Proceedings of the V Brazilian Conference on Neural Networks - V Congresso Brasileiro de Redes Neurais
pp. 577–580, April 2–5, 2001 - Rio de Janeiro - RJ - Brazil

Classifier scheme for clustered microcalcifications in digitized mammograms by using Artificial Neural Networks

Ana Claudia Patrocinio[1]; Homero Schiabel[1];
[1]Departamento de Engenharia Elétrica - EESC - Universidade de São Paulo
Av. Trabalhador Sãocarlense, 400 - 13566-590 - São Carlos, SP - Brasil
e-mails: acpatroc@sel.eesc.sc.usp.br, homero@sel.eesc.sc.usp.br

Abstract

Computer-Aided Diagnosis (CAD) schemes have presented good results in aiding the early diagnosis of breast cancer. The detected signals classification demands multi-works investigations, since cytological characteristics concerning the mammographic findings have to be investigated in addition to computer techniques. Artificial neural networks (ANN) have been successfully used in CAD classifiers, with success in the classification in CAD. For example, the classification of clustered microcalcifications has been made from individual microcalcifications analysis. In this work, regarding characteristics determined only from the cluster itself, and discarding the characteristics analysis and extraction from individual microcalcifications, such a classification was made in two classes: “non- suspect” and “suspect”. Dismissing microcalcifications individual features for the network input has allowed to eliminate procedures intended to separate each structure from the whole image. The classifier using ANN has shown the geometric descriptors efficiency for characterizing microcalcifications clusters as well as the influence of features extracted from images reports, as "age" and "density". The best data have shown 92% of correct results, with $A_z = 0.96$.

1.Introduction

Many models using pattern recognition and artificial intelligence have been contributing for investigations in medical images processing and evaluation with purposes of aiding in diagnosis ([1] to [3]). Computer-aided diagnosis (CAD) schemes have been proposed in order to provide a “second opinion” to the breast cancer detection by using mammographic images ([2] to [4]). These procedures have aided not only in detecting microcalcifications, but also in classification processes, by addressing a mammography finding to a malignant or benign tumor ([2], [4] to [6]). Microcalcifications identification, mainly if clustered, is an important information to diagnosis and to the future procedure to take; thus, it is an object of interest for these schemes. By the digitized mammogram, features are extracted from the images, which characterize clusters of microcalcifications, so that they are classified according to shape, size, number of microcalcifications in a cluster, among other factors. Information can be obtained by mathematical morphology in image processing, for example [5].

Artificial Neural Networks (ANNs) are often used as classifiers in CAD schemes, with good results ([2], [5] to [7]). Characteristics extracted from digitized mammograms are the input data for lesion classification. The classifier performance is dependent upon two previous steps: the pattern acquisition (mammographic image) and selection and extraction of features which better represent each class [8]. As part of a mammography CAD scheme, our purpose in this work was to develop a classifier based on ANN in order to classify clustered microcalcifications as “suspect” and “non suspect” regarding a better addressing of the medical diagnosis, avoiding, for example, unnecessary requests for biopsies.

The images features used can be commonly ambiguous due to the frequent small sizes of mammographic findings. Thus these measures should be normalized so that even small values variations among the different classes can be distinguished by the classifier scheme.

For classifying the cases in “suspect” and “non-suspect” from the clusters analysis, geometric descriptors, as first, second and third orders moments [9], were extracted, the Hu Invariant Moments ([7], [9], [10]) and the cluster radius gyration were calculated, and also some information from the patient medical report was used, as patient age and a “density” measure. Characteristics extracted from the clusters did not have considered the individual microcalcifications, but only the cluster itself.

2.Methodology

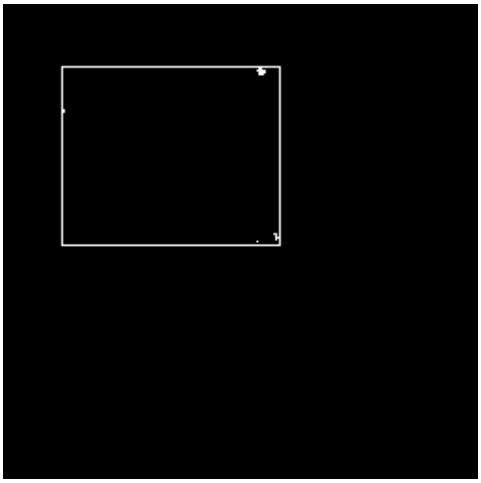
In this work 86 regions of interest (ROIs) were used, each one with 1 cm^2, all of them containing clustered microcalcifications. Mammograms were obtained from exams at the Hospital das Clínicas at Ribeirão Preto, SP, Brazil, together with the corresponding medical reports. The images were digitized in a laser scanner (Lumisys) with 0.15 mm of spatial resolution and gray scale corresponding to 8 bits.

ROIs were selected according to the medical reports, and a segmentation process technique [11] was applied to each image. The characteristics were calculated from

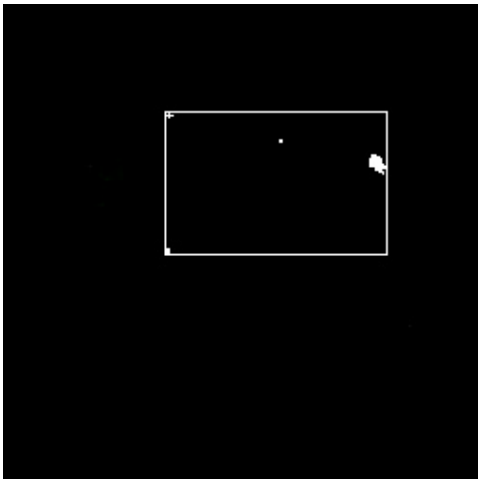
the segmented images, by considering only one cluster by ROI. Then, the cluster was detached and set in a rectangle with vertices corresponding to the microcalcifications located in the extremities. This means to determine the smallest rectangle containing all the microcalcifications in each ROI, as shown by Figure 1.

The rectangles shown in Fig. 1 were considered the objects for moments calculation. The Hu Invariant Moments ([9], [10]) – Hu1 to Hu7 – the radius gyration and the surface of every white signals (the possible microcalcifications) found in the clusters ($moment_{00}$), in addition to the second and third orders moments [9], were the characteristics extracted from each region.

The selection of the characteristics used to the classifier input was made by analysis of gaussian curves, verifying overlapping of the characteristics for two differents classes as showed by Figure 2, and they are described in Table 1. For training the network, such curves were compared to those more used according to the literature ([2], [3], [4], [5] and [7]) in order to select the most adequate characteristics to be used.



(a)



(b)

Figure 1: (a) and (b) Clusters detached in ROIs

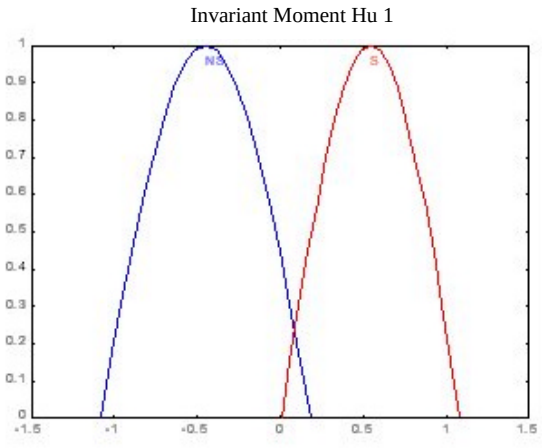


Figure 2: The red curve present the “suspect” class and the blue curve present the “non-suspect” class.

Table 1- Geometric descriptors (GD) as ANN input features

Features	Description
ϕ_1	Invariant Moment Hu 1
ϕ_2	Invariant Moment Hu 2
ϕ_3	Invariant Moment Hu 3
ϕ_4	Invariant Moment Hu 4
RG	Radius Gyration
I	Irregularity
C	Compactness
$Momento_{00}$	Area of microcalcifications
P	Perimeter
A	Area of Cluster

The Radius Gyration can be understood as the minimum radius including the most or all the signals inside the area it limits. It can represent the spatial distribution of significant image pixels [9]. The compactness and irregularity are measures that show no similarity of the object, this measure are extract through area and perimeter [4].

The system for clusters features extraction was developed in *Delphi 4.0*. The ANN with backpropagation algorithm [12] was used for classifying the ROIs containing mammographic findings in two classes: “suspect” and “non-suspect”. Also the ANN was implemented in *Delphi 4.0*.

The transfer function used in the neurons of ANN layers was the sigmoid, since this function has shown the best performance during the network training. The ANN was considered trained when the global error reached a valor the less of 0.0008.

The network was trained and tested with different configurations regarding the number of neurons in the hidden layer, the number of neurons in the hidden layer was empirical changed . The number of neurons in the input layer corresponds to the number of characteristics to be presented to the network and the number of neurons in the output layer corresponds to the number of classes to be recognized.

Tests were performed by dividing the features in four sets: (a) only geometric descriptors; (b) geometric descriptors plus patient age; (c) geometric descriptors plus "density"; and (d) geometric descriptors plus "density" plus patient age.

The samples set composed by 86 images was divided so that 70% of images were used for training and the other 30% for tests. From these images, 50% was correspondent to "non-suspect" cases and other 50% to "suspect" cases, according to the medical reports.

3. Results and discussion

The best result obtained during the tests corresponded to a combination of 10 input network characteristics, which were the four first Hu Invariant Moments, $Moment_{00}$, Radius Gyration, Irregularity, Compactness, Perimeter and Area.

For the ANN trained according to the "standard" mode [12], 86 examples – divided in training set and test set – were used. During the training, 70% of all data were presented to the ANN, being half of each class ("suspect" and "non-suspect"), as previously mentioned. The other 30% was used for tests, from which 48% were stated as "suspect" and 52% as "non-suspect".

The variation in the number of neurons in the hidden layers has registered influence on the classifier performance, as shown in Figure 3, for two groups: Group I presents 35 hidden neurons, divided for 25 neurons to first hidden layer and 10 neurons to second hidden layer and Group II with 45 hidden neurons, divided for 30 neurons to first hidden layer and 15 neurons to second hidden layer.

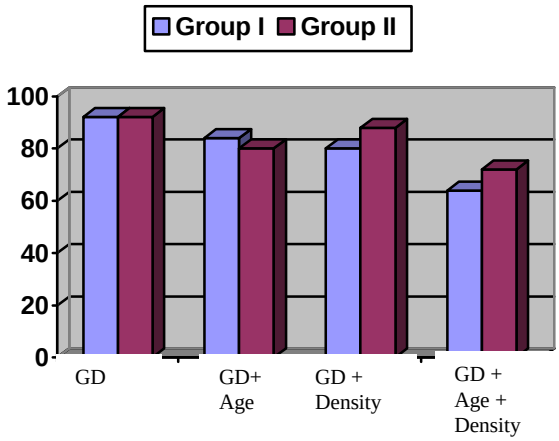


Figure 3: Classifier performance as a function of the number of neurons in the hidden layer.

During the training, the learning rate has ranged from 0.3 to 0.8 and the moment from 0.08 to 0.50. The ANN has converged faster in training with learning rates close to 0.5 and moment values close to 0.1. The average time for training was 11.3 h, and the largest times were registered for networks trained with larger number of input neurons.

Results were analyzed by ROC curves [13] as well as by percentage rates. The best results have reached 92% of true data. Table 2 describes in detail such results, expressed as percentage of correct and incorrect classifications according to the four features set considered.

Table 2 – Best results for the ANN implemented

FEATURES	Number Neurons	TP (%)	FP (%)	TN (%)	FN (%)	Total (%)
Geometric descriptors	35 and 45	48	8	44	0	92
Geometric descriptors Plus age	35	48	16	36	0	84
Geometric descriptors Plus "density"	45	48	12	40	0	88
Geometric descriptors Plus age and "density"	45	48	27.5	24.5	0	72,5

In Table 2, the terms TP, FP, TN and FN corresponding, respectively, to the true positive, false positive, true negative and false negative responses given by the classifier scheme output.

In Figure 4, the best result obtained from the classifier evaluation – and according to the data shown in Table 2 – is expressed as a ROC curve, showing essentially the relationship between the correct and incorrect rates yielded by the classifier.

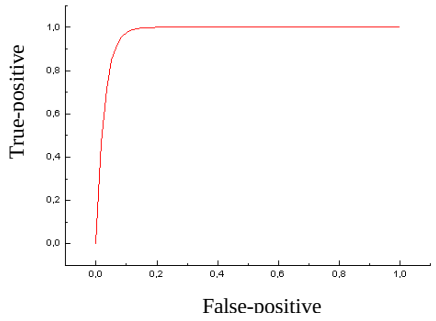


Figure 4: ROC curve corresponding to the test with the largest percentage of correct data.

The area under the ROC curve represents the classifier accuracy index, that is, the larger the area, the larger the number of correct outputs. This means that in a very accurate system, the curve should be as close as possible to the left superior corner in a graph as that in Fig. 3 – larger area with value close to 1.0. Therefore, this analysis leads to low false negative rates when the true positive is high, since the ideal system should have a false positive and false negative probabilities rates equal to zero. Thus, in an actual system, the aim is obtaining these two indexes as low as possible. In this

work the best area under the ROC curve was 0.96 as presented in figure. 4.

Therefore, one can see from Table 2 that the classifier scheme has returned no false negative rate, which corresponds to a good and expected result. Nevertheless, classifying correctly all the positive cases has a cost reflected in the false positive data, which is increased in some cases (mainly in the last line in Table 2). The good results are due to the fact that the images database is well representative for the type of diagnosis studied, since clusters, when present in the image, correspond already to a higher level of suspicion, suggesting a further and better investigation for that breast.

4.Conclusions

We could note from this work that only with the clusters characteristics – without considering individual microcalcifications features – good results could be obtained in testing the classifier, which are comparable to several others presented in literature ([2], [5], [6], [7] and [13]).

We also could observe the influence of the number of neurons in the hidden layers on the ANN performance, as well as the influence of the set of input features on the classifier.

The classification in "suspect" and "non-suspect" is also a manner of not stating or suggesting the diagnostic by the computational scheme, but it is a way of provide an objective information to the radiologist about the effective attention to be given to each case. Indeed, we consider that defining a case as positive ("the patient has breast cancer") or negative ("the patient does not have breast cancer"), although used by some CAD schemes, is speculative, since it is not so simple and, at the end, a responsibility of the physician. Therefore, our main purpose in the implementation of this classifier was to help the medical decision, as to aid the radiologist in stressing details which could allow, for example, to avoid unnecessary biopsies, as well as to save life by an accurate early diagnostic.

Acknowledgments

The authors are grateful to CAPES and FAPESP by the financial support, and also to Hospital das Clínicas at Ribeirão Preto, SP, Brazil, which has provided the mammograms used for the tests.

References

[1] Doi, K. "Perspectives on digital image analysis in medical image. Potencial usefulness of computer-aided diagnosis". *Anais do III Fórum Nacional de Ciência e Tecnologia em Saúde*, vol. 2, pp. 795-796, 1996.

[2] Jiang, Y.,et al., "Requirement of microcalcification for computerized classification of malignant and benign clustered microcalcifications". *Proceedings of SPIE*, vol. 3338, pp. 313-317, 1998.

[3] Veldkamp, W. J. H., et al., "An improved method for detection of microcalcification clusters in digital mammograms", *Proceedings of SPIE*, vol. 3661 pp. 512-522, 1999.

[4] Tao, E., et al., "Automatic detection of microcalcifications in digital mammography". *Proceedings of SPIE*, vol. 3338, pp. 1450-1458, 1998.

[5] Jiang, Y., et al., "Malignant and benign clustered microcalcifications: automated feature analysis and classification". *Radiology*, vol. 198, pp. 671-679, 1996.

[6] Yu.; et al., "Automatic detection of clustered microcalcifications in digital mammograms based on wavelet features and neural network classification". *Proceedings of SPIE*, vol. 3338, pp. 1540-1546, 1998.

[7] Shen,L., et al., "Detection and classification of mammographic calcifications". *Intern. J. Pattern Recogn. and Artif. Intelligence*, vol. 7 (6), pp. 1403-1416, 1993.

[8] Duda, R.O, et al., *Pattern Classification and Scene Analysis.* John Wiley & Sons, New York, 1973.

[9] Masters, T., "*Signal and Image Processing with Neural Networks"* - John Wiley & Sons, Inc. - 1994.

[10] Hu, M. K., "Visual Pattern Recognition by Moments Invariants" - *IRE Trans. Inf. Theory*, vol. 8 - pp. 179-187, 1962.

[11] Schiabel, H., et al., "A computerized scheme for detection of clusters of microcalcifications by mammograms image processing" - *World Congress On Medical Physics And Biomedical Engineering*, Nice(France), pp. 705, 1997.

[12] Hertz, J., et al., *Introduction to the Theory of Neural Computation.* Addison_Wesley Publishing Company, 1991.

[13] Chan, H-P., et al., "Computerized classification of malignant and benign microcalcifications on mammograms: Texture analysis using articificial Neural network". *Phys. Med. Biol.*, vol. 42, pp. 549-567, 1997.