10 x 10 |      | 95  | 97  | 92  | 92  | 95  | 90  |
| 20 x 20 |      | 85  | 95  | 93  | 93  | 96  | 88  |
| 30 x 30 |      | 90  | 90  | 93  | 88  | 95  | 89  |

4. Conclusion
Extracting features based on ROI without the need for accurate mass segmentation is an important part of a breast cancer diagnosis with the CAD system. In this paper, we used the feature extraction technique for analyzing mass tissue to reduce FP in breast mass detection. We presented (LBP) as a method of texture feature extraction. Correlation, homogeneity, energy, and contrast are textural features derived from LBP. We utilized these features as the input to the classifier to identify the abnormality masses in mammograms as benign and malignant. Best results offered with logistic regression classifier as the accuracy 85.5%, specificity 95%, sensitivity 73%. Lastly, the LBP method can be consolidated and matched with other feature extraction techniques to increase the efficiency of the CAD system for breast cancer detection.

References
[1] Kamil M Y 2016 Morphological gradient in brain magnetic resonance imaging based on intuitionistic fuzzy approach. In: 2016 Al-Sadeq International Conference on Multidisciplinary in IT and Communication Science and Applications (AIC-MITCSA): IEEE) pp 1-3
[2] Nabeel F. Lattoofi, Israa F. Al-sharuee, Mohammed Y. Kamil, Ayoob H. Obaid, Aya A. Mahidi, Ammar A. Omar and Saleh A k 2019 Melanoma Skin Cancer Detection Based on ABCD Rule First International Conference of Computer and Applied Sciences (1st CAS2019)
[3] Youssef Y B, Rabeh A, Zbitou J and Belaguid A 2014 Statistical features and classification of normal and abnormal mammograms. In: 2014 International Conference on Multimedia Computing and Systems (ICMCS): IEEE) pp 448-52
[4] Salih A M and Kamil M Y 2018 Mammography Image Segmentation Based on Fuzzy Morphological Operations. In: 2018 1st Annual International Conference on Information and Sciences (AICIS): IEEE) pp 40-4
[5] Abbas Q J C 2016 DeepCAD: A computer-aided diagnosis system for mammographic masses using deep invariant features 5 28
[6] Görgel P, Serbvas A and Uçan O N J E S 2015 Computer-aided classification of breast masses in mammogram images based on spherical wavelet transform and support vector machines 32 155-64
[7] Abdel-Nasser M, Moreno A, Abdelwahab M A, Saleh A, Abdulwahab S, Singh V K and Puig D 2019 Matching Tumour Candidate Points in Multiple Mammographic Views for Breast Cancer Detection. In: 2019 International Conference on Innovative Trends in Computer Engineering (ITCE): IEEE pp 202-7

[8] Torrents-Barrena J, Puig D, Melendez J, Valls A J J o E and Intelligente T A 2016 Computer-aided diagnosis of breast cancer via Gabor wavelet bank and binary-class SVM in mammographic images 28 295-311

[9] Tasdemir S B Y, Tasdemir K and Aydin Z 2018 ROI Detection in Mammogram Images Using Wavelet-Based Haralick and HOG Features. In: 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA): IEEE pp 105-9

[10] Gardezi S J S, Faye I J A M and Sciences I 2015 Fusion of completed local binary pattern features with curvelet features for mammogram classification 9 3037

[11] Král P and Lenc L 2016 LBP features for breast cancer detection. In: 2016 IEEE International Conference on Image Processing (ICIP): IEEE pp 2643-7

[12] Abdel-Nasser M, Moreno A, Puig D J J o E and Intelligente T A 2016 Towards cost reduction of breast cancer diagnosis using mammography texture analysis 28 385-402

[13] Ponraj N and Mercy M 2017 Texture analysis of mammogram for the detection of breast cancer using LBP and LGP: A Comparison. In: 2016 Eighth International Conference on Advanced Computing (ICoAC): IEEE pp 182-5

[14] Touahri R, AzizI N, Hammami N E, Aldwairi M and Benaida F 2019 Automated Breast Tumor Diagnosis Using Local Binary Patterns (LBP) Based on Deep Learning Classification. In: 2019 International Conference on Computer and Information Sciences (ICCIS): IEEE pp 1-5

[15] Matos C E F, Souza J C, Diniz J O B, Junior G B, de Paiva A C, de Almeida J D S, da Rocha S V, Silva A C J M T and Applications 2019 Diagnosis of breast tissue in mammography images based local feature descriptors 78 12961-86

[16] Mehta R and Egiazarian K O 2013 Rotated Local Binary Pattern (RLBP)-Rotation Invariant Texture Descriptor. In: ICPRAM, pp 497-502

[17] Ojala T, Pietikäinen M and Harwood D J P r 1996 A comparative study of texture measures with classification based on featured distributions 29 51-9

[18] Ojala T, Pietikäinen M, Mäenpää T J I T o P A and Intelligence M 2002 Multiresolution gray-scale and rotation invariant texture classification with local binary patterns 971-87

[19] Pawar M M, Talbar S N and Dudhane A J o h e 2018 Local binary patterns descriptor based on sparse curvelet coefficients for false-positive reduction in mammograms 2018

[20] Rabidas R, Midya A, Sadhu A and Chakraborty J 2016 Benign-malignant mass classification in mammogram using edge weighted local texture features. In: Medical Imaging 2016: Computer-Aided Diagnosis: International Society for Optics and Photonics) p 97851X

[21] SUCKLING J P J D M 1994 The mammographic image analysis society digital mammogram database 375-86

[22] Kamil M Y and Salih A M 2019 Mammography Images Segmentation via Fuzzy C-mean and K-mean International Journal of Intelligent Engineering and Systems 12 22-9