

# Categorization of Carcinogenic Abnormalities in Digital Mastography Using Deep Learning Algorithms

Wafaa Razzaq

College of Nursing, University of Thi-Qar, Nasiriyah, Iraq  
Corresponding Author Email: [wafaa\\_razzaq@utq.edu.iq](mailto:wafaa_razzaq@utq.edu.iq)

**Abstract** - Cancer has been considered a serious disease for centuries and globally is one of the most prevalent conditions, which has been reinforced in recent years; breast cancer is the most common type of cancer in women and the second leading cause of cancer death worldwide. This mortality rate has been reduced thanks to various early detection techniques, mainly mastography and correct analysis. Currently, digital mastography can be computer assisted and this research takes as a reference the application of image preprocessing and various assembled algorithms in conjunction with Deep learning to improve the efficiency of detection. Through datasets generated and applying LogitBoost and AttributeSelectedClassifier algorithms in conjunction with Deep Learning, it analyzes the histogram of the images belonging to MIAS Dataset, obtaining competitive results of 88.37%.

**Keywords:** breast cancer, micro-calcification, classification, deep learning.

## 1. Introduction

Cancer is a process of growth and spread of cells in an uncontrolled way, a normal cell divides from a stem cell, this, in turn, divides once more; If during the process any of them are damaged or aged, they are replaced and the cycle restarts. However, in cancer cells they grow abnormally and survive to also divide again. Because the human body is made up of millions of cells, this process can appear anywhere, even the tumor often invades surrounding tissue and can metastasize to distant sites in the body.

Breast cancer is the most common form of cancer among women and is also considered to be associated with the highest mortality rate. Mammography has been considered the most viable option for its early detection due to the cost-benefit ratio it offers. The detection of cancer in its primary stage leads to an effective treatment in patients, however, in some countries Existence, there are few specialists in the area of identification of possible malignant micro-calcifications in digital mammograms, which leads to inaccurate predictions of this abnormality specific.

Micro-calcifications in the breasts (see Fig. 1.) are frequent findings in digital mammography, most of them caused by benign pathologies, which can become malignant, especially in carcinomas, whose detection is difficult due to its small size and lack of expertise in taking and considering accurate readings of shape, texture, size, and location. However, the detection rate can be improved with computer aided programs that implement Machine Learning classification algorithms.

## 2. Related Works

The following works show the recent development in the medical and social field of breast cancer, emphasizing the application of artificial intelligence techniques to provide solutions such as: prediction and classification. These investigations were based on artificial intelligence techniques to provide solutions to the aforementioned problems.

Article – Peer Reviewed

Received: 20 July 2023

Accepted: 17 September 2023

Published: 12 October 2023

**Copyright:** © 2023 RAME Publishers

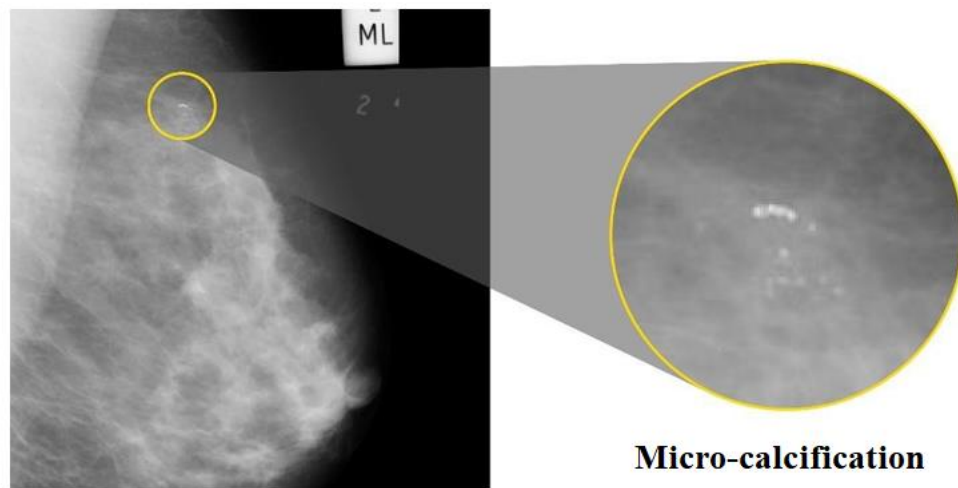
This is an open access article under the  
CC BY 4.0 International License.



<https://creativecommons.org/licenses/by/4.0/>

**Cite this article:** Wafaa Razzaq, "Categorization of Carcinogenic Abnormalities In Digital Mastography Using Deep Learning Algorithms", *International Journal of Computational and Electronic Aspects in Engineering*, RAME Publishers, vol. 4, issue 4, pp. 119-126, 2023. <https://doi.org/10.26706/ijceae.4.4.20231004>

Moradkhani et al, [6] based on the MIAS image extraction, cut and removed the additional information to then use a filter on the image, obtaining data to later be classified and obtaining a method that provides correct classification of 97%.



**Figure 1:** Micro-calcification present in a Digital Mammography.

Arafi et al, [7] implemented a method for cancer detection based on the Support Vector Machine, as a supervised learning technique to classify empirical data. Thus, they optimized the performance of the resulting classifier and obtained 94.74%.

Carreras et al, [3] to address the problem of classifying cancer abnormalities, used the MIAS dataset, in which they implemented a partial k-means clustering algorithm and as a result a single false positive, from the mdb026 image. The result was 95% confidence in the classification of cancer types in mammographic images.

Neto et al, [8] to automate the segmentation of masses in mammograms, particle swarm optimization (PSO) and graph clusters were used, achieving 95.2% effectiveness.

Arevalo et al, [9] used a hybrid approach where convolutional neural networks are used to learn the representation in a supervised way, obtaining a percentage of 82%.

Levy et al, [10] implemented a comprehensive deep learning model to classify pre-detected mammary masses from mammograms, using the AlexNet and GoogLeNet architecture, obtaining the highest accuracy percentage of 92.9% with the latter.

Gerazov et al, [11] applied deep learning methods to a time domain data set on homogeneous breast adipose tissue. They used convolutional neural networks as well as the Support Vector Machine input classifier for an accuracy of 93.44%.

Al-Masni et al, [12] used the computer-aided design (CAD) system for the detection of breast masses and cancer classification, implemented a Convolutional Neural Network, achieving an efficiency of 93.20% when classifying benign images, while malignant images were 78% effective, and their overall percentage of 85.52% to classify abnormalities.

Cruz et al, [5] the approach used was to evaluate the accuracy and robustness of a method based on deep learning to automatically identify the extension of the invasive tumor in the digitized images, this yielded a 75.86% of detected instances and a positive predictive value of 71.62%.

Camacho et al, [4] implemented the heuristic method based on data mining to extract essential information from mammographic images and transform them into patterns.

Pedraza et al., trained a convolutional neural network based on the GoogLeNet architecture, to develop the model, after which a cross-validation process was carried out. Thus, the algorithm provides an accuracy of 95.62% for a set of 5750 instances [13].

Dalmi et al, [14] used the Random forests assembled algorithm, combining it with a CNN, for the classification of lesions, obtaining 85% effectiveness in distinguishing the different types of abnormalities.

### 3. Motivation and Problems

The important points that motivated this research were: 1) According to the World Health Organization (WHO) [19], the female population has a high rate of breast cancer, and the first place of cancer mortality in the countries. 2) The uncertainty of a correct interpretation of digital mammograms for an efficient early diagnosis is still offered as a niche opportunity and 3) the scarcity of certified radiologists to be able to interpret mammography with malignant microcalcifications.

As has been mentioned, worldwide, cancer is one of the most serious diseases that has spread in recent years, seriously affecting the population, according to the WHO, it ranks second in cause of death, causing 8, 8 million deaths in 2015 [19].

Mammography is a tool used for the timely detection of breast cancer, it is the only study that has shown that its use reduces mortality from breast cancer by up to 30% [1]. But, it is essential that the evaluations of the mammograms are evaluated by certified experts to interpret these studies, which, according to [18],

Therefore, it can be deduced that early detection by mammography can reduce mortality from this disease.

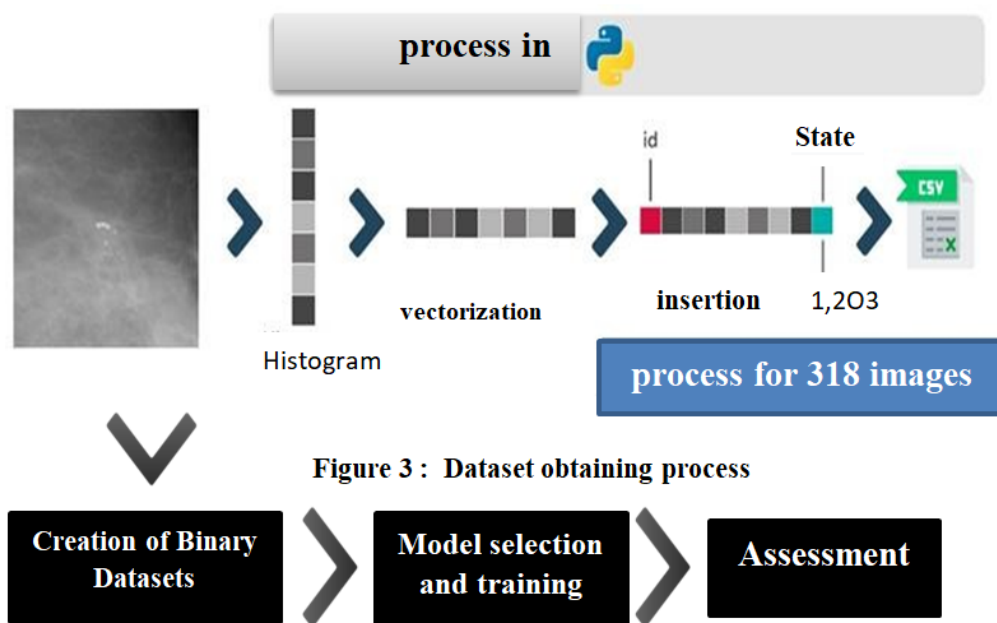


Figure 2. Methodology based on Moradkhani et al. [6].

### 4. Proposed Solution

The objective of this research is to identify microcalcifications that may exist within digital mammography and their probable status (Normal, Benign or Malignant) using assembled algorithms to minimize the margin of error, since competitive and highly reproducible detection rates can be achieved to facilitate timely detection of breast cancer. Section 5 analyzes in more detail the methodology that must be followed to solve the problem.

### 5. Methodology

For the development of this work, two methodologies are used to reach the final analysis of the results, the first is described as the general process to obtain the data of interest derived from the images belonging to the MIAS dataset, and the second to be able to carry out the evaluation of effectiveness according to established parameters that correspond to the initial tests.

### 5.1 Description of the overall process methodology

A classifier model is associated with pattern recognition. A brief description of the methodology is given below (see Fig. 2.):

- Original Database. It consists of the use of a database with 322 images, with dimensions of 1024x1024 pixels, of which 207 are normal mammograms, 60 have microcalcifications detected as benign and 51 malignant that present an atypical concentration of cells in the breast.
- Zoom 250px. A crop of the selected images is made with a height and width of 250 pixels, based on the methodology of [2] because images 226 and 239 have more than one group of microcalcifications and the size of the crop directly covers those present.
- 180 degree turn Non. The procedure is carried out for the images with odd numbers so that they have the same direction as the even images, this to have a series of images with similar characteristics, in addition to this, images 133, 134, 151 and 152 are discarded due to excess dimensions, this to have a series of elements with similar characteristics, in order to analyze and compare with other cropping systems that may exist in the future where a higher pixelage is handled.
- Obtaining the dataset. Through an algorithm developed in the Python language, the images are subjected to obtaining their histogram and subsequent vectorization of the same. Said algorithm adds an identifier (id) to the beginning of the vector, going through the information of the histogram. Also, at the end of the data obtained from the image, three identifiers are added to the end of the vector, which are: 1 (normal), 2 (benign) and 3 (malignant), this according to the data provided by the official page of the MIAS dataset. That is to say, 318 files with a .csv extension which are subsequently processed.
- Selection of characteristics. In obtaining the dataset, an id is integrated, 250 elements of the histogram, plus the Normal, Benign or Malignant state, that is, 252 characteristics, the last one acting as the main one.
- Creation of multiclass dataset. From the 318 files generated in obtaining the dataset, these are compiled into a single dataset, which contains 318 lines, through an enclosure in a folder and using native Python functions. However, in order to use it in subsequent analysis, it is necessary to insert the corresponding text headers as well as the conversion of the last values that the states contain, that is; 1 (replace with n), 2 (replace with b) and 3 (replace with m).
- Creation of binary datasets. Taking into account that there are three types of classes, these will be tested by placing versus between them, for this the multiclass dataset is divided into subdatasets with combinations of; Malignant-Benign which contains 111 vectorized histogram records, Normal-Benign, which will contain 267 lines and Normal-Malignant with 257 lines.

Model selection and training. For this, the sampling criteria are used: 1/3-2/3, Cross validation with 10 iterations and representative sample. The best positioned algorithms, that is, the ones that generated the best results, were LogitBoost+Dlj4Mlp, AttributeSelectedClassifier+Dlj4Mlp, FilteredClassifier+Dlj4Mlp and Staking+Dlj4Mlp, which are described below:

LogitBoost [17]. Also known as additive logistic regression, it optimizes the probability directly. From a practical point of view, LogitBoost uses a baseline regression scheme. This algorithm can be seen as a convex optimization, specifically, since an additive model of the form  $f =$

$$\sum_t a_t h_t, \text{ where the LogitBoost algorithm minimizes logistic losses by } \sum_i \log(1 + e^{-y_i f(x_i)}).$$

AttributeSelectedClassifier [15]. This algorithm uses ranking with InfoGainAttributeEval and Ranker lookup and can remove less useful attributes. This algorithm must be used to transform the data before passing it to your process.

FilteredClassifier [15]. This is a class that runs an arbitrary classifier on data that has passed a filter. Like a classifier, the filter's structure is based exclusively on the training data, and test instances can be processed by the filter without changing its structure. If there are unequal instance weights or attribute weights and the filter or classifier is unable to handle them, the instances and/or attributes are resampled with a replacement based on the weights before being passed to the filter or classifier (as appropriate).

Stacking [16]. Algorithm where there exists a set of n members. Each of these members is trained on a given set of training data. Members of this set can share the same type of classifier (homogeneous) or use different classifiers (heterogeneous). The diversity of data encourages among members for each member to generate different estimates.

The algorithm allows configuring the following layers to build more sophisticated architectures: Subsampling layer,

which subdivides groups of units from the mother layer by different strategies (mean, maximum, etc.); BatchNormalization, which applies the common batch normalization strategy on activations in the parent layer; OutputLayer, which generates classification/regression outputs, among a few others to enhance learning. Assessment. Having executed the classification algorithms one by one together with the DLj4Mlp algorithm, the best results were validated using the following metrics:

- confusion matrix,
- sensitivity, which is the ability to predict positive cases when they are really diseased or with the presence of microcalcifications, that is, the ability to detect disease in mammograms with signs of microcalcifications,
- specificity, which will provide information on negative cases of those that are truly healthy and the proportion of correctly identified healthy ones, that is, the ability to detect the disease in mammograms of healthy characteristics.

### 6. Experiments and Analysis of Results

The experiments were carried out on a computer with the following characteristics: Windows 10 Home Single Language, Intel(R) Core i7-6500U CPU 2.50 GHz, Ram 8.00 GB, 480Gb Solid State HDD, 64-bit operating system, x64 processor, image processing was performed with XnView software.

**Table 1.** Algorithms applied to the multiclass dataset.

Assembled(Meta)	Deep learning	2/3 – 1/3	CV-10	MR(45)
	Dlj4Mlp	47.1698	54.0881	48.5714
LogitBoost	Dlj4Mlp	65.1509	65.0945	63.4286
Stacking	Dlj4Mlp	65.1509	65.0945	63.4286
FilteredClassifier	Dlj4Mlp	17.9245	51.2579	65.1429

**Table 2.** Algorithms applied to the Normal-Malignant binary dataset.

Assembled(Meta)	Deep Learnig	2/3 – 1/3	CV-10	MR(45)
	Dlj4Mlp	66.2791	71.5963	60
AttributeSelectedClassifier	Dlj4Mlp	88.3721	80.9339	80
LogitBoost	Dlj4Mlp	82.5581	80.5447	77.4194
FilteredClassifier	Dlj4Mlp	86.9465	75.4864	32.2581

Algorithms were programmed in Spyder (Python 3.6) and classification in Weka 3.8.2.

To carry out the tests, assembled algorithms were applied to the generated datasets, that is, Multiclass, Normal-Malignant Binary, Normal-Benign Binary, and Benign-Malignant Binary, taking the criteria of 1/3 2/3, cross validation of 10 iterations and representative sample.

#### 6.1. Dataset 1: multiclass

For this test, a dataset containing 318 elements and three different classes was taken, for which we observed that, according to Table 1. The best algorithms are LogitBoost+Dlj4Mlp and Staking+Dlj4Mlp in the criteria 2/3 1/3 and cross validation of 10 iterations with equal percentages of 65.1509% and 65.0945% respectively, while for the representative sample criterion there is an algorithm with an effectiveness rate of 65.1429% despite not being benefited in the first two criteria.

#### 6.2. Dataset 2: normal-malignant binary

For this test, the dataset has 257 elements, with two classes Normal and Malignant. The best algorithm is AttributeSelectedClassifier+Dlj4Mlp, offering a result of 88.3721% effectiveness, see Table 2. However, there is the

presence of the algorithms LogitBoost+Dlj4Mlp and FilteredClassifier+Dlj4Mlp, which are also listed in Table 1. This can give us a reference for future tests with these datasets applying some different preprocessing.

**Table 3.** Algorithms applied to the Normal-Benign binary dataset.

Assembled(Meta)	Deep Learnig	2/3 – 1/3	CV-10	MR(45)
	Dlj4Mlp	69.6629	75.2809	62.6582
AttributeSelectedClassifier	Dlj4Mlp	78.6517	79.7753	81.0127
Staking	Dlj4Mlp	78.6517	77.5281	75.9494
FilteredClassifier	Dlj4Mlp	78.6517	74.5318	75.9494

**Table 4.** Algorithms applied to the Benign-Malignant binary dataset.

Assembled(Meta)	Deep Learnig	2/3 – 1/3	CV-10	MR(45)
	Dlj4Mlp	69.6629	75.2809	62.6582
LogitBoost	Dlj4Mlp	51.3514	54.5455	54.6512
AttributeSelectedClassifier	Dlj4Mlp	48.6486	51.8182	55.8114
FilteredClassifier	Dlj4Mlp	51.3514	50.9091	54.6512

### 6.3 Dataset 3: normal-benign binary

In this test, the dataset has 267 elements, with the Normal and Benign classes, according to Table 3. The most competitive algorithm is AttributeSelectedClassifier+Dlj4Mlp with a yield of 78.6517%, and looking back at Table 2, we can define that the indicated algorithm is good for identifying and performing an effective segmentation of a normal mammogram to one that may present a degree of lesion or presence of micro -calcifications.

It is important to note that the Staking+Dlj4Mlp and FilteredClassifier+Dlj4Mlp algorithms in the 2/3-1/3 criterion show a competitive effectiveness equaling the result of the first algorithm.

### 6.4. Dataset 4: benign-malignant binary

In the analysis of the dataset and counting with 110 resulting lines and two possible states Malignant and Benign, we have Table 4 in which the LogitBoost algorithm can be seen again in the first place, sharing effectiveness with the FilteredClassifier algorithm in the 2/3-1/3 criterion, which in retrospect to the previous analyzes we can declare that for data that may present confusion or multiclass it is better to apply the LogitBoost+Dlj4Mlp algorithm.

Comparing these results with previous works, the classifiers proposed in this work surpass the results obtained in Arevalo et al. [9] with 71.62 % effectiveness and Dalmi et al. [14] with an effectiveness rate of 85.52%. However, Arafı et al. [7] obtained 94.74%, Carreras et al. [3] got 95%. Net et al. [8] and Pedraza et al. [13] obtained a percentage of 95.2 and 95.62% respectively.

**Table 5:** Confusion matrix and values of sensitivity, specificity, false negative and false positive of the Normal-Malignant binary dataset.

	a=N	b=M		Sensitivity	0.957746479
Normal	68	3	71	Specificity	0.46666667
Evil one	8	7	15	False Negative	0.042253521
Total			86	False Positive	0.533333333

**Table 6.** Confusion matrix and values of sensitivity, specificity, false negative and false positive of the Normal-Benign binary dataset.

	a=B	b=N		Sensitivity	0.971428571
Normal	68	2	70	Specificity	0.894736842
Benign	2	17	19	False Negative	0.028571429
Total			89	False Positive	0.105263158

Lastly, it is worth mentioning Moradkhani et al. [6] who obtained an effectiveness of 97%.

Finally, the confusion matrices of the experiments with the best results are attached, that is, Binary Normal Malignant (Fig. 4) and Normal Benign (Fig. 5) taking the 2/3-1/3 criterion that offered the highest percentage as well as sensitivity and specificity together with false positive and false negative values.

## 7. Conclusion

After an exhaustive search and with a classification rate of 88.37% correct in a binary dataset, the following efforts were determined to expand this study in a second stage:

- The LogitBoost+Dlj4Mlp algorithm is generally good for classifying datasets that are multiclassified or that may present confusing data in the analysis of our histogram-derived data.
- The AttributeSelectedClassifier+Dlj4Mlp algorithm proved to be good in datasets whose characteristics can be separated substantially, that is, according to the results obtained, it can be used to define whether a mammogram presents normal characteristics or some microcalcification.
- The FilteredClassifier+Dlj4Mlp and Staking+Dlj4Mlp algorithm show good effectiveness rates in some sampling criteria, which can be used as reinforcements to define the final diagnosis.

It should be noted that in order to improve the results obtained, an image thresholding or segmentation process must be carried out, since in this work raw data and images were analyzed without prior processing.

## References

1. Cunha, P., Nunes, M., Patrocínio, A.: Breast density pattern characterization by histogram features and texture descriptors. *Research on Biomedical Engineering* 33(1), pp. 69–77 (2017) <https://doi.org/10.1590/2446-4740.07916>
2. Melladoa, M., Osab, M., Murillo, A.: Influence of digital mammography in the detection and management of microcalcifications. *Radiology: Official publication of the Spanish Society of Medical Radiology* 55(2), pp.142–147 (2013) <https://doi.org/10.1016/j.rxeng.2011.06.005>
3. Carreras, M., Martínez, M., Rosas, K.: Mass segmentation in digital mammograms. *Ambient Intelligence for Health* 9456(1), p. 110–115 (2015) <https://doi.org/10.1007/978-3-319-26508-7>
4. Camacho, S.: Heuristic Method for the Diagnosis of Breast Cancer based on Data Mining. *PGI Magazine - Research, Science and Technology* 1, pp. 97–101 (2014) <https://www.mdpi.com/2076-3417/9/21/4492#>
5. Cruz, A., Gilmore, H., Basavanahally, A.: Accurate and reproducible invasive breast cancer detection in whole-slide images: A deep learning approach for quantifying tumor extent. *Scientific Reports*, pp. 97–101(2017) doi: 10.1038/srep46450.
6. Moradkhani, F., Sadeghi, B.: A New Image Mining Approach for Detecting Micro-Calcification in Digital Mammograms. *Applied Artificial Intelligence* 31(5), p. 411–424 (2017) <http://dx.doi.org/10.1080/08839514.2017.1378082>
7. Arafí, A., Fajr, R., Bouroumi, A.: Breast cancer data analysis using support vector machines and particle swarm optimization. In: *Complex systems (WCCS), Second world conference*, pp. 1–6 (2014) <http://dx.doi.org/10.1109/ICoCS.2014.7060900>
8. Neto, O., Carvalho, O., Sampaio, W.: Automatic segmentation of masses in digital mammograms using particle swarm optimization and graph clustering. In: *International Conference on Systems, Signals and Image Processing (IWSSIP)*, pp.109–112 (2015) <https://doi.org/10.1016/j.jksuci.2018.01.004>
9. Arevalo, J., González, F., Ramos, R. et al.: Representation learning for mammography mass lesion classification with convolutional neural networks. *Computer methods and programs in biomedicine* 127(1), pp. 248–257 (2016) <https://doi.org/10.1016/j.cmpb.2015.12.014>
10. Lévy, D., González, F.: Breast mass classification from mammograms using deep convolutional neural networks. In: *CoRR* (2016) <https://doi.org/10.48550/arXiv.1612.00542>
11. Gerazov, B., Conceicao, R.: Deep learning for tumor classification in homogeneous breast tissue in medical microwave imaging. In: *IEEE (EUROCON'17) - 17th International Conference on Smart Technologies*, pp. 564–569 (2017) <http://dx.doi.org/10.1109/EUROCON.2017.8011175>
12. Al-masni, M., Al-antari, M., Park, J. et al.: Detection and classification of the breast abnormalities in digital mammograms via regional Convolutional Neural Network. In: *39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 1230–1233 (2017) <https://doi.org/10.1016/j.compbimed.2021.104248>
13. Pedraza, A., Serrano, I., Fernández, M., et al.: Automatic Diagnosis of HER2 with Deep Learning. *Google Scholar* (2016) <https://doi.org/10.1016/j.compbimed.2019.05.020>
14. Dalmi, M., Gubern, A., Vreemann, S. et al.: A computer-aided diagnosis system for breast dce-mri at high spatiotemporal

- resolution. *Medical physics* 43(1), p. 84–94 (2016) <http://dx.doi.org/10.1118/1.4937787>
15. Witten, I., Frank, E., Hall, M., et al.: *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufman (2017) <https://doi.org/10.1016/C2009-0-19715-5>
  16. Wolpert, D.: Stacked generalization. *Neural Networks Journal* 5, pp. 241–259 (1992) [https://doi.org/10.1016/S0893-6080\(05\)80023-1](https://doi.org/10.1016/S0893-6080(05)80023-1)
  17. Li, Q.: ABC-LogitBoost for Multi-Class Classification. Department of Statistical Science, Cornell University (2012) <https://www.mdpi.com/2227-9709/7/4/50#>
  18. CONACYT: Algorithm for early detection of breast cancer developed, [newsnet.conacytprensa.mx/index.php/documentos/36532-desarrollan-algoritmo-para-la-deteccion-precoz-de-cancer-de-mama](http://newsnet.conacytprensa.mx/index.php/documentos/36532-desarrollan-algoritmo-para-la-deteccion-precoz-de-cancer-de-mama) (2018) <https://www.mdpi.com/2411-9660/2/2/13#>
  19. WHO: Position paper on mammography screening, <http://www.who.int/cancer/publications> (2018) <https://doi.org/10.1016/j.ejca.2017.11.002>
  20. INEGI: Statistics about the World Cancer Day, [http://www.beta.inegi.org.mx/contenidos/saladeprensa/aproposito/2018/cancer2018\\_Nal](http://www.beta.inegi.org.mx/contenidos/saladeprensa/aproposito/2018/cancer2018_Nal) (2018)