



Validação de um programa semiautomática para a avaliação de microcalcificações em um phantom de mama

Validation of a semiautomatic program for evaluation of microcalcifications on a breast phantom

C. M. Vitoria^{1*}; L. Paixão,^{2,3}; P. Santana³; F. Leyton^{2,4}; M. S. Nogueira^{1, 2}.

¹Departamento de Engenharia Nuclear, Universidade Federal de Minas Gerais (UFMG), CEP: 31270-970, Belo Horizonte/MG, Brasil

²Pós-Graduação em Ciências das Radiações, Centro de Desenvolvimento da Tecnologia Nuclear (CDTN), CEP: 30161-970, Belo Horizonte/MG, Brasil

³Departamento de Anatomia e Imagem/Faculdade de Medicina, Universidade Federal de Minas Gerais (UFMG), CEP: 30130-100, Belo Horizonte/MG, Brasil

⁴Faculty of Health and Odontology, Diego Portales University, Santiago, Chile and Radiological Sciences Center, Health Sciences Faculty, Tarapacá University, Arica, Chile.

*cmvb@cdtn.br

(Recebido em 06 de outubro de 2015; aceito em 14 de abril de 2016)

A fim de garantir de garantir que o rastreamento e diagnóstico mamográfico são feitos corretamente e de forma eficiente, a Vigilância Sanitária do governo de Minas Gerais, no Brasil criou o 'Programa de Monitoramento Mensal de Qualidade em Mamografia dos serviços públicos e privados do Estado de Minas Gerais' no Brasil, Imagens do phantom do Colégio Brasileiro de Radiologia são enviadas para a Agência Governamental onde são avaliadas visualmente. O objetivo deste trabalho é desenvolver um programa para a avaliação semiautomática das microcalcificações presentes nas imagens do phantom. Técnicas de processamento digital foram aplicadas nas imagens, com o objetivo de caracterizar cada objeto de interesse presente no phantom de mama. Os algoritmos foram desenvolvidos em MATLAB e a avaliação dos objetos teste foi feita de acordo com parâmetros estabelecidos pelo Ministério da Saúde do Brasil. O nível de concordância entre a avaliação visual e computacional é considerado "razoável", de acordo com o estatístico "coeficiente kappa" o qual determina a força de concordância entre os métodos, o valor de kappa foi de 0,291. O programa pode se tornar uma boa ferramenta na automação das avaliações das imagens do phantom.

Palavras-chave: Mamografia, Controle de Qualidade, Porcessamento Digital de Image

In order to ensure to ensure that screening and mammography diagnosis are made properly and efficiently, the Sanitary System of the government of Minas Gerais in Brazil created the Program of Monthly Monitoring of Quality in Mammography of public and private facilities State of Minas Gerais in Brazil, Images of the Brazilian College of Radiology phantom are sent to the Governmental Agency where they are assessed visually. The objective of this work is to develop a program to semiautomatic the assessment of the microcalcifications present in the images. Techniques of digital processing were applied to the images, aiming to characterize each object of interest present in the breast phantom. The algorithms were developed in MATLAB and assessment of test objects were made according to parameters established by the Ministry of Health of Brazil. The level of concordance between visual and computational assessment is considered 'fair', according to the statistical 'kappa coefficient' which determines the strength of agreement between the methods, the value was 0.29. The program can become a good tool in the automation of the assessments of phantom images.

Keywords: Mammography, Quality Control, Digital Image Processing.

1. INTRODUCTION

Breast cancer is the second most common type of cancer in the world and the most common amongst women. In Brazil, mortality rates related to breast cancer remain high, most likely because the disease is still diagnosed in advanced stage [1]. Mammography is the most effective method in the diagnosis of illnesses in the breast tissue.

Some important signs of breast cancer sought by radiologists are clusters of microcalcifications, masses, and architectural distortions [2]. One of the main alterations in the breasts found through mammography are the microcalcifications that are tiny calcium deposits which appear as small white spots on the film [3].

Aiming at keeping the quality standards indispensable in mammography, The Superintendence of Sanitary Vigilance (SVS) instituted in 2007 the 'Program of Monthly Monitoring of Quality in Mammography Facilities for State and Private Facilities in the State of Minas'. This monitoring consists in assessing, on a monthly basis, the quality of the radiographic image of a breast phantom for each mammography equipment. Each facility must submit a phantom image to the SVS. The images are then analyzed and the results publicized by the end of each month. The assessment is made in accordance with the criteria defined by the guidelines "Medical Radiodiagnosis Safety and Performance of Equipment" of National Health Agency [4]. The certified phantom by the Brazilian College of Radiology (BCR) is used for this purpose.

Some authors have developed programs to analyze the quality of digital images in various mammography phantoms in the marketplace through techniques of automatic processing. [5], [6], [7] and [8].

In some studies, as the one done by Chakraborty and Eckert, [9] and Lee et al. [10], readings done by observers of mammographic images at the American College of Radiology (ACR) were compared with a computerized analysis, the results showed that accuracy rises when computerised analysis is used and demonstrated inconsistency in the analysis done by the human observer, given to differences in the interpretation of the criteria.

The aim of this work is to validate the algorithms to carry out an semiautomated assessment of the quality of the images obtained through a BCR phantom, and compare it with the visual assessment of the same images done by a human trained by the Monthly Monitoring program.

2. MATERIAL AND METHODS

The Brazilian College Radiology (BCR) breast phantom.

The phantom is made up of acrylic plates which simulate a compressed breast of 5 cm of thickness, and mean composition (50% of glandular tissue and 50% of adipose tissue). It consists of three plexiglass (Polymethylmethacrylate – PMMA) plates of 10 x 120 x 160 mm³, and one plate of 20 x 120 x 160 mm³. This last plate contains a 5 x 70 x 140 mm³ inserted wax and a step wedge that produces five optical densities used to assess image contrast. In the inserted wax are embedded the following structures: four metal grids of 4, 6, 8 and 12 lp/mm (G1 to G4) to evaluate spatial resolution; five groups of Al₂O₃ specks of 0.45, 0.35, 0.30, 0.25 and 0.18 mm diameter (m1 to m5) to simulate microcalcifications; eight polyester discs of 6 mm diameter with thickness ranging from 0.1 to 0.8 mm (D1 to D8) to simulate low contrast areas; six nylon fibers of 1 cm length and 1.4, 1.2, 0.8, 0.7, 0.6 and 0.4 mm diameters (F1 to F6) to simulate fibrils; and five nylon spherical caps of 2.00, 1.5, 1.0, 0.75 and 0.50 mm heights (M1 to M5) to simulate tumor-like masses [11]. A schematic diagram of the BCR breast phantom is showed in Figure 1. In each category of test objects, the size and/or contrast of objects gradually decreases. For example, microcalcifications group m1 have larger grains than m5 group.

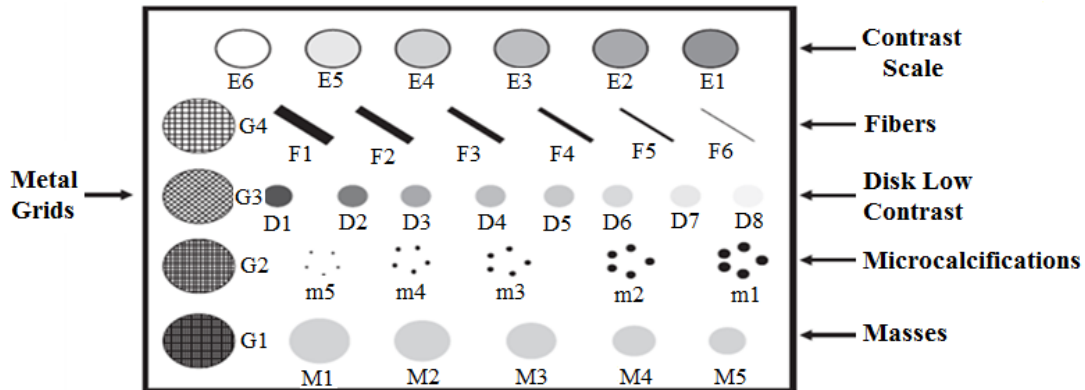


Figure 1. Schematic diagram of BCR breast phantom

The 50 images of the breast phantom studied were chosen randomly from image database of the Superintendence of Sanitary Vigilance in Minas Gerais state. In order to provide semiautomatic assessment, these images were digitalized at scanner ScanMaker 9800XL from Microtek at a resolution of 1200 dpi and saved in TIFF (Tagged Image File Format), which does not presents loss of information. In the visual assessment, the images were analyzed on radiologic film.

The algorithms used in the assessment of the microcalcifications were implemented in a MATLAB (Matrix Laboratory) R2012b[®] platform. It has Image Processing Toolbox[™], which provides a comprehensive set of reference-standard algorithms, functions, and apps for image processing, analysis, visualization, and algorithm development.

Semiautomatic and Visual Assessment

Specific software tools for application in the analysis of the BCR breast phantom were developed. The algorithms are based on standard digital image processing techniques, such as the softening methods, spatial filtering to remove noise, image thresholding to detect objects, filtering in the frequency domain, application of morphological operators, and so on [12].

The regions of interest (ROI) were selected manually, taking into interest each microcalcifications groups. The segmentation process in each one of ROIs was carried out; the segmentation consists in detecting and isolating each of the test objects so that they can be assessed. Wiener filter was applied between pixels, using neighborhoods of size 3 by 3 to estimate the local image mean and standard deviation. The Wiener filter is a 2-D adaptive noise-removal filtering. After, a thresholding was applied, for it was calculated the average and the standard deviation of ROI, and was established as threshold the sum of the pixel average value plus three standard deviations [13].

Morphological operators were applied based on the principle of Mathematical Morphology, which consists in extracting the geometry and topology of an unknown image through the transformation of an image fully defined – namely structuring element, which in such case are disk-shaped structuring element with the radius of 3 pixels. The applied morphological operators were in sequence: Dilate (Expanding the foreground), Erode (Shrinking the foreground), Open (Removing stray foreground pixels in background), Close (Removing holes in the foreground) and finally, Clean to removes isolated pixels.

Then, an algorithm for counting structures in the image was applied. It consists in the identification and counting of centroids in the white areas of the image.

In order to prevent any sort of subjectivity, an algorithm for counting structures in the image was applied. This algorithm consists in the identification and counting of centroids in the white areas of the image, classifying the image assessed as approved if the identification of the fourth group of microcalcifications is possible.

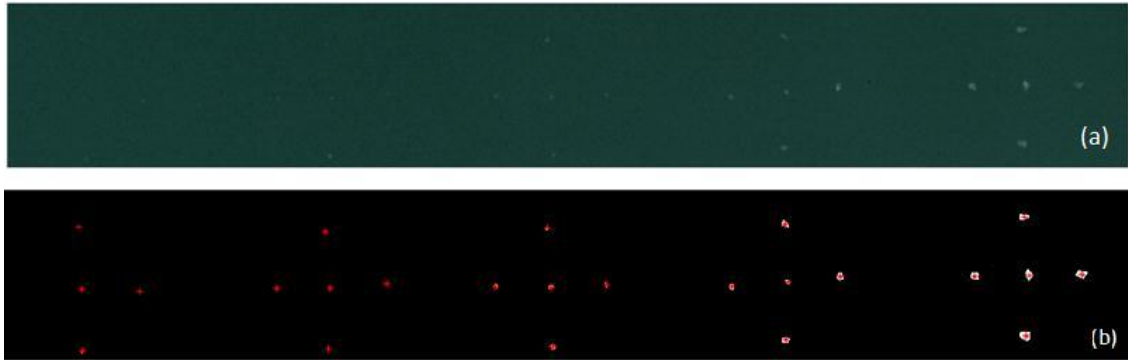


Figure 2. Groups of microcalcifications before (a) and after processing (b).

To visual assessment, the radiography of BCR breast phantom was positioned on negatoscope; the number of structures visible in the image were analyzed by four appraisers using the aid of a magnifying glass, when judged necessary. The assessment criteria adopted were the same as used in automatic evaluation.

3. RESULTS AND DISCUSSION

The results show the assessment of 50 images studied, the comparison between the mean visual assessment the group of microcalcifications made by trained appraisers with the assessment obtained by the algorithms (Figure 3). The images were coded MG followed by the image number. It is possible to observe a disagreement about the visual assessment.

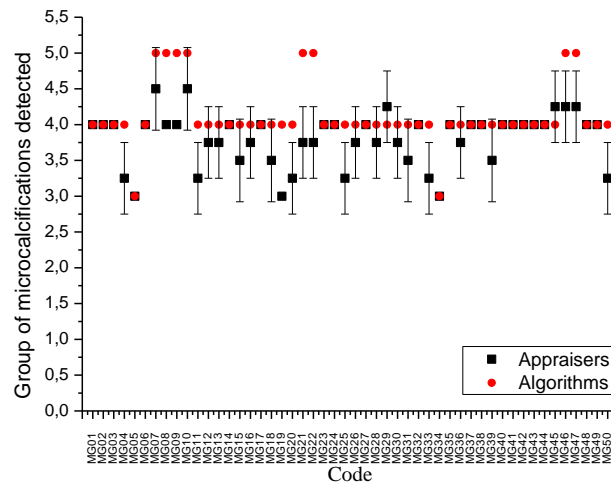


Figure 3. Comparison of visual and computerized assessment of group of Microcalcifications in BCR phantom images.

Human limitations created subjectivity in the visual analysis of the images; the observed differences between the appraisers would be attributed potentially inherent differences in their corresponding visual perception of contrast and spatial resolution. As shown Ginesu [14], based on the Human Visual System (HVS), the three subjective factors of distortion that affect image quality assessment are Block distortion (blockiness), edge errors and visual disparities, and these factors may affect the visual assessment.

The number of detected microcalcifications computerized is significant respect the number that was given by visual assessment, this can be explained by the fact that a computer can

process much more information than what a human eye can see, which was also reported by Brooks et al. [15] and Chakraborty and Eckert [9].

The degree of concordance between the methods used is a measure used to analyze the agreement between appraisers and the computerized assessment. This procedure is generally referred to as interjudge reliability and interobserver reliability. The statistical tool of reliability is the kappa Coefficient (κ), described in 1960 by Cohen [16]. The kappa is a measure of agreement intra and inter-observer. It adopts a value between -1 to +1, being the strongest degree of interobserver concordance equal to +1. On the contrary, a value of $\kappa = 0$ reflects that the observed concordance is precisely the one that is expected by chance. The interpretation of Kappa Coefficient is performed by correlating its value with a qualitative scale, which includes six level of strength of agreement ("poor", "slight", "fair", "moderate", "substantial" and "almost perfect"). Table 3 corresponds to the qualitative scale used frequently to express the strength of the agreement, according to Landis and Koch [17]

Table 1. Evaluation of the kappa coefficient (Landis and Koch, 1977).

Kappa (κ)	Force concordance
< 0,00	Poor
0,01 – 0,20	Slight
0,21 – 0,40	Fair
0,41 – 0,60	Moderate
0,61 – 0,80	Substantial
0,81 – 1,00	Almost perfect

The kappa value was calculated using the online free software 'Graphpad' [18]. Five categories were defined one for each group calcifications, comparing the visual and computerized methods as shown in Table 2.

Table 2. Input data comparing the concordance of visual and computational method.

		Visual Assessment				
		m1	m2	m3	m4	m5
Computerized Assessment	m1	0	0	0	0	0
	m2	0	0	0	0	0
	m3	0	0	4	7	0
	m4	0	0	1	34	4
	m5	0	0	0	0	0

The kappa value obtained was 0.291 with a standard deviation of 0.145, at a confidence interval of 95% from 0.006 to 0.575, the strength of agreement is considered to be 'fair'.

Although visual and computerized assessments are concordants, the result is considerably satisfactory, computerized analysis proved to be a good alternative in terms of reducing or even eliminating the subjectivity inherent in visual methods of assessment of images of a breast phantom.

4. CONCLUSION

An algorithm was developed to assess images of test objects in a BCR breast phantom. The efficiency of this algorithm was evaluated by comparing its performance to the concordance with the average performance of the four appraisers.

When developing this algorithm, it was possible to notice that after comparing various types of filters, the Wiener filter was more adequate in the removal of the noise created by the digitalization of the images of the phantom, as in Pragathi et al. [13].

This algorithms turned out to be useful in the reduction of subjectivity, the results of estimated assessment by the semiautomated method was ‘fair’ according the kappa coefficient, used to calculate the concordance between the two methods of assessment.

Although the program is still under construction, not entirely automatic yet, since in this work it was necessary to select the ROI manually to perform the assessment, it demonstrated the feasibility of replacing the assessors with an automatic program, thus reducing the subjectivity, time and number of people involved in the assessment of an image, making the expansion and improvement of the Program of Monthly Monitoring of Quality in Mammography for State and Private Facilities in the State of Minas Gerais is possible.

5. ACKNOWLEDGES

The authors are thankful to the Coordination of Improvement of Higher Education Personnel, CAPES, for the incentive with a doctorate’s fellowship. This work was supported by FAPEMIG and the Ministry of Science and Technology - MCT/Brazil, through the Brazilian Institute of Science and Technology (INCT) for Radiation Metrology in Medicine.

6. REFERENCES

1. Brazil. Brazilian National Cancer Institute (INCA). Mammography: From practice to control. Rio de Janeiro. 2007. [in Portuguese]
2. Tang J, Rangayyan RM, Xu J, El Naqa I, Yang Y. Computer-aided detection and diagnosis of breast cancer with mammography: recent advances. *IEEE Trans. Inf. Technol. Biomed.* 2009 ;13(2): 236-251, doi: 10.1109/TITB.2008.2009441
3. American Cancer Society. Breast Cancer: Early Detection. 2013. Accessed at <http://www.cancer.org/acs/groups/cid/documents/webcontent/003165-pdf.pdf> on October 22, 2013.
4. Brazil. Radiological Protection Directives for Medical and Dental Radiodiagnostic. 1998. [in Portuguese]
5. Brooks KW, Trueblood JH, Kearfott KJ, Lawton DT. Automated analysis of the American College of Radiology mammographic accreditation phantom images. *Med. Phys.* 1997; 24: 709–723, doi: 10.1118/1.597992
6. Chakraborty DP. Physical measures of image quality in mammography. *Proc. SPIE* 2708.1996. pp. 179–193. doi: 10.1117/12.237781
7. Dougherty G. Computerized evaluation of mammographic image quality using phantom images. *Comput. Med. Imaging Graphics.*1998; 22: 365 – 438, doi: 10.1016/S0895-6111(98)00043-3
8. Mayo P, Rodenas F, Campayo JM, Verdu G., 2010. Quality assurance applied to mammographic equipments using phantoms and software for its evaluation. *Nuclear Instrum. Methods.* 2010; 74: 619 – 372, doi: 10.1016/j.nima.2010.01.015
9. Chakraborty DP, Eckert MP. Quantitative versus subjective evaluation of mammography phantom images. *Med. Phys.* 1995; 22: 133–143, doi: 10.1118/1.597463
10. Lee Y, Tsai DY, Shinohara N. Computerized quantitative evaluation of mammographic accreditation phantom images. *Med. Phys.* 2010; 37: 6323 – 6354, doi: 10.1118/1.3516238
11. Oliveira M, Nogueira MS, Guedes E, Andrade MC, Peixoto JE, Joana GS, Castro JG. Average glandular dose and phantom image quality in mammography. *Nuclear Instruments and Methods in Physics Research A.* 2007; 580: 574–577.
12. Gonzalez RC, Woods RE, Eddins SL. *Digital Image Processing Using MATLAB.* Pearson Education, Inc. 2004.
13. Pragathi J, Patil H.. Segmentation Method for ROI Detection in Mammographic Images using Wiener Filter and Kittler’s Method. *IJCA Proceedings on International Conference on Recent Trends in Engineering and Technology*, 2013 May, ICRTET(4). p. 27-33. Published by Foundation of Computer Science, New York, USA.
14. Ginesu G, Massidda F, Giusto D. A multi-factors approach for image quality assessment based on a human visual system model. *Elsevier Signal Processing: Image Communication.* 2006; 21: 316–333, doi: 10.1016/j.image.2005.11.005
15. Brooks KW, Trueblood JH, Kearfott KJ. Subjective evaluations of mammographic accreditation phantom images by three observer groups. *Inv. Radiol.* 1994; 29: 42–47. doi: 10.1097/00004424-199401000-00006

16. Cohen J. A coefficient of agreement for nominal scales. *Educational and psychological measurement*. 1960; 20: 37-46.
17. Landis J, Koch G: The measurement of observer agreement for categorical data. *Biometric*. 1977; 33:159-74.
18. Graphpad Software. Quantify agreement with kappa. Available in: <<http://graphpad.com/quickcalcs/kappa1/>>. Access in: 28 Sep. 2015.