

# Multiclassification of mammographic density levels with genetic programming

Ninderlisbhet Vargas Cardenas<sup>1</sup>, María Guadalupe Sánchez Cervantes<sup>2</sup>, Daniel Fajardo-Delgado<sup>3</sup>

<sup>1,2,3</sup>(Department of Systems and Computing, Instituto Tecnológico de Ciudad Guzmán, México)

**ABSTRACT:** Breast cancer is the most frequent in women, as well as cervical and skin cancer. Medical research for the prevention of breast cancer has shown that breast density is a strong indicator of cancer risk. The density can be evaluated through the classification proposed by the American College of Radiology (ACR). The objective of this article is to show the results obtained from the multiclassification of mammographic density levels using genetic programming. The multiclass classification starts from a set of texture characteristics of the mammographic images. Various methods of feature extraction and classification have been implemented. The features of the images are the input for the classifiers. For the experiments, the INbreast mammography database was used. The results show good classification and will help doctors to make a reliable diagnosis of cancer risk.

{vargascardenasn, magusace, dfajardod}@gmail.com

**KEYWORDS**—Breast Density, Genetic Programming, Multiclassification.

## I. INTRODUCTION

Breast cancer is a disease with a complex natural evolution that, despite the advances of modern oncology, is the second cause of death by neoplasia in women worldwide, with nearly 500 thousand deaths each year, from which 70% occurs in developing countries [1]. Breast cancer can be detected through microcalcifications and mammographic breast densities. Approximately 55% of non-palpable breast cancers present with visible microcalcifications. For better study over time, some methods of mammary density classification have been implemented, such as the "TNM" system of the American Joint Committee on Cancer (AJCC), and the BIRADS classification system (Breast Imaging Reporting and Data Systems). However, the detection of breast cancer with conventional methods applied by specialists has not been enough, now some dynamic methods have been implemented. In the process of providing a solution, emphasis has been placed on various classification techniques and, based on this, [2] they investigated the evolutionary algorithms of Genetic Programming (GP) and Learning Classification Systems (LCS), that work in such a way that they allow to classify the mammographic images by their levels of density under a binary classification environment. However, the proposed method allows a multiclass classification of mammographic density levels. To carry out this classification through GP, there are a series of steps to follow, such as the pre-processing of the image, the extraction of the characteristics and the implementation of GP for the multiclassification.

## II. ENVIRONMENT

This research is based on various concepts such as the ones mentioned below, which are important for the understanding of this article.

### 2.1. Binary classification and multiclassification

According to [3], the binary classification is based on classification algorithms that assign a set of labels to a bifurcated categorical target value and mention the constant implementation of machine learning, because this often involves ordering situations between several categories. However, [4] he indicates that when you have multiple classes, the main problem can be broken down into several binary classification tasks and is one of the common ways of solving multi-class classification problems. In addition to this, there are different methods of multiclass classification such as One Versus All (OVA), All Versus All (AVA), Correction Error Output Coding (ECOC) and the Generalized Codification.

### 2.2. Genetic programming

The Genetic programming (GP) is based and inspired by the Darwinian theory of evolution, which is why it is based on the following aspects:

- The population represents the group of individuals that are candidate solutions for solving problems.
- The method of inheritance by a generation that is implemented with evolution.
- The selection method (Fitness) of the best candidate solutions based on the search space.
- The probability of crossing between individuals, which consists basically of the selection of a random point in each of the two individuals to cross and only exchange that characteristic between them.
- The probability of mutation of an individual, where from a single individual a random point is selected that will be modified without following a defined pattern but, respecting the structure of the tree.

## III. RELATED WORK

There are several works that precede the present investigation, however, each of them has a specific orientation as you can see in this section. Table 3.1 shows some works related to the classification of levels of mammographic density based on the MIAS data set, in addition, the number of images used, the number of classes to be classified, the type of ROI (Region of Interest), the classifiers implemented and the percentage of accuracy.

Researchers	Details				
	Images	Class Número	ROI Size	Classifier	Accuracy (%)
[5]	265	3	SBT (segmented Breast Tissue)	kNN	63.0
[6]	322	3	SBT (segmented Breast Tissue)	SVM	91.3
[7]	321	3	SBT (segmented Breast Tissue)	DAG-SVM	77.5
[8]	43	3	SBT (segmented Breast Tissue)	SVM	95.4
[9]	322	3	SBT (segmented Breast Tissue)	SVM	84.1
[10]	42	3	SBT (segmented Breast Tissue)	KSFD	94.4
[11]	322	3	512 X 384	IB1	82.0
[12]	320	3	300 X 300	SVM	77.1
[13]	322	3	200 X 200	SVM	86.3
[14]	322	3	200 X 200	SVM	87.5
[15]	322	2	200 X 200	SMO-SVM	96.4
[16]	212	2	200 X 200	kNN	97.2
[17]	322	2	200 X 200	SVM	94.4
[18]	322	2	200 X 200	kNN	95.6
[19]	322	2	200 X 200	kNN	96.2
[20]	322	2	200 X 200	SSVM	94.4

Table 3.1 Previous research for classification of densities in mammograms.

Once each of the works in Table 3.1 is reviewed, it is understood that they are focused on the extraction of the characteristics rather than on the classification, because the desired results of the classification depend on a correct extraction. Likewise, each of these works is mainly based on the MIAS data set and in some cases on DDSM, but so far none report results on the use of the INbreast database.

An important precedent for the classification of mammographic density levels is that of [2], who investigated the GP and Learning Classification Systems, which work in such a way that they allow to classify the mammographic images by their density levels. They studied in detail the utilities of different local binary patterns and statistical characteristics using GP and LCS using four conventional classification techniques, such as Naive, decision trees, nearest neighbor to K, and support vector machines.

#### IV. MULTICLASSIFICATION OF MAMMOGRAPHIC DENSITY LEVELS

The multiclassification is developed under a genetic program, which is defined by a set of individuals in the form of a tree, which make up a population. The population can reproduce, cross and mutate in each generation to preserve the best characteristics, in such a way that they can improve the descendants progressively in each epoch. On the other hand, an implementation of the Stratified Cross Validation is made to maintain a load balance between the samples per class.

##### 4.1. Programa genético

The multiclassification concerning the genetic program considers that individuals are evaluated as a regression tree from which four subtrees emerge, where each one of them represents a classifying model for each level of mammographic density, as shown in the Fig. 4.1. The objective of assessing individuals in the form of a tree allows that, at the end of the evolution, in addition to being generated in a single execution, the solution is obtained from the root node of the best tree, and then used in the multiclassification. In other words, when the evolution of individuals is finished, they can evaluate themselves in the root node to determine a single classifier model. Within this point, each of the subtrees is composed of a set of leaves that correspond to a set of terminals  $T = \{IN0, IN1, IN2, \dots\}$ , denoted by random numbers between 0 and 1, and that they contain the characteristics extracted from the mammographic images. On the other hand, the internal nodes are composed of the set of primitives or functions  $F = \{+, -, \times, \text{mydiv}, \text{mypower2}, \text{cos}, \text{sin}, \text{mysigmoid}, \text{switch}\}$ . The evaluation of each tree returns a floating value that represents each of the different classifications.

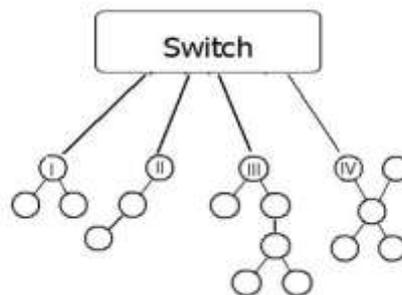


Fig. 4.1 Representation of the individual classifier.

Under the framework of the genetic program, the multiclassification is developed according to the following evolutionary parameters, which correspond to the crossover rate  $CXPB = 0.8$ , mutation  $MUTPB = 0.15$ , number of generations  $NGEN = 50$ , frequency  $FREQ = 10$  and number of Folds  $n\_splits = 10$  that are also used in the implementation of the parallelization with the model by islands. Also, considering that the algorithm implemented in this project is based on  $\mu + \lambda$  that evaluates individuals with an invalid fitness, it is thus necessary to define the evolutionary parameters such as,  $MU = 100$  that represents the number of individuals to be selected in each generation and  $LAMBDA = 200$  as the number of children produced in each generation. Starting from  $LAMBDA$ , the descendants are generated by a variation function that performs crossing, mutation

or reproduction, and in turn, manages the number of generations defined to return a tuple with the final population of evolution finally.

**4.2. Cross-validation**

In cross-validation, the data set is randomly divided into subsets called in this environment such as k-Folds themselves that are mutually exclusive and most of the same size. For this work we use k = 10 Folds, following the assessment of [21], which will be trained and tested k times to obtain a better precision when classifying, precision that is given mainly in the validation as the division of the total of correct classifications by the number of entities in the data set.

The implementation of the stratified cross-validation allows, during the process of splitting the data to define the test and training sets, these are reorganized in such a way that each of the Folds represents a subset of data with at least one characteristic of each type, to avoid the bias of the information that could well occur in a regular cross-validation. Fig. 4.2 shows the distribution of the data through the stratified cross-validation.

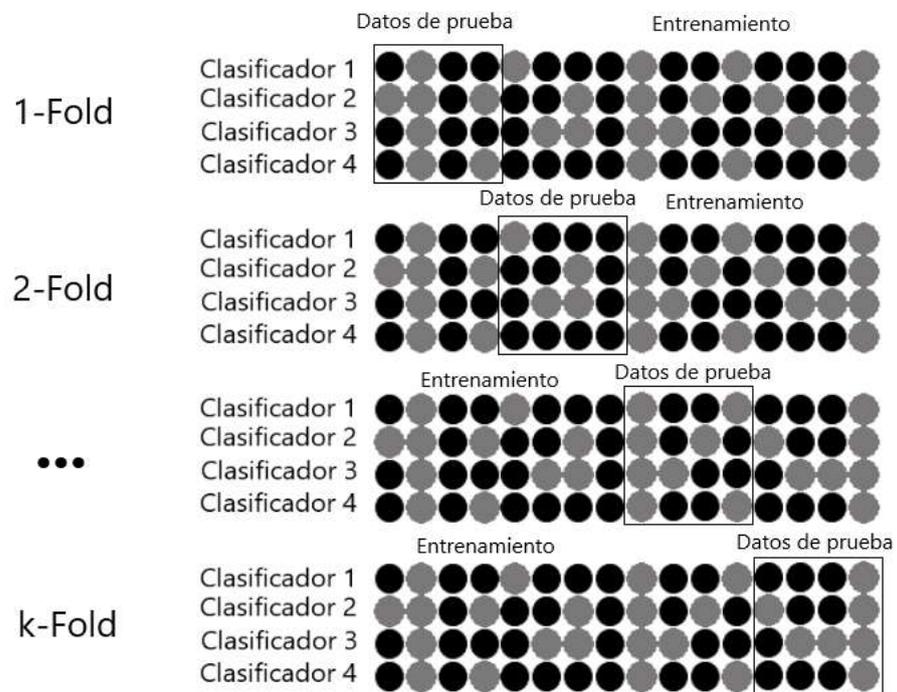


Fig. 4.2 Diagram of the stratified cross-validation

**4.3. Evaluation function**

The evaluation is defined as a maximization problem because it consists of the sum of coincidences; therefore, the individual is a tree to maximize. For the Fitness function, the expression of the tree is transformed into a function that evaluates whether the input data coincides with the training data in each level and, when corresponding, an increase is generated to the variable designated for the counting of coincidences of the individuals with a base to the confusion matrix.

Specifically, the evaluation of the matches is based on the concept of the F1-Score precision measure, which is directly linked to two metrics, such as the accuracy and recall that are part of the statistical measures of performance for the classification.

Finally, once the size of the population, the number of generations, the probabilities of mutation and crossing are defined, the execution of the genetic program is done for all the individuals waiting for the best classifier. Experimentally, other classifiers are also executed to generate a comparison, such as Nearest Neighbors, Decision Tree, Random Forest, SVM, and Naive Bayes.

## V. RESULTS AND DISCUSSION

As proof of the work developed, the representation of the results is fundamental, which is expressed based on the multiclassification of mammographic density levels with genetic programming, the values obtained in comparison with the other classifiers.

### 5.1. Results of the multiclassification

Throughout this section, we will be able to appreciate the results of the multiclassification of mammographic density levels with genetic programming for the INbreast data set. Next, Table 5.1 and Graph 5.1 show the performance of our multiclass classifier in comparison with the performance of other classifiers for the INbreast data, and the result is exposed based on the measures of F1-score, Precision, Accuracy, Recall and Confusion matrix.

Classifier	F1-Score	Precision	Accuracy	Recall	Confusion_Matrix
<b>Our_GP</b>	0.4599	0.4473	0.4736	0.4736	[7 5 1 0] [5 6 3 0] [1 3 5 0] [1 0 1 0]
<b>Nearest Neighbors</b>	0.2532	0.2526	0.2631	0.2631	[3 7 2 1] [5 5 3 1] [3 5 1 0] [0 1 0 1]
<b>Decision Tree</b>	0.2648	0.2286	0.3157	0.3157	[8 2 3 0] [6 0 6 2] [3 2 4 0] [0 1 1 0]
<b>Random Forest</b>	0.4225	0.4621	0.4473	0.4473	[8 5 0 0] [7 7 0 0] [2 5 2 0] [0 1 1 0]
<b>SVM</b>	0.1983	0.1357	0.3684	0.3684	[ 0 13 0 0] [ 0 14 0 0] [0 9 0 0] [0 2 0 0]
<b>Naive Bayes</b>	0.4667	0.4879	0.5263	0.5263	[13 0 0 0] [7 5 2 0] [2 5 2 0] [2 0 0 0]

Table 5.1 Results of the multiclassification for the INbreast data set

Based on Table 5.1, it is possible to see that, in the best case of cross-validation, Our\_GP is only below Naive Bayes, but it is just above the rest of the classifiers, a condition that makes it competent to solve the multiclassification problem of mammographic density levels.

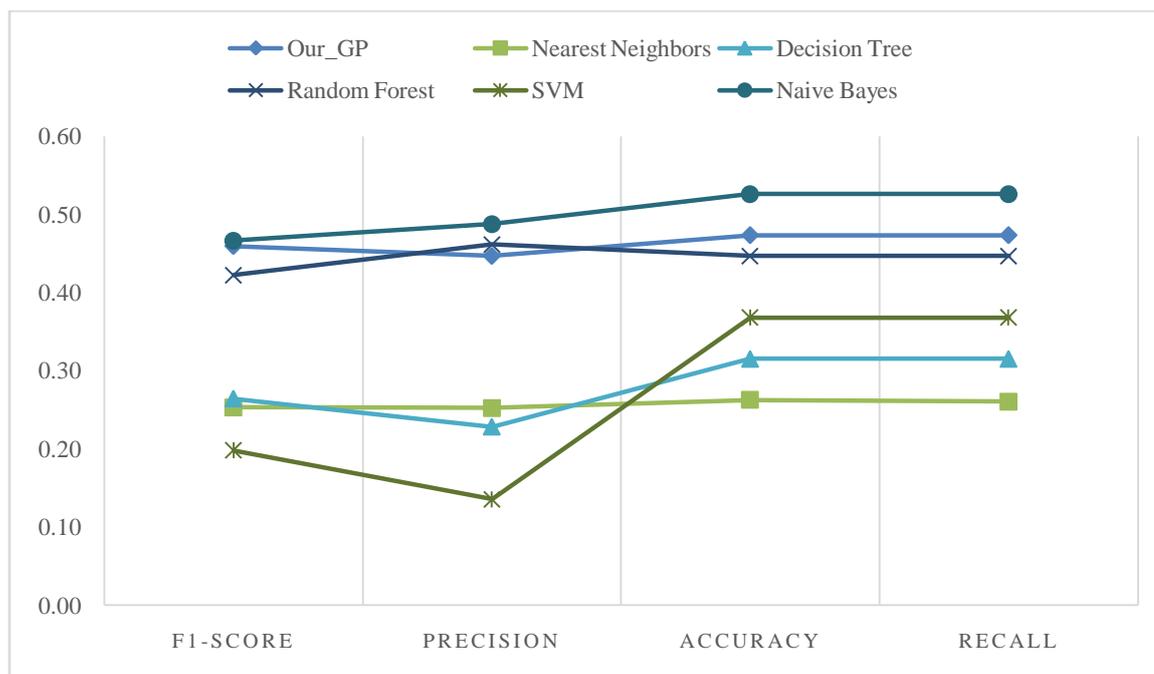


Fig. 5.1 Graph of results of the multiclassification for INbreast data

Graphically in Fig. 5.1, we can see the performance of Our\_GP based on the information in Table 5.1, specifying that Our GP is among the best classifiers in question.

## VI. CONCLUSION

Throughout this article, the proposed method for the multiclassification of mammographic density levels with genetic programming was presented, starting with the use of a vector of image characteristics for the effective multiclass classification. The result shows a good performance with respect to other classifiers.

Multiclassification has been carried out with the INbreast database.

The INbreast images are images that until now have not been implemented within any multiclassification method. Our multiple classification is competitive with Naive Bayes classifier. In addition to this, it is necessary to mention that it is expected to improve the results obtained for this type of images in future work because the genetic programming applied to the multiclass classification represents an opportunity to compete and overcome other classifiers for this specific problem.

## REFERENCES

- [1] Dr. Jesús Cárdenas Sánchez, Dr. Enrique Bargalló Rocha, Dra. Aura Erazo Valle, Dr. Eduardo Maafs Molina, Dra. Adela Poitevin Chacón, Primer Consenso Nacional sobre Tratamiento del Cáncer Mamario, Rev Inst Nal Cancerol (Mex)2013;41(3):136-145.
- [2] Burling-Claridge, F., Iqbal, M., Zhang, M. (2016). Evolutionary Algorithms for Classification of Mammographic Densities using Local Binary Patterns and Statistical Features. Proceedings of 2016 IEEE Congress on Evolutionary Computation (CEC 2016), 3847–3854. doi: <https://doi.org/doi:10.1109/CEC.2016.7744277>
- [3] Frank Eibe, Hall Mark. (2001). A Simple Approach to Ordinal Classification. Lecture Notes in Computer Science. 2167. 145-156. 10.1007/3-540-44795-4\_13.
- [4] Aly, M. (2005). Survey on multiclass classification methods. Survey on Multiclass Classification Methods, Technical Report, Caltech, USA, 2005.
- [5] Blot L, Zwiggelaar R (2001) Background texture extraction for the classification of mammographic parenchymal patterns. In: Proceedings of conference on medical image understanding and analysis, 145–148, July 2001.
- [6] Bosch A, Munoz X, Oliver A, Martí J (2006) Modeling and classifying breast tissue density in mammograms. In: Proceedings of IEEE computer society conference on computer vision and pattern recognition, New York, USA, 15552–15558, June 2006.

- [7] Muhimmah I, Zwiggelaar R (2006) Mammographic density classification using multiresolution histogram information. In: Proceedings of 5th international IEEE special topic conference on information technology in biomedicine, Ioannina, Greece, 1–6, Oct 2006.
- [8] Subashini TS, Ramalingam V, Palanivel S (2010) Automated assessment of breast tissue density in digital mammograms. *Comput Vis Image Und* 114:33–43.
- [9] Tzikopoulos SD, Mavroforakis ME, Georgiou HV, Dimitropoulos N, Theodoridis S (2011) A fully automated scheme for mammographic segmentation and classification based on breast density and asymmetry. *Comput Meth Prog Bio* 102:47–63.
- [10] Li JB (2012) Mammographic image-based breast tissue classification with kernel self-optimized fisher discriminant for breast cancer diagnosis. *J Med Syst* 36:2235–2244.
- [11] Muštra M, Grgić M, Delac K (2012) Breast density classification using multiple feature selection. *Automatika* 53:362–372.
- [12] Silva WR, Menotti D (2012) Classification of mammograms by the breast composition. In: Proceedings of the 2012 international conference on image processing, computer vision, and pattern recognition, Athens, 1–6.
- [13] Kriti, Virmani J, Thakur S (2016) Application of statistical texture features for breast tissue density classification. In: Awad AI (ed) *Image feature detectors and descriptors*. Springer, Switzerland.
- [14] Kriti, Virmani J (2016) Comparison of CAD systems for three class breast tissue density classification using mammographic images. In: Dey N (ed) *Medical imaging in clinical applications*. Springer, Switzerland.
- [15] Sharma V, Singh S (2014) CFS–SMO based classification of breast density using multiple texture models. *Med Biol Eng Comput* 52:521–529.
- [16] Sharma V, Singh S (2015) Automated classification of fatty and dense mammograms. *J Med Imaging Health Inform* 5:520–526.
- [17] Kriti, Virmani J, Dey N, Kumar V (2015) PCA-PNN and PCA-SVM based CAD systems for breast density classification. In: Hassanien AE, Grosan C, Tolba MF (eds) *Applications of intelligent optimization in biology and medicine*. Springer, Switzerland.
- [18] Kriti, Virmani J (2015) Breast density classification using Laws' mask texture features. *Int J Biomed Eng Technol* 19:279–302.
- [19] Virmani J, Kriti (2015) Breast tissue density classification using wavelet-based texture descriptors. In: Proceedings of 2nd international conference on computer and communication technologies (IC3T-2015), Hyderabad, India, 539–546, Jul 2015.
- [20] Zemmal N, Azizi N, Sellami M, Dey N (2016) Breast abnormalities diagnosis based on transductive SVM a mutual information-based feature selection. In: Proceedings of international conference for engineering and sciences, Barcelona, 23–27 July 2016.
- [21] Markatou, M., Tian, H., Biswas, S., Hripcsak, G. (2005). Analysis of variance of cross-validation estimators of the generalization error. *J. Mach. Learn. Res.*, 6:1127–1168 (electronic). MR2249851.