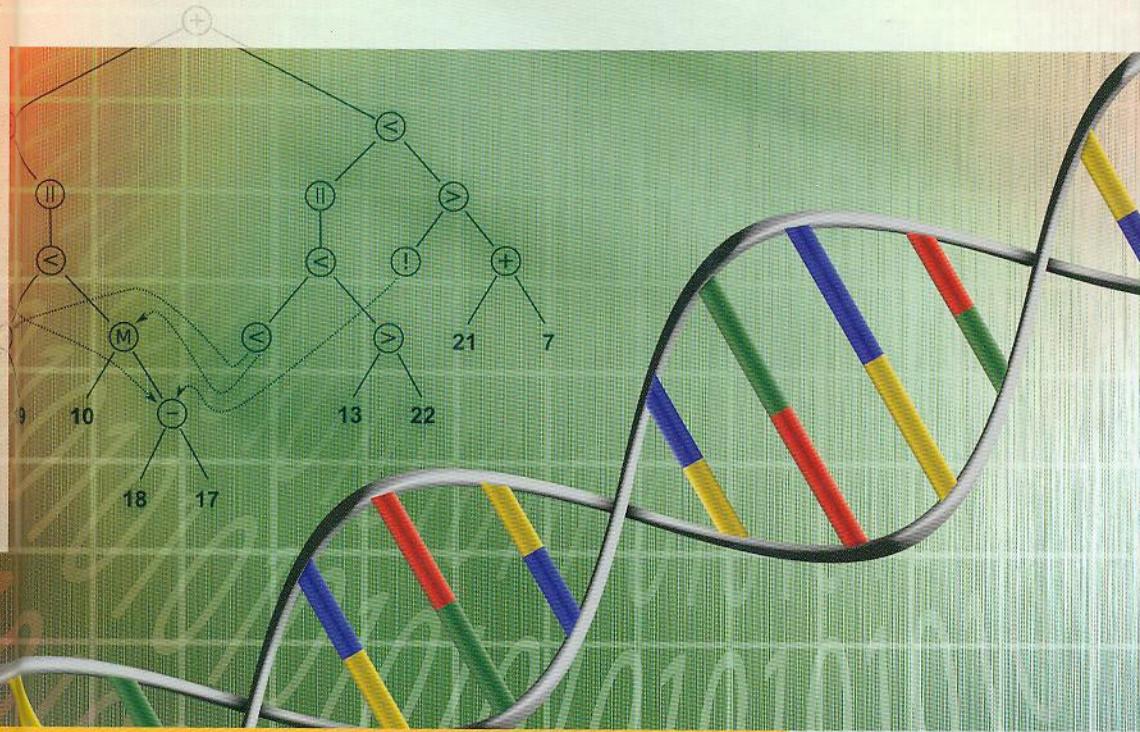


EDITORS

STEPHEN L. SMITH AND STEFANO CAGNONI

# GENETIC AND EVOLUTIONARY COMPUTATION

Medical Applications



 WILEY

# **Genetic and Evolutionary Computation**

Volume 10 Number 1

January 2006

ISSN 1063-6536

# **Genetic and Evolutionary Computation Medical Applications**

Edited by

**Stephen L. Smith**

*Department of Electronics, University of York, York, UK*

**Stefano Cagnoni**

*Dipartimento di Ingegneria dell'Informazione, Università degli  
Studi di Parma, Parma, Italy*



**WILEY**

A John Wiley and Sons, Ltd., Publication

This edition first published 2011  
© 2011 John Wiley & Sons, Ltd  
Chapter 5.3 2011 © the Crown in right of Canada

*Registered office*  
John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ,  
United Kingdom

For details of our global editorial offices, for customer services and for information about how  
to apply for permission to reuse the copyright material in this book please see our website at  
[www.wiley.com](http://www.wiley.com).

The right of the author to be identified as the author of this work has been asserted in  
accordance with the Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval  
system, or transmitted, in any form or by any means, electronic, mechanical, photocopying,  
recording or otherwise, except as permitted by the UK Copyright, Designs and Patents Act  
1988, without the prior permission of the publisher.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in  
print may not be available in electronic books.

Designations used by companies to distinguish their products are often claimed as  
trademarks. All brand names and product names used in this book are trade names, service  
marks, trademarks or registered trademarks of their respective owners. The publisher is not  
associated with any product or vendor mentioned in this book. This publication is designed to  
provide accurate and authoritative information in regard to the subject matter covered. It is  
sold on the understanding that the publisher is not engaged in rendering professional services.  
If professional advice or other expert assistance is required, the services of a competent  
professional should be sought.

*Library of Congress Cataloguing-in-Publication Data*

Genetic and evolutionary computation : medical applications / edited by Stephen L. Smith,  
Stefano Cagnoni.

p. ; cm.

Includes index.

ISBN 978-0-470-74813-8 (cloth)

1. Medicine—Data processing. 2. Genetic programming (Computer science)

3. Evolutionary programming (Computer science) I. Smith, Stephen L., 1962–

II. Cagnoni, Stefano, 1961–

[DNLM: 1. Computational Biology—methods. 2. Models, Genetic. 3. Algorithms.]

4. Evolution, Molecular. QU 26.5]

R858.G457 2010

610.285—dc22

2010025421

A catalogue record for this book is available from the British Library.

Print ISBN: 9780470748138

ePDF ISBN: 9780470973141

eBOOK ISBN: 9780470973134

Set in 10/12.5pt Palatino by Aptara Inc., New Delhi, India.

Printed and bound in Singapore by Markono Print Media Pte Ltd.

# Contents

## About the Editors

xi

## List of Contributors

xiii

### 1 Introduction

1

### 2 Evolutionary Computation: A Brief Overview

3

*Stefano Cagnoni and Leonardo Vanneschi*

#### 2.1 Introduction

3

#### 2.2 Evolutionary Computation Paradigms

4

##### 2.2.1 Genetic Algorithms

5

##### 2.2.2 Evolution Strategies

7

##### 2.2.3 Evolutionary Programming

8

##### 2.2.4 Genetic Programming

8

##### 2.2.5 Other Evolutionary Techniques

10

##### 2.2.6 Theory of Evolutionary Algorithms

11

#### 2.3 Conclusions

12

### 3 A Review of Medical Applications of Genetic and Evolutionary Computation

17

*Stephen L. Smith*

#### 3.1 Medical Imaging and Signal Processing

18

##### 3.1.1 Overview

18

##### 3.1.2 Image Segmentation

18

##### 3.1.3 Image Registration, Reconstruction and Correction

21

##### 3.1.4 Other Applications

24

#### 3.2 Data Mining Medical Data and Patient Records

25

#### 3.3 Clinical Expert Systems and Knowledge-based Systems

27

3.4	Modelling and Simulation of Medical Processes	29
3.5	Clinical Diagnosis and Therapy	34
<b>4</b>	<b>Applications of GEC in Medical Imaging</b>	<b>45</b>
<b>4.1</b>	<b>Evolutionary Deformable Models for Medical Image Segmentation: A Genetic Algorithm Approach to Optimizing Learned, Intuitive, and Localized Medial-based Shape Deformation</b>	
	<i>Chris McIntosh and Ghassan Hamarneh</i>	
4.1.1	Introduction	47
4.1.1.1	Statistically Constrained Localized and Intuitive Deformations	54
4.1.1.2	Genetic Algorithms	57
4.1.2	Methods	58
4.1.2.1	Population Representation	58
4.1.2.2	Encoding the Weights for GAs	58
4.1.2.3	Mutations and Crossovers	59
4.1.2.4	Calculating the Fitness of Members of the GA Population	60
4.1.3	Results	62
4.1.4	Conclusions	63
<b>4.2</b>	<b>Feature Selection for the Classification of Microcalcifications in Digital Mammograms using Genetic Algorithms, Sequential Search and Class Separability</b>	<b>69</b>
	<i>Santiago E. Conant-Pablos, Rolando R. Hernández-Cisneros, and Hugo Terashima-Martín</i>	
4.2.1	Introduction	69
4.2.2	Methodology	71
4.2.2.1	Pre-processing	71
4.2.2.2	Detection of Potential Microcalcifications (Signals)	72
4.2.2.3	Classification of Signals into Microcalcifications	74
4.2.2.4	Detection of Microcalcification Clusters	76
4.2.2.5	Classification of Microcalcification Clusters into Benign and Malignant	77
4.2.3	Experiments and Results	77
4.2.3.1	From Pre-processing to Signal Extraction	77
4.2.3.2	Classification of Signals into Microcalcifications	79

4.2.3.3	Microcalcification Clusters Detection and Classification	81
4.2.4	Conclusions and Future Work	82
<b>4.3</b>	<b>Hybrid Detection of Features within the Retinal Fundus using a Genetic Algorithm</b>	<b>85</b>
	<i>Vitoantonio Bevilacqua, Lucia Cariello, Simona Cambò, Domenico Daleno, and Giuseppe Mastronardi</i>	
4.3.1	Introduction	85
4.3.2	Acquisition and Processing of Retinal Fundus Images	88
4.3.2.1	Retinal Image Acquisition	89
4.3.2.2	Image Processing	90
4.3.3	Previous Work	91
4.3.4	Implementation	93
4.3.4.1	Vasculature Extraction	94
4.3.4.2	A Genetic Algorithm for Edge Extraction	99
4.3.4.3	Skeletonization Process	103
4.3.4.4	Experimental Results	104
<b>5</b>	<b>New Analysis of Medical Data Sets using GEC</b>	<b>111</b>
<b>5.1</b>	<b>Analysis and Classification of Mammography Reports using Maximum Variation Sampling</b>	<b>113</b>
	<i>Robert M. Patton, Barbara G. Beckerman, and Thomas E. Potok</i>	
5.1.1	Introduction	113
5.1.2	Background	114
5.1.3	Related Works	116
5.1.4	Maximum Variation Sampling	118
5.1.5	Data	122
5.1.6	Tests	124
5.1.7	Results & Discussion	124
5.1.8	Summary	129
<b>5.2</b>	<b>An Interactive Search for Rules in Medical Data using Multiobjective Evolutionary Algorithms</b>	<b>133</b>
	<i>Daniela Zaharie, D. Lungeanu, and Flavia Zamfirache</i>	
5.2.1	Medical Data Mining	133
5.2.2	Measures for Evaluating the Rules Quality	134
5.2.2.1	Accuracy Measures	135
5.2.2.2	Comprehensibility Measures	135
5.2.2.3	Interestingness Measures	136
5.2.3	Evolutionary Approaches in Rules Mining	137
5.2.4	An Interactive Multiobjective Evolutionary Algorithm for Rules Mining	138

5.2.4.1	Rules Encoding	139
5.2.4.2	Reproduction Operators	139
5.2.4.3	Selection and Archiving	140
5.2.4.4	User Guided Evolutionary Search	141
5.2.5	Experiments in Medical Rules Mining	143
5.2.5.1	Impact of User Interaction	144
5.2.6	Conclusions	146
<b>5.3</b>	<b>Genetic Programming for Exploring Medical Data using Visual Spaces</b>	<b>149</b>
<i>Julio J. Valdés, Alan J. Barton, and Robert Orchard</i>		
5.3.1	Introduction	149
5.3.2	Visual Spaces	150
5.3.2.1	Visual Space Realization	150
5.3.2.2	Visual Space Taxonomy	150
5.3.2.3	Visual Space Geometries	151
5.3.2.4	Visual Space Interpretation Taxonomy	151
5.3.2.5	Visual Space Characteristics Examination	153
5.3.2.6	Visual Space Mapping Taxonomy	154
5.3.2.7	Visual Space Mapping Computation	155
5.3.3	Experimental Settings	157
5.3.3.1	Implicit Classical Algorithm Settings	158
5.3.3.2	Explicit GEP Algorithm Settings	159
5.3.4	Medical Examples	161
5.3.4.1	Data Space Examples	161
5.3.4.2	Semantic Space Examples	164
5.3.5	Future Directions	170
<b>6</b>	<b>Advanced Modelling, Diagnosis and Treatment using GEC</b>	<b>173</b>
<b>6.1</b>	<b>Objective Assessment of Visuo-spatial Ability using Implicit Context Representation Cartesian Genetic Programming</b>	<b>175</b>
<i>Michael A. Lones and Stephen L. Smith</i>		
6.1.1	Introduction	175
6.1.2	Evaluation of Visuo-spatial Ability	176
6.1.3	Implicit Context Representation CGP	178
6.1.4	Methodology	180
6.1.4.1	Data Collection	181
6.1.4.2	Evaluation	181
6.1.4.3	Parameter Settings	182
6.1.5	Results	184
6.1.6	Conclusions	186

<b>6.2</b>	<b>Towards an Alternative to Magnetic Resonance Imaging for Vocal Tract Shape Measurement using the Principles of Evolution</b>	<b>191</b>
<i>David M. Howard, Andy M. Tyrrell, and Crispin Cooper</i>		
6.2.1	Introduction	191
6.2.2	Oral Tract Shape Evolution	194
6.2.3	Recording the Target Vowels	195
6.2.4	Evolving Oral Tract Shapes	196
6.2.5	Results	199
6.2.5.1	Oral Tract Areas	200
6.2.5.2	Spectral Comparisons	201
6.2.6	Conclusions	204
<b>6.3</b>	<b>How Genetic Algorithms can Improve Pacemaker Efficiency</b>	<b>209</b>
<i>Laurent Dumas and Linda El Alaoui</i>		
6.3.1	Introduction	209
6.3.2	Modeling of the Electrical Activity of the Heart	211
6.3.3	The Optimization Principles	213
6.3.3.1	The Cost Function	213
6.3.3.2	The Optimization Algorithm	213
6.3.3.3	A New Genetic Algorithm with a Surrogate Model	214
6.3.3.4	Results of AGA on Test Functions	215
6.3.4	A Simplified Test Case for a Pacemaker Optimization	216
6.3.4.1	Description of the Test Case	216
6.3.4.2	Numerical Results	218
6.3.5	Conclusion	220
<b>7</b>	<b>The Future for Genetic and Evolutionary Computation in Medicine: Opportunities, Challenges and Rewards</b>	<b>223</b>
7.1	Opportunities	224
7.2	Challenges	224
7.3	Rewards	226
7.4	The Future for Genetic and Evolutionary Computation in Medicine	227
<b>Appendix: Introductory Books and Useful Links</b>		<b>229</b>
<b>Index</b>		<b>231</b>

## 4.2

# Feature Selection for the Classification of Microcalcifications in Digital Mammograms using Genetic Algorithms, Sequential Search and Class Separability

Santiago E. Conant-Pablos, Rolando R. Hernández-Cisneros,  
and Hugo Terashima-Marín

*Centro de Computación Inteligente y Robótica, Tecnológico de Monterrey, Monterrey, Mexico*

### 4.2.1 Introduction

Breast cancer is one of the main causes of death in women, and early diagnosis is an important means to reduce the mortality rate. Mammography is one

of the most common techniques for breast cancer diagnosis, and calcifications one of the findings that can be seen on mammograms. These are very small calcium deposits that can appear within the soft tissue of the breast. Calcifications, that appear as white dots on the mammogram, are not always a sign of breast cancer, but sometimes an indication of a precancerous condition. Calcifications are divided into two types. Macrocalcifications are larger deposits of calcium, not usually linked to breast cancer. Microcalcifications are typically 100 microns to several millimeters in diameter, may occur in clusters of three or more and usually appear in areas less than  $1\text{ cm}^2$ , or in patterns such as circles and lines. These microcalcifications are associated with extra cell activity in breast tissue, and tight clusters of microcalcifications can indicate early breast cancer.

However, the predictive value of mammograms is relatively low, compared with biopsy. This low sensitivity [8] is caused by the low contrast between the cancerous tissue and the normal parenchymal tissue, the small size of microcalcifications and possible deficiencies in the image digitalization process. The sensitivity may be improved through having each mammogram checked by two or more radiologists, with the consequence of making the process more resource intensive. A viable alternative is replacing one of the radiologists by a computer system, giving a second opinion [1, 19].

The computer system should process the mammograms as digitized images for the detection of suspicious points as being microcalcifications. Then, microcalcifications have to be separated through a classifications process and grouped into clusters. Finally, the obtained microcalcification clusters have to be classified as indicative of benign (harmless) breast tissue or malignant, a sign of breast cancer.

We decided to use artificial neural networks (ANNs) to construct the classifiers for the recognition of individual microcalcifications and benign/malignant microcalcification clusters. ANNs have been used successfully for classification purposes in medical applications. Unfortunately, for an ANN to be successful in a particular domain, its architecture, training algorithm and the domain variables selected as inputs must be adequately chosen. Designing an ANN architecture is a trial-and-error process requiring the selection and tuning of several parameters that depend strongly on the features selected to describe the training data. The classification problem could involve too many features (variables), most of them not relevant at all for the classification process itself. Genetic algorithms (GAs) may be used to address both problems mentioned above, helping to obtain more accurate ANNs with better generalization abilities. An exhaustive review of evolutionary artificial neural networks (EANNs) is presented by Yao [20] and Balakrishnan and Honavar [2].

In particular, this chapter describes and compares three methods for selecting the most relevant features extracted from both individual microcalcifications and microcalcification clusters, which provide the inputs of two simple feedforward ANNs for their classification, with an expectation of improved accuracy. The first method works by ordering features according to a class separability metric and then selecting the most relevant ones as inputs to an ANN. The second method implements a forward sequential search algorithm that sequentially adds features to the ANN while its classification error decreases, and stops when this error starts to increase. Finally, the third method uses a GA to evolve good feature subsets for the classification tasks.

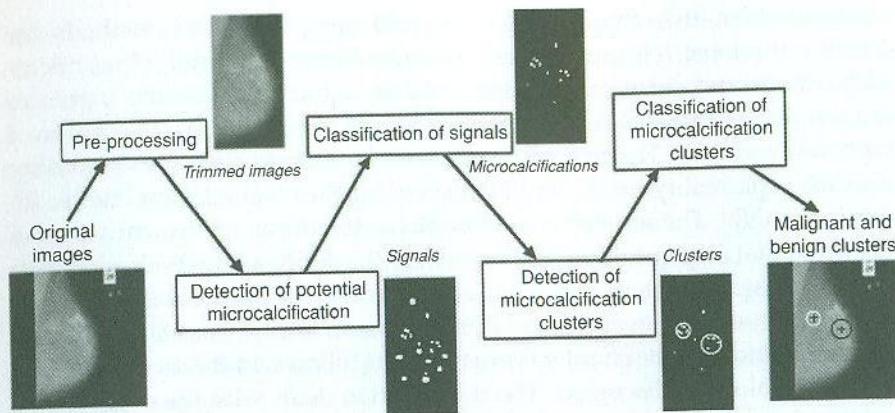
The remainder of the chapter is organized as follows. In the second section, the methodology is discussed. The third section deals with the experiments and the main results of this work. Finally, in the fourth section, the conclusions are presented, and some comments regarding future work are also made.

## 4.2.2 Methodology

The mammograms used in this project were provided by the Mammographic Image Analysis Society (MIAS) [18]. The MIAS database contains 322 images, all medio-lateral (MLO) view, digitized at resolutions of 50 microns/pixel and 200 microns/pixel. In this work, images with a resolution of 200 microns/pixel were used. The data has been reviewed by a consultant radiologist, and all the abnormalities have been identified and marked. The truth data consists of the location of the abnormality and the radius of a circle which encloses it. From the totality of the database, only 25 images contain microcalcifications. Among these 25 images, 13 cases are diagnosed as malignant and 12 as benign. Some related works have used this same database [6, 9, 11, 14]. The general procedure, as shown in Figure 4.2.1, receives digital mammograms as input, and consists of five stages: pre-processing, detection of potential microcalcifications (signals), classification of signals into microcalcifications, detection of microcalcification clusters, and classification of microcalcification clusters into benign and malignant.

### 4.2.2.1 Pre-processing

This stage has the aim of eliminating those elements in the images that could interfere in the process of identifying microcalcifications. A secondary goal is to reduce the work area only to the relevant region that exactly contains the breast.



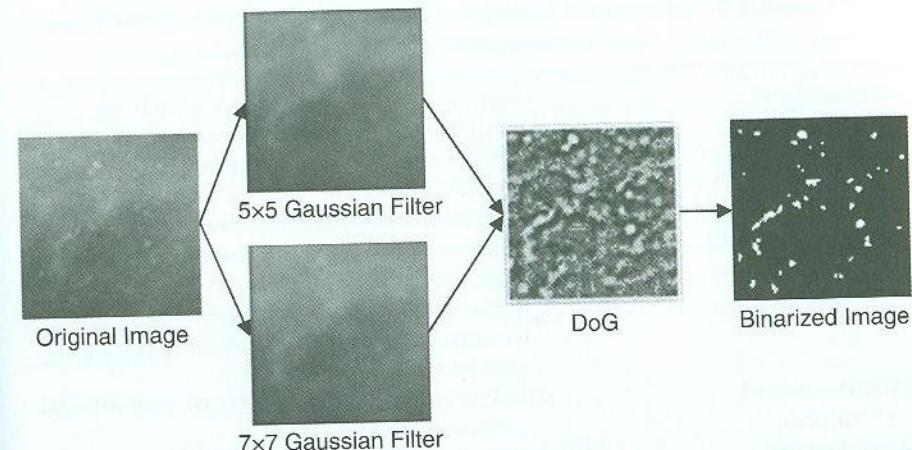
**Figure 4.2.1** General procedure followed by the classification system

The procedure receives the original images as input. A median filter is applied in order to eliminate the background noise and then an automatic cropping procedure deletes the background marks and the isolated regions, so the image will contain only the region of interest. The result of this stage is a smaller image, with less noise.

#### 4.2.2.2 Detection of Potential Microcalcifications (Signals)

The main objective of this stage is to detect the mass centers of the suspicious points as being microcalcifications in the images (signals). The pre-processed images of the previous stage are the inputs of this procedure. A method employing difference of gaussian filters (DoG) is used for enhancing those regions containing bright points. The use of DoG for detection of potential microcalcifications has been addressed successfully by Dengler, Behrens and Desaga [7] and Ochoa [15]. A gaussian filter is obtained from a gaussian distribution and when applied to an image, eliminates high-frequency noise, acting like a smoothing filter. The DoG filter is obtained from the difference of two gaussian functions, as shown in equation (4.2.1), where  $x$  and  $y$  are the coordinates of a pixel in the image,  $k_1$  and  $k_2$  are the heights of the functions, and  $\sigma_1$  and  $\sigma_2$  are the different standard deviations of the two gaussian filters that construct the DoG filter.

$$DoG(x, y) = k_1 \exp\left(\frac{x^2+y^2}{2\sigma_1^2}\right) - k_2 \exp\left(\frac{x^2+y^2}{2\sigma_2^2}\right) \quad (4.2.1)$$



**Figure 4.2.2** Example of the application of a DoG filter ( $5 \times 5$ ,  $7 \times 7$ )

The resultant image after applying a DoG filter is globally binarized, using a certain threshold. In Figure 4.2.2, an example of the application of a DoG filter is shown. A region-labeling algorithm allows the identification of each one of the points. Then, a segmentation algorithm extracts small  $9 \times 9$  windows, containing the regions of interest whose centroids correspond to the centroids of the identified points. The size of the windows is adequate for containing the signals given that, at the current resolution of 200 microns, the potentially malignant microcalcifications have an area of  $5 \times 5$  pixels on average [16].

In order to detect the greatest possible number of points, several DoG filters are applied several times, varying the binarization threshold. The points obtained by applying each filter are added to the points obtained by the previous one, deleting the repeated points.

All of the points obtained by the DoG filters are subsequently passed to three selection procedures in order to transform them into signals (potential microcalcifications). The first method performs selection according to the object area. The second method performs selection according to the gray level of the points. Finally, the third selection method uses the gray gradient (or absolute contrast, the difference between the mean gray level of the point and the mean gray level of the background). The result of these three selection processes is a list of signals (potential microcalcifications) represented by their centroids.

**Table 4.2.1** Summary of features extracted from the signals (potential microcalcifications)

Signal contrast (7 features)	Maximum gray level, minimum gray level, median gray level, mean gray level, standard deviation of the gray level, gray level skewness, gray level kurtosis.
Background contrast (7 features)	Background maximum gray level, background minimum gray level, background median gray level, background mean gray level, standard deviation of the background gray level, background gray level skewness, background gray level kurtosis.
Relative contrast (3 features)	Absolute contrast, relative contrast, proportional contrast.
Shape features (20 features)	Area, convex area, background area, perimeter, maximum diameter, minimum diameter, equivalent circular diameter, fiber length, fiber width, curl, circularity, roundness, elongation1, elongation2, eccentricity, aspect ratio, compactness1, compactness2, compactness3, solidity.
Contour sequence moments (7 features)	CSM1, CSM2, CSM3, CSM4, mean radii, standard deviation of radii.
Invariant moments of Hu (4 features)	IM1, IM2, IM3, IM4.

#### 4.2.2.3 Classification of Signals into Microcalcifications

The objective of this stage of processing is to identify if an obtained signal corresponds to an individual microcalcification or not. With this in mind, the 47 features shown in Table 4.2.1 are extracted from each individual signal.

In order to process the signals and accurately classify the microcalcifications, we decided to use ANNs as classifiers. In the first section, we mentioned that one of the difficulties of working with conventional feedforward ANNs is that a classification problem could involve too many variables (features), and most of them may not be relevant at all for the classification process itself. For this reason, instead of using the full set of 47 features to construct the signal's classifier, we experimented and compared three methods for the selection of the subset of features to use: class separability, forward sequential search, and genetic algorithms.

**Class Separability:** As in [4] and [5] we decided to use the Kullback–Leibler (KL) distance between histograms of feature values to estimate how well a feature separates the data into different classes and to define an ordering of the 47 features. Two histograms are created for each feature, one for microcalcifications and another for other signals. For this, features are discretized using  $b = \sqrt{|D|}$  equally spaced bins, where  $|D|$  is the size of the training data. The histograms are later normalized by dividing the number of elements in each bin by the total number of elements, so the probability  $p_j(d = i | n)$  that the  $j$ th feature takes a value in the  $i$ th bin of the histogram, given a class  $n$ , is obtained. For each feature  $j$ , the class separability  $\Delta_j$  is calculated as

$$\Delta_j = \sum_{m=1}^c \sum_{n=1}^c \sum_{i=1}^b p_j(d = i | c = m) \log \left( \frac{p_j(d = i | c = m)}{p_j(d = i | c = n)} \right) \quad (4.2.2)$$

where  $c$  is the total number of classes, in our case two, one for microcalcifications and one for other signals. The features are finally ranked by sorting them in descending order of the distances  $\Delta_j$  (larger distances mean better separability). In the application of this method, two features were heuristically assumed redundant if their distances differ by less than 0.0001, and the feature with the smallest distance is eliminated. Other irrelevant non-discriminative features with  $\Delta_j$  distances less than 0.001 are eliminated also.

**Forward Sequential Search:** As in [17], we decided to use a method that implements a forward sequential search (FSS) algorithm [1].

A FSS algorithm starts with an empty feature set and in each iteration adds a new feature to the ANN used as the microcalcification's classifier while the error is decreasing, and stops when the error starts increasing again. The feature added in each iteration is chosen based on its ability to reduce the error in the classifier.

Our FSS algorithm uses the information gain of each feature as the ordering metric. The features are selected in the descending order of their gain values. To compute its information gain, the feature data is discretized, and the gain values are computed from the difference of entropy measures as

$$\text{gain}(S, A) = \text{entropy}(S) - \sum_{v \in \text{values of } (A)} \frac{|S_v|}{|S|} \text{entropy}(S_v) \quad (4.2.3)$$

where  $S$  are the signals to be classified,  $A$  is the feature for which the gain is computed,  $S_v$  is a subset of  $S$  where the  $A$  feature takes the  $v$  value, and the entropy function of a set of signals  $S$  is computed as

$$\text{entropy}(S) = \sum_{i \in C} -p_i(S, S_i) \log_2 p_i(S, S_i) \quad (4.2.4)$$

where  $C$  is the set of class values, in our case if it is or is not a calcification, and  $p_i$  is the proportion of occurrences of class  $i$  in the set  $S$  obtained as  $p_i(S, S_i) = |S_i|/|S|$ .

**Genetic Algorithms:** Expecting to achieve greater accuracy in the classification, we use a third method, this being based on a GA for selecting features. The chromosomes of the individuals in the GA contain 47 bits, one bit for each extracted feature; the value of the bit determines whether that feature will be used in the classification or not [3].

Each individual is evaluated by constructing and training a feedforward ANN (with a predetermined structure). The number of inputs of this ANN is determined by the subset of features to be included, coded in the chromosome. For solving nonlinearly separable problems it is recommended that at least one hidden layer is provided in the network, and according to Kolmogorov's theorem [13], considering the number of inputs ( $n$ ), that the hidden layer contains  $2n + 1$  neurons. The output layer has only one neuron. The accuracy of each network is used to determine the fitness of each individual.

When the GA stops, either because the generations limit has been reached or because improvements in the evaluation of the best individual have not been observed during five consecutive generations, the subset of features of the ANN with the best performance in terms of overall accuracy is obtained, and its ANN used as the microcalcification's classifier.

#### 4.2.2.4 Detection of Microcalcification Clusters

During this stage, the microcalcification clusters are identified. The detection and posterior consideration of every microcalcification cluster in the images may produce better results in a subsequent classification process, as shown in [17]. Because of this, an algorithm for locating microcalcification cluster regions where the quantity of microcalcifications per square centimeter (density) is higher, was developed.

The basic clustering process starts by obtaining the set of all pairs formed with the detected microcalcifications, eliminating all pairs with a Euclidian distance between the centroids of its microcalcifications larger than an empirically defined threshold of 100 pixels. Then it enters a loop for creating the set of microcalcification clusters. In each iteration, the best new cluster, formed with microcalcifications that remain ungrouped, is added. For this, a list of neighbors for each remaining microcalcification is obtained containing all the remaining microcalcifications to a distance smaller than a chosen threshold. The density of each of these potential clusters is computed from its points count and the size of the area computed from the convex polygon it forms. Finally, the group with the maximum density is chosen as a new cluster, the microcalcifications that it contains are eliminated, and a new iteration is started. This process goes on until all the detected microcalcifications are added to a cluster. Every detected cluster is then labeled.

#### 4.2.2.5 Classification of Microcalcification Clusters into Benign and Malignant

This stage has the objective of classifying each cluster into one of two classes: benign or malignant (information provided by the MIAS database). With this in mind, the 30 features shown in Table 4.2.2 are extracted from every microcalcification cluster detected in the previous stage.

In order to process microcalcification clusters and accurately classify them into benign or malignant, we decided again to use ANNs as classifiers. To determine which of the 30 extracted features from the clusters are relevant for their classification, we applied the same three methods described in a previous section: using the same class separability criteria, deriving a forward sequential search, and using genetic algorithms. The main difference in the implementation of the methods for classifying clusters resulted from their adaptation to the use of the 30 different features computed for the clusters.

### 4.2.3 Experiments and Results

#### 4.2.3.1 From Pre-processing to Signal Extraction

As mentioned in the previous section, only 25 images from the MIAS database contain microcalcifications. Among these 25 images, 13 cases are diagnosed as malignant and 12 as benign. Three images were discarded because the positions of the microcalcification clusters, marked in the additional data that comes with the database, were outside the boundaries of the breast. So,

**Table 4.2.2** Summary of features extracted from the microcalcification clusters

Cluster shape (14 features)	Number of calcifications, convex perimeter, convex area, compactness, microcalcification density, total radius, maximum radius, minimum radius, mean radius, standard deviation of radii, maximum diameter, minimum diameter, mean of the distances between microcalcifications, standard deviation of the distances between microcalcifications.
Microcalcification area (6 features)	Total area of microcalcifications, mean area of microcalcifications, standard deviation of the area of microcalcifications, maximum area of the microcalcifications, minimum area of the microcalcifications, relative area.
Microcalcification contrast (10 features)	Total gray mean level of microcalcifications, mean of the mean gray levels of microcalcifications, standard deviation of the mean gray levels of microcalcifications, maximum mean gray level of microcalcifications, minimum mean gray level of microcalcifications, total absolute contrast, mean absolute contrast, standard deviation of the absolute contrast, maximum absolute contrast, minimum absolute contrast.

only 22 images were finally used for this study, and they were passed through the pre-processing stage first (application of a median filter and trimming).

In the second phase, six gaussian filters of sizes  $5 \times 5$ ,  $7 \times 7$ ,  $9 \times 9$ ,  $11 \times 11$ ,  $13 \times 13$ , and  $15 \times 15$  were combined, two at a time, to construct 15 DoG filters that were applied sequentially. Each one of the 15 DoG filters was applied 51 times to the pre-processed images, varying the binarization threshold in the interval [0, 5] in increments of 0.1. The points obtained by applying each filter were added to the points obtained by the previous one, deleting the repeated points. The same procedure was repeated with the points obtained by the remaining DoG filters. These points passed through the three selection methods for selecting signals (potential microcalcification), according to region area, gray level, and the gray gradient. For this work, a minimum area of 1 pixel ( $0.0314 \text{ mm}^2$ ) and a maximum of 77 pixels ( $3.08 \text{ mm}^2$ ) were considered. Studying the mean gray levels of pixels surrounding identified microcalcifications, it was found that they have values in the interval [102, 237] with a mean of 164. For this study, we set the minimum gray level for points to be selected to 100. Again, studying the mean gray gradient of points

surrounding identified microcalcifications, it was found that they have values in the interval [3, 56] with a mean of 9.66. For this study, we set the minimum gray gradient for points to be selected to 3. The result was a list of 1,242,179 signals (potential microcalcifications) represented by their centroids.

The additional data included in the MIAS database define, with centroids and radii, the areas in the mammograms where microcalcification clusters are located. It is supposed that signals within these areas are mainly microcalcifications, but there are many signals that lie outside the marked areas. With these data and the support of expert radiologists, each of the signals located in these 22 mammograms was pre-classified as a microcalcification, or as non-microcalcification. Out of the 1,242,179 signals, only 4,612 (0.37%) were microcalcifications, and the remaining 1,237,567 (99.63%) were not. Because of this imbalanced distribution of examples of each class, an exploratory sampling was made. Several samplings with different proportions of each class were tested and finally we decided to use a sample of 10,000 signals, including 2,500 microcalcifications (25%).

#### 4.2.3.2 Classification of Signals into Microcalcifications

After the 47 microcalcification features were computed from each signal, the first method for feature selection, based on class separability for ranking the features, reduced the relevant features to five: median gray level, mean gray level, minimum gray level, background maximum gray level, and background mean gray level. A transactional database was obtained, containing 10,000 signals (2,500 of them being microcalcifications randomly distributed) and five features describing each signal.

The second approach, based on the forward sequential search, reduced the relevant features to only three: absolute contrast, standard deviation of the gray level of the signal, and the third-order moment of the contour sequence [10]. Again, a transactional database was obtained, containing 10,000 signals including 2,500 microcalcifications randomly distributed, and three features describing each signal. For the third approach, using the GA, the original transactional database with all 47 features was used.

To test these two feature selection methods, simple feedforward ANNs with the corresponding number of inputs were trained and tested. The architecture of these ANNs consisted of five and three inputs respectively,  $2n + 1$  neurons in the hidden layer (where  $n$  is the number of inputs), and one output. All the units had the hyperbolic tangent function as the transfer function. The data (inputs and targets) were scaled in the range  $[-1, 1]$  and divided into 10 non-overlapping splits, each one with 90% of the data for

**Table 4.2.3** Results of the classification of individual microcalcifications

Method	Accuracy (%)	Sensitivity (%)	Specificity (%)
CS	84.56	50.91	95.94
FSS	81.33	76.21	81.92
GA	95.40	83.33	94.87

training and the remaining 10% for testing. Ten fold cross-validation trials were performed; that is, the ANNs were trained 10 times, each time using a different split on the data, and the averages of the overall performance, sensitivity, and specificity were reported. These results are shown in Table 4.2.3, representing the ANNs that had the best performance in terms of overall accuracy (percentage of correctly classified microcalcifications). The sensitivity (percentage of true positives or correctly classified microcalcifications) and specificity (percentage of true negatives or correctly classified objects that are not microcalcifications) of these ANNs are also shown.

For the third method, a GA was combined with ANNs to select the features to train them, as described earlier. The GA had a population of 50 individuals, each one with a length of  $l = 47$  bits, representing the inclusion (or exclusion) of each one of the 47 features extracted from the signals. We used a simple GA, with Gray encoding, stochastic universal sampling selection, single-point crossover, fitness-based reinsertion, and a generational gap of 0.9. The probability of crossover was 0.7 and the probability of mutation was  $1/l$ , where  $l$  is the length of the chromosome (in this case,  $1/l = 1/47 = 0.0213$ ). The initial population of the GA was always initialized uniformly at random. All the ANNs constructed by the GA are feedforward networks with one hidden layer. All neurons have biases with a constant input of 1.0. The ANNs are fully connected, and the transfer functions of every unit is the hyperbolic tangent function. The data (input and targets) were normalized to the interval  $[-1, 1]$ . For the targets, a value of “−1” means “non-microcalcification” and a value of “1” means “microcalcification”. For training each ANN, backpropagation was used, only one split of the data was considered (90% for training and 10% for testing) and the training stopped after 20 epochs. The GA ran for 50 generations, and the results of this experiment are also shown in Table 4.2.3 in the row labeled “GA”.

The best solution found with the GA method was an ANN with 23 inputs (five related to the contrast of the signal, four related to the background contrast, two related to the relative contrast, seven related to the shape, four moments of the contour sequence, and only one of the invariant geometric

moments), corresponding to 48.94% of the original 47 extracted features. The ANNs coded in the chromosomes of the final population of the GA use 20.02 inputs on average, that is, the ANNs with the best performance need only 42.60% of the original 47 features extracted from the microcalcifications.

#### 4.2.3.3 Microcalcification Clusters Detection and Classification

The process of cluster detection and the subsequent feature extraction phase generates another transactional database, this time containing the information on every microcalcification cluster detected in the images. A total of 40 clusters were detected in the 22 mammograms from the MIAS database that were used in this study. According to MIAS additional data and the advice of expert radiologists, 10 clusters are benign and 30 are malignant.

After the 30 features were computed from each microcalcification cluster, the first method for feature selection, based on class separability for ranking the features, reduced the relevant features to five: maximum radius, convex perimeter, standard deviation of the distances between microcalcifications, minimum absolute contrast, and standard deviation of the mean gray level of the microcalcifications in a cluster. The second approach, based on the forward sequential search, reduced the relevant features to only three: minimum diameter, minimum radius and mean radius, of the clusters.

The same procedure for the evaluation of the first two feature selection methods for the classification of individual microcalcifications was applied for these methods for the case of classification of clusters. The results in terms of the overall accuracy, sensitivity, and specificity of the best ANNs are shown in Table 4.2.4.

A GA method was also used to select the features for training ANNs, as described earlier. In this case, the transactional database containing the 30 features extracted from the clusters was used. The GA used in this case was similar to that used for the classification of signals, differing basically in the size of the chromosome and the mutation probability determined by the total number of features (30). For the targets, a value of “−1” means that the

**Table 4.2.4** Results of the classification of microcalcification clusters

Method	Accuracy (%)	Sensitivity (%)	Specificity (%)
CS	84.56	50.91	95.94
FSS	77.5	53.85	88.890
GA	100.00	100.00	100.00

cluster is "benign" and a value of "1" means "malignant". The results of this experiment are also shown in Table 4.2.4 in the row labeled "GA".

The best ANN obtained with the GA method had nine inputs, corresponding to 30% of the original cluster feature set (five features related to the shape of the cluster, one related to the area of the microcalcifications, and three related to the contrast of the microcalcifications). On average, the chromosomes of the last generation coded 14.03 inputs; that is, the ANNs with the best performance only receive 46.76% of the original features extracted from the microcalcification clusters.

#### 4.2.4 Conclusions and Future Work

We found that the use of GAs combined with ANNs greatly improves the overall accuracy, specificity, and sensitivity of the recognition of microcalcifications. We also found that all the ANNs coded in the chromosomes of the final population of the GA use 20.02 inputs on average; that is, the ANNs with the best performance need only 42.60% of the original 47 features. As an additional note, the first method, based on ranking features by class separability, had similar results in the case of the specificity, and a good overall performance, but had poor accuracy for the specificity of the classification.

In the case of the classification of microcalcification clusters, we also observed that the use of a GA for feature selection greatly improved the overall accuracy, sensitivity, and specificity, achieving values of 100%. On average, the best ANN architectures receive 14.03 inputs on average; that is, they only receive 46.76% of the 30 original cluster features as inputs. Nevertheless, only 40 microcalcification clusters were detected in the 22 mammograms used in this study. The test sets used in the ten fold cross-validation trial were very small and, in some splits, all the examples belonged to only one of the two classes, so either sensitivity or specificity could not be calculated. These splits were ignored in the calculation of the respective mean. On the other hand, the first two methods, based on ranking features by class separability and forward sequential search, had similar performance in terms of specificity and overall performance, but both showed deficient results for the specificity of the classification.

As future work, it would be useful to include and process other mammography databases, in order to have more examples and produce transactional feature databases that are more balanced and complete, and also test how different resolutions could affect system effectiveness. The size of the gaussian filters could be adapted depending on the size of the microcalcifications

to be detected and the resolution of images. Different new features could be extracted from the microcalcifications in the images and tested, too. In this study, simple GAs and ANNs were used, and more sophisticated versions of these methods could produce better results. The inclusion of simple backpropagation training in the EANNs has consequences in terms of longer computation times, so alternatives to backpropagation should be tested in order to reduce time costs.

#### References

- [1] Anttilinen, I., Pamilo, M., Soiva, M., and Roiha, M.: Double reading of mammography screening films: one radiologist or two? *Clinical Radiology* **48** (1993) 414–421.
- [2] Balakrishnan, K. and Honavar, V.: Evolutionary design of neural architectures. A preliminary taxonomy and guide to literature. Iowa State University, Department of Computer Sciences. Technical Report CS TR 95-01 (1995).
- [3] Cantú-Paz, E. and Kamath, C.: Evolving neural networks for the classification of galaxies. In Proceedings of the Genetic and Evolutionary Computation Conference (GECCO 2002), San Francisco, CA, USA (2002), pp. 1019–1026.
- [4] Cantú-Paz, E.: Feature subset selections, class separability and genetic algorithms. Lawrence Livermore National Laboratory, Center for Applied Scientific Computing (2004).
- [5] Cantú-Paz, E., Newsam, S., and Kamath, C.: Feature selection in scientific applications. Lawrence Livermore National Laboratory, Center for Applied Scientific Computing (2004).
- [6] Chandrasekhar, R. and Attikiouzel, Y.: Digitization regime as a cause for variation in algorithm performance across two mammogram databases. The University of Western Australia, Centre for Intelligent Information Processing Systems, Department of Electrical and Electronic Engineering. Technical Report 99/05 (1999).
- [7] Dengler, J., Behrens, S. and Desaga, J.F.: Segmentation of microcalcifications in mammograms. *IEEE Transactions on Medical Imaging* **12**(4) (1993) 634–642.
- [8] Ganott, M.A., Harris, K.M., Kaman, H.M., and Keeling, T.L.: Analysis of false-negative cancer cases identified with a mammography audit. *The Breast Journal* **5**(3) (1999) 166–175.
- [9] Gulsrød, T.O.: Analysis of mammographic microcalcifications using a computationally efficient filter bank. Stavanger University College, Department of Electrical and Computer Engineering, Technical Report (2001).
- [10] Gupta, L. and Srinath, M.D.: Contour sequence moments for the classification of closed planar shapes. *Pattern Recognition* **20**(3) (1987) 267–272.
- [11] Hong, B.W. and Brady, M.: Segmentation of mammograms in topographic approach. IEE International Conference on Visual Information Engineering, Guildford, UK (2003).

- [12] Kozlov, A. and Koller, D.: Nonuniform dynamic discretization in hybrid networks. In Proceedings of the 13th Annual Conference of Uncertainty in AI (UAI), Providence, RI, USA (2003), pp. 314–325.
- [13] Kurkova, V.: Kolmogorov's theorem. In *The Handbook of Brain Theory and Neural Networks*. MIT Press, Cambridge, MA, USA (1995), pp. 501–502.
- [14] Li, S., Hara, T., Hatanaka, Y., Fujita, H., Endo, T., and Iwase, T.: Performance evaluation of a CAD system for detecting masses on mammograms by using the MIAS database. *Medical Imaging and Information Science* **18**(3) (2001) 144–153.
- [15] Ochoa, E.M.: Clustered microcalcification detection using optimized difference of gaussians. Air Force Institute of Technology, Wright-Patterson Air Force Base. Master Thesis (1996).
- [16] Oporto-Díaz, S.: Automatic detection of microcalcification clusters in digital mammograms. Master Thesis, Center for Intelligent Systems, Tecnológico de Monterrey, Campus Monterrey, Monterrey, Mexico (2004).
- [17] Oporto-Díaz, S., Hernández-Cisneros, R.R., and Terashima-Marín, H.: Detection of microcalcification clusters in mammograms using a difference of optimized gaussian filters. In Proceedings of the Second International Conference on Image Analysis and Recognition, ICIAR 2005, Toronto, ON, Canada (2005), pp. 998–1005.
- [18] Suckling, J., Parker, J., Dance, D., et al.: The Mammographic Image Analysis Society digital mammogram database. *Excerpta Medica International, Congress Series* **1069** (1994) 375–378.
- [19] Thurfjell, E.L., Lernevall, K.A., and Taube, A.A.S.: Benefit of independent double reading in a population-based mammography screening program. *Radiology* **191** (1994) 241–244.
- [20] Yao, X.: Evolving artificial neural networks. *Proceedings of the IEEE* **87** (1999) 1423–1447.