Diagnosis of microcalcifications using Case-Based Reasoning and Genetic Algorithms

Xavier Llorà[a] Elisabet Golobardes[a] Maria Salamó[a] Joan Martí[b]

[a]Computer Science Department, Enginyeria i Arquitectura La Salle, Universitat Ramon Llull
Passeig Bonanova 8, 08022 Barcelona, Catalonia, Spain
{xevil,elisabet,mariasal}@salleURL.edu
[b]Institute of Informatics and Applications, Universitat de Girona
Avda. Lluís Santaló s/n, 17071 Girona, Catalonia, Spain
joanm@eia.udg.es

Abstract

This paper describes the application of Case-Based Reasoning and Genetic Algorithms to diagnose a mammogram in cancerous or not. Our work is based on a previous one, which detects a set of microcalcifications that appear in a mammogram. This paper is focused on the automatic classification of the different sets of microcalcifications using machine learning techniques. Our goal is to improve the previous results obtained and propose new points of view into the Case-Based Reasoning and the Genetic Algorithms usage.

Keywords: Machine Learning, Case-Based Reasoning, Genetic Algorithms, Diagnosis, Human-Medicine and Healthcare

1 Description of the problem

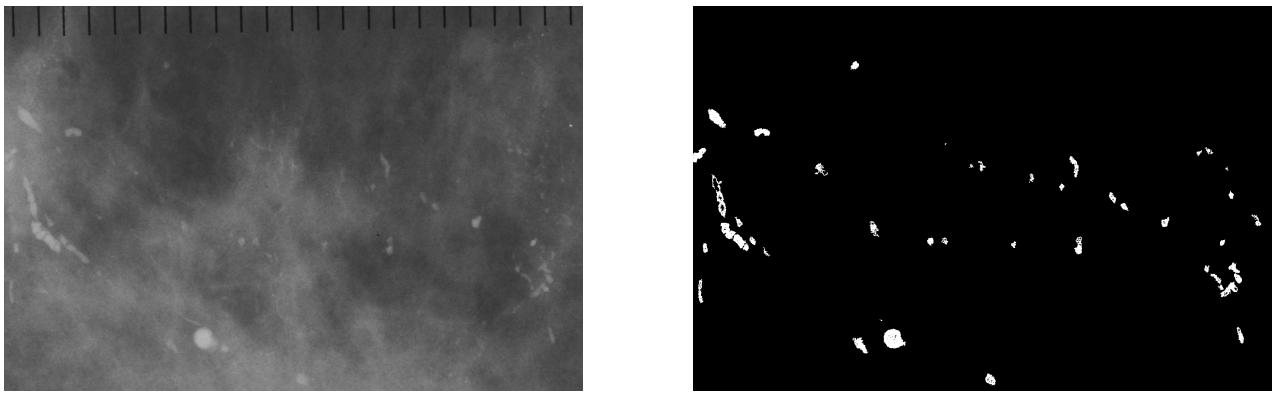
The incidence of breast cancer varies greatly among countries, but recent statistics show that every year 720.000 new cases will be diagnosed world-wide. Breast cancer screening has been proved as a good practical tool for detecting and removing breast cancer prematurely and also for increasing the survival percentage in women [15]. However, a low percentage of women that suffers breast cancer can be detected using mammography methods. Therefore, it is necessary to develop new strategies to detect breast cancer formation in early stages.

The main idea is to introduce CAD systems (Computer Aided Diagnosis) in the preliminary diagnosis. The work presented in this paper is based on microcalcifications. A microcalcification ($Ca++$) usually appears, in the mammographies, as small, bright, arbitrarily shaped regions on the large variety of breast texture background. Thus their analysis and characterisation are performed throughout the extraction of features and visibility descriptors by means of several image processing techniques [12], such as grey-level image analysis, signal processing algorithms or morphological methods.

The main guidelines of the CAD system used can be described as: (1) digitising the mammography image, (2) processing the image, (3) doing microcalcification identification and feature extraction, and finally (4) using machine learning techniques in order to diagnose automatically the processed mammography. Figure 1 shows an original mammographic image and the clustered microcalcifications after segmentation.

This paper focuses its work on the last part of the CAD system. We present here two machine learning techniques, *Case-Based Reasoning* (CBR) and *Genetic Algorithms* (GA), applied to the automatic diagnosis of the processed mammography images. The previous image processing phases can be found in [9]. Both systems use as input information, a set of processed images (or samples). Each sample contains the description of several Ca++ present in the image.

For each of these microcalcifications there are 23 real valued features related to the shape of individual microcalcifications (see the table 1). Shape of individual microcalcifications as long as shape of the cluster and number of microcalcifications have been pointed out as the three main indicators for malignacy. In other words, the input information used is



(a) Original

(b) Segmented

Figure 1: Digitisation and segmentation processes transform the original grey-level image into a binary image, where the background tissue has been removed and clustered microcalcificacions appear.

a set of $m \times 23$ real valued matrixes, where m (we want to remark that the number of Ca++ (m) can be different for each mammogram) is the number of Ca++ present on the image. Using this input information, CBR and GA play the backend role of diagnosing a sample into one of the following classes: *malign*, *benign*, *do not know*.

The following two sections describe the machine learning techniques used, *Cased-Based Reasoning* and *Genetic Algorithms*, and the systems that implement those techniques.

2 CaB-CS: Case-Based Classifier System

Case-Based Reasoning (CBR) integrates in one system two different characteristics: machine learning capabilities and problem solving capabilities. CBR uses a similar philosophy to that which humans sometimes use: it tries to solve new cases (examples) of a problem by using old previously solved cases [10]. The process of solving new cases contributes with new information and new knowledge to the system. This new information can be used for solving other future cases. The basic method can be easily described in terms of its four phases [1]. The first phase *retrieves* old solved cases similar to the new one. In the second phase, the system tries to *reuse* the solutions of the previously retrieved cases for solving the new case. The third phase *revises* the proposed solution. Finally, the fourth phase *retains* the useful information obtained when solving the new case.

In a Case-Based Classifier System, it is possible to simplify the reuse phase. Reuse can be done by classifying the new case with the same class as the most similar retrieved case.

2.1 CaB-CS and extensions

We use CaB-CS (Case-Based Classifier System) [2, 3, 5] and some extensions [11]. CaB-CS allows the user to test several variants of CBR. The variants presented in this paper are focused on the *retrieval* phase (phase 1).

Phase 1 retrieves the *most similar* case or cases to the new case. Obviously, the meaning of *most similar* will be a key concept in the whole system. Similarity between two cases is computed using different similarity measures.

For the problem that we present in this paper, we use the main similarity functions of the CaB-CS [5], and some extensions presented in [11]. The different similarity functions can be classified in two groups: 1) Similarity functions based on the distance concept: Minkowski's metric (Hamming, Euclidean and Cubic distance), Clark's distance, and Cosine distance; and 2) Similarity functions based on spheres: Sphere of Proximity, MinMax Sphere and Mean Sphere (these functions were proposed by Golobardes in [5]).

2.2 Similarity functions based on distance

The most used similarity function is the *Nearest Neighbour (NN)* algorithm, which computes the similarity between two cases using a global similarity

Feature	Description
Area	The number of pixels in the microcalcification
Perimeter	The total length of boundaries of the microcalcification
Compactness	Derived from the perimeter (P) and area (A) of a microcalcification, it is equal to $\frac{P^2}{4\pi A}$
Box Min. X,Y; Max. X,Y	The coordinates of the extreme left, top, right, and bottom pixels, respectively, of the microcalcification
Feret X,Y	The dimensions of the minimum bounding box of the microcalcification in the horizontal and vertical directions, respectively
Feret Minimum Diameter	The smallest Feret diameter found after checking a certain number of angles (maximum 64)
Feret Maximum Diameter	The largest Feret diameter found after checking a certain number of angles
Feret Mean Diameter	The average Feret diameter at all angles checked
Feret Elongation	A measure of the shape of the microcalcification, it is equal to $\frac{FeretMax.Diameter}{FeretMin.Diameter}$
Number of Holes	The number of holes in the microcalcification
Convex Perimeter	An approximation of the perimeter of the convex hull of the microcalcification
Roughness	A measure of the roughness, it is equal to $\frac{Perimeter}{ConvexPerimeter}$
Length	A measure of the true length of the microcalcification
Breadth	A measure of the true breadth
Elongation	Equal to $\frac{Length}{Breadth}$
Centroid X,Y	The (x, y) position of the center of gravity of the microcalcification
Principal Axis	The angle at which a microcalcification has the least moment of inertia (the axis of symmetry). For elongated microcalcifications, it is aligned with the longest axis
Secondary Axis	The angle perpendicular to the principal axis

Table 1: Initial feature set used to characterise the segmented microcalcifications.

measure. The practical implementation (used in CaB-CS) of this function is based on the *Minkowski's metric* [5], and some extensions of CaB-CS [11] includes the *Clark's distance* and the *Cosine distance*.

2.2.1 Minkowski's metric

The Minkowski's metric is defined as:

$$Similarity(Case\_x, Case\_y) = \sqrt[r]{\sum_{i=1}^{F} w_i \times |x_i - y_i|^r} \quad (1)$$

Where $Case\_x$ and $Case\_y$ are two cases, whose similarity is computed; F is the number of features that describes the case; x_i, y_i represent the value of the *ith* feature of cases $Case\_x$ and $Case\_y$ respectively; and w_i is the weight of the *ith* feature.

In this study we test the Minkowsky's metric for three different values of r: *Hamming distance* for $r = 1$, *Euclidean distance* for $r = 2$, and *Cubic distance* for $r = 3$.

2.2.2 Clark's distance

The Clark's distance is defined as:

$$Similarity(Case\_x, Case\_y) = \sqrt[2]{\sum_{i=1}^{F} \frac{|\ (x_i - y_i)\ |^2}{|\ (x_i + y_i)\ |^2}} \quad (2)$$

Where $Case\_x$ and $Case\_y$ are two cases, whose similarity is computed; F is the number of features that describes the case; and x_i, y_i represent the value of the *ith* feature of cases $Case\_x$ and $Case\_y$ respectively.

2.2.3 Cosine distance

The Cosine distance is based on vector properties in an Euclidean space. It measures the Cosine angle in an n-dimensional vector space. This metric is defined

as:

$$Similarity(Case\_x, Case\_y) = \frac{\sum_{i=1}^{F}(x_i \cdot y_i)}{\sqrt[2]{\sum_{i=1}^{F} x_i^2 \cdot \sum_{i=1}^{F} y_i^2}} \quad (3)$$

Where F represents the number of features that describes the cases; and x_i, y_i represent the value of the *ith* feature of cases $Case\_x$ and $Case\_y$ respectively.

2.3 Similarity functions based on spheres

CaB-CS proposes other similarity functions based on the *sphere* concepts [5]. These functions search some sphere able to *explain* the new case -that we want to solve-. The first and the second function proposed, the *Sphere of Proximity* and the *MinMax Sphere*, compute the similarity between two cases using a local similarity measure, but the third function, the *Mean Sphere*, computes the similarity using a global similarity measure.

2.3.1 Sphere of Proximity

The Sphere of Proximity searches cases from the case memory that are into a delimited sphere that describes the new case, feature by feature. So, we say that two cases are similar if they are also similar feature by feature. The sphere boundaries are computed using the variance -of the class which belongs to the retrieval case- for each feature. In this sense, we select the cases from the case memory if they satisfy the following condition:

$$\text{If } \forall a_i : \triangle^2 a_i \leq threshold \times \left(\frac{variance_{(a\,i)}}{iteration}\right) \text{ then include this case in the } list\_of\_selected\_cases; \quad (4)$$

Where a_i is the *ith* feature; $\triangle^2 a_i$ is the squared difference between both values of the *ith* feature -for the new case and the retrieved case-; the *threshold* weighs the relevance of the *ith* feature; the *iteration* represents the number of tries that the function computes in order to obtain a correct classification; the $list\_of\_selected\_cases$ is the list where the function retains the "similar" cases; and the *variance* of the feature a_i is computed as:

$$Variance(a_i) = \frac{\sum_{j=1}^{N}(x_{ij} - \overline{x_i})^2}{N-1} \quad (5)$$

Where N is the cardinality of the case memory (the number of cases); x_{ij} is the value of the feature i for the case j; and $\overline{x_i}$ is the mean of the *ith* feature.

If we obtain an empty $list\_of\_selected\_cases$ then we can not classify the new case, otherwise we can use different criteria in order to choose the most similar case to the new case.

2.3.2 MinMax Sphere

The MinMax Sphere computes one *sphere* for each class -in which we can classify the new case-. Each sphere -of any class C- contains information for each feature about the minimum and maximum values, based on the cases of the case memory that belong to this class:

$$\textit{MinMax Sphere - Class C} \begin{cases} feature\_1 \begin{cases} Minimum \\ Maximum \end{cases} \\ \cdots \\ feature\_F \begin{cases} Minimum \\ Maximum \end{cases} \end{cases} \quad (6)$$

In this sense, this similarity function classifies a new case in the class C if, for all features, it satisfies that:

$$\forall a_i : (value\_min(C, a_i) \times threshold\_min) \leq value(New\_case, a_i) \leq (value\_max(C, a_i) \times threshold\_max) \quad (7)$$

Where F is the number of features that describes the case; a_i is the feature i; $value\_min(C, a_i)$ and $value\_max(C, a_i)$ are the minimum value and the maximum value of the sphere of the class C for the *ith* feature; $value(New\_case, a_i)$ represents the value of the *ith* feature of the new case; and the $threshold\_min$ and the $threshold\_max$ weighs the relevance of the *ith* feature for the minimum value and the maximum value respectively.

2.3.3 Mean Sphere

The Mean Sphere also computes one *sphere* for each class -in which we can classify the new case-. Each sphere -of any class C- contains information for each feature about the mean value based on the cases of the case memory that belongs to this class:

$$\textit{Mean Sphere - Class C} \begin{cases} feature\_1 \ \{ \ Mean \\ \cdots \\ feature\_F \ \{ \ Mean \end{cases} \quad (8)$$

Now, the Mean Sphere function uses a similarity function based on distance (e.g. Hamming distance) in order to retrieve the "most similar sphere" to the new case. In this sense, we say that this function uses a global similarity measure.

3 GENIFER: GENetic based classIFiER system

GENIFER [8] uses a Genetic Algorithm (GA) [7, 4] in order to obtain a set of rules that solves our classification problem. The application of Genetic Algorithms to Machine Learning problems has been addressed from two different points of view: the Pittsburgh approach and the Michigan approach, early exemplified by LS-1 [13] and CS-1 [6] respectively.

In the Pittsburgh approach, each individual of the population represents a complete solution to the problem, which is a whole set of rules. In contrast, the Michigan approach codifies only one rule in each individual. Therefore, the solution consists on all the population. This difference in representation leads to significant differences between the two systems. Using the first approach, the GA can be applied directly. But in the Michigan approach, the GA is limited to the exploration of new points of the search space (new rules) and the learning process is performed by other algorithms (e.g. Bucket Brigade Algorithm [6], Q-Learning technique [14], etc.).

3.1 GENIFER overview

GENIFER [8] is a general purpose classifier system based on GAs. It is designed to be applied to problems with real-valued attributes. The starting point is the GeB-CS (*Genetic-Based Classifier System*) [2, 3]. GENIFER aim is to obtain a set of classification rules that solves the classification problem described by a set of examples. Like GeB-CS, GENIFER is also a *Pittsburgh* based classifier system, but it is designed to face problems with real-valued features.

The GeB-CS ideas are the base for the GENIFER system. The aim is to look for a change in knowledge representation of classification rules. In GeB-CS a binary codification of PC_0 (*Predicate Calculus of zero order*) rules is used as GA individuals. These $Condition \rightarrow Action$ rules are redefined in GENIFER. The sought goal is to adapt those rules to the real-value nature of features. The implications of this idea are: (1) looking for a new rule representation (and its genetic codification), (2) choosing a good matching function, (3) adapting the GA fitness functions, and (4) designing new genetic operators capable to deal with the new genetic rule codification.

From the set of current GENIFER [8] variants, we choose two of them in order to solve the mammography classification problem presented in this paper. They can be found, in bold font, in table 2. These two variants were chosen in order to obtain a first evaluation of the system. GENIFER-MDA obtains classification rules, so a pure performance evaluation can be obtained from the problem. On the other hand, GENIFER-RA builds classification rules and, at the same time, it is able to select the most relevant features involved in each rule.

GENIFER-MDA	**Minimal Distance Activation**
GENIFER-MDAA	Minimal Distance Adapt. Activation
GENIFER-RA	**Representative Attributes**
GENIFER-DIA	Diploid based Incremental Approach

Table 2: GENIFER used variants

Both, GENIFER-MDA and GENIFER-RA variants, are based on a two-phase approach to the classification problem. They are divided in the *training* phase, where rules are obtained using a *training* set of correctly classified samples, and the *test* phase, where rules are exploited. An incremental approach, where *train* and *test* have been merged, can be found in [8].

3.2 GENIFER-MDA

3.2.1 System overview

In a classification problem where all features belong to $\mathcal{R}$, an n-dimensional space can be defined, so all the examples belong to it. The question is: Can any affinity be defined in this space? In other words, is it possible to identify space regions that share the same classification concept?

GENIFER-MDA (*Minimal Distance Activation*) searches a way of splitting the n-dimensional space, described by the problem features, into space regions that share the same classification concept. In order to reach this aim, we use what we call *significant points*. These points are linked to a classification concept, in this paper: a *class*. If we want to classify an example m_i, the process of obtaining its associated class can be seen as the process of identifying the concept/class region where it belongs. This process can be defined easily as finding which is the nearest *significant point* to the m_i sample. Once it is obtained, the class where m_i belongs is the class linked to the nearest *significant point* recovered. This process can be seen as an analogy of some *similarity functions* used in Case Based Reasoning Systems [5].

Under this new point of view, a redefinition of rule representation used by the GA is needed. As it can be seen, GENIFER-MDA does not look for PC_0 rules. Instead, it looks for *significant points* of the n-dimensional space defined by the set of features.

This revision modifies: (1) rule representation and its associated matching function (the key of the

classification process) and (2) the GA structure and the codification of individuals.

Representation

In GENIFER-MDA a classification rule has the form: $Condition \rightarrow Concept$. The condition part is an ordered set of real values (as many as the number of problem features). This condition expresses a *significant point* of the features in the n-dimensional space. The concept identifies the class linked to the *significant point* described in the *condition* part of the rule.

Matching

The rule matching process can be described as:

1. Let m_i be the sample to classify.
2. Let R be the rule set that solves the problem, and x a rule that $x \in R$.
3. Let $Dist$[1] be the similarity function between a sample and a *significant point*. In GENIFER-MDA, $Dist$ is the Euclidean distance:

$$Dist(m_i, x) = \sqrt{\sum_{j=1}^{F} (m_{ij} - x_j)^2} \qquad (9)$$

4. Find the rule r that satisfies:

$$Dist(m_i, r) \leq \min_{\forall x \in R} (Dist(m_i, x)) \qquad (10)$$

5. Classify m_i as a member of the associated class to rule r.

3.2.2 GA modifications

Due to the new proposal in rule representation, some modification must be introduced in the GA. They can be summarised as:

1. The introduction of a new operator that prunes the rules not used in the training set classification. Useless rules can easily appear due to the proximity between *significant points*.
2. The fitness function used is:

$$fitness(ind_i) = (\%CorrectClassified)^2 \qquad (11)$$

3. The crossover and mutation operators were slightly modified to enable them to manipulate real coded individuals.

[1]The $Dist$ function used in the GA approach is equivalent as the $Similarity$ function used in the CBR approach.

3.3 GENIFER-RA

3.3.1 System overview

The aim of GENIFER-RA (*Representative Attributes*) is to exploit the adaptive behaviour of GAs. GA is modified in order to obtain: (1) a rule set that solves the classification problem and (2) a representative set of features for each rule. GENIFER-RA proposes a GA that is in charge of choosing which features are used in the *nearest neighbour* metric function. In other words, the choice to be done for each feature involved in a rule is a binary decision: use it or do not use it.

In order to include the previous considerations, the $Dist$ function must be slightly modified. It is defined as follows:

$$Dist(m_i, x, w) = \sqrt{\sum_{j=1}^{F} val\,(w_j, m_{ij}, x_j)}$$

$$val(w_j, m_{ij}, x_j) = \begin{cases} 0 & \text{if } w_j = 0, \\ (m_{ij} - x_j)^2 & \text{if } w_j = 1. \end{cases} \qquad (12)$$

In the $Dist$ function, m_i is the example to be classified and x the *significant point*. A $\overline{w}$ vector is added to discard which features are not representative. The contribution of each feature is computed using the val function. As it can be observed in equation 12, $\overline{w}$ is defined by setting $\forall w_i \in \{0, 1\}$. So $\overline{w}$ becomes binary valued.

3.3.2 GA modifications

The main modification affects the individuals codification. In order to use the adaptive behaviour of the GA to adjust the $\overline{w}$ vector [2], it must be codified in the *genotype* of an individual.

4 Results

This section describes the results obtained from the application of CaB-CS and GENIFER to the mammography classification problem. First we present the testbed and, second, the results obtained using the CBR and the GA approaches respectively.

4.1 Testbed

The information used to feed the machine learning systems, can be summarised as follows. After the image processing phases, for each mammography, an $m \times 23$ real valued matrix is obtained. This matrix contains as many rows, m, as the number of

[2]In other words, to choose which are the representative features.

Sim. Function	%Correct	%Incorrect
Hamming	72.857	27.143
Euclidian	72.857	27.143
Cubic	**74.286**	25.714
Clark	**74.286**	25.714
Cosine[3]	64.286	25.714
Proximity	72.857	27.143
MinMax	72.857	27.143
Mean	72.857	27.143

Table 3: Results using the CBR approach.

microcalcifications presents in the image. In order to feed this information to the machine learning systems (CaB-CS and GENIFER), the matrix is flattened into a vector. This process is achieved computing the mean value of each feature of the microcalcifications present in the image. So an image can be reduced to a real-valued vector with 23 features.

The human experts also decided which *training* and *test* sets must be used. The *training* set contains 146 samples, while the *test* set has 70 samples.

4.2 Previous results

In [9] a statistical prediction model was developed. This statistical model was based on regression, and a *logit* function was used in order to obtain which features are relevant to the classification process. The results obtained never outperformed the 51% of success that human experts were able to reach.

4.3 Results using CaB-CS

In this subsection we present the results using a CBR approach. In fact, we present the results using the CaB-CS system and their extensions.

The table 3 shows the results using the different similarity functions: Hamming distance, Euclidean distance, Cubic distance, Clark's distance, Cosine distance, Sphere of Proximity, MinMax Sphere and Mean Sphere.

We want to remark that the different similarity functions retrieve the most similar case to the new case from the case memory, using very different policies. On one hand, we use the more classical view: the similarity functions based on the distance. On the other hand, we present the similarity functions based on spheres, which retrieve the most similar case using -again- different criteria. For example, the function Mean Sphere, use the cases of the case memory in order to construct the spheres that represent the different classes, so these spheres do not represent a real case. Although we use very different points of view, we obtain -as table 3 shows - the same results: 72.857% of *Prediction Accuracy* (PA). And punctually, Cubic and Clark's distances reach the 74.286% of PA. These results show that these different criteria have similar behaviour on this problem.

Also, we want to remark that these results are the best results after trying about 500 different options of the CaB-CS and extensions. But, almost all results are very very close. The different options consist, for instance, on using different criteria in order to weigh the features; or training previously the initial case memory or not; or using different policies in the retain phase; etc. [5].

[3]The Cosine distance diagnose 64.286% correctly, 25.714% incorrectly, and for a 10% is not clear their diagnostic. So it diagnoses a 71.42% correctly among all the classified cases.

Variant	%CA	%PA
MDA	69.178	72.857
RA	69.178	**74.286**

Table 4: GENIFER results using MDA and RA

4.4 Results using GENIFER

GENIFER is divided in two different working phases. The first one, the *training* phase, is in charge of obtaining the classification rule that solve the classification problem. When the rules are obtained, the second phase, *test* phase, checks them using the test set.

In table 4, the results obtained with GENIFER MDA and RA are presented. For each GENIFER variant, two results are presented. The *Classification Accuracy* (CA) is the percentage of samples correctly classified in the *training* phase. On the other hand, the *Prediction Accuracy* (PA) is the percentage of correctly classified samples in the *test* phase. As it can be seen, the maximum system performance is obtained using the RA variant, where the PA raises up to a 74%. These results clearly outperform the ones presented in section 4.2.

5 Conclusions and further work

If we analyse the results reached for both approaches (Case-Based Reasoning and Genetic Algorithms) some conclusions and further work can be deduced.

We must point out that our techniques *outperform* the prediction accuracy (51%) obtained in [9]. Both approaches (*Case-Based Reasoning* and *Genetic Algorithms*), implemented by CaB-CS and GENIFER

systems respectively, move about the 72.286% of PA, and both -punctually- raise their PA up to 74.286% when using Cubic and Clark's distances for the CBR approach and the GENIFER-RA variant for the GA approach. So, we can conclude that we obtain successful results.

On the other hand, the previous results presented in [9] only classified a 55% of cases, where a 51% are classified correctly and only a 4% are classified incorrectly. This means that a 92.7% have been correctly diagnosed over the classified ones. Although the results obtained by the Cosine distance include this idea, we are working to improve the reliability of our results.

Acknowledgements

We would like to thank the *Institute of Informatics and Applications* of the *University of Girona* for their co-operation. We would also like to thank the *Comissionat per a Universitats i Recerca* of the *Generalitat de Catalunya* for its support under grant number 1999FI-00719. Finally, we also like to thank the *Fondo de Investigación Sanitaria* for its support under grant numbers 00/0033-01 and 00/0033-02.

References

[1] A. Aamodt and E. Plaza. Case-based reasoning: Foundations issues, methodological variations, and system approaches. *AI Communications*, 7:39–59, 1994.

[2] J. M. Garrell, E. Golobardes, E. Bernadó, and X. Llorà. Automatic Classification of mammary biopsy images with machine learning techniques. *Proceedings of Engineering of Intelligent Systems (EIS'98)*, 1998.

[3] J.M. Garrell, E. Golobardes, E. Bernadó, and X. Llorà. Automatic diagnosis with genetic algorithms and case-based reasoning. *Artificial Intelligence in Engineering*, 13(4):367–362, October 1999. Elsevier Science Ltd., ISSN 0954-1810.

[4] D. E. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning.* Addison-Wesley Publishing Company, Inc., 1989.

[5] E. Golobardes. *Aportacions al raonament basat en casos per resoldre problemes de classificació.* PhD thesis, Enginyeria i Arquitectura La Salle, Universitat Ramon Llull, July 1998.

[6] J. H. Holland. Escaping Brittleness: The Possibilities of General Purpose Learning Algorithms Applied to Parallel Rule-Based Systems. *Machine Learning: An Artificial Intelligence Approach, Vol. II*, pages 593–623, 1986.

[7] J. H. Holland. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence.* MIT Press/ Bradford Books edition, 1992.

[8] X. Llorà and J. M. Garrell. GENIFER: A Nearest Neighbour based Classifier System using GA. *Proceedings of the Genetic and Evolution Computation Conference (GECCO99), Morgan Kaufmann*, 1999.

[9] J. Martí, X. Cufí, J. Regincós, and et al. Shape-based feature selection for microcalcification evaluation. *Proceedings of the SPIE Medical Imaging Conference on Image Processing*, 3338:1215–1224, 1998.

[10] C.K. Riesbeck and R.C. Schank. *Inside Case-Based Reasoning.* Lawrence Erlbaum Associates, Hillsdale, NJ, US, 1989.

[11] M. Salamó, E. Golobardes, J.M. Garrell, and J. Martí. Clasificación de microcalcificaciones usando razonamiento basado en casos. In *III Jornadas de Transferencia Tecnológica de Inteligencia Artificial (TTIA'99)*, Murcia, Spain, November 1999. ISDN 931170-0-5.

[12] L. Shen, R.M. Rangayan, and J.L. LeoDesautels. Detection and classification of mammographic calcifications. *K.W. Bowyer and S.Astley editors, State of the Art in Digital Mammographic Image Analysis*, pages 198–212, 1994.

[13] S. F. Smith. Flexible Learning of Problem Solving Heuristics through Adaptive Search. In *Proceedings of the 8th International Joint Conference on Artificial Intelligence*, pages 422–425, 1983.

[14] Stewart W. Wilson. Classifier Fitness Based on Accuracy. *Evolutionary Computation*, 3(2):149–175, 1995.

[15] D. Winfields, M. Silbiger, G.S. Brown, and et al. Technology transfer in digital mamography. *Report of the Joint National Cancer Institute. Workshop of May19-20, Invest. Radiol.*, pages 507–515, 1994.