

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

185,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Aid System to Evaluation of Breast Calcification by Case-Based Reasoning According to BI-RADS Category

Fatima G.G. Elpidio, Lourdes M. Brasil and Janice M. Lamas

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/59939>

1. Introduction

Breast cancer is a multi-factor disease. It is heterogeneous neoplasm with histopathological changes that has been registered with a high occurrence in women in the last decades according to advance of female age [1].

Early detection and right diagnosis of breast cancer are complex processes that are derived among other factors from reasoning and experience of expert. The Breast Imaging Reporting and Data System (BI-RADS) aims at standardizing mammographic reporting in order to reduce differences in the subjective interpretation of mammographic images and to facilitate the control of the results [2].

Calcifications can be characterized as small radiopaque deposits with high sensitivity to X-rays. Generally, they are associated with benign cases, however they have been found in about 30 to 50% of clinically undetectable lesions [2].

The malignancy level from suspicious breast calcifications can increase or decrease according to their characteristics such as type, size, quantity, and distribution density. Usually have been considered suspicious for malignancy if were grouped and have irregular or linear appearance. Most researchers say that carcinomas are associated with calcification clusters that have more than 5 elements and it are in a larger area than 3 mm from breast [3] [2].

Classifying calcification clusters as benign or malignant is a complex task, often requiring a biopsy (removal process of tissue fragment for microscopic analysis) to do a definitive

conclusion. In spite of high probability to diagnose cancer in calcification regions that are suspicious of malignancy [4], researches [5, 12] shows that a amount of problems including false diagnose could be avoided if there was more precise analysis before surgery.

The use of computational techniques on the decision support helps medical professionals to disseminate the implicit knowledge and to map computationally the reasoning processes that lead them to make a decision. In this context, a system of Case-Based Reasoning (CBR) [6] is a computational technique derived from Artificial Intelligence (AI) which uses tacit knowledge to model it into an Expert System able to use previous solutions on case solutions.

The result of this research is also, part of a greater extension project entitled "3D Anatomical Atlas Applied to Breast" with the National Laboratory for Scientific Computing of Rio de Janeiro (NLSC/RJ) that received collaboration between the University of Brasília at Gama (FGA/UNB) and Janice Lamas Clinical Radiology (JLCR).

2. Materials and methods

The system was developed using PHP software language [7], PostgreSQL version 9.0 as relational database and Structured Query Language to access data handling (SQL) [8]. The information requirements for CBR development was done by knowledge elicitation techniques [9] under expert supervision and following similar cases of study. The expert contributed with definition of characteristics and more relevant indices to evaluate the global and local similarity and also to setting the weights for the calculation of the overall similarity. So the main strategies adopted for acquisition and mapping information were as follows:

- *Brainstorming* – It is a technique of meetings where participants suggest and explore ideas [9]. This technique was applied for widely discuss the knowledge domain to the specialist explains the main points that should be considered in problem definition.
- Questionnaires - this approach is useful because it allows to obtain the knowledge in a targeted, practical and concise way. The forms adopted had multiple choice questions, checklist and descriptive questions in order to bunch information on specific topics questions.
- The questionnaires were developed based on lists of importance to the specialist indicates the weights of indices. Interviews – it was followed a script guidelines according to knowledge items that were being analyzed to obtain information efficiently.
- Prototyping - this technique allowed to monitore critical aspects of the requirements, expediting the development and minimizing the risks of system construction [9].

In the end of this step was generated the formal documentation of system requirements and it was defined the scope of the research in order to develop CBR system.

In CBR systems the cases representation can be performed using data directly from the structure where they are or by the generation of a second structure containing only data

relevant to the cases composition. The cases representation was done using attribute-value [6] through relationship model entities in table form on relational database [8]. So each case was represented by a record in the table and, consecutively, each field corresponds to a feature of the case. The data were extracted from the database through domain information interpretation, according to methods established by the expert and BI-RADS standard.

These interpretation methods were translated to a computational algorithm to do the cases base automatically. In the cases base algorithm, each instance of case and attribute values were calculated using combinatorics [10] method to appoint the BI-RADS class according to case characteristics. The Figure 1 shows a representation overview of CBR cases structure. The similarity algorithm adopted in CBR system retrieves correlated cases according to characteristics weights and their importance level for BI-RADS classification. Details about analysis of calcifications done by CBR system designed to suggest the BI-RADS classification will be explained at following.

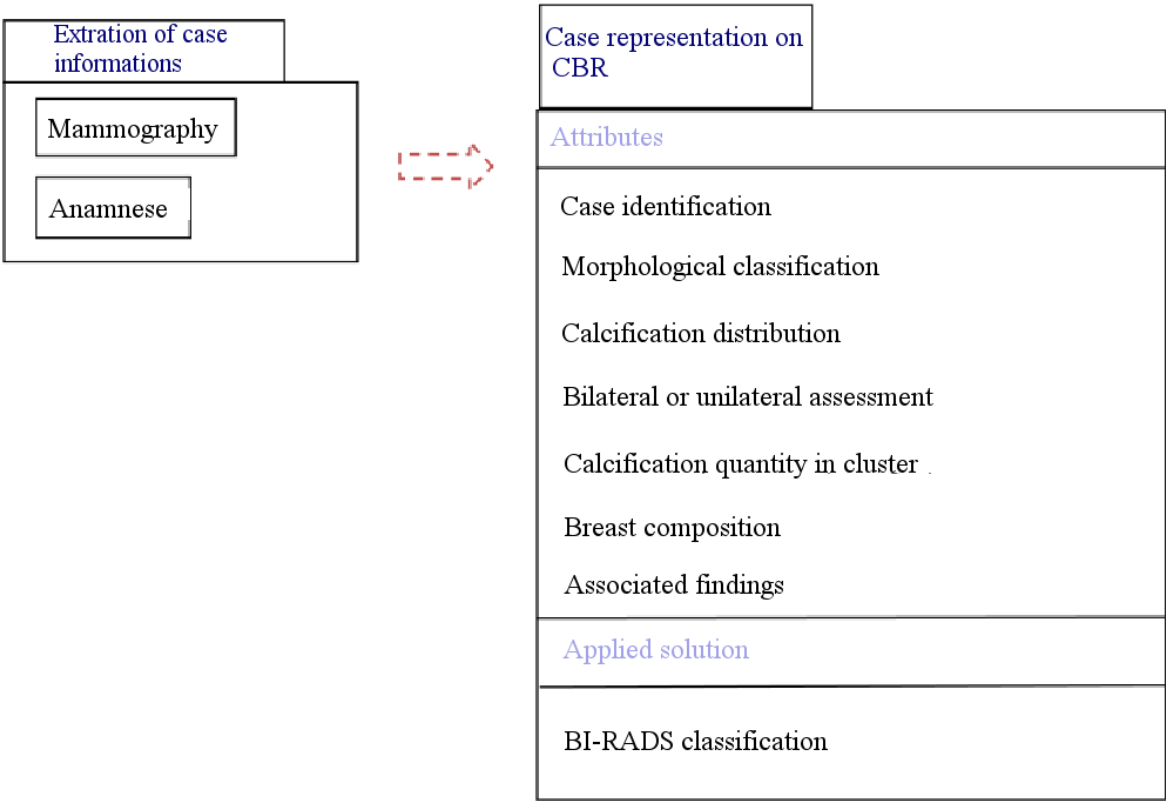


Figure 1. Cases representation structure.

3. Morphology

Breast calcifications are classified in one of those cases BI-RADS described:

- Typically benign if they were skin calcifications, vascular calcifications, coarse or "popcorn-like" calcifications, large rod-like calcifications, round calcifications, lucent-centered calcifications, "eggshell" or "rim" calcifications, milk of calcium calcifications, suture calcifications, dystrophic calcifications.
- Intermediate concern, suspicious calcifications - if they were amorphous or indistinct calcifications or coarse heterogeneous calcifications.
- Higher probability malignancy - if they here fine pleomorphic calcifications, fine linear or fine-linear branching calcifications.

4. Distribution

Distribution modifiers indicate the disposition and consecutive breast calcifications suspicion degree. It was adopted six possibilities proposed by BI-RADS: Diffuse/scattered, Regional, Clustered grouped, Linear or Segmental.

5. Bilateral or unilateral assessment

In general it is considered whether the grouping of calcifications has the same distribution and the same morphology in both mammas, because generally the cases where the features are just one mamma (unilateral) are more associated with suspected malignancy.

6. Breast composition

According to the breast kind, mammography may have a decreased detection sensitivity. The Table 1 describes the possibilities for classification adopted according to BI-RADS.

Breast density	Breast density
The breasts are almost entirely fatty	Glandular tissue <25%
Fibroglandular density	Glandular tissue between 25% and 50%
Heterogeneously dense	Glandular tissue between 51% and 75%
The breasts are extremely dense	Glandular tissue >75%

Table 1. Patterns of breast density by BI-RADS

7. Stability level from calcifications

One way to evaluate the stability of breast calcifications is following previous exams where it is possible to analyze if there was any change in the number of particles, distribution, morphology or associated findings.

8. Calcification quantity in cluster

The reference literature and the expert requirements from area show that above 5 calcifications grouped have a higher degree of suspicion to cancer.

9. Associated findings

They are other structures which may increase the level of malignance suspicion if it were together calcifications. Some associated findings receive a higher priority degree than calcifications for BI-RADS classification. e.g, nodules, focal asymmetry, architectural distortion, etc.

Therefore, BI-RADS class suggestion was done by system based on analysis of the calcification majority found in mammogram using as decision criteria the weight of each risk factor like: breast composition, associated findings, distribution pattern of calcification, bilateral or unilateral assessment, calcification quantity in cluster and the predominant morphological classification which was indicated by Artificial Neural Networks (ANN) specifically, the MultiLayer Perceptron (MLP) [11].

CBR knowledge base had 78,336 cases relating to the classification criteria referenced in the BI-RADS. Beyond this sample, it was used fifty real cases (they came directly from structure), which were used in the validation of the system. The features weight of the represented cases in the CBR system were defined assigning 70% of the weight between the morphology and distribution of calcifications, besides considering if the breast evaluation was bilateral or unilateral. The remaining 30% were distributed between the standard of breast composition, evaluation of particles amount in the calcifications cluster and associated findings.

The similarity measure is a function that measures the similarity between two cases. It is used to define the most similar cases ordering them according to the level of higher similarity. This metric is usually normalized in a range from 0 (total dissimilarity) to 1 (absolute coincidence) and it can be evaluated in a global and local context. Global similarity analyzes the usefulness of a case for a given question, where the similarity between the question and the case should be determined. On the other hand, Local Similarity evaluates similarity in terms of relevant attributes or features of a case. Hence measure of Global similarity adopted by the CBR system was developed using Nearest Neighbor with feature weighting method which it is a technique that determines the nearest neighbor geometrically considering the distance measurement definition and weight (importance degree) for each feature as Equation 1:

$$(T, S)=\sum_{i=1}^n SIM_i(T_i, S_i)*i$$

Where:

- a. T is the input case;
- b. S is a base case
- c. n is attributes quantity of each case;
- d. i is a individual attribute-value;
- e. SIM is similarity function for the attribute I on T and S cases;
- f. weighted nearest neighbor - is the nearest neighbor changed by the weight of the attributes of the case, as expressed in Equation 2.
- g. Where w is the weight given to attribute i.

The Step function was adopted to calculate the Local similarity. This method consists in calculating the distance between two values as specified by Equation 2. The result of the local similarity is 1 when the result of the difference is less than the threshold otherwise the result is 0.

$$(v_i, v_j)=\begin{cases} 1, & \text{if } V(v_i - v_j) \leq S \\ 0, & \text{if } V(v_i - v_j) > S \end{cases}$$

The Table 2 shows a set of case examples that will be used to explain how the system works during a search execution.

Characteristics	Weight	Case 1	Case 2	Query cases
Morfology	0.35	0	0	0
Distribution	0.35	0.8	0.8	0.8
Bilateral/unilateral assessment	0.02	0.5	0.8	1
Calcification quantity in cluster	0.03	0.2	0.3	0
Breast composition	0.05	0.1	0.1	0.1
Associated findings	0.20	0	0	0

Table 2. Comparing query cases with case bases from CBR system.

Using as an example the case of Table 2 and any query case (C), the calculations performed by algorithm developed to recover the cases of higher similarity with the problem will be explained below. Considering: X=Case 1 and C=Query case. Distance measure between the case X and the query case $C = (0.35 \times 0 - 0 \vee^2) + (0.35 \times 0.8 - 0.8 \vee^2) + (0.02 \times 1 - 0.5 \vee^2) + (0.03 \times 0 - 0 \vee^2) + (0.05 \times 0.01 - 0.01 \vee^2) + (0.20 \times 0 - 0 \vee^2) 1/2$;

Distance measure between the case X and the query case $C = (0.35 \times 0) + (0.35 \times 0.25) + (0.02 \times 1) + (0.03 \times 0.04) + (0.05 \times 0) + (0.20 \times 0) 1/2$;

Therefore, the distance measure between the case X and the query case $C \approx 0.33$; Considering: X=Case 2 and C=Query case.

Distance measure between the case X and the query case $C = (0.35 \times |0 - 0|2) + (0.35 \times |0.8 - 0.8|2) + (0.02 \times |1 - 0.8|2) + (0.03 \times |0 - 0.3|2) + (0.05 \times |0.01 - 0.01|2) + (0.20 \times |0 - 0|2) 1/2$;

Distance measure between the case X and the query case $C = (0.35 \times 0) + (0.35 \times 0) + (0.02 \times 0.04) + (0.03 \times 0.09) + (0.05 \times 0) + (0.20 \times 0) 1/2$;

Therefore, the distance measure between the case X and the query case $C \approx 0.06$;

According to the calculations made, the most similar case to the new problem is one that has the smallest distance [6]. Therefore, in the example presented the case that has the largest similarity to the query case is Case 2. The attributes used to compare the similarity between input cases and the case bases was done using explanation technique according to expert requirements. Because despite the existence of several methods for automatic cases indexing, the researches shows that the best strategy is still the choice of manual indexes, as adopted in this work.

The recovery of similarity measure between cases is calculated sequentially for all the case bases in order to determining the x most similar cases. Search by global similarity between new cases related to previous cases inserted at the base is done using the nearest neighbor algorithm weighted normalized besides metrics for measuring contrast. Storage of new cases is done manually, so it is necessary human intervention to register new cases in the system. In each case of base is analyzed successively the ratio of specific preference for the situation by measures of similarity of the system. All cases of the base are arranged according to the result of the similarity function. The local similarities set was calculated for each case attribute using the standard BI-RADS applied in a staircase function.

In this work it was chosed the zero adjustment, which is generally used in situations of complex problems, but with simple solutions. As discussed previously, the tacit reasoning involved to distinguish benign calcifications of suspected malignancy is a complex process, however the resolution applied in the scope of this proposal is a bit simple: it consist in to indicate the BI-RADS category most applicable based on risk factors referenced in the literature.

The solution evaluation was performed by an expert, in addition, we adopted the technique of predictive validation [2] using the concepts of sensitivity and specificity used with historical cases in known diagnostic tests. The Figure 2 summarizes the reasoning structure adopted in the CBR functional components scope.

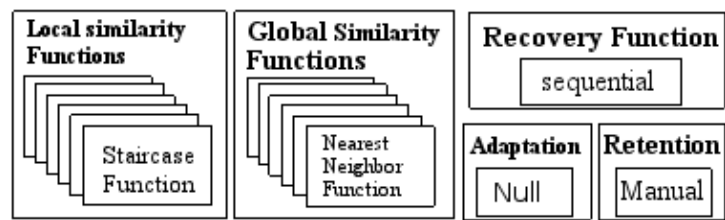


Figure 2. Reasoning Structure adopted in this research.

10. Results

So if a query is done in system developed which has X number of cases, it shows an electronic form to be filled with important characteristics from risk factor to cancer diagnosis.

The system conducts a search by measure of contrast and returns a set of cases where it is applied to the nearest neighbor algorithm to show the most applicable BI-RADS class according to the solution of similar previous cases present in the CBR database system.

The following figures show system interfaces with its steps to recover similarity cases.

The Figure 3 shows the system interface used to do a searching in its database.

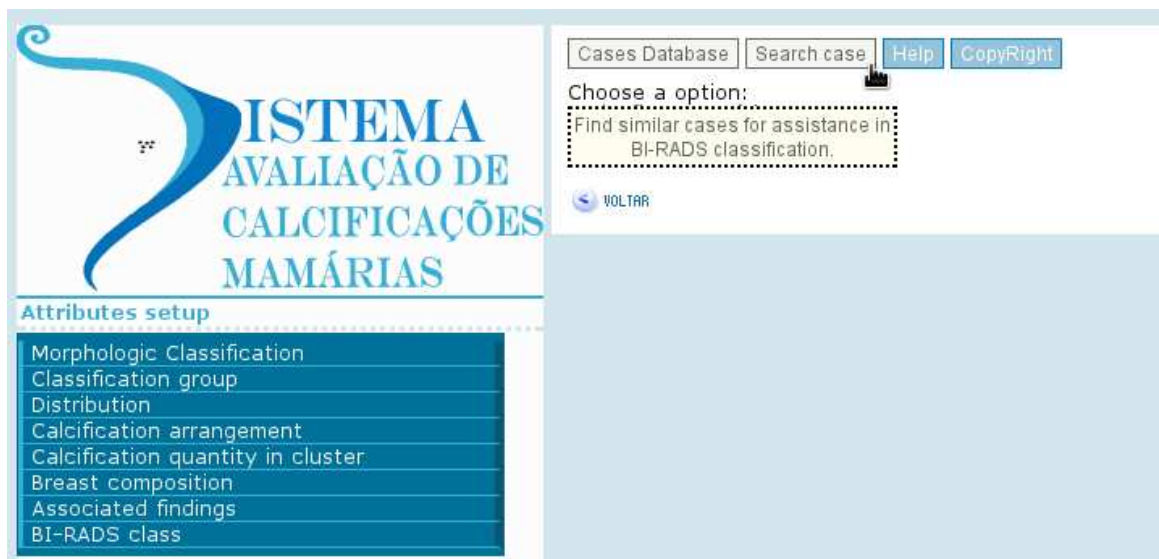


Figure 3. System interface to access the option for query cases.

The Figure 4 shows the system interface to fill characteristics of case searched.

This screen allows a similarity search in the RBC system by fill in a form with the calcification risk factors previously described. After it to fill feature of query case the CBR developed performs a similarity search based in its base cases and shows a suggestion of

BI-RADS class as illustrated Figure 5.

Figure 4. Search form used in CBR system.

Figure 5. interface of CBR developed that shows the BI-RADS class based on features of case query.

11. Discussion and conclusion

Both ANNs as CBR are AI techniques of enshrined in the acquisition and utilization of information to generate solutions similar to human reasoning in specialized fields.

In general, the ANN is an alternative to the traditional computing model based on the model of Von Neumann Machines. They are proven effective in solving computational problems, such as learning by examples or patterns of input data, whose main line of resolution is the inference rules of entry set to absorb the learning and consecutively reaching generalization of knowledge. The ANNs also differ from conventional programming by not requiring explicit programming. They do not adopt the paradigm of imperative programming, they use the connectionist paradigm which learns by examples using the analogy of solving problems based on previous solutions. Another ANN's advantage is their ability to test all possible reactions to parallel stimuli and their generalization ability if there are incompleteness of the data presented for the problem.

Whereas the CBR methodology characterizes the paradigm of symbolic AI with emphasis on mathematical logic to knowledge representation at high level. Unlike conventional expert systems using the Rules Based Reasoning with logical sequences composed of premises and conclusions, CBR modeling knowledge from a database of cases that constitute the knowledge previously experienced. Usually each case has a problem description and its solution, which may contain a set of successes and failures related cases resolved. CBR case base uses inferences strategies to identify the current situation and find a similar experience on its memory to solve a new problem.

The advantage of using a hybrid system, although in different tasks, is to expand and complement the potential of each approach. If on one hand the ANNs are powerful tools in the knowledge acquisition for other way they have difficulties to explain the proposed solution because they are as a "black-box" which maps the weight of connections between neurons. But in general, CBR besides to map the knowledge by its cases base, it is able to show the way of its process decision.

The ANN MPL adopted indicated the predominant morphological type of calcifications detected on mammography. However, for effective analysis in order to establish a diagnosis of mammography, it is necessary a complete analysis also reviewing the patient's anamnese. So CBR and ANN were adopted together to assist in the analysis and BI-RADS classification of breast calcification.

Although there are some solid solutions of CBR, the most of them has some restrictions of use that could affect the experiment, so it was decided to develop the algorithm, in order to ensure the requirements of this research. It was filtered the risk factors from patient anamnesis to keep just the parameters that has higher importance in breast calcification analysis according to BI-RADS.

One of the problems faced in breast calcifications system developing was that, in general, the BI-RADS protocol says that the radiologist should describe in his mammography report just

the greater suspicions found or those that may raise doubts by other professional interpretations. However, this proposal performs an independent assessment of suspicion level from breast calcifications. Hence there was difficulty in collecting information about typically benign calcifications, which are usually not referenced in the reports because they have a low degree of suspicion. So combinatorics technique was used based on indices analyzed in order to generate the CBR base case and to cover these situations creating a broad enough base, covering all possible cases under expert orientation and BI-RADS protocol.

For while the calcifications stability level is not used by CBR developed, but on next CBR algorithm versions it will be increase this parameter, which also has great relevance in the analysis of breast calcifications.

This proposal was done to provide a system for aiding the evaluation of breast calcifications. With the methodology and results is possible to validate that this proposal provides a consistent process for analyzing breast calcifications that can be used like a second opinion to the experts.

Author details

Fatima G.G. Elpidio^{1*}, Lourdes M. Brasil¹ and Janice M. Lamas²

*Address all correspondence to: fgizele@hotmail.com

1 University of Brasília at Gama (UnB /FGA) / Post-Graduation Program in Biomedical Engineering, Gama, Brazil

2 Janice Lamas Radiology Clinic, Brazil

References

- [1] WHO. Breast Cancer. <http://www.who.int> Accessed 2014-04-07.
- [2] A.C.R: Breast Imaging Reporting and Data System (BI-RADS), 3rd edn. American College of Radiology, VA (2003).
- [3] Martins, L.A.L., Barra, A.A., Lucena, C.E.M.: Microcalcificacoes mamárias suspeitas de malignidade. *Revista Brasileira de Cancerologia* 27(3), 262–6 (2006).
- [4] Kouskos, E.e.a.: Wire localisation biopsy of non-palpable breast lesions: reasons of unsuccessful excision. *Eur J Gynaecol Oncol* 27(3), 262–6 (2006).
- [5] Biffi, R.G.: A dinâmica familiar de um grupo de mulheres com câncer de mama. Tese de doutorado. PhD thesis, Escola de Enfermagem de Ribeirão Preto, USP (2003).

- [6] Michael, M.R., Weber, R.O.: Case-Based Reasoning. A Textbook, 1st ed. Springer, New York (2013).
- [7] Reiersol, D., Baker, M., Shiflett, C.: PHP in Action: Objects, Design, Agility., 1st ed. Manning Publications Co, Greenwich, Conn (2007).
- [8] Date, C.J.: Database Design and Relational Theory, 1st ed. O'Reilly Media, CA (2012).
- [9] Pressman, R. S. Software engineering: a practitioner's approach, McGraw-Hill Higher Education, 2010.
- [10] Brualdi, R.A.: Introduction to Combinatorics, 5th ed. Pearson, New York (2009).
- [11] Elpídio, F.G.G., et al: Artificial neural network to classify breast calcification in accordance with morphological analysis described by birads and le gal (in portuguese: Classificação de calcificações mamárias baseado na análise morfológica descrita por birads e le gal através de rede neural artificial). In: XXIII Congresso Brasileiro de Engenharia Biomedica (2012).
- [12] Elpidio, F.G.G. Aid System to Evaluation of Breast Calcification by Digital Processing of Images and Artificial Intelligence (in portuguese: Sistema de Auxílio na Avaliação de Calcificações Mamárias por Processamento Digital de Imagens e Inteligência Artificial), RIUnB, 2012.