An Accurate Computer Method to Assist Phisycians for Breast Cancer Detection

Maria Rizzi, Matteo D'Aloia and Beniamino Castagnolo

Politecnico di Bari - Dipartimento di Elettrotecnica ed Elettronica, Via E. Orabona, 4,70125 Bari, Italy

Abstract: Mammography is a not invasive diagnostic technique widely used for an early detection of breast cancer. Unfortunately, mammographic image analysis is a complex and difficult task which requires the services of specialized radiologists. In fact, the intrinsic difficulty in detecting signs of cancer makes image study particularly tiring, especially in mass screening that requires radiologists to examine a high number of mammograms. Therefore, the risk that radiologists may miss some subtle abnormalities exists especially when microcalcifications are present. The difficulty in microcalcification detection is due both to their small size and to low contrast between microcalcifications and surrounding tissues especially in raw images. The high correlation between the appearance of microcalcification clusters and the presence of cancer shows that Computer Aided Detection systems of microcalcifications are extremely useful and helpful in an early detection of breast cancer. In this paper, a three stage method for microcalcification computer aided detection and localization is indicated. In the first stage, standard methods are used for background removing and for tissue enhancement, in the second stage, wavelet filters are adopted for true microcalcification recognition, while in the third section, clusters are localized for subsequent diagnosis. Tests performed on standard database have confirmed the approach effectiveness and the method high accuracy.

Key words: Microcalcifications · Computer aided detection · Mammography · Wavelet transforms

INTRODUCTION

Microcalcifications are considered as early signs of breast cancer which is a leading cause of mortality among women. They are tiny calcium deposits accumulated in breast tissue that appear in mammograms as small bright spots embedded within an inhomogeneous background [1]. The size, shape, density, distribution pattern and number of microcalcifications are analyzed in the benign and malignant classification phase [2]. Benign microcalcifications typically have a diameter of 1 - 4 mm and are coarse, round or oval and uniform in size and shape. Their distribution pattern is scattered or diffuse. Malignant microcalcifications generally have a diameter of less than 0.5mm and are fine, linear branching, stellate modeled and varying in size and shape [3]. Generally, their distribution pattern is clustered with a number of microcalcifications usually more than 3.

Computer aided analysis could be extremely useful for radiologists to improve both diagnosis sensitivity (i.e. accuracy in recognizing all malignant pathologies) and specificity (i.e. possibility of classifying benign pathologies as malignant). Therefore, the development of adequate computational tools that are able both to focus the physician attention on suspect image regions and provide quantitative image descriptions, is very important for cancer detection at an early stage.

Although Computer Aided Detection (CAD) systems for mammographic images have been studied over the last two decades, automated microcalcification detection and interpretation still remains very difficult. The fundamental problems are [4, 5].

- Small dimensions of the objects of interest which can lead to potential misidentification;
- Different sizes, various shapes and variable distributions of microcalcifications to detect;
- Low contrast between microcalcifications and surrounding tissues in raw images;
- Presence of high-dense breast tissues for fibroglandular tissue predominance such as mammograms of young women which make the distinction between normal glandular tissues and malignant disease, difficult. In fact, in a breast that is particularly dense, mammography sensitivity for early malignancy detection is reduced as a result of the effort required in locating cancer within an opaque, uniform background.

Corresponding Author: Prof. Maria Rizzi, Politecnico di Bari - Dipartimento di Elettrotecnica ed Elettronica,
Via E. Orabona, 4,70125 Bari, Italy, Fax: +39 080 5963680, E-mail: rizzi@deemail.poliba.it.

To deal with these problems, it is very important for CAD systems to suppress noise, to enhance contrast between region of interest and background, to extract and select microcalcification features and hence, to detect/classify microcalcifications more accurately. Many approaches for enhancement of microcalcification clusters have been adopted such as global and local thresholding methods, histogram equalization, mathematical morphology transformations, statistic methods, wavelet transformations, neural networks, stochastic models, fractal models, high-order statistic methods, fuzzy logic approaches, etc. [6-17].

All the methods mentioned above are characterized by various problems such as the selection of regions of interest inside the mammogram and the manual calibration of certain parameters for the procedure test which could cause an increase in false negative/positive detections.

In this paper an efficient procedure for the automatic detection/localization of microcalcification clusters in mammograms with various densities is presented. The proposed algorithm consists of three phases: preprocessing, feature extraction and classification phases. Adopting this tool, all suspicious microcalcifications are enhanced and background noise reduced by applying point processing operators according to image statistical parameters (i.e. mean gray level pixel value and standard deviation). Moreover, for microcalcification localization, the image under test is decomposed adopting wavelet filters and each decomposition level is processed using a hard threshold technique.

Wavelet Approach: From an image-processing point of view, microcalcifications are relatively high-frequency components embedded both in the background of low-frequency components and in noise signals characterized by high frequency detail. With the aim of extracting microcalcifications from background and noise, the wavelet technique can be adopted. In fact, both the property of time-frequency localization (which allows us to obtain a signal at a particular time and frequency or to extract features at various locations in space) and the multirate filtering option (which permits the differentiation of signals with different frequencies) make the wavelet transform an effective tool in image processing analysis. It decomposes the mammographic image into several components with various scales or resolutions. Therefore, it can identify useful information for microcalcification detection and discard signal bands which provide scant contribution to the study [5]. Since wavelet functions are

compact, wavelet coefficients only measure the variations around a small region of data array. This property makes wavelet analysis very useful for image or signal processing [18, 19]; the "localized" nature of wavelet transform allows to easily pick out features in analyzed data such as spikes (i.e. noise or discontinuities), discrete objects (for instance in medical images or satellite photos), edges of objects, etc. Moreover, wavelet coefficients at one location are not affected by coefficients at other locations in data under test. This makes it possible to remove "noise" of all different scales from a signal, simply by discarding the lowest wavelet coefficients.

There are essentially two types of wavelet decompositions: the redundant ones (generally continuous wavelet transform) and the nonredundant ones (orthogonal, semi-orthogonal, or biorthogonal wavelet bases) [20]. The first type is preferable for feature extraction because it provides a description that is truly shift-invariant. The latter type is better for data reduction or when orthogonality of representation is an important factor.

Implemented Detection System: The procedure aim is the development of a tool for automatic detection of microcalcification cluster in mammograms with various tissue density, reducing false positive and false negative rates.

The first algorithm step is the preprocessing phase whose purpose is the contrast enhancement of suspicious mammographic images, removing both background information and non pathological breast tissue. Therefore, the method starts by separating mammary tissue from non-mammary tissue regions inside the image under test (Figure 1). To reject areas that do not contain tissue pixels, the mammographic image is subdivided into 16 x 16 pixel windows and is segmented [21, 22]. The chosen window dimension is a compromise between the detection rate and processing time parameters. Windows containing only background pixels are excluded from further analysis. The obtained image is analyzed applying point processing operators such as hard thresholding and full scale histogram stretch (Figure 2).

In the second step, called feature extraction phase, suspicious zones are detected for true microcalcification identification. As the mammographic image is generally composed of low frequency features while microcalcifications appear as high frequency details, a wavelet approach is adopted allowing the separation of high resolution mammogram components from the low
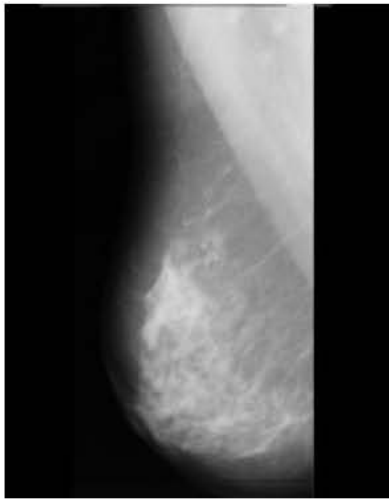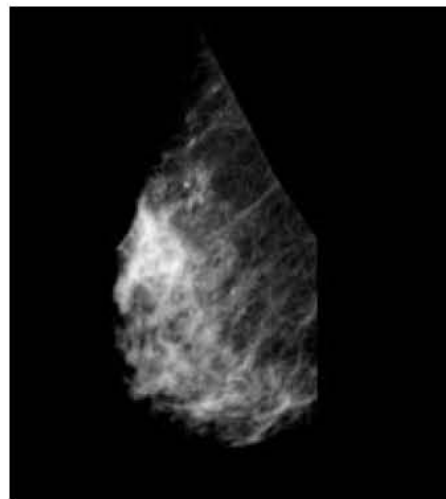
Fig. 1: Mammogram n.249 of MIAS database

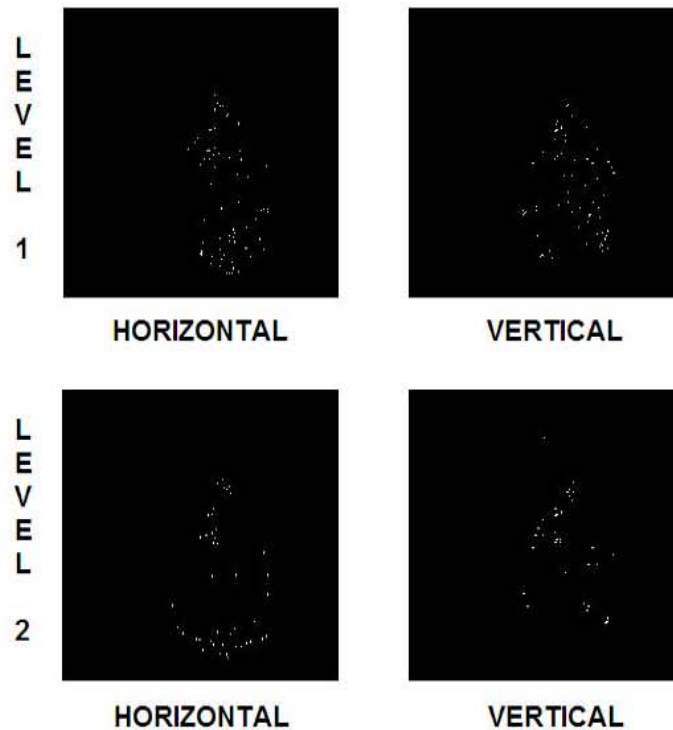Fig. 2: Mammogram after hard thresholding and FSHS

Fig. 3: Procedure result

resolution ones. Biorthogonal wavelet filter coefficients were chosen for this study as they cause low image distortions due to their symmetrical properties [1]. In particular, the Bior 2.6 mother wavelet is used and the image under study is decomposed at different levels. As the mammographic image was previously processed through linear functions so as to discard tissue regions without critical characteristics, coefficients from levels one and two are taken into account since they contain important information about microcalcifications. Therefore, the de-noising and enhancement procedures analyze only levels one and two of wavelet decomposition and adopt a Donoho thresholding technique [23]. The method performs a reconstruction which considers the coefficient values of all decomposition levels. As a biorthogonal wavelet family is adopted, a perfect reconstruction of the original image is possible.

Fig. 4: Identified Cluster

The aim of the successive step, called feature extraction, is the detection of suspected zones where microcalcifications can be localized. Since microcalcification tissue has a higher intensity than the surrounding pixels, a wavelet filter can be adopted to analyze the reconstructed image. Indeed, suspect microcalcifications can be detected by discarding the lowest frequency sub-band in the image wavelet decomposition. In the proposed method, mammograms are decomposed into two-level wavelet representations adopting the Haar mother wavelet.

To recognize true microcalcifications, after the adoption of a hard thresholding method, the implemented algorithm employs a 2D nonlinear filtering procedure for each level.

In the classification phase, the last algorithm step, the procedure analyzes the whole tested mammogram and discriminates the normal image from the abnormal one in which microcalcification clusters have been detected. A special custom filter is used which is characterized by a 6x6 fixed window for the first level wavelet decomposition and by a 3x3 fixed window for the second level wavelet decomposition. Filtered points will have maximum value if there are at least one local maximum, at least one local minimum and at least one zero-crossing point in the corresponding window; otherwise a zero value is assigned.

Moreover, a pixel quantization change is implemented which transforms the original gray level image into a black and white mammogram. Therefore, pixels with a value above zero become white, while all the others turn black (Figure 3).

In particular, microcalcifications are localized making pixel by pixel product of horizontal and vertical detail coefficients at each decomposition level. All suspicious objects localized in the two levels are considered real microcalcifications. A cluster is identified if more than 3 microcalcifications are detected in $1\,cm^2$ square area (Figure 4) [24-26].

RESULTS AND DISCUSSION

The data collection used to test the procedure has been adopted from the MiniMammographic database provided by the MIAS [27]. It consists of 322 different images belonging to three categories: normal, abnormal benign and abnormal malign. Abnormal cases are further classified in six categories, according to different pathological types: circumscribed mass, spiculate mass, microcalcification, ill-defined mass, architectural distortion and asymmetry. All mammograms are digitized at $200\,\mu m$ pixel edge, with 1024x1024 pixel resolution and 8 bit accuracy. For each image, position, size and type of pathological changes as well as breast tissue type (i.e. fatty, fatty glandular and dense) are recorded.

The adoption of a standard, public available database as procedure testing bench makes the comparison with other computer aided detection systems previously presented in literature, possible.

To evaluate the procedure performance, the following well-known coefficients have been calculated, characterizing the diagnostic system behaviour [28]:

- The Sensitivity (Se) defined as the probability of detecting a microcalcification cluster when it exists really.

$$Se = \frac{TP}{TP + FN} \qquad (1)$$

- The Accuracy (Ac) defined as the observed agreement between the procedure results and the physicians opinion about the mammogram under test reported inside the MIAS database.

$$Ac = \frac{TP + TN}{TP + TN + FP + FN} \qquad (2)$$

- The Cohen's Kappa coefficient (K) which indicates how reliably can the system be trusted if the system makes a decision like a physician.

$$K = \frac{Ac - C}{1 - C} \qquad (3)$$

with:

$$C = \frac{TP + FN}{TP + TN + FP + FN} x \frac{TP + FP}{TP + TN + FP + FN} +$$
$$\frac{TN + FN}{TP + TN + FP + FN} x \frac{TN + FP}{TP + TN + FP + FN}$$

- TP (number of true positives) defined as the number of correct identifications of microcalcification clusters inside the mammogram under test;
- FN (number of false negatives) that is the number of microcalcification clusters present in the image that algorithm is not able to detect;
- FP (number of false positives) which represents the number of microcalcification clusters detected by the algorithm but really not present in the mammogram.

High value of sensitivity is desirable for CAD systems, above all during the screening phase, when the diagnostic system should ensure that most cases are detected. In particular, the implemented tool gets a sensitivity of about 97% at a rate of 0.6 FP/image which is better than the value obtained with some other recent methods adopting the same database [5, 25, 29-33]. Moreover, the obtained accuracy value of about 98.5% at a rate of 0.6 FP/image shows the great agreement between the system results and the physician opinion, indicated in the MIAS database. In Figure 5 the shape of accuracy vs FP is indicated.

The Cohen's Kappa coefficient, evaluated for the adopted procedure considering 0.6 FP/image, is 0.83 (Figure 6) that proves the method validity and the possibility to adopt this procedure as "second opinion" during the mammographic screening phase. In fact a C value higher than 0.81 shows an almost perfect agreement between the CAD system results and the physician opinion as regards the mammogram under test [34].
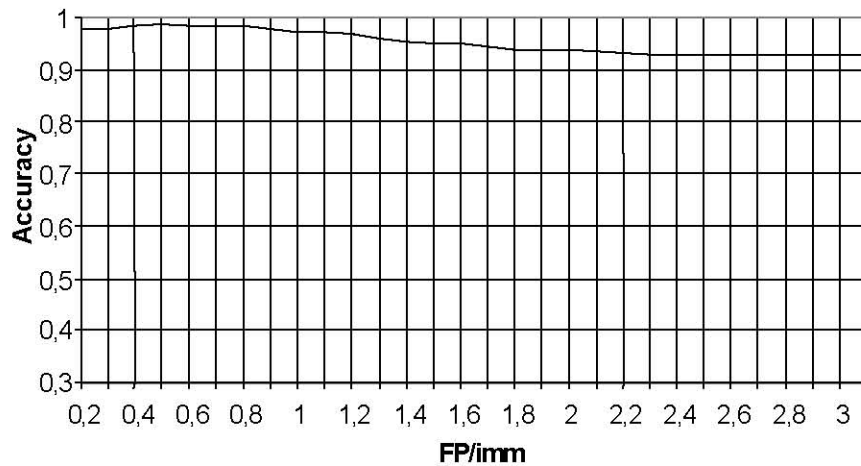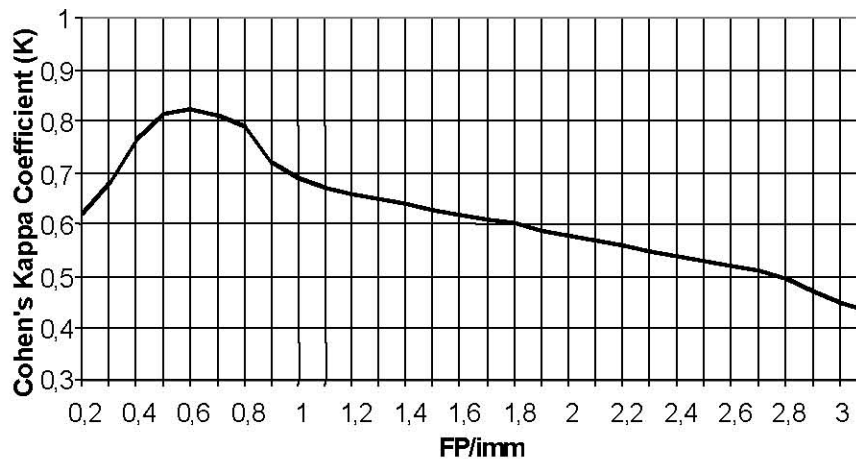

Fig. 5: Accuracy

Fig. 6: Cohen's Kappa Coefficient (K)

CONCLUSIONS

In this paper, a computer aided detection tool devised to support radiologists in digital mammography analysis has been developed.

The procedure focuses on microcalcification cluster detection and localization in mammograms with various densities. It consists of several phases which perform different operations such as image decomposition in different wavelet levels, feature extraction, peak localization, processed image assignment to the class of microcalcifications or of normal breast tissue, highlight of suspiciousareas inside the original mammogram.

The method analyzes the entire mammary tissue, so guaranteeing high accuracy in recognizing all malignant pathologies by avoiding multiple cases of false negative detection. In fact, the common starting assumption of other published procedures which select a region of interest (inside which microcalcifications may be detected) and discard all other parts in the tested image, is rejected.

The proposed procedure was tested adopting the whole MIAS database. The obtained performance show the method accuracy and its efficiency to operate as "second opinion" in detecting and localizing microcalcification clusters. In fact accuracy and k coefficients of 98.5% and 0.83 at a rate of 0.6FP/image are obtained, respectively. Moreover, the algorithm sensitivity of 97% with only 0.6 FP/image ensures the capacity of the procedure to detect most cases and therefore it is very useful during the screening phase.

REFERENCES

1. Boccignone, G., A. Chainese and A. Picariello, 2000. Computer aided detection of microcalcifications in digital mammograms, Computers in Biology and Medicine, 30(5): 267-286.

2. Mousa, R., Q. Munib and A. Moussa, 2005. Breast cancer diagnosis system based on wavelet analysis and fuzzy-neural, Expert systems with Applications, 28(4): 713-723.

3. Sickles, E.A., 1986. Breast calcifications: Mammographic evaluation, Radiology, 160: 289-293.

4. Numes, F.L.S., H. Schiabel and C.E. Goes, 2007. Contrast enhancement in dense breast images to aid clustered microcalcifications detections, Journal of Digital Imaging, 20(1): 53-66.

5. Yu, S.N., K.Y. Li, and Y.K. Huang, 2006. Detection of microcalcifications in digital mammograms using wavelet filter and Markov random field model, Computerized Medical Imaging and Graphics, 30(3): 163-173.

6. Chang, C.M. and A. Laine, 1999. Coherence of multiscale features for enhancement of digital mammograms, IEEE Trans. Inform. Technol. Bio. Med., 3(1): 32-46.

7. Gonzales, R.C., and R.E. Woods, 2002. Digital Image processing, 2nd (ed.) Englewood Cliffs, NJ: Prentice Hall.

8. Laine, A., J. Fan and W. Yang, 1995. Wavelets for contrast enhancement of digital mammography, IEEE Eng. Med. Biol., 14(5): 536-550.

9. Gurcan, M.N., Y. Yardimci, A.E. Cetin and R. Ansari, 1997. Detection of microcalcifications in mammograms using higher order statistics, IEEE Signal Process. Lett, 4(8): 213-216.

10. Cheng, H.D., Y.M Lui. and R.I. Freimainis, 1998. A novel approach to microcalcification detection using fuzzy logic technique, IEEE Trans. Med. Imag, 17: 442-450.

11. Li, L., W. Qian and L.P. Clarke, 1997. Digital mammography: directional wavelet analysis for feature extraction and mass detection, Acad. Radiol. 4(11): 724-731.

12. Nam, S.H. and J.Y. Choi, 1998. A method of image enhancement and fractal dimension for detection of microcalcifications in mammogram, Proc. of the 20th Annual Int. Conf. of IEEE Engineering in Medicine and Biology Society, 2, Hong Kong (China), 29 oct-1 Nov. 1998, 1009-1012.

13. Pfistere, R. and F. Aghdasi, 2001. Tumour detection in digitalized mammograms by image texture analysis, Opt. Eng., 40(2): 209-216.

14. Kim, J.K. and H.W. Park, 1999. Statistical texture features for detection of microcalcifications in digitized mammograms, IEEE Trans. Med. Imag., 18(3): 231-238.

15. Zhang, W., K. Doi, M.L. Giger, R.M. Nishikawa and R.A. Schmidt, 1996. An improved shift invariant artificial neural network for computerized detection of clustered microcalcifications in digital mammograms, Med. Phys., 23: 595-601.

16. Zheng, B., Y.H. Chang, M. Staiger, W. Good and D. Gur, 1995. Computer aided detection on clustered microcalcifications in digitalized mammograms, Acad. Radiol., 2(8): 655-662.

17. Zheng, B., W. Quian and L.P. Clarke, 1996. Digital mammography: mixed feature neural network with spectral entropy decision for detection of microcalcifications, IEEE Trans. Med. Imag., 15: 589-597.

18. Rizzi, M., M. D'Aloia and B. Castagnolo, 2008. Fast parallelized algorithm for ECG analysis, WSEAS Trans. on Biology and Biomedicine 5: 210-219.

19. Rizzi, M., M. D'Aloia and B. Castagnolo, 2009. High sensitivity and noise immune method to detect impedance cardiography characteristic points using wavelet transform, J. Appl. Sci., 9: 1412-1421.

20. Unser, M. and A. Aldroubi, 1996. A review of Wavelets, Biomedical Application in Proceedings of the IEEE, 84(4): 626-638.

21. Sentelle, S., C. Sentelle and M.A. Sutton, 2002. Multiresolution-based segmentation of calcifications for early detection of breast cancer, Real-Time Imag., 8(3): 237-252.

22. Jiang, J., B. Yaoa and A.M. Wason, 2007. A genetic algorithm design for microcalcification detection and classification in digital mammograms, Computerized Medical Imag. and Graphics, 31(1): 49-61.

23. Donoho, L., 1995. De-noising by soft-thresholding, IEEE Trans on Information Theory, 41: 612-627.

24. Cheng, H.D., X. Cai., X. Chen, L. Hu and X. Lou, 2003. Computer-aided detection and classification of microcalcifications in mammograms: a survey, Pattern Recognition, 36(12): 2967-2991.

25. Regentova, E., L. Zhang, J. Zheng and G. Veni, 2006. Detecting Microcalcifications in Digital Mammograms using Wavelet Domain Hidden Markov Tree Model, 28th IEEE Engineering in Medicine and Biology Society (EMBS) Annual International Conference, New York City, USA, Aug.30-Sept.3 2006, 1972-1975.

26. Salvado, J. and B. Roque, 2005. Detection of calcifications in digital mammograms using wavelet analysis and contrast enhancement, IEEE Int. Workshop on Intelligent Signal Processing, Faro (Portugal), 1-3 sept. 2005, 200-205.

27. MIAS Database http://peipa.essex.ac.uk/ipa/pix/mias

28. Miaou, S.G., F.L. Chang, I.K. Timotious, H.C. Huang, J.L. Su, R.S. Liao and T.Y. Lin, 2009. A multi-stage recognition system to detect different types of abnormality in capsule endoscope images, J. Medical and Biological Eng., 29(3): 114-121.

29. Retico, A., P. Delogu, M.E. Fantacci, A. Preite Martinez, A. Stefanini and A. Tata, 2006. A scalable computer aided detection system for microcalcification cluster identification in a pan-European distributed database of mammograms, Nuclear Instruments and Methods in Physics Res. A, 569: 601-605.

30. Rezai-rad, G. and S. Jamaran, 2005. Detecting Microcalcification Clusters in Digital Mammograms Using Combination of Wavelet and Neural Network, Proceedings of the Computer Graphics, Imaging and Vision: New Trends (CGIV'05) 26-29 July, 197-201.

31. Song, L., Q. Wang and J. Gao, 2006. Microcalcification detection using combination of wavelet transform and morphology, Proc. 8th Int. Conf. on Signal Processing (ICSP2006), 4: 16-20.

32. Jin Y., Z. Sheng and Y. Changcai, 2008. SVM-based microcalcification detection in digital mammograms, Int. Conf. on Computer Science and Software Engineering, 6: 12-14 Dec., 89-92.

33. Regentova, E., L. Zhang, J. Zheng and G. Veni, 2007. Microcalcification detection based on wavelet domain hidden Markov tree model: study for inclusion to computer aided diagnostic prompting system, Medical Physics 34: 2206-2219.

34. Landis, J.R. and G.G. Koch, 1977. The measurement of observer agreement for categorical data, Biometrics, 33: 159-174.