

Analysis of Shape-Based and Texture-Based Attributes in Classification of Mammographic Findings by Machine Learning Algorithms

Matheus C. de Melo¹, Andy A. T. Gajadhar¹, Hugo N. de Oliveira¹,
Arnaldo G. de Andrade e Silva¹, Leonardo V. Batista¹

¹Departamento de Informática – Universidade Federal da Paraíba (UFPB)
João Pessoa – PB – Brazil

{matheus.melo, andygaajadhar, hugoneves, arnaldo.gualberto, leonardo}@ci.ufpb.br

Abstract. *Breast cancer is the most frequent cancer type among women. We present a method of classification of nodules (malignant or benign) found in mammograms using shape-based attributes and texture-based ones. Firstly, we built a test database, then we segmented and extracted a Gray Level Co-occurrence Matrix (GLCM) from each mammographic finding and analyzed texture-based and shape-based attributes. Finally, classification was performed through machine learning algorithms. Tests reached a maximum Correct Classification Rate (CCR) of 93.75%, when performed with the Radial Basis Function Network algorithm. The largest area under the ROC curve (AUC), 0.964, was achieved with the Multilayer Perceptron algorithm.*

1. Introduction

Breast cancer is the second most common cancer type in the human species and the most frequent among women, corresponding to 22% of new cancer cases every year [Wang et al. 2014]. Breast cancer is also the second most fatal cancer type in women, causing less deaths than lung cancer only [Society 2014]. In this scenario, development and improvement of methods, such as ultrasound and infrared imaging, to detect this kind of disease on early stages is crucial. According to [Society 2014], early cancer detection with mammography saves lives and increases treatment options. Steady declines in breast cancer mortality among women since 1989 have been attributed to a combination of early detection and improvements in treatment.

Mammographic Computer-Aided Diagnosis (CAD) systems aim to segment breast image findings in order to separate them from normal breast tissue, extract diagnostic characteristics from them and apply machine learning techniques to provide a clinic suggestion or a second opinion. This may be used to assist medical professionals in detection and classification of potential nodules. The analysis of mammograms is a susceptible error-prone task and factors like fatigue and lack of experience of the professional may account for unnecessary biopsies or wrong diagnosis [Lan et al. 2012]. Thus, there has been a raise in the use of computational systems as support tools in the diagnostic process. In fact, only 15 to 30% biopsy cases evaluate that there are cancerous lumps [Ke et al. 2010] and 10 to 30% of cancers remains undetected [Fraschini 2011].

One of the possible goals of a mammographic CAD system is to separate malignant and benign findings. Benign masses often appear in mammograms with well

defined borders, while malignant ones generally present ill-defined borders with sharp ramifications [Braz Junior et al. 2013]. Therefore, most CAD systems look to explore differences between forms of masses often extracting from them shape-based attributes in order to aid the classification procedure. In [Weiqiang et al. 2008] was used shape and boundaries properties for masses classification, achieving an AUC of 0.9265. In [Boujelben et al. 2009] follows the same approach and constructs feature vectors based on the shape of the mass, resulting in a sensitivity of 94.2% with the Multilayer Perceptron algorithm.

As can be seen, shape characteristics alone may provide very good feature vectors. However, according to [Mudigonda et al. 2000], while attributes related to the shape of mammographic findings are effective in differentiating spiculated malignant tumors from circumscribed benign masses, texture-based attributes are essential when shape-based fail. Benign masses can eventually appear with spiculated border, while, on the other hand, malignant tumors may appear with rounded one. In this case, the use of texture-based features combined with shape-based ones helps to overcome the problem of similarity between masses and tumors shapes. In [Alto et al. 2005] was executed analysis procedures using Haralick [Haralick et al. 1973] texture features and shape features, stating their importance in a CAD system.

This paper proposes an investigation of shape-based attributes and texture-based ones to classify mammographic findings in malignant or benign. In order to achieve a good classification, it was necessary to use an effective segmentation, extracting relevant attributes and to use machine learning algorithms.

The next sections of this paper are organized as follows: Section 2, where we present databases, used features, overview of our method and the the softwares used; Section 3, where the tests results will be exposed and discussed for classification algorithms; and Section 4, where we present our conclusions about this work.

2. Methodology

Our method is composed by the following steps: segmentation of Regions of Interest (ROIs) from the Digital Database for Screening Mammography (DDSM) [Braz Junior et al. 2013, Duarte et al. 2013, Li et al. 2014], shape features extraction, texture features extraction, feature vector composition and classification.

2.1. Mammography Image Database

The test results, which will be presented in Section 3, were obtained using the DDSM database. DDSM is composed of about 2,600 cases, each one containing two images from each breast, corresponding to the craniocaudal and mediolateral oblique views [Braz Junior et al. 2013]. The database also contains metadata about the patient, the mammography records and the diagnosis confirmed by pathological and anatomical exam.

We chose to use the DDSM due to its large popularity in the mammography analysis literature [Braz Junior et al. 2013, Duarte et al. 2013, Li et al. 2014]. There are other mammography databases and they could have been used to evaluate the proposed method, but they are either not available for free or not available on the Internet.

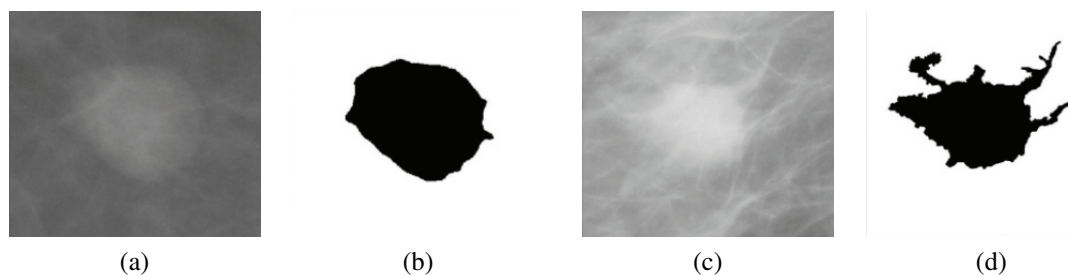


Figure 1. Segmentation of mammographic findings using ImageJ software. (a) Benign mammographic finding. (b) Segmented image of (a). (c) Malignant mammographic finding. (d) Segmented image of (c).

2.2. Segmentation

Image segmentation is a process that aims to segment an image in regions, contours, segments or in a set of pixels in order to change its representation, making it easier to analyze [Shapiro and Stockman 2001].

Segmentation of mammographic images is a complex process largely due to breast's anatomy, low contrast and non homogeneous exposition [Martí et al. 2007]. Normally, masses appear as brighter regions with well defined borders in a mammography. This feature permits further analysis of these images in order to identify possible ROIs.

For the segmentation stage of our method, we implemented an algorithm based on region growing. It was parametrized for usage on mass segmentation during the pre-processing phase of images. Part of the segmentation phase is shown on Figure 1. The segmented images were evaluated by one doctor, who marked them as satisfactory or not. During training and test phases, only images considered satisfactory were used, since this work's objective consists in analyzing the influence of shape-based attributes and texture-based ones in the classification process, not focusing on the segmentation algorithm.

Our segmentation method tries to slice the mammographic findings, keeping them as similar as possible to the norms adopted by the Breast Imaging-Reporting and Data System (BI-RADS).

2.3. Shape-based Attributes

Shape-based attributes of segmented masses were extracted using Matlab [Gonzalez et al. 2003]. The set of shape attributes is an important data to classification of mammographic findings [Weiqiang et al. 2008, Boujelben et al. 2009]. Following earlier proposals [Weiqiang et al. 2008, Boujelben et al. 2009], shape-based attributes used to form the vector of attributes are described as follows:

2.3.1. Solidity

Solidity represents the ratio between the area of the original object and the area of its convex contour. This attribute is calculated using Eq. 1, where A is the polygon area given by the contour of the object and A_C is the polygon area given for the convex contour of the object. The areas shown in Figure 2 (a) and (c) are the numerator and denominator, respectively.

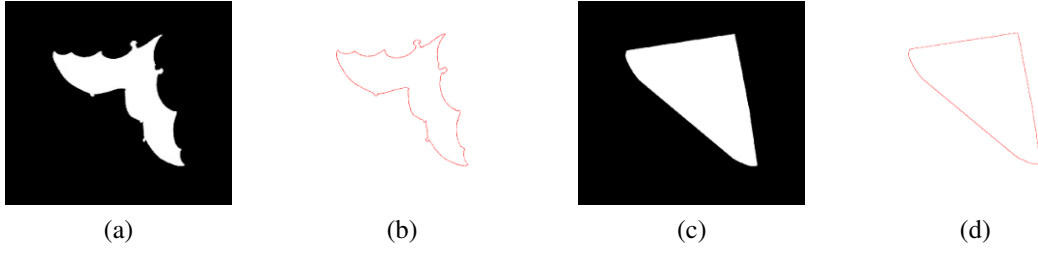


Figure 2. Graphical representation of the solidity. (a) Object. (b) Contour of the object. (c) Area delimited by the convex contour of the object. (d) Convex contour of the object.

$$Solidity = \frac{A}{A_C} \quad (1)$$

2.3.2. Eccentricity

The eccentricity of an image is defined as a scalar that specifies the eccentricity of an ellipse which has the same area of that image. This attribute can be calculated by Eq. 2, where c is the distance between the ellipse focus and a is its major axis. Eccentricity assumes values between 0 and 1 (0 meaning the ellipse is a circle and 1 meaning it is a line segment).

$$Eccentricity = \frac{c}{a} \quad (2)$$

2.3.3. Circularity

This attribute is computed as shown in Eq. 3, where A is the area of object and p is the perimeter of object with circular form has unitary circularity. A object which has an oblong form has circularity smaller than the unitary circularity.

$$Circularity = \frac{4\pi A}{p^2} \quad (3)$$

2.3.4. Convex area of segmented mass

The convex area attribute describes the number of pixels of the region formed by the smallest convex polygon that contains the segmented mass. Figure 2 (a) and (c) show an object and the smallest convex polygon that encompasses the object.

2.3.5. Aspect ratio

Aspect ratio is the ratio between the minor axis (R_{min}) and the major axis (R_{max}) and can assume values between 0 and 1. It is expressed in Eq. 4. When width and height of an

object have significantly different values, aspect ratio gets closer to 0 and when the object is a circle it is 1.

$$AspectRatio = \frac{R_{min}}{R_{max}} \quad (4)$$

2.4. Texture Attributes

Appearances of tumors and calcifications distort the image of the breast tissue. As a result, the textural features become significant in the detection of a tumor and furnishes a reliable basis for classification [Wang et al. 2014]. Texture attributes can be extracted from mammographic findings using Gray Level Co-occurrence Matrix (GLCM) [Alto et al. 2005, Ke et al. 2010]. Each element (i, j) in a GLCM matrix represents how often a grayscale pixel value i occurs horizontally adjacent to a pixel with value j . Following earlier proposals [Wang et al. 2014], the Haralick texture-based attributes used in this work are: contrast, correlation, energy, homogeneity and entropy.

2.4.1. Contrast

It reflects the extent of images' textures. If the texture is deep, the contrast will become larger to yield good visual effects. Otherwise visual effects will be unsatisfactory. A significant difference between gray levels means notable elements with big contrast [Wang et al. 2014]. This feature can be defined by Eq. 5.

$$Contrast = \sum_{i,j} |i - j|^2 p(i, j) \quad (5)$$

2.4.2. Correlation

It is the measurement of similarities on the directions of column and row [Wang et al. 2014]. Correlation may assume values in the interval between -1 and 1 , depending on how much the pixels are intrinsically similar or dissimilar. Calculations of correlation features are executed using Eq. 6.

$$Correlation = \sum_{i,j} \frac{(i - \mu_i)(j - \mu_j)p(i, j)}{\sigma_i \sigma_j} \quad (6)$$

2.4.3. Energy

Energy is the sum of squared elements in the GLCM. It reveals the fineness of texture and uniformity of a gray-scale distribution. If every value in GLCM is equal, values of energy are the lowest. In contrast, if all values in GLCM are uneven, the values of energy are larger [Wang et al. 2014]. Energy is defined by Eq. 7.

$$Energy = \sum_{i,j} p(i, j)^2 \quad (7)$$

2.4.4. Homogeneity

It reveals the homogeneity of texture in an image to show the extent of change in texture. The bigger the value is, the smaller the change in texture [Wang et al. 2014]. This feature can be defined by Eq. 8.

$$Homogeneity = \sum_{i,j} \frac{p(i, j)}{1 + |i - jx|} \quad (8)$$

2.4.5. Entropy

Entropy is a statistical measure of randomness that can be used to characterize the texture of a image. This feature is defined as in Eq. 9.

$$Entropy = - \sum_{i,j} p(i, j) \log(p(i, j)) \quad (9)$$

2.5. Proposed Method

Figure 3 shows an overview of the method presented in this paper. In the segmentation phase, ROIs in mammograms selected by expert doctors from DDSM are segmented using the previously mentioned region growing algorithm. In the attributes extraction step, the shape-based attributes are extracted from segmented images and the texture-based attributes are extracted from the GLCM of each image. For each image there is an vector of attributes comprising solidity, eccentricity, circularity, convex area of segmented mass, aspect ratio, contrast, correlation, energy, homogeneity and entropy.

After the extraction of the attributes, a set of attributes vectors is built to be used in the classification step. In the last step, the Weka tool is used, providing several machine learning algorithms [Witten et al. 2011]. We used the algorithms Radial Basis Function (RBF) Network, Multilayer Perceptron and K-Nearest Neighbors (K-NN) because are well known in the literature [Arora and Suman 2012, Duarte et al. 2013].

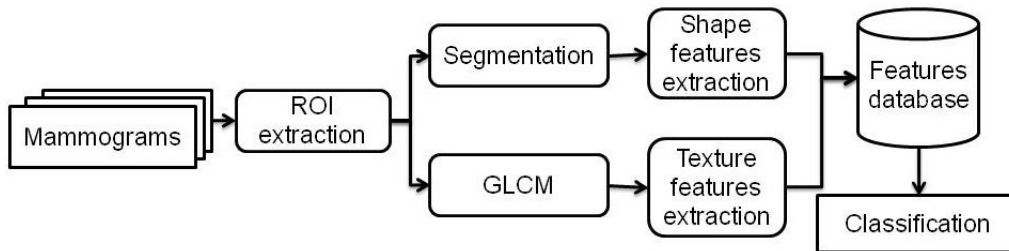


Figure 3. Sequence of performed steps.

For the experimental evaluation, each algorithm previously described was executed using the 10-fold cross-validation method [Acharya et al. 2012]. The following metrics were adopted to evaluate the classifiers: CCR, sensitivity and specificity [Wang et al. 2014, Braz Junior et al. 2013]. The sensitivity is defined, as the ability of a test to properly identify those objects that are malignant mammographic mass, while the specificity is defined as the ability of the test to correctly identify those that are benign mammographic mass [Braz Junior et al. 2013]. Generally, receiver operating characteristic (ROC) curves are used to demonstrate the performance of a CAD system too. The largest area under the curve (AUC) is defined as the evaluation criteria [Jalalian et al. 2013].

3. Results and Discussion

For the evaluation of our method, 160 images were selected randomly from DDSM, where 103 images are benign and 57 images are malignant. For each image, the steps shown in Figure 3 were executed. It is important to note that the training set is not included in the set used to test the classification.

Classification results are shown in Table 1, Table 2 and Table 3. Table 1 shows the classification result using the RBF Network algorithm; Table 2 shows the results using MultiLayer Perceptron; and Table 3 show the classification results using K-NN. It is important to note that the same image set is used for each machine learning algorithm.

The values of some algorithms' parameters were modified to generate an analysis of how much the parametrization affects the classification procedure. As shown in Table 1 to Table 3, the modified parameters are the number of clusters in the RBF Network algorithm, learning rate (LR) and momentum (M) in the MultiLayer Perceptron algorithm, and the K parameter in the K-NN algorithm. The values of the parameters were initialized with the default value of Weka, and they were modified while the CCR of the algorithm increased.

Table 1. RBF Network Algorithm

Clusters	CCR	Sensitivity	Especificity
1	90.63%	86.0%	93.2%
2	90.63%	86.0%	93.2%
3	93.75%	89.5%	96.1%
4	91.88%	89.5%	93.2%
5	90.63%	84.2%	94.2%

Table 2. Multilayer Perceptron Algorithm

LR	m	CCR	Sensitivity	Especificity
0.3	0.2	91.250%	89.50%	92.2%
0.3	0.4	91.875%	89.50%	93.2%
0.3	0.6	90.630%	87.70%	92.2%
0.4	0.2	91.875%	89.50%	93.2%
0.4	0.4	91.250%	87.70%	93.2%
0.4	0.6	90.000%	82.25%	94.2%

Table 3. K-NN Algorithm

K	CCR	Sensitivity	Especificity
1	89.4%	78.9%	95.1%
3	91.3%	87.7%	93.2%
5	92.5%	89.5%	94.2%
7	91.9%	86.0%	95.1%
9	93.1%	87.7%	96.1%
11	91.9%	82.5%	97.1%

The best CCR was 93.75%, achieved with the RBF Network algorithm with 3 clusters, as shown in Table 1. Besides CCR, we evaluated the classifiers by sensitivity and specificity metrics. RBF Network, Multilayer Perceptron and K-NN achieved the best sensitivity, 89.50%. The algorithm that achieved the best specificity, 97.1%, was the K-NN, with K = 11, as shown in Table 3.

Our results are competitive with other state of the art methods in the literature. In [Rangayyan et al. 1997] was reached 95% of sensitivity, but that work used only 39 images from MIAS database. In [Kallergi 2004] was achieved a sensitivity of 80% with only 60 images. However, a fair comparison of different methods is extremely difficult as they are evaluated on various databases [Jalalian et al. 2013].

In [Alam et al. 2013] and [Huang and Chen 2005], CAD systems were developed for detecting and diagnosing breast cancer where the AUC was 0.95 and 0.97, respectively. Figure 4 shows the ROC curves of RBF Network, Multilayer Perceptron and K-NN. RBF Network has AUC equal to 0.942, Multilayer Perceptron has AUC equal to 0.964 and K-NN has AUC equal to 0.961.

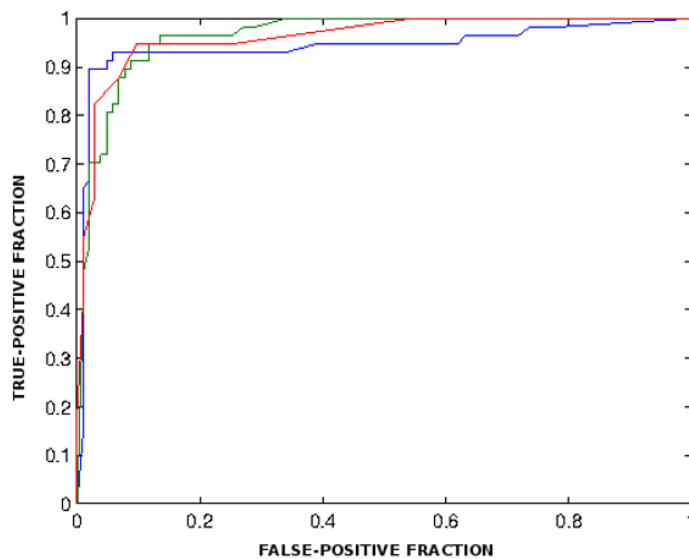


Figure 4. Roc curves of RBF Network (blue line), Multilayer Perceptron (green line) and KNN (red line) algorithms.

4. Conclusion

Using artificial intelligence algorithms, this paper showed how a CAD system can present good results in breast cancer diagnosis, providing a second opinion to doctors. Since early diagnosis are an important tool for the treatment of breast cancer it is expected that an automation of the detection method for this disease provide a better recovery chance for patients. Thanks to advances in technology, imaging processes have improved over the last decades, which may improve the performance of CAD systems, allowing them to contribute even more as a second opinion alternative.

The method presented in this article used shape and texture features in order to perform a classification process over 160 images belonging to the DDSM Database. As presented in Section 3, we obtained competitive CCRs using supervised machine learning algorithms. The higher CCR, 93.75%, was achieved with the RBF algorithm. Another measure used as performance rating was the AUC, which achieved a maximum value of 0.964 with the Multilayer Perceptron. Future improvements to the method presented in this paper may include the utilization of new texture and shape features; utilization of algorithms for attributes selection, image enhancement for improving the segmentation process and thereafter the subsequent steps of the algorithm; and, finally, the execution of the method on other mammograms databases.

References

- Acharya, U., Vinitha Sree, S., Saba, L., Molinari, F., Guerriero, S., and Suri, J. (2012). Ovarian tumor characterization and classification: A class of gynescan systems. In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pages 4446–4449. IEEE.
- Alam, S., Feleppa, E., Rondeau, M., Kalisz, A., and Garra, B. (2013). Computer-aided diagnosis of solid breast lesions using an ultrasonic multi-feature analysis procedure. *Bangladesh Journal of Medical Physics*, 4(1):1–10.
- Alto, H., Rangayyan, R. M., and Desautels, J. L. (2005). Content-based retrieval and analysis of mammographic masses. *Journal of Electronic Imaging*, 14(2):023016–023016.
- Arora, R. and Suman, S. (2012). Comparative analysis of classification algorithms on different datasets using weka. *International Journal of Computer Applications*, 54(13):21–25.
- Boujelben, A., Chaabani, A. C., Tmar, H., and Abid, M. (2009). Feature extraction from contours shape for tumor analyzing in mammographic images. In *Digital Image Computing: Techniques and Applications, 2009. DICTA'09.*, pages 395–399. IEEE.
- Braz Junior, G., Da Rocha, S. V., Gattass, M., Silva, A. C., and De Paiva, A. C. (2013). A mass classification using spatial diversity approaches in mammography images for false positive reduction. *Expert systems with applications*, 40(18):7534–7543.
- Duarte, Y. A. S., Do Nascimento, M. Z., and Oliveiras, D. L. L. (2013). Algoritmo de extração de textura baseado em wavelet e clbp para classificação de lesões em mamogramas. In *XIII Workshop de Informática Médica*, pages 174–183.
- Fraschini, M. (2011). Mammographic masses classification: novel and simple signal analysis method. *Electronics letters*, 47(1):14–15.

- Gonzalez, R. C., Woods, R. E., and Eddins, S. L. (2003). *Digital Image Processing Using MATLAB*. Pearson Prentice Hall, New Jersey, United States.
- Haralick, R. M., Shanmugam, K., and Dinstein, I. H. (1973). Textural features for image classification. *Systems, Man and Cybernetics, IEEE Transactions on*, (6):610–621.
- Huang, Y. and Chen, D. (2005). Support vector machines in sonography: application to decision making in the diagnosis of breast cancer. *Clinical imaging*, 29(3):179–184.
- Jalalian, A., Mashohor, S. B., Mahmud, H. R., Saripan, M. I. B., Ramli, A. R. B., and Karasfi, B. (2013). Computer-aided detection/diagnosis of breast cancer in mammography and ultrasound: a review. *Clinical imaging*, 37(3):420–426.
- Kallergi, M. (2004). Computer-aided diagnosis of mammographic microcalcification clusters. *Medical Physics*, 31(2):314–326.
- Ke, L., Mu, N., and Kang, Y. (2010). Mass computer-aided diagnosis method in mammogram based on texture features. In *Biomedical Engineering and Informatics (BMEI), 2010 3rd International Conference on*, volume 1, pages 354–357. IEEE.
- Lan, Y., Ren, H., Zhang, Y., and Yu, H. (2012). Similarity in mammography cad using cbir approach: A validation study. In *Intelligent Human-Machine Systems and Cybernetics (IHMSC), 2012 4th International Conference on*, volume 2, pages 373–376. IEEE.
- Li, X., Williams, S., and Bottema, M. J. (2014). Texture and region dependent breast cancer risk assessment from screening mammograms. *Pattern Recognition Letters*, 36:117–124.
- Martí, R., Oliver, A., Raba, D., and Freixenet, J. (2007). Breast skin-line segmentation using contour growing. In *Pattern Recognition and Image Analysis*, pages 564–571. Springer.
- Mudigonda, N. R., Rangayyan, R. M., and Desautels, J. L. (2000). Gradient and texture analysis for the classification of mammographic masses. *Medical Imaging, IEEE Transactions on*, 19(10):1032–1043.
- Rangayyan, R. M., El-Faramawy, N. M., Desautels, J. L., and Alim, O. A. (1997). Measures of acutance and shape for classification of breast tumors. *Medical Imaging, IEEE Transactions on*, 16(6):799–810.
- Shapiro, L. G. and Stockman, G. (2001). *Computer Vision*. Pearson Prentice Hall, New Jersey, United States.
- Society, A. C. (2014). Cancer effects & figures. *American Cancer Society*.
- Wang, Z., Yu, G., Kang, Y., Zhao, Y., and Qu, Q. (2014). Breast tumor detection in digital mammography based on extreme learning machine. *Neurocomputing*, 128(0):175 – 184.
- Weiqiang, Z., Xiangmin, X., and Wei, H. (2008). Shape and boundary analysis for classification of breast masses. In *Computational Intelligence and Design, 2008. ISCID'08. International Symposium on*, volume 2, pages 42–46. IEEE.
- Witten, I. H., Frank, E., and Hall, M. A. (2011). *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann.