15th International Symposium on Computer Methods in Biomechanics and Biomedical Engineering
and
3rd Conference on Imaging and Visualization
CMBBE 2018
P. R. Fernandes and J. M. Tavares (Editors)

EXTRACTING AND EVALUATING TEXTURE FEATURES FROM BINARY GRADIENT CONTOURS OF MICROCALCIFICATIONS CLUSTERS IN BREAST MAMMOGRAMS CMBBE 2018

Duarte, M.A.[1], Alvarenga, A. V.[2], Pereira, W. C. A.[3]

[1] Electronic Engineering Department/UniCarioca University Centre, Rio de Janeiro, Brazil
[2] Laboratory of Ultrasound/National Institute of Metrology, Quality and Technology, Brazil
[3] Biomedical Engineering Program/COPPE, Federal University of Rio de Janeiro (UFRJ), Brazil

Keywords: CAD systems, microcalcifications, breast cancer, mammograms, texture features.

Abstract: *The binary gradient contours and local binary pattern techniques were applied to calculate 991 texture features from microcalcifications clusters in digital mammograms. A feature selection method based on mutual information ranked the features, and a Linear Discriminant Analysis used a forward procedure to search among the ranked features, the best subset. The best classification performance was achieved with 21 texture features. The results achieved using just texture features are encouraging.*

1 INTRODUCTION

Women death rates related to breast cancer are high worldwide, mainly because of late diagnosis, i.e., when the disease is already in an advanced stage [1]. Early detection and tumor removal in the initial stages of the disease are efficient strategies to reduce such high death rates. The best methods for detecting the early signs of breast cancer are clinical and mammographic examinations. The main objective of mammographic examinations is to identify non-palpable breast lesions [2].

Although the mammography is the best kind of exam for detecting breast cancer, several factors may affect the diagnostic accuracy in breast cancer screening, such as equipment quality, breast densities and physician knowledge and experience. The higher is the density, the more difficult is the analysis and diagnosis, and then the final mammography quality is highly dependent on the breast tissue itself [3].

Microcalcifications (MCs) can be found in mammographic routine exams, and they are considered as significant signs of the existence of malignant lesions. MCs are small granular deposits of calcium that appear in a mammogram as small bright dots and their detection is often difficult, requiring a radiologist to carefully examine the mammogram, since they may be hidden, especially in dense tissues [4]. Despite their frequent occurrence in mammograms, although 60%–80% are detected via histological examination, only 30%–50% of MCs in breast carcinoma are detected via mammographic examination [4]. The difficulty in detecting MCs in mammograms is related to their variation in orientation, brightness, shape (from granular to rod shapes), and diameter, as well as the density of the surrounding tissue [5].

Computer-aided detection (CADe) and computer-aided diagnosis (CADx) systems have been developed [6][7] in an effort to assist MC detection and diagnosis. CADx systems are

used to provide a second opinion, thereby increasing the accuracy of a radiologist's final diagnosis [3][8][9][10], and basically involve three steps[11][12]: (i) segmentation, (ii) features extraction and selection from the segmented MCs and their clusters, and (iii) classification. Such systems are usually based on features extracted from MCs, as compactness, roughness, number of MCs in a cluster, average area, orientation [13][14], and they can help minimizing false-positive and false-negative rates in breast cancer diagnosis.

Texture features combined to morphological features have been studied to characterize MCs clusters [15], as well as texture features alone [16]. Chan and Sahiner [15] have pointed out that the texture features were more effective than morphological features in distinguishing malignant and benign MCs, while their best classification performance was obtained combining texture and morphological features. Karahaliou et al. [16], using just texture features, have indicated that texture analysis of the tissue surrounding MCs shows promising results in diagnosis of breast cancer.

In this work, the binary gradient contours (BGC) and local binary pattern (LBP) techniques [17][18] were applied to calculate 991 texture features from MCs clusters presented on 190 images from the Digital Database for Screening Mammography (DDSM) [19]. Hence, texture features were ranked based on mutual information technique. Finally, an incremental procedure sequentially adds the top m-ranked features to the Fisher discriminant analysis to identify the best set of texture features in classifying benign or malignant clusters.

2 MATERIAL AND METHODS

The database is composed by 190 digital mammograms (300 dpi, 8 bits), from 78 patients, selected from the DDSM, being 140 images with MCs benign clusters and 50 images with malignant ones. The segmentation procedure applied to the digital mammograms to segment MCs is based on morphological operations as described in [20][21]. The segmented MCs were used just to define the MCs clusters of each image, and the MCs most external to the cluster were used as vertices for the definition of the convex polygon that delimits the cluster. The texture characteristics of the region defined by the convex polygon over the original digital mammogram were quantified by calculating the binary gradient contours (BGC) and local binary pattern (LBP).

BGC is a family of texture descriptors that computes a set of eight binary gradients between pairs of pixels all along a closed path around the central pixel of a 3×3 grey scale image. In [18], three different paths are presented, producing three versions of BGC, namely single-loop (BGC1), double-loop (BGC2) and triple-loop (BGC3) (Fig. 1). Considering LBP, a texture operator 3×3 is applied to the image through the concept of local thresholding, in which the grey scale values of the periphery of the 3×3 window are converted into a set of binary values using the grey scale value of the central pixel as a threshold. For the BGC1 and BGC3 paths, a set of 255 features was determined, while 225 features were calculated for BCG2, and 256 features were estimated for LBP, totaling 991 features. Equations (1), (2), (3) and (4) show the texture features calculation. More details about BGC and LBP can be found in [18].

$$\mathrm{BGC1}_{3\times3} = \sum_{n=0}^{7} \xi\left(I_n - I_{(n+1) \bmod 8}\right) \times 2^n - 1 \tag{1}$$

$$\mathrm{BGC2}_{3\times3} = 15 \times \sum_{n=0}^{3} \xi\left(I_{2n} - I_{2(n+1) \bmod 8}\right) \times 2^n + \sum_{n=0}^{3} \xi\left(I_{2n+1} - I_{(2n+3) \bmod 8}\right) \times 2^n - 16 \tag{2}$$

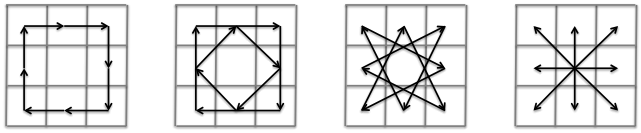
$$BGC3_{3\times3} = \sum_{n=0}^{7} \xi\left(I_{3n \bmod 8} - I_{3(n+1) \bmod 8}\right) \times 2^n - 1 \tag{3}$$

$$LBP3_{3\times3} = \sum_{n=0}^{7} \xi(I_n - I_c) \times 2^n \tag{4}$$

where I_c is the grey-level of the central pixel, I are the grey-levels of the peripheral pixels over the specific path of the 3×3 kernel, and ξ is defined as:

$$\xi(x) = \begin{cases} 1 \text{ if } x \geq 0 \\ 0 \text{ if } x < 0 \end{cases} \tag{5}$$

For selecting and assessing the performance of BGC and LBP texture feature subsets in distinguishing between benign or malignant clusters, a three-step procedure was implemented based on the one proposed in [17].



(a) (b) (c) (d)

Figure 1. Schematic example of paths proposed by [18] for (a) single-loop, (b) double-loop, and (c) triple-loop of the BGC texture descriptor, and (d) the layout of the 3×3 LBP.

First, the texture feature set is ranked using the minimal-redundancy-maximal-relevance criterion (MI-mrMR) based on mutual information technique [22]. In the second step, an incremental procedure sequentially adds the top m-ranked features to the classification step. The classification performance is assessed using the .632+ bootstrap method [23], considering, 500 bootstrap samples for each feature subset and the local Fisher discriminant analysis (LFDA), with linear kernel, as classifier [24]. The performance index was estimated by calculating the average over the 500 bootstrap of the area under ROC curve ($AUC_{.632+}$). In the third step, the feature subset with the best classification performance is considered as the "reference subset", and a search process tries to find a subset with fewer features that performs statistically similar to the reference subset given a 95% confidence interval. The one-way analysis of variance (ANOVA1) tested whether the mean values between compared groups were different, and a post-hoc analysis was made, using the Scheffé's method to determine if there was significant difference among particular pairs of groups [25]. Details about all the procedure steps are found in [17].

3 RESULTS

Figure 2 presents the performance of the classifier as the number of features ranked by the MI-mrMR technique is increased. The maximum $AUC_{.632+}$ value (0.928 ± 0.022) is achieved considering 21 features. Figure 3 presents the result of the post-hoc analysis using the Scheffé's method. One can observe that four groups are clearly presented. The best result considering the set of 21 texture features (star group in Figure 3) is statistically different from all other sets with less texture features. The second best performance presenting the less number of features is achieved with 9 texture features (blue line with circle in Figure 3). One can note that the set with 9 texture features is not statistically different from the other sets with 10 to 20 texture features (grey groups in Figure 3). The third best result is reached with

5 texture features (blue line with diamond in Figure 3). The groups with 1, 2, 3 and 4 texture parameters (red line groups in Figure 3) are not statistically different among them.

Table 1 presents the $AUC_{.632+}$ values, and respective standard deviations, for the sets presenting the best performance and less number of texture parameters. Table 2 shows the list of the 21 best features ranked.

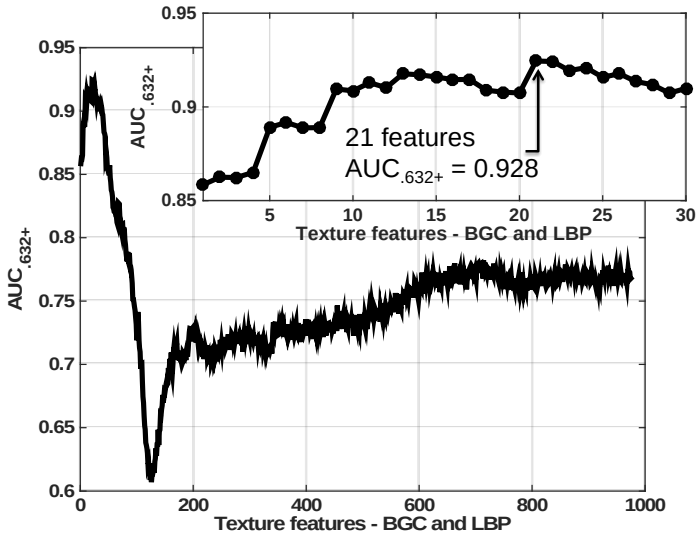


Figure 2. Estimation of $AUC_{.632+}$ values as the number of ranked features increases. The region in which the highest $AUC_{.632+}$ value is achieved is presented in detail.

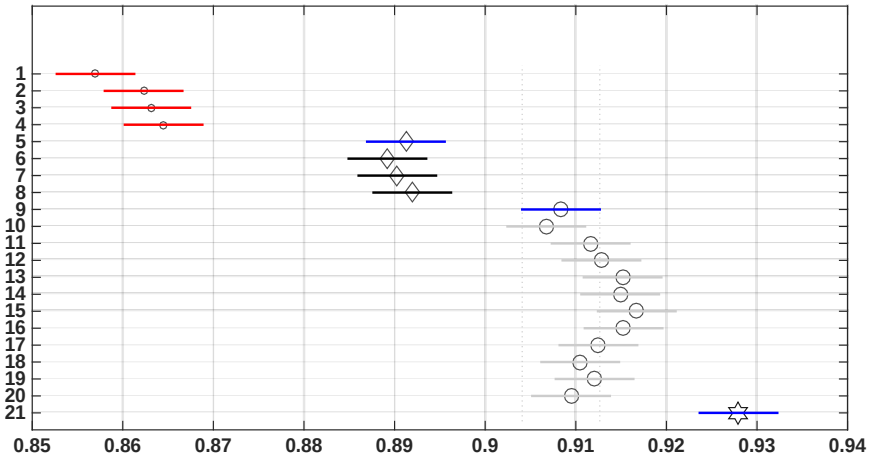


Figure 3. Result of the post-hoc analysis using the Scheffé's method. The star group indicates the best result considering the set of 21 texture features.

Number of texture features	$AUC_{.632+}$ values
21	0.928 ± 0.022
9	0.908 ± 0.024
5	0.891 ± 0.021
1	0.857 ± 0.022

Table 1. Values of AUC.632+ (± standard deviation) for the sets presenting the best performance and less number of texture parameters.

Rank	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Feature	BGC1_36	BGC3_245	BGC1_141	BGC3_4	LBP_21	BGC3_106	BGC3_129
Rank	**8**	**9**	**10**	**11**	**12**	**13**	**14**
Feature	BGC1_7	BGC3_89	BGC1_166	LBP_40	LBP_68	BGC3_114	BGC3_74
Rank	**15**	**16**	**17**	**18**	**19**	**20**	**21**
Feature	BGC2_122	LBP_5	LBP_100	BGC2_29	LBP_231	BGC1_162	LBP_187

Table 2. The best twenty-one texture features ranked.

4 DISCUSSIONS

The experimental results suggested that the feature selection method proposed by [24] is useful in finding the reduced subset of features for improving the classification performance in terms of AUC index. The 991 texture features were reduced to a set of 21 with a performance of AUC.632+= 0.928 ± 0.022. However, using the post-hoc Scheffé's method, no set with similar performance and a smaller number of features was found.

To warrant a consistent statistical analysis, the heuristic rule of a minimum of 30 sample images for each feature added was assumed. Therefore, considering the 190 mammograms used in this work, the limit for the number of texture features is 6 (190/30). Based on this assumption, the set with 5 texture features was the one with the best performance (AUC.632+= 0.891 ± 0.021) that it is not statistically different from the one with 6 features.

As far as we know, it is the first time that BGC and LBP are used to quantify texture of region of MCs cluster in mammograms, and the performance achieved is encouraging. Chan and Sahiner [15] stated an AUC value of 0.84 using just texture features calculated based on the Grey Level Co-occurrence Matrix. Combining texture and morphological features, the AUC value was increased to 0.89. That result is similar to the one achieved herein using just 5 texture features (Table 1). Karahaliou et al. [16], using Laws' Texture Energy Measure Features, achieved an accuracy of 89 %.

5 CONCLUSIONS

The binary gradient contours and local binary pattern techniques were applied to calculate texture features from microcalcifications cluster regions in digital mammograms. A total of 991 texture features were computed. A feature selection method, based on mutual information and statistical tests, determined that 21 texture features are capable of attaining the best classification performance with AUC $= 0.928 \pm 0.022$. It was also found that five texture features are enough to reach AUC $= 0.891 \pm 0.021$, which represents the best classification performance considering a minimum of 30 sample images for each texture feature studied. The results achieved using just texture features are encouraging and in accordance with literature. In future, other texture features will be studied in order to improve classification performance.

ACKNOWLEDGEMENTS

Thanks to the financial support of the Brazilian National Council for Scientific and Technological Development (CNPq) (Grants:434.858/2016-1, 309717/2014-0,and 308.627/2013-0), and CAPES/PROEX.

REFERENCES

[1] Brasil, Ministério da Saúde, Instituto Nacional de Câncer (INCA), Tipos de Câncer, Câncer de Mama, http://www2.inca.gov.br/wps/wcm/connect/tiposdecancer/site/home/mama (accessed August 2017).

[2] Brasil, Ministério da Saúde, Instituto Nacional de Câncer (INCA), Câncer de Mama, Detecção Precoce, http://www2.inca.gov.br/wps/wcm/connect/tiposdecancer/site/home/mama/deteccao_precoce+ (accessed August 2017).

[3] A. Jalalian, S.B.T. Mashoror, H.R. Mahmudb, M.I.B. Saripan, A.R.B. Ramli, B. Karasfi, "Computer-aided detection/diagnosis of breast cancer in mammography and ultrasound: a review," *Clin. Imaging*, vol. 37, no. 3, pp. 420–426, Nov2013.

[4] S. Halkiots, T. Botsis, M. Rangoussi, "Automatic detection of clustered microcalcifications in digital mammograms using mathematical morphology and neural networks," *Signal Processing*, vol. 87, no. 7,pp. 1559–1568, Jul2007.

[5] L. Wei, Y. Yanga, and R.M. Nishikawa, "Microcalcification classification assisted by content-based image retrieval for breast cancer diagnosis," *Pattern Recognition*, vol. 42, no. 6, pp. 1126–1132, Jun 2009.

[6] R.M. Nishikawa, "Current status and future directions of computer-aided diagnosis in mammography," *Comput. Med.Imaging Graph*, vol. 31, no. 4-5,pp. 224–235, Jun2007.

[7] M. Elter, and A. Horsch, "CADx of mammographic masses and clustered microcalcifications: a review," *Medical Physics*, vol. 36, no. 6, pp. 2052–2068, Jun2009.

[8] M.J.G. Calas, B. Gutfilen, and W.C.A. Pereira, "CAD and mammography: why use this tool?," *Braz. J. Biomed. Eng.*,vol. 45, no. 1, pp.46–52, Jan/Feb 2012.

[9] C.H. Chen, and G.G. Lee, "On digital mammogram segmentation and microcalcification detection using multiresolution wavelet analysis," *Graph. Models Image Process.*, vol. 59, no. 5, pp.349–364, Sep 1997.

[10] H.D. Cheng, X. Cai, X. Chen, L. Hu, and X. Lou, "Computer-aided detection and classification of microcalcifications in mammograms: a survey," *Pattern Recogn.*, vol. 36, no. 12, pp.2967–2991, Dec 2003.

[11] N.S. Arikidis, A. Karajaliou, S. Skiadopoulos, P. Korfiatis, E.Likaki, G. Panayiotakis, et al., "Size-adapted microcalcification segmentation in mammography utilizing scale-space signatures," *Comput. Med. Imaging Graph.*,vol. 34, no. 6,pp. 487–493, Sep 2010.

[12] S. Paquerault, L.M. Yarusso, J. Papaioannou, and Y. Jiang, "Radial gradient-based segmentation of mammographic microcalcifications: observer evaluation and effect on CAD performance," *Med. Phys.*, vol. 31, no. 9,pp. 2648–2657, Sep 2004.

[13] M. De Santo, M. Molinara, F. Tortorella, and M. Vento, "Automated classification of clustered microcalcifications by a multiple expert system," *Pattern Recogn.*, vol. 36,no. 7, pp.1467–1477, Jul 2003.

[14] W.J. Veldkamp, N. Karssemeijer, J.D.M. Otten, and J.H.C.L. Hendriks, "Automated classification of clusters microcalcifications into malignant and benign types," *Med. Phys.*, vol. 27, no. 11, pp.2600–2608, Nov 2000.

[15] H.P. Chan and B.Sahiner, "Computerized analysis of mammographic microcalcifications in morphological and texture feature spaces," *Med. Phys.*, vol. 25, no. 10, pp. 2007-2019, Oct 1998.

[16] A.N. Karahaliou, I.S. Boniatis, S.G. Skiadopoulos, F.N. Sakellaropoulos, E. Likaki, G.S. Panayiotakis and L.I. Costaridou, "A texture analysis approach for characterizing microcalcifications on mammograms," In: *Proceeding of IEEE International Special Topic Conference on Information Technology in Biomedicine (ITAB 2006)*, pp, 251–257,Oct 2006.

[17] W.G. Flores, W.C.A. Pereira, A.F.C. Infantosi, "Improving classification performance of breast lesions on ultrasonography," *Pattern Recognition*, vol. 48, no. 4, pp. 1125-1136, April 2015.

[18] A. Fernández, M.X. Álvarez, F.Bianconi, "Image classification with binary gradient contours," *Optics and Lasers in Engineering*, vol. 49, no. 9–10, pp. 1177-1184, September–October 2011.

[19] University of South Florida, DOD Breast Cancer Research Program, US Army Research and Material Command, Digital Database for Screening Mammography (DDSM), in http://marathon.csee.usf.edu/Mammography/Database.html(accessed 3.01.2014).

[20] M.A. Duarte, A.V. Alvarenga, C.M. Azevedo, M.J.G. Calas, A.F.C. Infantosi, and W.C.A. Pereira, "Segmenting mammographic microcalcifications using a semi-automatic procedure based on Otsu's method morphological filters," *Braz. J. Biomed. Eng.*, vol. 29, no. 4, pp. 377–388, Dec 2013.

[21] M.A. Duarte, A.V. Alvarenga, C.M. Azevedo, M.J.G. Calas, A.F.C. Infantosi, W.C.A. Pereira, "Evaluating geodesic active contours in microcalcifications segmentation on mammograms," Computer Methods and Program in Biomedicine, vol. 122, no. 3, pp. 304-315, Dec 2015.

[22] H. Peng, F. Long, C. Ding, "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 8, pp. 1226–1238, Aug 2005.

[23] B. Sahiner, H.P.Chan, L.Hadjiiski, "Classifier performance prediction for computer-aided diagnosis using a limited dataset,"*Med.Phys.*, vol. 35, no. 4, pp.1559–1570 Apr2008.

[24] M. Sugiyama, "Dimensionality reduction of multimodal labeled data by local Fisher discriminant analysis," *J. Mach. Learn. Res.*, vol. 8,pp. 1027–1061, May 2007.

[25] M.R. Stoline, "The status of multiple comparisons: simultaneous estimation of all pairwise comparisons in one-way ANOVA designs," *Am.Stat.*, vol. 35, no. 3, pp. 134–141, Aug 1981.