Breast Tissue Classification using Local Binary Pattern variants: A Comparative study

Minu George and Reyer Zwiggelaar

Department of Computer Science, Aberystwyth University, Wales, UK
{mig24,rrz}@aber.ac.uk

Abstract. Mammographic tissue density is considered to be one of the major risk factors for developing breast cancer. In this paper we use quantitative measurements of Local Binary Patterns and its variants for breast tissue classification. We compare the classification results of LBP, ELBP, Uniform ELBP and M-ELBP for classifying mammograms as fatty, glandular and dense. A Bayesian-Network classifier is used with stratified ten-fold cross-validation. The experimental results indicate that ELBP patterns at different orientations extract more relevant elliptical breast tissue information from the mammograms indicating the importance of directional filters for breast tissue classification.

Keywords: Breast density tissue classification, risk estimation, Local Binary Patterns

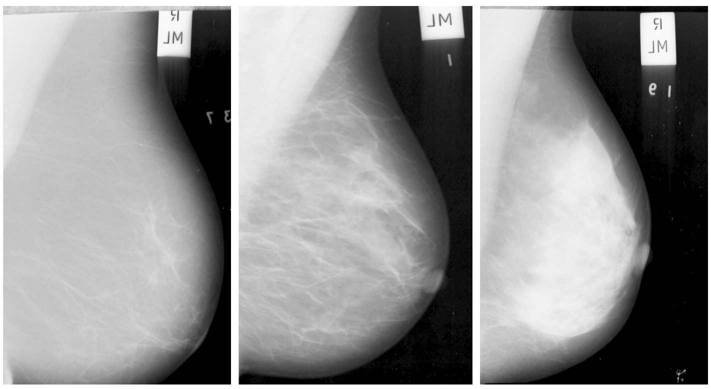
1 Introduction

Breast cancer is one of the most life threatening diseases among women in developed countries [20]. It is estimated that between one in eight women have the chance of getting breast cancer in their life time. It is a challenging task to find breast cancer at its early stage through self-assessment. A variety of medical imaging modalities like mammography, MRI, ultrasound and tomosynthesis are widely used for breast cancer detection. Among all these, mammography is still considered the primary imaging modality. Mammography provides the radiologist visualization of the internal structure of the breast along with an indication of the amount of glandular and connective tissue relative to the fatty tissue in the breast [1–3]. Recent studies have shown that Computer Aided Diagnosis (CAD) systems performance decreases with increasing breast density by decreasing sensitivity [4]. It has been shown that there is a strong correlation between the breast parenchymal patterns and the risk of developing breast cancer [5]. Therefore, developing automatic methods for breast density estimation and classification are appropriate. According to American College of Radiology (ACR), the latest classification and evaluation of breast tissues uses four categories as follows [21]: BIRADS I: the breast is entirely fatty, BIRADS II: scattered areas of fibro-glandular density, BIRADS III: heterogeneously dense breast, BIRADS IV: extremely dense breast.

It is interpreted that glandular tissue in mammographic images are represented as brighter areas and the darker regions represent fatty tissue in the breast. The classification of mammographic images based on tissue type density and mammographic risk estimation was initiated by Wolfe [5]; which was followed by different automatic classification techniques for breast tissue classification. Most CAD systems have used either segmented breast tissue or a pre-selected ROI based breast tissue regions for further feature extraction and classification. Diverse feature extraction techniques have been used such as histograms [7–9], intensity based [6], texture based approaches [10–12]. While Oliver et al. [1, 12] used features based on texture and morphology of tissue patterns, Mustra et al. [13] used GLCM features, and Petroudi et al. [14] used textons to capture mammographic appearance.

Subsequently, Zwiggelaar et al. [15] combined Local Binary Pattern (LBP) texture features and texture features extracted from grey level co-occurrence matrices to classify mammograms. Ojala et al. [18] introduced a powerful and computationally simple rotation invariant texture classification based on local binary patterns. LBP and its variants have proven to be useful [15, 16] in various medical image analysis applications. The extraction of elliptical micro-pattern features from breast tissue is important in classification. So in our approach we used variants of Local Binary Patterns [1], namely Elliptical Local Binary Pattern (ELBP), which have been used for extracting facial features using horizontal and vertical elliptical patterns.

In addition, we have used Uniform ELBP and Mean-Elliptical Local Binary Pattern (M-ELBP) [22] to capture intrinsic and detailed micro-pattern features from the mammographic images for breast tissue classification into fatty, glandular and dense breast tissue as shown in Fig. 1.



(a) (b) (c)

Fig. 1. Example mammographic tissue types, where (a) fatty, (b) glandular, and (c) dense tissue.

2 Methodology

In the proposed method, we use a ROI based approach for feature extraction and classification of tissue. The breast tissue region was segmented from the fibroglandular disk region and was followed by noise removal filtering [22]. The denoised image was used for texture feature extraction by LBP variants at different orientations to obtain the microstructure and pattern details of the breast tissue. Subsequently, the LBP variants features were used to classify the mammographic images as fatty, glandular and dense by using a Bayesian Network (other classifiers were investigated and reported in Sec.3).

The mammographic images are pre-processed to identify the breast tissue region. The pectoral muscle and artifacts were removed from the mammographic images. As most of the dense tissues and parenchymal patterns are located within the breast fibroglandular disk area, we extracted a ROI from this region of the mammogram. An ROI was extracted from each mammogram image with size equal to 256×256 pixels as shown in Fig. 2. Thereafter noise reduction was performed on the extracted ROI using a median filter of 3×3 size.

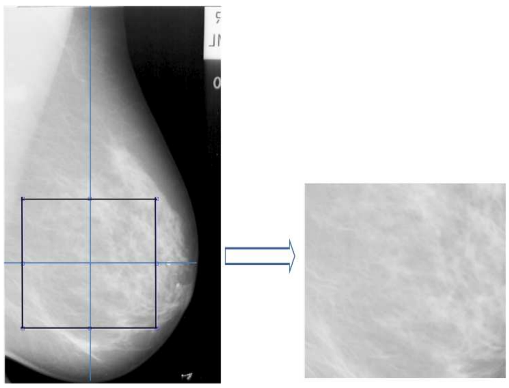


Fig. 2. ROI extraction from the fibro-glandular disk region.

The Local Binary Pattern (LBP) operator was introduced by Ojala et al. [18] for texture feature extraction with low computational complexity and low sensitivity to illumination changes in the images. For each central pixel (x_c, y_c) of the input image with a grey value g_c, its LBP value is estimated by comparing the g_c value with the grey values of pixels at R distance within its surrounding P neighbourhood pixels following the pixels along a circular path either clockwise or counter-clockwise as in Fig. 3. If the central pixel value has a higher grey level value than the neighbouring pixel, the neighbour pixel is assigned a value 0 else 1 giving an 8-bit binary number. The histogram for the whole ROI were calculated to extract the histogram texture features. The $LBP^{P,R}$ is calculated as follows:

$$LBP^{P,R}(x_c, y_c) = \sum_{i=1}^{P} s(g_i^{P,R} - g_c)2^{i-1} \tag{1}$$

where s(x) is defined as

$$s(x) = \begin{cases} 1, & \text{if } x >= 0 \\ 0, & \text{if } x < 0 \end{cases} \tag{2}$$

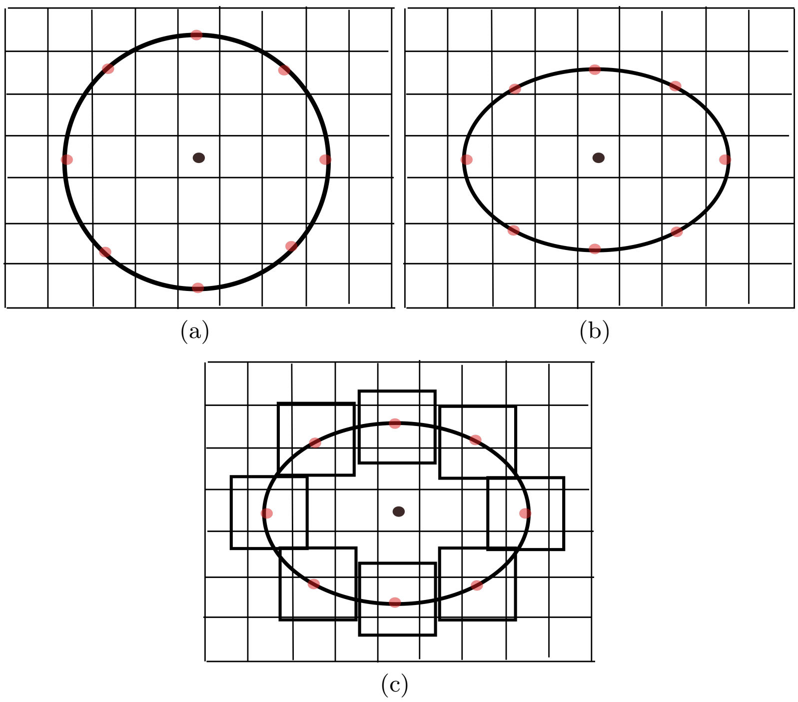


Fig. 3. LBP pattern for (a) $LBP^{P,R=3}(x_c, y_c)$, (b) $ELBP^{P,R_1=3,R_2=2}(x_c, y_c)$ (c) $M-ELBP^{P,R_1=3,R_2=2}(x_c, y_c)$ where P=8.

In Elliptical Local Binary Patterns (ELBP), for each central pixel (x_c, y_c), we consider the neighbouring pixels P which lie at radius R_1 and R_2 on an ellipse.The ELBP label of pixel (x_c, y_c) with radius distance of R_1 and R_2 with P surrounding pixels are calculated as :

$$LBP^{P,R_1,R_2}(x_c, y_c) = \sum_{i=1}^{P} s(g_i^{P,R_1,R_2} - g_c)2^{i-1} \tag{3}$$

where the i^{th} neighbouring pixel coordinate of (x_c, y_c) , is computed as follows:

$$angle_{step} = 2 * \frac{\pi}{P} \tag{4}$$

$$x_i = x_c + R_1 * \cos((i-1) * angle_(step)) \tag{5}$$

$$y_i = y_c - R_2 * \sin((i-1) * angle_(step)) \tag{6}$$

Compared to the LBP patterns, ELBP descriptors help in extracting more specific spatial features from the mammographic images.

In order to extract more features from the mammogram images, we perform ELBP descriptor at eight different orientations. When $(R_1 = R_2)$, the ELBP reduces to the LBP descriptor while $(R_1 < R_2)$, we have vertical ellipse and if $(R_1 > R_2)$ we get horizontal ellipse structure. This will help in capturing more intrinsic features in detail from the mammograms. The detailed overview of the method is described in Fig. 3

In order to extract the most intrinsic microstructure directional patterns from the mammogram ROI areas, the ELBP descriptors are computed on the ROI at different orientations $\theta = \{0°, 45°, 90°, 135°, 180°, 225°, 270° \text{ and } 315°\}$.

Later, the ELBP histogram labels for the ELBP descriptor at different angles $\theta = \{0°, 45°, 90°, 135°, 180°, 225°, 270° \text{ and } 315°\}$ are concatenated to get the final histogram feature vector. Fig. 4. summarises the classification of breast tissue using the ELBP descriptor variants at eight different orientation with an elliptical radius of $(R_1 = 4, R_2 = 7)$ for the ELBP descriptor for a neighbourhood range of eight pixel.

To extract more intrinsic features, we include intensity features using M-ELBP[22]. Similar to the ELBP operation, M-ELBP extract texture and intensity features of the mammographic tissue pattern in different directions.The M-ELBP descriptor is represented in Fig. 3. The local window around each neighbouring pixel will calculate the mean intensity of region around it and then this mean intensity value is compared with the central pixel for generating the binary pattern.

Once the feature extraction has been completed, feature selection is performed in-order to retain only the prominent features and to reduce the computational cost due to large feature length/dimensionality. A correlation based feature subset selection method with a best first search method was used. Highly correlated feature subsets that are highly correlated with the group class while having lower inter correlation among feature subsets are preferred for attribute selection. It calculates the individual predictive ability of each attribute/feature in the dataset along with the redundancy between each feature. The selected features are then classified into fatty, glandular and dense breast tissue using a Bayesian network.

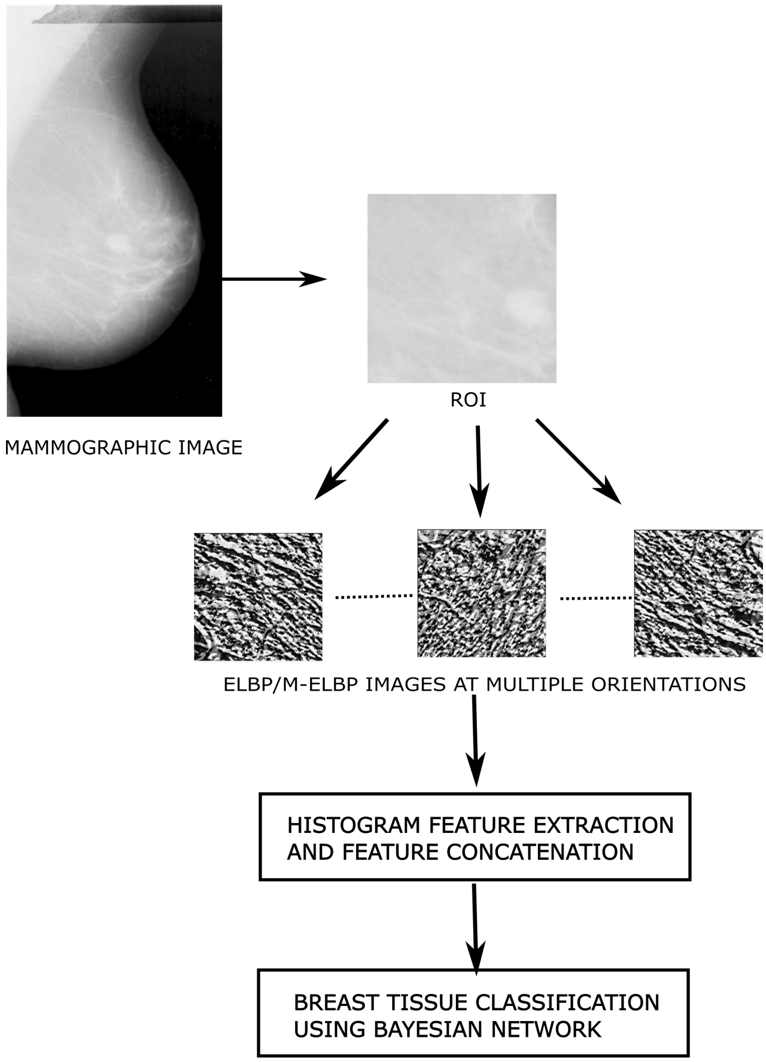


Fig. 4. Summary of ROI selection, feature extraction and classification using ELBP variants

3 Experimental Results

The methodology presented in this paper was applied to the Mammographic Image Analysis Society (MIAS) database [19]. The MIAS database contains 161 pairs of mediolateral oblique view mammograms. Each image was annotated by expert radiologists according to breast density into three distinct classes: Fatty (F), Fatty-Glandular (G) and Dense-Glandular (D) images. The whole MIAS database of mammographic images contains 106 fatty images, 104 fatty-glandular images, 112 dense-glandular images.

With the experimental results, it is suggested that feature extraction from a single orientation is not sufficient for estimating or defining the tissue type of a mammographic image due to the complex and multidimensional appearance of parenchymal patterns in the breast. In our approach to classify mammograms into different classes based on texture features, we apply the ELBP descriptors at eight different orientation $\theta = \{0°, 45°, 90°, 135°, 180°, 225°, 270°$ and $315°\}$. The ELBP binary patterns at each orientation were computed for the images and then histograms for each pattern were calculated for ELBP binary images. All the texture features obtained using the ELBP operator at different orientations are combined as a single feature representing the microstructure patterns in the mammograpic images.

In order for the comparison of classification of different Local Binary Pattern variants, we performed LBP, Elliptical Local Binary Pattern (ELBP), Uniform-ELBP and Mean-Elliptical Local Binary Pattern (M-ELBP). While LBP extracted features only from circular neighbourhood, the ELBP was able to extract more structural features at different orientations extracting multidimensional micropattern features of the breast tissue. In order to incorporate the intensity features along with the textural features into the histogram, we used M-ELBP. In M-ELBP each of the neighbour pixels mean value is estimated with its neighbouring pixels of a window size equal to 3×3. This mean value is compared with the central pixel for generating the binary pattern and M-ELBP image histogram. Similarly, the effect of Uniform patterns in tissue classification is estimated using Uniform ELBP [17] .

Table 1. Confusion matrix for automatic tissue classification using the Local Binary Pattern (LBP) descriptor.

	Automatic Classification		
Truth Data	**Fatty**	**Glandular**	**Dense**
Fatty	86	19	1
Glandular	16	73	15
Dense	7	38	66

Tab. 1 shows the classification results on the selected ROIs from the MIAS database. The classification accuracy is 70%. Since the LBP operator consider a circular pattern of neighbourhood, it cannot capture directional features from the breast tissue.

Table 2. Confusion matrix for automatic tissue classification using the Elliptical Local Binary Pattern (ELBP) descriptor.

	Automatic Classification		
Truth Data	**Fatty**	**Glandular**	**Dense**
Fatty	86	20	0
Glandular	11	68	25
Dense	3	22	86

Tab. 2 shows the classification results on the selected ROIs from the MIAS database using the Elliptical Local Binary Pattern descriptor. The classification accuracy has increased to 75% which shows that the ELBP could extract more relevant information from the breast tissue considering the multidimensional structure of the tissue. Since the ELBP operator is computed at different orientations, it could extract directional breast tissue features at multiple dimensions.

Table 3. Confusion matrix for automatic tissue classification using the Uniform Elliptical Local Binary Pattern (uniform-ELBP) descriptor.

	Automatic Classification		
Truth Data	**Fatty**	**Glandular**	**Dense**
Fatty	86	19	1
Glandular	11	69	24
Dense	2	27	82

Tab. 3 shows the classification results performed by ELBP when the uniform patterns are selected. The classification remained similar to the ELBP operator with an accuracy of 74% .

In order to incoperate more features using the ELBP descriptors, the mean intensity value of the neighbouring pixels are included to consider the intensity aspects along with texture features. Tab. 4 indicates the effect of considering the intensity along with texture features. The classification results has improved to approximatly 80%. This indicates that the M-ELBP descriptor performed better than the other variants of LBP for the extracted ROIs from MIAS.

Table 4. Confusion matrix for automatic tissue classification using the Mean Elliptical Local Binary Pattern (M-ELBP) descriptor.

	Automatic Classification		
Truth Data	**Fatty**	**Glandular**	**Dense**
Fatty	92	14	0
Glandular	13	76	15
Dense	0	23	88

Through the experiments conducted, the feature selection has improved the classification accuracy by 3-4%. By adopting different classifiers for comparing the classification accuracy, the Bayesian Network algorithm using a hill climbing algorithm restricted by an order on the variables performs better than others as illustrated by Tab. 5.

Table 5. Classification results by various classifiers for LBP variants.

Classifier	**LBP**	**ELBP**	**Uniform ELBP**	**M-ELBP**
Bayesian Network	70.1	75.0	74.0	80.0
KNN	69.2	69.2	71.3	75.7
SVM	66.0	67.3	70.1	69.8
Random Forest	67.3	73.5	72.9	72.0

4 Conclusion

In this paper, we have quantitatively compared the texture features generated by variants of Local Binary Patterns for classifying breast tissue. For the MIAS database, divided into the three density categories, the best classification result of 80% was obtained by M-ELBP followed by ELBP, Uniform-ELBP and LBP with classification accuracy of 75%, 74%, and 70%, respectively by the Bayesian network. M-ELBP performed better than LBP and ELBP as it included intensity features along with the texture feature of the mammographic image. Similarly the ELBP descriptors outperformed LBP as it was able to extract elliptical features at different orientations taking into account the multidimensional tissue structure. So the results gives the indication that directional filters improves the classification result considerably leading to the scope of building better directional filters. In the future, we will evaluate extracting masses, linear structure features and microcalcifications from the mammographic images for abnormality classification using the ternary ELBP and M-ELBP variants.

References

1. Oliver, A., Freixenet Bosch, J., Mart, R., Pont, J., Prez, E., Denton, E.R. and Zwiggelaar, R.: A Novel Breast Tissue Density Classification Methodology.,IEEE Transactions on Information Technology in Biomedicine, vol. 12, 55–65(2008).
2. McCormack, V.A. and dos Santos Silva, I.: Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. Cancer Epidemiology and Prevention Biomarkers, 15(6), 1159-1169(2006).
3. George, M., Rampun, A., Denton, E. and Zwiggelaar, R., Mammographic Ellipse Modelling Towards Birads Density Classification. In International Workshop on Digital Mammography, pp. 423-430. Springer (2016).
4. Obenauer, S., Sohns, C., Werner, C. and Grabbe, E., Impact of breast density on computer-aided detection in full-field digital mammography. Journal of digital imaging, 19(3), 258(2006).
5. Wolfe, J.N., Risk for breast cancer development determined by mammographic parenchymal pattern. Cancer, 37(5), 2486-2492(1976).
6. Kallenberg, M.G., Lokate, M., Van Gils, C.H. and Karssemeijer, N.,Automatic breast density segmentation: an integration of different approaches. Physics in Medicine and Biology, 56(9), 2715(2011).
7. Muhimmah, I. and Zwiggelaar, R., October. Mammographic density classification using multiresolution histogram information. In Proceedings of the International Special Topic Conference on Information Technology in Biomedicine, ITAB (2006).
8. Zwiggelaar, R., Muhimmah, I. and Denton, E.R.E., Mammographic density classification based on statistical grey-level histogram modeling. Proceedings of the Medical Image Understanding and Analysis (MIUA'05), pp.183-186(2005).
9. Hadjidemetriou, E., Grossberg, M.D. and Nayar, S.K., Multiresolution histograms and their use for recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 26(7), pp.831-847(2004).
10. Zhou, C., Chan, H.P., Petrick, N., Helvie, M.A., Goodsitt, M.M., Sahiner, B. and Hadjiiski, L.M., Computerized image analysis: estimation of breast density on mammograms. Medical physics, 28(6), pp.1056-1069 (2001).
11. He, W., Denton, E.R. and Zwiggelaar, R., Mammographic segmentation and risk classification using a novel binary model based bayes classifier. In International Workshop on Digital Mammography (pp. 40-47). Springer, Berlin, Heidelberg (2012).
12. Oliver, A., Freixenet, J. and Zwiggelaar, R., Automatic classification of breast density. In Image Processing, 2005. ICIP 2005. IEEE International Conference Vol. 2, pp. II-1258, IEEE (2005).
13. Mutra, M., Grgi, M. and Dela, K., Breast density classification using multiple feature selection. automatika, 53(4), pp.362-372 (2012).

14. Petroudi, S. and Brady, M., Breast density segmentation using texture. In International Workshop on Digital Mammography (pp. 609-615). Springer, Berlin, Heidelberg(2006).
15. Oliver, A., Llad, X., Marti, R., Freixenet, J. and Zwiggelaar, R., Classifying mammograms using texture information. In Medical Image Understanding and Analysis, Vol. 223, (2007).
16. Nanni, L., Lumini, A. and Brahnam, S., Local binary patterns variants as texture descriptors for medical image analysis. Artificial intelligence in medicine, 49(2), pp.117-125 (2010).
17. Nguyen, H.T. and Caplier, A., Elliptical local binary patterns for face recognition. In Asian Conference on Computer Vision (pp. 85-96). Springer, Berlin, Heidelberg (2012).
18. Ojala, T., Pietikinen, M. and Menp, T., Gray scale and rotation invariant texture classification with local binary patterns. In European Conference on Computer Vision (pp. 404-420). Springer, Berlin, Heidelberg (2000).
19. Suckling, J., Parker, J., Dance, D., Astley, S., Hutt, I., Boggis, C., Ricketts, I., Stamatakis, E., Cerneaz, N., Kok, S. and Taylor, P., The mammographic image analysis society digital mammogram database. In Exerpta Medica. International Congress Series ,Vol. 1069, pp. 375-378(1994).
20. Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., Parkin, D.M., Forman, D. and Bray, F.,. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. International journal of cancer, 136(5),(2015).
21. American College of Radiology , Illustrated Breast Imaging Reporting and Data System BIRADS, PA, Philadelphia:Amer. College of Radiol, 1998.
22. George, M., Denton, E. and Zwiggelaar, R., Mammogram Breast Density Classification Using Mean-Elliptical Local Binary Patterns. In International Workshop on Breast Imaging, (2018).