Investigations on the Effect of Different Characteristics of Images Sets on the Performance of a Processing Scheme for Microcalcifications Detection in Digital Mammograms

Homero Schiabel, Fátima L.S. Nunes, Mauricio C. Escarpinati, and Rodrigo H. Benatti

The performance of a computer-aided diagnosis (CAD) scheme is closely dependent on the database used for its development and tests. The scheme sensitivity can be reduced by 15% to 25%, with only 20% of changes in the database cases. Previously, we have developed a processing scheme in order to detect clustered microcalcifications in digital mammograms, and we have tested such a procedure with two different databases. Further evaluations in developing a CAD scheme for mammography have indicated the need for more extensive investigation on the effects resulting from different characteristics of the images bank used for tests. Therefore, this work reports some results regarding such an investigation, with a further discussion over characteristics that can affect the performance of a CAD scheme.

Copyright © 2001 by W.B. Saunders Company

THE PERFORMANCE of a computer-aided diagnosis (CAD) scheme is strongly dependent on the database used for testing the procedures.¹ In a previous work,² we developed a computerized procedure as part of a mammography CAD scheme intended to detect clustered microcalcifications in digitized mammograms. The first tests with this scheme indicated an efficacy rate of 94% for a particular set of images. With the goal to study the effect of the characteristics of different sets of images on the performance of such a procedure, we have checked the results yielded from five different sets of mammograms.

METHODS

The following sets of images were used for the investigation:

(1) a set of actual mammograms, obtained from the archives of Hospital das Clínicas, at Ribeirão Preto (Brazil), digitized by a UMAX Powerlook II (UMAX Technologies, Fremont, CA) scanner (light + charge-coupled device [CCD]), with 0.05 mm and 8 bits of spatial and contrast resolution, respectively; (2) the

same set digitized by a LUMISCAN (Lumisys, Sunnyvale, CA) laser scanner, with 0.15 mm and 8 bits; (3) another set of mammograms obtained from that same hospital in 1995, and digitized in a UMAX UC1260-Pro, with the same resolutions as the first group; (4) a set of mammograms obtained from recent examinations performed at the Santa Casa hospital at S. Carlos (Brazil), only corresponding to dense breast cases and digitized by the LUMISCAN scanner, with the same resolutions as the second group above; and (5) a previously used set obtained off the internet from the National Expert and Training Center for Breast Cancer Screening (the University of Nijmegen, the Netherlands), also with the same resolutions as the second and fourth groups. The steps of the processing scheme for clusters detection were as follows: identification of regions of interest (ROIs); ROI segmentation; area-point transformation,³ which converts each detected signal in a unique pixel; and microcalcification grouping, where clusters are identified and marked in the final image. Initially, the procedure was applied to all of the image sets with no changes in the processing parameters, in order to investigate the effect of the acquisition and digitization processes. Then, some parameters were changed according to the database characteristics under test in order to determine the best result possible for that set of images.

RESULTS

Figure 1 illustrates the resultant images after some steps during the processing. Table 1 shows results from the procedures for the different image sets. The total of positive and negative cases (that is, with and without microcalcifications clusters, respectively) according to medical reports for each set of images was as follows: set 1—30 positive, 30 negative; set 2—30 positive, 30 negative; set 3—25 positive, 25 negative; set 4—20 positive, 30 negative; and set 5—34 positive, 16 negative.

CONCLUSIONS

The results have shown significant variations, mainly if we compare the values for true and false rates among the image sets. The best performance revealed an efficacy of 95%, and a variation of about 15% in the correct results could be detected by comparing the scheme performance among the databases; in addition, the variations were more significant in the analysis of the errors rates, mainly the false-positive cases, showing important dependency on the image set cases. An important result can be seen comparing the first two lines in Table 1: the same set of original mammograms yielded different results after we applied the pro-

From the Departamento de Engenharia Elétrica EESC, and the Departamento de Física e Informática IFSC, Universidade de São Paulo, Brasil.

Address reprint requests to Homero Schiabel, Departamento de Engenharia Elétrica EESC, Universidade de São Paulo, Av. Trabalhador São-carlense, 400-13566-590 São Carlos, SP, Brasil. E-mail: homero@sel.eesc.sc.usp.br.

Copyright © 2001 by W.B. Saunders Company

0897-1889/01/1402-1065\$35.00/0

doi:10.1053/jdim.2001.23885

Fig 1. Example of results from the processing (image from the set 2): (a) original ROI; (b) ROI after segmentation; (c) ROI after area-point transformation²; (d) ROI with a detected cluster.

cessing for detecting clusters when digitized by different scanners. As expected, when processing images from the laser scanner (set 2), we obtained no false-negative results, although for set 1, we registered a 3% false-negative rate. This shows that the scanner type and technology are critical features in defining the performance of not only a particular processing scheme, but also the whole CAD process.

Furthermore, factors such as structures in the breast, image acquisition, and film development pa-

rameters, as well as spatial and contrast resolution, the digitization process, in addition to the parameters of image processing, could be responsible for such variations. Therefore, we think changes in some of these factors during image acquisition and improved image processing techniques—mainly when the database is composed of dense breast images—could minimize the effects of those variations for a CAD scheme performance.

REFERENCES

1. Nishikawa RM, Giger ML, Doi K, et al: Effect of case selection on the performance of computer-aided detection schemes. *Med Phys* 21:265-269, 1994
2. Schiabel H, Nunes FLS, Marques PMA, et al: A computerized scheme for detection of clusters of microcalcifications by mammograms image processing. *Med Biol Eng Comput* 35: 705, 1997 (suppl 2, abstr)
3. Nishikawa RM, Giger ML, Doi K, et al: Computer-aided detection and diagnosis of masses and clustered microcalcifications from digital mammograms, in Bowyer KW, Astley S (eds): *State of the Art in Digital Mammographic Image Analysis*. Singapore, World Scientific Publishing, 1993, pp 82-102

Table 1. Results From Tests of Processing for the Five Image Sets

Set	TP (%)	FP (%)	TN (%)	FN (%)
1	47	2	48	3
2	50	5	45	0
3	94	6	0	0
4	38	8	52	2
5	70	12	18	0

NOTE: Rates were calculated relative to the total number of cases tested.

Abbreviations: TP, true-positive rate; FP, false-positive rate; TN, true-negative rate; FN, false-negative rate.