

Mammographic Images Analysis by Use the RMRBF Neural Network

José A. Moreno-Escobar¹, Francisco J. Gallegos-Funes¹,
Rene Cruz-Santiago¹ and Volodymyr I. Ponomaryov²

Instituto Politécnico Nacional, Escuela Superior de Ingeniería Mecánica y Eléctrica

¹Av. IPN s/n, U.P.A.L.M. SEPI-ESIME, Edif. Z, Acceso 3, Tercer Piso,
Col. Lindavista, 07738, Mexico, D. F., Mexico.

²Av. Santa Ana 1000, Col. San Francisco Culhuacan, 04430, Mexico, D. F., Mexico.
j.augusto.moreno@gmail.com, fgalegosf@ipn.mx

Abstract. The use of Computer Aided Diagnosis (CAD) for medical analysis is taking relevance in areas as ECG and EEG, but also in cancer detection. In here is proposed the use of the Rank M-Type Radial Basis Function (RMRBF) Neural Network for mammographic images analysis. The proposed neural network uses a proposed RM-estimator in the scheme of Radial Basis Function to train the neural network. To improve the efficiency of the RMRBF, the parameters used to train the network were manipulated in accordance with the RM-estimator theory. From simulation results we observe the classification capabilities of the proposed neural network.

1 Introduction

The artificial neural networks are nonparametric pattern recognition systems that can generalize by learning from examples [1, 4, 0]. They are particularly useful in problems where decision rules are vague and there is no explicit knowledge about the probability density functions governing sample distributions. Therefore, breast cancer detection, in particular mammogram screening, make ideal candidates for application of neural networks [2, 3]. Since the beginning of '90s, different neural networks applications have been considered in breast cancer detection. The neural networks have a potential to improve the performance of computer-based algorithms, especially when used in conjunction with other algorithms. The increase in availability of quality data through publicly accessible databases will provide in the near future more conclusive evidence on utility of neural networks in solving the difficult problem of breast cancer detection. Most frequently, the network architecture of choice in computer-aided mammography is multilayer feed forward, i.e., multilayer perceptron (MP) trained by supervised learning in a form of a backpropagation learning (BPL) law. Typically, neurons are fully connected and employ standard forms of transfer functions. The number of hidden layers is predominantly one, and the number of nodes in the hidden layer is also relatively small. The subject of MP and BPL has been extensively studied; general discussion can be found in text books, and comparative studies and improvements are frequent topics in conferences and journals [10].

In this paper is proposed the Rank M-Type Radial Basis Function (RMRBF) Neural Network for breast cancer detection purposes. The neural network uses a RM-estimator in the scheme of radial basis function to train the neural network according with the schemes found in the references [9, 10]. The use of robust RM-estimators has been introduced for image denoising applications [0, 7, 8]. The combined RM-estimators use different rank estimators such as the median, Wilcoxon and Ansari-Bradley-Siegel-Tukey estimators, and the M-estimator with different influence functions to provide better robustness. The performances of the RM-estimators are better in comparison with original R- and M- estimators [7]. The RMRBF-based training is less biased by the presence of outliers in the training set and was proved an accurate estimation of the implied probabilities.

2 Radial Basis Function Neural Networks

Radial Basis Functions (RBF) have been used in several applications for pattern classification and functional modeling. These functions have been found to have very good functional approximation capabilities [4, 9, 10]. It has been proven that any continuous function can be modeled up to a certain precision by a set of radial basis functions [10]. RBFs have their fundamentals drawn from probability function estimation theory. The structure of the RBF network is depicted in Figure 1. Each network input is assigned to a vector entry and the outputs correspond either to a set of functions to be modeled by the network or to several associated classes.

2.1 Radial Basis Functions Network

Several functions have been tested as activation functions for RBF networks. In pattern classification applications the Gaussian function is preferred, and mixtures of these functions have been considered in various scientific fields.

The Gaussian activation function for RBF networks is given by [9]:

$$\phi_j(\mathbf{X}) = \exp\left[-(\mu_j - \mathbf{X})^T \Sigma_j^{-1} (\mu_j - \mathbf{X})\right] \quad (1)$$

where $\mathbf{X}$ is the input feature vector, μ_j is the mean vector and Σ_j is the covariance matrix of the j th Gaussian function. Geometrically, μ_j represents the center or location and Σ_j the shape of the basis functions. Statistically, an activation function models a probability density function where μ_j and Σ_j represent the first and second order statistics. A hidden unit function can be represented as a hyper-ellipsoid in the N -dimensional space.

The output layer implements a weighted sum of hidden-unit outputs [4, 9, 10]:

$$\psi_k(\mathbf{X}) = \sum_{j=1}^L \lambda_{jk} \phi_j(\mathbf{X}) \quad (2)$$

where L is the number of hidden units, M is the number of outputs with $k=1, \dots, M$. The weights λ_{kj} show the distribution of the hidden unit j for modeling the output k .

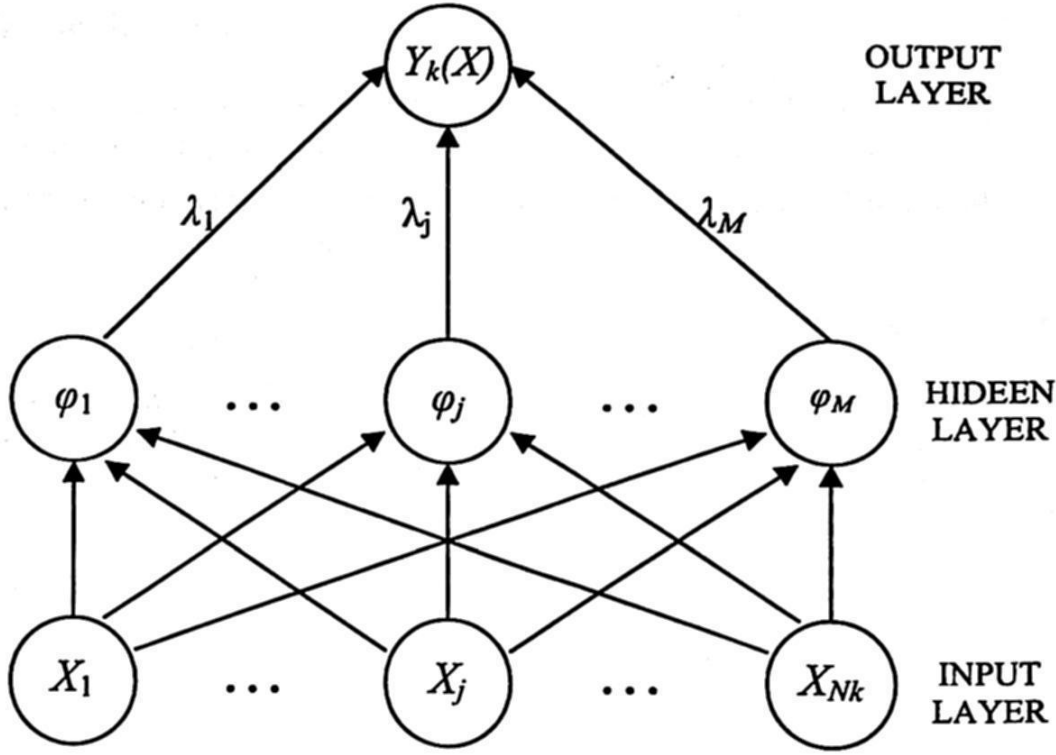


Figure 1. Traditional radial basis function network. Each of N_k components of the input vector $\mathbf{X}$ feeds forward to M basis functions whose outputs are linearly combined with weights $\{\lambda_j\}_{j=1}^M$ into the network output $Y_k(\mathbf{X})$

2.2 Learning Techniques of RBF Networks

Radial Basis Functions have interesting properties which make them attractive in several applications. A combined unsupervised-supervised learning technique has been used in order to estimate the RBF parameters [9]. In the unsupervised stage, k -means clustering algorithm is used to find the pdf's parameters, LMS or instead pseudo-inverse matrix can be used in the supervised stage to calculate the weights coefficients in the neural network [4, 9].

3 Rank M-Type Radial Basis Function Neural Network

In here, we present the use of the RM-estimator as statistic estimation in the Radial Basis Function network architecture. The combined RM-estimators can use different rank estimators such as the median, Wilcoxon or Ansari-Bradley-Siegel-Tukey [0, 7, 8]. The M-estimator uses different influence functions to provide better robustness.

3.1 Activation Function

The Gaussian activation function is the most used function in the RBF networks. In our case we use the inverse multiquadratic function [9]:

$$\phi_j(\mathbf{X}) = \frac{1}{\sqrt{\mathbf{X}^2 + \beta_j^2}} \quad (3)$$

where $\mathbf{X}$ is the input feature vector, β_j is a real constant. In our simulation results $\beta_j=1$.

3.2 K-means Algorithm

In our case we used the clustering k -means algorithm to estimate the parameters of the RBF neural network [4, 9]. The k -means algorithm is used in the unsupervised stage. The input feature vector $\mathbf{X}$ is classified in k different clusters. A new vector $\mathbf{x}$ is assigned to the cluster k whose centroid μ_k is the closest one to the vector. The centroid vector is updated according to,

$$\mu_k = \mu_k + \frac{1}{N_k}(\mathbf{x} - \mu_k) \quad (4)$$

where N_k is the number of vectors already assigned to the k -cluster. The centroids can be updated at the end of several iterations or after the test of each new vector. The centroids can be calculated with or without the new vector. By other hand, the steps for the k -means algorithm are the following:

- Step 1. Select an initial partition with k clusters. Repeat steps 2 through 4 until the cluster membership stabilizes.
- Step 2. Generate a new partition by assigning each pattern to its closest cluster center.
- Step 3. Compute new cluster centers as the centroids of the clusters.
- Step 4. Repeat steps 2 and 3 until an optimum value of the criterion function is found.

3.3 Rank M-type (RM) Estimator

The RM-estimator that is used in the proposal RBF network is the Median M-type (MM) estimator [0, 7]. The non-iterative MM-estimator used as robust statistics estimate of a cluster center is given by,

$$\mu_k = \text{med}\{\mathbf{X} \varphi(\mathbf{X} - \theta)\} \quad (5)$$

where $\mathbf{X}$ is the input data sample, φ is the normalized influence function ψ : $\psi(\mathbf{X}) = \mathbf{X} \varphi(\mathbf{X})$, $\theta = \text{med}\{X_k\}$ is the initial estimate, and $k=1, 2, \dots, N_k$. The presented estimator is the combined RM-estimator. The R -estimator provides good properties of impulsive noise suppression and the M -estimator uses different influence functions according to the Huber scheme, providing better robustness. So, it is expected that the performances of combined RM-estimator can be better in comparison with original R - and M - estimators [7].

3.4 Influence Functions

In our experiments we used the following influence functions [7]:

The simple cut (skipped mean) influence function,

$$\psi_{\text{cut}(r)}(X) = X \cdot 1_{[-r,r]}(X) = \begin{cases} X, & |X| \leq r \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

and the Tukey biweight influence function,

$$\psi_{\text{bi}(r)}(X) = \begin{cases} X^2(r^2 - X^2), & |X| \leq r \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

where X is a data sample and r is a real constant. The parameter r depends of the data to process and can be change for different influence functions.

4 Segmentation and Feature Extraction

4.1 Image Collection

To have access to real medical images for experimentation is a very difficult undertaking due to the privacy issues. The data collection that was used in our experiments was taken form many internet sources, but the most important was the MIAS (Mammographic Image Analysis Society) image data base [11, 12, 13]. The MIAS image collection has been used in other studies of automatic mammography classification. Its corpus consists of 322 images, which belong to 3 big categories: normal, benign and malign. There are 208 normal images, 63 benign and 51 malign.

4.2 Segmentation Stage

The first step in image analysis generally consists in a segmentation phase [14]. In this stage the image is divided in regions of interest that contain relevant information for a specific purpose. In our case, due to the irregularity of mammography images, a combination of morphology and threshold methods was used [14, 15, 16]. This way we could divide the mammography in two main regions: a strange object (possible tumor or cancer) and the breast.

4.2.1 Morphology

Morphology techniques offer a powerful method to segment images with irregular shapes or figures as the ones were are using in this work.

The most important morphology operations defined for two sets A and B are [14]:

- Dilation- Consists in growing the geometrical area of a region of interest in an image and can be defined as $A \oplus B$.

- Erosion- Consists in the reduction of the geometrical area of a region of interest in an image, and can be defined as $A \ominus B$.

Combining these techniques the following operations can be implemented [14]:

- Opening, used to eliminate small objects for smoothing a region of interest in an image.

$$A \circ B = (A \ominus B) \oplus B \quad (8)$$

- Closing, also used for smoothing, eliminates small separations or holes of a region of interest in an image. .

$$A \bullet B = (A \oplus B) \ominus B \quad (9)$$

4.2.2 Thresholding

Thresholding is useful to distinguish pixels that are located in different gray levels (values) and can be considered part of an object. Thresholding values are obtained according to the processed image. In our case, we use two thresholding values: one for the detected object and other one for the breast. The background is not used because it does not offer relevant information, and considering that it is a big part of the whole image and the processing time is reduced.

4.3 Feature Extraction

There are techniques of geometric measurement that allow the evaluation of some characteristics associated to objects detected in an image. These techniques give us an idea of how compact, bright and smooth an object in an image is [15, 16]. Two of these characteristics are:

Compactness: Is a measure of an object distribution,

$$compactness = \frac{perimeter^2}{area} \quad (10)$$

Contrast: Is the difference between the average gray levels of two objects in an image,

$$contrast = \frac{object_average - breast_average}{object_average + breast_average} \quad (11)$$

Also, some basic statistics quantities were used:

- Average value (detected object and breast),
- Standard deviation (detected object and breast),
- Range (detected object and breast),

making a total of 8 characteristics used in the proposed RMRBF neural network.

5 Experimental Results

The first step to be done was to select the type of classification that the neural network was going to make. Because all of the literature related focuses only in microcalcification detection, the purpose here was to find also cancer abnormalities. That was the reason to purpose 2 main groups of classification: the first one (Group 1) will be constituted of normal images and benign abnormalities; the second one (Group 2), of images with any kind of microcalcification and malign abnormalities. The classification process is presented in Figure 2.

To train the network for getting the appropriate pdf's parameters were used 32 images (8 normal, 8 benign abnormalities, 8 malign abnormalities, and 4 benign and 4 malign microcalcifications), and to probe the efficiency were used 125 images (40 normal, 38 benign abnormalities, 30 malign abnormalities, 8 benign microcalcifications and 9 malign microcalcifications), all of them of the MIAS image collection.

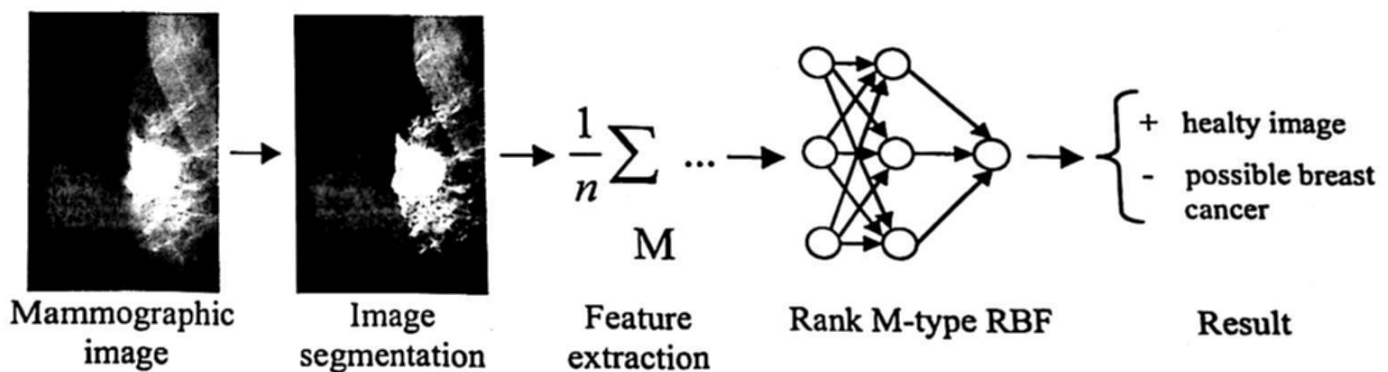


Figure 2. Classification process

The neural network RMRBF propose here is evaluated using the influence functions mentioned in section 3.4, and its performance is compared with the simple RBF algorithm, which was implemented according to its references [9, 10].

Tables 1 to 7 show the experimental results, for being the most important ones, in terms of efficiency, uncertainty and error for the image collection in the case of normal (NORMAL), benign abnormalities (AN_BEN), malign abnormalities (AN_MAL), benign microcalcifications (uC_BEN), and malign microcalcifications (uC_MAL). Table 1 shows the results obtained using the simple RBF neural network, and Tables 2 to 7 show the results obtained with the proposed RMRBF neural network.

Table 1. Results obtained by the simple RBF algorithm

SIMPLE RBF	NORMAL	AN_BEN	AN_MAL	uC_BEN	uC_MAL	TOTAL
Efficiency	52.50%	47.37%	30.00%	12.50%	55.56%	39.59%
Uncertainty	1.33%	0.00%	0.00%	0.00%	0.00%	0.27%
Error	46.17%	52.63%	70.00%	87.50%	44.44%	60.15%

The described RMRBF neural network with different influence functions has been evaluated with the simple cut and Tukey influence functions. The data that was used to get the pdf's parameters change in accordance with the variation of the r parameter value that was calculated as a factor of the difference between the mean and the minimum value of the data obtained for each characteristic in the training stage. For this reason the maximum value of $r \approx 2$ (*mean-minimum value*).

Tables 2 to 4 show three relevant results obtained from the proposed RMRBF using Simple Cut influence function with different values of r .

Table 2. Results obtained using the simple function with $r=0.3$

SIMPLE CUT	NORMAL	AN BEN	AN MAL	uC BEN	uC MAL	TOTAL
Efficiency	70.00%	68.42%	30.00%	37.50%	66.67%	54.52%
Uncertainty	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Error	30.00%	31.58%	70.00%	62.50%	33.33%	45.48%

Table 3. Results obtained using the simple function with $r=0.4$

SIMPLE CUT	NORMAL	AN BEN	AN MAL	uC BEN	uC MAL	TOTAL
Efficiency	72.50%	65.79%	23.33%	62.50%	66.67%	58.16%
Uncertainty	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Error	27.50%	34.21%	76.67%	37.50%	33.33%	41.84%

Table 4. Results obtained using the simple function with $r=0.5$

SIMPLE CUT	NORMAL	AN BEN	AN MAL	uC BEN	uC MAL	TOTAL
Efficiency	77.50%	63.15%	36.67%	50.00%	55.56%	56.58%
Uncertainty	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Error	22.50%	36.85%	63.33%	50.00%	44.44%	43.42%

From Tables 2-4 is observed that the results obtained with simple cut RMRBF are better than results shown in Table 1. Also one can see that the change of r value helps us to obtain a higher efficiency.

Tables 5-7 show the results obtained from RMRBF using Tukey influence function with the same values of r used with Simple Cut influence function.

We notice that the best results are given by the Simple Cut influence function that in one result approaches to 60% of efficiency, but there is still a big percentage of error for medical purposes. One important reason might be the irregularity of mammography images that makes difficult the segmentation stage. Examples of proper and improper results can be seen in Figures 3 and 4.

Table 5. Results obtained using the Tukey function with $r=0.3$

TUKEY	NORMAL	AN BEN	AN MAL	uC BEN	uC MAL	TOTAL
Efficiency	70.00%	63.16%	30.00%	37.50%	66.67%	53.46%
Uncertainty	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Error	30.00%	36.84%	70.00%	62.50%	33.33%	46.54%

Table 6. Results obtained using the Tukey function with $r=0.4$

TUKEY	NORMAL	AN BEN	AN MAL	uC BEN	uC MAL	TOTAL
Efficiency	75.00%	63.16%	26.67%	50.00%	55.56%	54.08%
Uncertainty	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Error	25.00%	36.84%	73.33%	50.00%	44.44%	45.92%

Table 7. Results obtained using the Tukey function with $r=0.5$

TUKEY	NORMAL	AN BEN	AN MAL	uC BEN	uC MAL	TOTAL
Efficiency	72.50%	65.79%	36.67%	37.50%	33.33%	49.16%
Uncertainty	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Error	27.50%	34.21%	63.33%	62.50%	66.67%	50.84%

One improvement that could be made to the RMRBF is taking the best efficiency of each one of the 8 characteristics used for all the values of r used in simulations, and combines them in a new RMRBF Neural Network for trying to get better results.

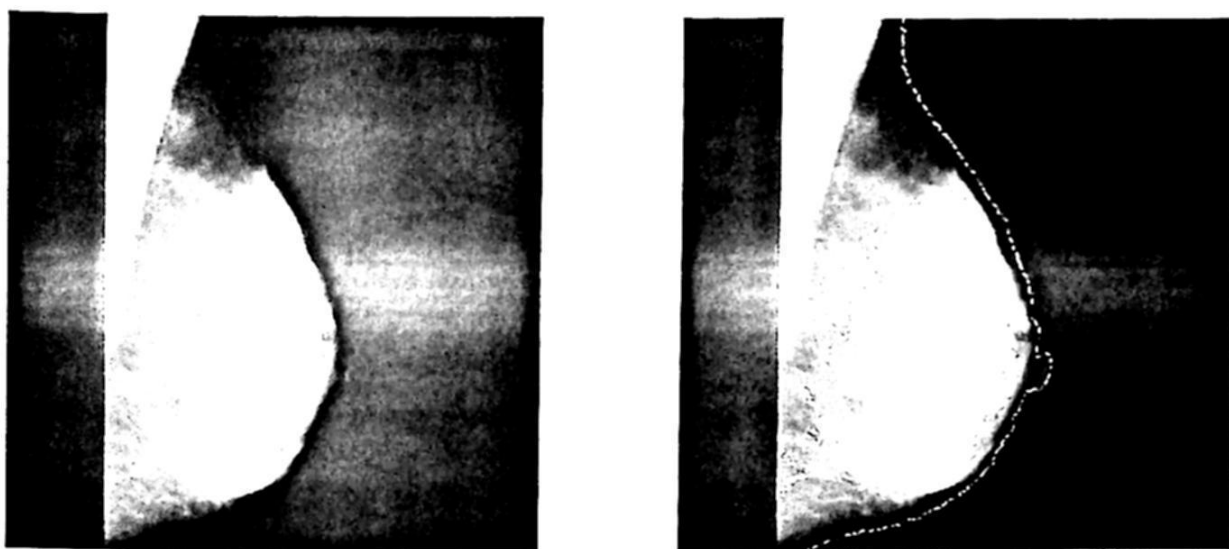


Figure 3. Mammography image with a proper result

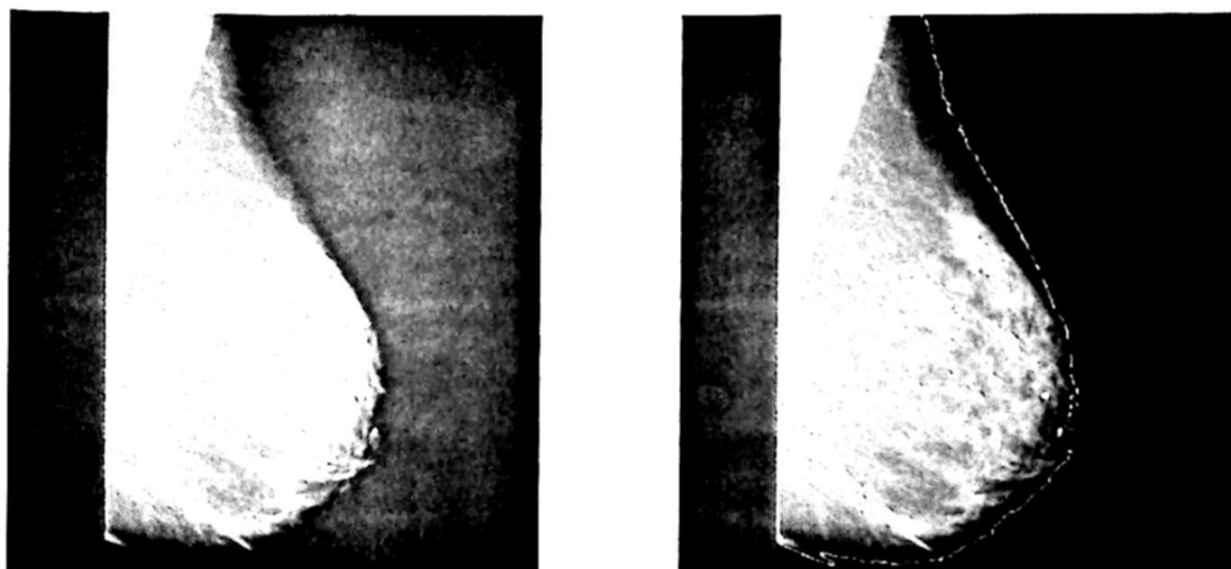


Figure 4. Mammography image with an improper result

6 Conclusions

We present the RMRBF Neural Network, it uses the RM-estimator in the scheme of radial basis function to train the proposed neural network. The results obtained with the use of the proposed RMRBF are better than others results obtained with simple RBF algorithms. Unfortunately the error is still big. The implementation of the Neural Network mentioned at the end of previous section could help, but also another segmentation algorithm should be implemented to see if there is a better differentiation between the regions of interest purposed in this paper.

6.1 Acknowledgements

The authors thank the National Polytechnic Institute of Mexico for its support.

References

1. Hudson D. L., Cohen M. E.: *Neural Networks and Artificial Intelligence for Biomedical Engineering*. Wiley-IEEE Press, Hoboken, New Jersey (1999)
2. Suri J. S., Rangayyan R. M.: *Recent Advances in Breast Imaging, Mammography, and Computer-Aided Diagnosis of Breast Cancer*. SPIE Press, Bellingham (2006)
3. Sajda P., Spence C., Pearson J.: Learning contextual relationships in mammograms using a hierarchical pyramid neural network. *IEEE Trans. Medical Imag.*, 21(3) (2002), 239-250
4. Haykin S.: *Neural Networks: a Comprehensive Foundation*. Prentice Hall, Upper Saddle River, NJ (1994)
5. M. Egmont-Petersen, D. de Ridder, H. Handels, "Image processing with neural networks - a review," *Pattern Recognition*, vol. 35, (2002) 2279-2301
6. Gallegos-Funes F. J., Ponomaryov V., Sadovnychiy S., Nino-de-Rivera L.: Median M-type K-nearest neighbour (MMKNN) filter to remove impulse noise from corrupted images, *IEEE Electronics Letters*, 38(15), (2002) 786-787
7. Gallegos-Funes F. J., Ponomaryov V.: Real-time image filtering scheme based on robust estimators in presence of impulsive noise, *Real Time Imag.*, 10(2), (2004) 69-80
8. Gallegos-Funes F. J., Varela-Benitez J. L., Ponomaryov V.: Real-time image processing based on robust linear combinations of order statistics, *Proc. SPIE 6063, Real-Time Image Processing 2006*, 177-187, San Jose, USA, (2006)
9. Buhmann M. D.: *Radial Basis Functions: Theory and Implementations*, Cambridge University Press, Cambridge, (2003)
10. Park J., Sandberg J.W.: Universal approximation using radial basis functions network, *Neural Computation*, vol. 3, (1991) 246-257
11. <http://www.wiau.man.ac.uk/services/MIAS/MIAScom.html>
12. www.opolanco.es
13. members.tripod.com/~gineco
14. González R. C., Woods R. E.: *Tratamiento Digital de Imágenes*, Addison Wesley/Díaz de Santos, (1996)
15. Myler H. R., Weeks A. R.: *The Pocket Handbook of Image Processing Algorithms in C*, Prentice Hall (1993)
16. Ritter G.: *Handbook of Computer Vision Algorithms in Image Algebra*, CRC Press, Boca Raton-New York (2001)