IEEE International Workshop on Imaging
Systems and Techniques – IST 2008
Chania, Greece, September 10–12, 2008

Computer Aided Insights on Obscure Cases of Breast Cancer Diagnosis

Andreadis I., Nikita K.
Department of Electrical and Computer Engineering
National Technical University of Athens
Athens, Greece

Giannakopoulou G., Koulocheri D., Zografos G.
Breast Unit
Hippocration Hospital
Athens, Greece

Antaraki A., Ligomenides P., Spyrou G.
Informatics Laboratory-Biomedical Research Foundation
Academy of Athens
Athens, Greece

Abstract— **Breast cancer is a leading cause of deaths in women. Mammography is considered as the most effective technology presently available for breast cancer screening, being very effective in the detection of clustered microcalcifications which are considered as one of the most important findings associated to the existence of breast cancer. A computer aided diagnosis (CAD) system named "Hippocrates-mst" has been already developed in the lab based on detailed analysis and evaluation of related features of microcalcifications (individually and in clusters). Preliminary evaluation results have shown that the system achieves high levels of sensitivity, while suffering from low specificity. For this reason, our current studies aim to a methodology refinement which will lead to optimized classification results. In this paper, we focus on obscure diagnostic cases classified by the radiologists as BI-RADS 3. In such cases, although short-term re-examination is normally advised, radiologists and physicians usually have strong doubts about their recommendations. We tested the performance of two classifiers embedded in the proposed CAD system using a dataset of 63 (57 benign and 6 malignant) mammograms, all classified as BI-RADS 3 and biopsy proven. The sensitivity achieved by the first one (the default Hippocrates-mst classifier) is as high as 100%, classifying correctly all the malignant cases. As far as the benign cases are concerned, system's specificity is 35.09%. Using the second classifier (a rule based and SVM hybrid classifier) the specificity increases to 63.16% with a cost of sensitivity decrease to 66.67%.**

Keywords-breast cancer; microcalcifications; CAD; BI-RADS

I. INTRODUCTION

Breast cancer is the second leading cause of cancer deaths in women today, occurring in nearly one out of eight women [1]. According to the World Health Organization, more than 1.2 million people will be diagnosed with breast cancer every year worldwide [2]. Several studies have proved that the early detection of breast cancer can reduce mortality rates and improve the chances that the patient is treated successfully.

Many imaging techniques for the detection of breast cancer are currently being studied, such as magnetic resonance imaging (MRI), ultrasound, electrical impedance scanning and breast-specific positron emission tomography (PET). However, mammography is the most effective technology presently available for breast cancer screening, despite the fact that there are still some inherent limitations of the imaging technique, such as insufficient resolution, low local contrast and non-negligible noise combined with the subtle nature of the usual radiographic findings [3], [4].

One of the most important radiographic findings associated to the existence of breast cancer is clustered microcalcifications [5]-[11]. Especially, it has been shown that some characteristics concerning the clustering parameters of microcalcifications are of great diagnostic value. However, the aforementioned limitations of mammography make the detection and interpretation of microcalcifications a really complicated task. Therefore, in order to increase the efficiency and effectiveness of mammographic screening process, many scientific research groups have attempted to develop systems for computer-aided diagnosis (CAD) using various methods [12]-[19]. Such systems would be of great value if able to provide a reliable second opinion to the radiologist.

A CAD system, called "Hippocrates-mst", has been developed and is based on detailed analysis and evaluation of related features of individual microcalcifications and of formed clusters [20]-[23]. Up to now, the output of the system is considered to be very useful to the radiologists, giving them extra input before reviewing each case. Continuous research is being carried out, leading to new classification schemes which aim to the improvement of the diagnostic accuracy of the system [24].

In this paper, we describe the investigation of the potentiality of "Hippocrates-mst" system and other developed classification schemes to assist the radiologist's diagnosis in cases that the mammograms are classified as BI-RADS 3. When a mammogram is classified by the radiologist as BI-

978-1-4244-2497-9/08/$20.00 ©2008 IEEE

RADS 3, a follow-up in a short period of time is advised [25]. In many cases, the patient is also referred for a biopsy, as the physician has doubts whether there is the need or not for immediate actions.

We try to examine the behavior of the CAD system and the other classification schemes in such obscure diagnostic cases and their potentiality to provide a reliable second opinion to the doctor, evaluating a set of 63 mammograms, classified as BI-RADS 3 by the radiologists. The results from the evaluation phase indicated the potentiality of "Hippocrates-mst" system to correctly classify (i) with its default rule-based classifier all the malignant cases, achieving at the same time a reduction of 35.09% on the unnecessary biopsies, (ii) with an alternative hybrid classifier 66.67% of the malignant and 63.16% of the benign cases.

II. METHODS AND MATERIALS

A. "Hippocrates-mst" CAD system

The computer aided diagnosis system we have already developed and tested, named "Hippocrates-mst", is based on detailed analysis and evaluation of related features of individual microcalcifications and of formed clusters.

The user has the ability to examine digitised mammograms. Every selected image may be processed with the help of a digital lens, which leads to the effective investigation of small structures and suspicious regions on the mammogram. The user is also able to apply several techniques such as histogram equalization, zoom, edge detection, differentiation of contrast and brightness. Some of them are very effective visualization tools, especially in breast periphery where the tissue is overexposed and very dark. At the next step, the user may use the detection algorithm and reveal the microcalcifications existing in the selected region.

After the detection of the microcalcifications, image analysis is needed for the microcalcification feature extraction and quantification. The considered microcalcification features are: (i) size, (ii) circularity, (iii) existence of dark center, (iv) level of brightness, (v) irregularity regarding the shape, (vi) level of branching and (vii) circumvolution.

The above features are calculated for every microcalcification inside a sub region of the initial ROI selected by the user, comprising a cluster. Then, the system uses a two stage procedure for the risk assessment of the selected ROI. In the first stage, an empirical model, compromising of rules from the related literature and simulating the medical experts' way of classifying a microcalcification, is used in order to classify each microcalcification and categorize it according to its risk using the seven characteristics feature space. In the second stage, the final risk assessment for the selected cluster is obtained from a three characteristics feature space including the following parameters: (i) the risk distribution of microcalcifications, (ii) the number of microcalcifications at high risk, (iii) the cluster polymorphism.

The procedure just described estimates the region risk index from the classification of the selected cluster, exploiting morphological features of the detected microcalcifications. However, there are two possible shifts of the estimated risk which resulted from the location of the cluster or the patient's file. Specifically, the system invites the radiologist to give an extra risk percentage due to the relative position of the cluster inside breast, according to his expertise. Moreover, another risk percentage based on the data from the patient's file (age, history, family, other clinical examinations, etc) may be estimated in order to offer a more complete diagnosis. In Fig.1, the form of Hippocrates-mst for the final risk estimation is shown.

Finally, the risk estimation can be classified in the following four virtual zones of risk:

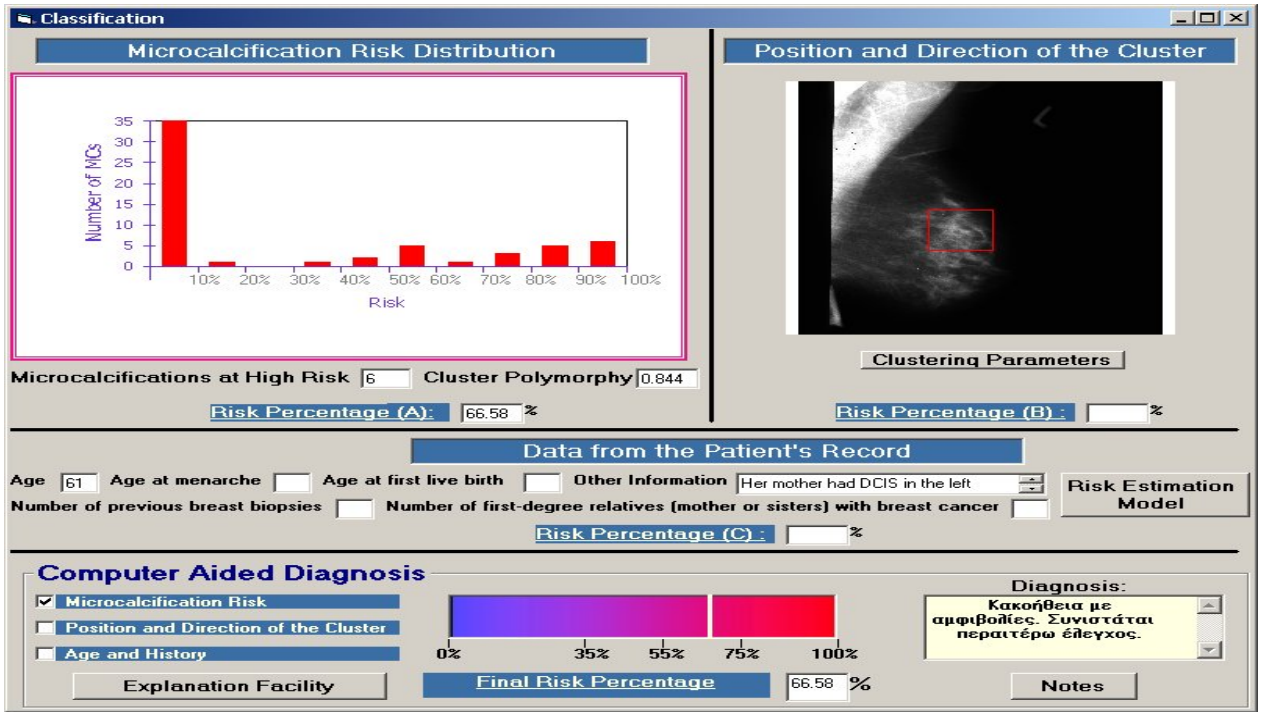

Figure 1. Form of "Hippocrates-mst" for the risk assesment of breast cancer

(i) Zone 1 with risk between 0 % and 35 % (definitely benign),

(ii) Zone 2 with risk between 35 % and 55 % (benign with doubts),

(iii) Zone 3 with risk between 55 % and 75 % (malignant with doubts),

(iv) Zone 4 with risk between 75 % and 100 % (definitely malignant).

When the estimated risk lies in the first zone, biopsy is discouraged and the examined region is considered definitely benign. In all other cases, there is possible evidence for malignancy and the CAD system refers the patient for further examinations or for surgical biopsy.

Preliminary laboratory tests have already been performed in order to evaluate the diagnostic accuracy of "Hippocrates-mst" system. Up to now the results from the current evaluation procedures indicate that the system estimates the risk of breast cancer towards the right direction, achieving high levels of sensitivity. However, although it presents high sensitivity, it suffers from relatively low specificity. This indicates that there is an overestimation of risk driven from the attempt to minimize the false negative results, which was our primary goal during the development phase of the system.

During the last two years, the system is undergoing an evaluation procedure in Hippocration Hospital of Athens in Greece. The initial examination by the radiologist is followed by an extra interpretation of the mammogram through the use of the "Hippocrates-mst" system. The radiologists evaluate the CAD system's contribution in the daily clinical practice. In order to have a measure for the classification performance of the system, the system's results must be compared versus the corresponding biopsy results.

B. Binary Logical Classifier

As we have already mentioned, preliminary evaluation results have indicated that the "Hippocrates-mst" CAD system achieves high level of sensitivity, while mainly suffering from low specificity. As a result, a methodology refinement is required in order to improve the system's diagnostic accuracy. Our current studies are focused on developing new classification schemes able to provide better classification results, optimizing that way the final diagnosis phase of the "Hippocrates-mst" CAD system.

A binary logical classifier (BLC) has already been developed for the classification and characterization of clustered microcalcifications in digitized mammograms, using the risk assessment of the CAD system, as described in the previous section. This classifier combines the rule-based classifier, used in the CAD system, with various classification schemes based on the support vector machines (SVM) binary methodology [26], [27]. We put an effort on keeping the prediction of the rule-based decision tree as the mainstream diagnosis mode, while giving to the aforementioned SVM classifiers the potentiality to act as a reliable second opinion to the "Hippocrates-mst" classifier, which may change the initial diagnosis only under special circumstances where there is strong evidence for benignity. All the classifiers are combined using simple logical gates. Details on the development of the SVM classifiers and the topology of the BLC may be found in [24]. Due to the use of logical gates and the binary nature of the SVM methodology, the final diagnosis of the BLC is a logical index which indicates whether the considered cluster of microcalcifications is associated or not with the existence of breast cancer.

The first evaluation procedure for the BLC system has indicated that the proposed classification scheme can be beneficial for the CAD system, by reducing the number of false positive diagnoses, achieving greater levels of specificity, and maintaining at the same time the specificity at high levels [24].

C. Data Collection

In this paper, we concentrated on obscure diagnostic cases classified as BI-RADS 3 by the radiologists. We aimed to investigate the effectiveness of the proposed CAD system to the classification of these cases and its potentiality to assist radiologists' diagnosis.

BI-RADS is a system developed by the American College of Radiology (ACR) for reporting mammogram results using a common language. After creating the report of a mammogram, one single digit BI-RADS score ranging from 0 to 5 is assigned. The score BI-RADS 3 means that the mammogram is probably benign but a short-term follow-up should be considered (after 6 months) until longer term stability is demonstrated (for almost 2 years). The risk of malignancy is less than 2% for this category and it is not expected to change over the follow-up intervals [25].

However, in many cases, the procedure followed in such cases is not like the one described above. This may be due to various reasons such as doctor's uncertainty, controversial diagnoses between radiologists, patient's anxiety etc. As a result, in many cases, despite classified as BI-RADS 3, the patient may be referred for a biopsy, as there are doubts whether there is the need or not for immediate actions.

We collected a set of mammograms classified as BI-RADS 3 by the radiologists of Hippocration Hospital of Athens for a period of eight months from November 2007 to June 2008. All cases were referred for a Stereotactic Vacuum-Assisted Breast Biopsy (SVABB), thus there is histological verification for their status. The images set contained 63 mammograms, of both craniocaudal and mediolateral views, with annotated regions of microcalcifications where the biopsy has been done. Each case was also accompanied with the patient's demographic data and medical history as well.

The mammoraphic images obtained through digitization using Microtek Scanmaker 9800XL scanner, with a resolution of 300 dots per inch (DPI) 16 bit greyscale colour depth and .tiff unzipped as the image format. The image size for each image was 2116X2791 while the size in megabytes was approximately 11 MB for each image.

III. RESULTS

Among the 63 cases collected, there were only six malignant cases and the rest were benign, according to the

biopsy. We used both "Hippocrates-mst" CAD system and the BLC classifier to evaluate the suspicious regions in the mammograms. BLC system directly assigns to each mammogram a binary value indicating whether it refers or not the case for further examinations. As long as "Hippocrates-mst" is concerned, a risk percentage related to the existence of breast cancer was estimated for each case. We have described earlier the methods used for the risk estimation of breast cancer. We should emphasize the fact that the risk estimated was based only on the evaluation of findings in the suspicious regions. In other words, the two possible shifts due to the location of the cluster or the medical data of the patient were not at all co-estimated and used. When the estimated risk was greater than 35% the system referred the patient for further examinations. Otherwise, the mammogram was considered by the system as benign. The systems' indications for all the 63 cases of the dataset are presented in table 1. To estimate performance, three objective indexes have been used: sensitivity, specificity and accuracy. In table 2 the measures for the classification performance of the systems are shown. These measures may be used for the evaluation and comparison of the performances of the two different approaches.

TABLE I. SYSTEMS' INDICATIONS

	True Negative (TN)	***False Positive (FP)***	***True Positive (TP)***	***False Negative (FN)***
"Hippocrates-mst"	20	37	6	0
BLC	36	21	4	2

TABLE II. SYSTEMS' PERFORMANCE MEASURES

	Performance Measures		
	Sensitvity	***Specificity***	***Accuracy***
"Hippocrates-mst"	1.00	0.3509	0.4127
BLC	0.6667	0.6316	0.6349

Concerning the six malignant cases, we observe that "Hippocrates-mst" assigns to all of them risk greater than 35%, encouraging biopsy to all of them. As a result, it achieves the maximum sensitivity (100%), as there are no false negative indications (missed cancers). The binary classifier BLC classifies correctly only 4 of the 6 malignant cases, achieving sensitivity as high as 66.67%. In other words, the BLC system fails to maintain the high levels of sensitivity provided by "Hippocrates-mst", as two more malignant cases are misclassified as benign, leading to a reduction of the sensitivity's value.

On the other hand, we observe that the BLC outperforms "Hippocrates-mst" system in terms of specificity. The binary classifier would discourage biopsy to 36 out of 57 cases, achieving that way a reduction of 63.16% on the unnecessary biopsies, while the corresponding percentage of "Hippocrates-mst" current classification scheme would be only 35.09%.

Both approaches present their own advantages. BLC classifier was developed aiming to improve the classification results of the CAD system. The current evaluation procedure using obscure cases classified as BI-RADS 3 by the radiologists, shows that the BLC is not able to improve the diagnostic accuracy, as it fails to maintain the high levels of sensitivity. Of course, the number of malignant cases is quite small and thus safe conclusions cannot be exported.

On the other hand, "Hippocrates-mst" mainly lacks in specificity (35.09%), as there are quite enough false positive indications. However, in 20 cases the biopsy is discouraged. That means, using the Hippocrates-mst system a reduction of 35.09% on the unnecessary biopsies would be achieved, as the 20 benign cases are correctly classified. At the same time, the maximum sensitivity (100%) assures us that there are no missed cancers, satisfying that way our primary goal. As a result, we can conclude that the overall performance of "Hippocrates-mst" system is better than the performance of binary classifier.

We should remind here that all the cases used in this evaluation process, were referred for biopsy by the radiologists. At least our CAD system had sent 37 women for biopsy although they were benign, while doctors referred 57 benign cases for biopsy.

In figure 2, we present the distribution of the estimated risk for all cases of the dataset using the "Hippocrates-mst" CAD system.

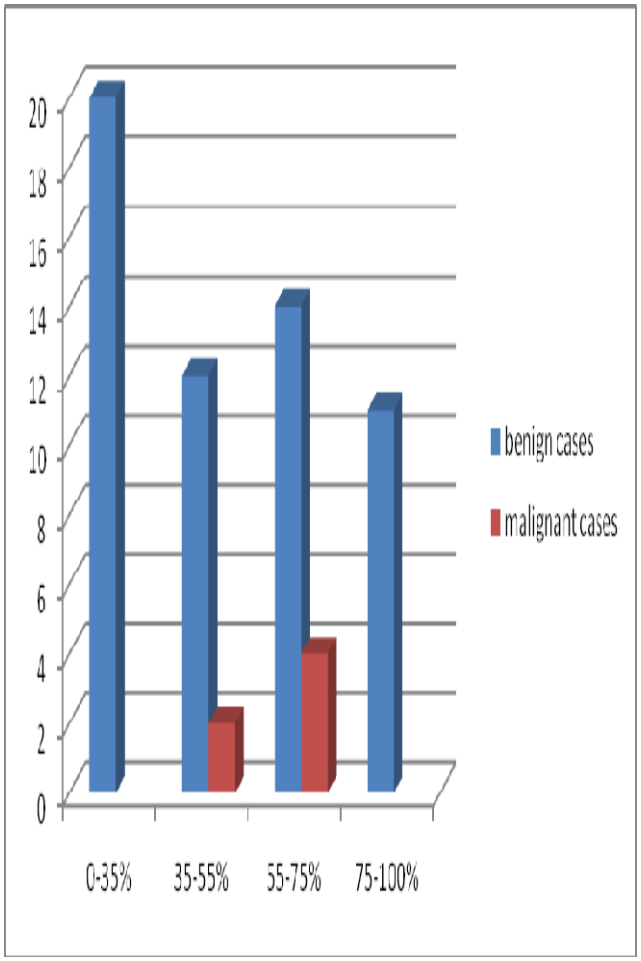

Figure 2.Classification in four virtual zones using "Hippocrates-mst"

IV. CONCLUSIONS

The aim of this paper was to investigate the potentiality of two classifiers embedded in an already developed computer aided diagnosis system named "Hippocrates-mst" to assist radiologists' diagnosis in BI-RADS 3 cases. For such cases, there are usually many doubts by the physicians whether there is need of immediate actions or short-term follow-up examinations should be followed.

A dataset of 63 cases, all classified as BI-RADS 3 by the radiologists of the Hippocration Hospital of Athens, has been used for the evaluation of the system and the comparison with the doctors' performance. Among the 63 cases, there were 6 malignant cases and 57 benign cases. Data analysis shows that the default classification scheme (rule-based) of the "Hippocrates-mst" system has enough satisfying performance, classifying correctly all the malignant cases, achieving that way a sensitivity as high as 100%. As long as the benign cases are concerned, the system classified correctly 20 out of 57 cases. It must be stated that all these benign cases have already gone either to surgery or biopsy so every successful indication of a benign case from the CAD system would save this woman from an unnecessary biopsy. As a result, the system would achieve a reduction of 35.09% on unnecessary biopsies. On the other hand, the second classifier (BLC) achieves a much higher specificity (63.16%) with a (not proportional) decrease of sensitivity at 66.67%.

Concluding, the use of the proposed CAD system in the evaluation of obscure mammographic cases seems to be beneficial in daily clinical practice, as it is able to act as a reliable second opinion in cases where the doctor has many doubts about his future recommendations, leading that way to an important reduction on the unnecessary biopsies. This first laboratory evaluation has shown encouraging performance. This is a work in progress, which will include in the near future evaluation with larger number of obscure cases of diagnosis and investigation of new classification schemes to refine the existing one.

REFERENCES

[1] American Cancer Society, Cancer Facts and Figures 2004.

[2] World Health Organization, WHO Statistical Information System.

[3] D.J. Vyborny, "Can computers help radiologists read mammograms", Radiology, vol. 191, p.p 315-317, 1994.

[4] M.L. Geiger, "Computer-aided diagnosis, AAPM/RSNA Categorical Course in Diagnostic Radiology Physics: Physical Aspects of Breast Imaging – Current and Future Considerations", Haus A.and Yaffe M.,eds., pp 249-272, 1999.

[5] Bassett, L.W., "Mammographic analysis of calcifications", Radiol. Clin. North Am. 30 1, pp. 93-105, 1992.

[6] Buchbinder, SS., Leichter, IS., Lederman, RB., Novak, B., Bamberger, PN., Coopersmith, H., Fields, SI., "Can the size of microcalcifications predict malignancy of clusters at mammography?" Acad Radiol.9(1):18-25, 2002.

[7] Laine, A.F., Schuler, S., Fan, J. and Huda, W., "Mammographic feature enhancement by multiscale analysis". IEEE Trans. Med. Imag. 13, pp. 725-740, 1994.

[8] Lanyi, M., "Differential diagnosis of microcalcifications, X-ray film analysis of 60 intraductal carcinoma, the triangle principle", Radiologe, 17(5):213-6, 1977.

[9] Lanyi, M., "Microcalcifications in the breast--a blessing or a curse? A critical review", Diagn. Imaging Clin Med. 54(3-4):126-45, 1985.

[10] Le Gal, M., Durand, JC., Laurent, M., Pellier, D., "Management following mammography revealing grouped microcalcifications without palpable tumor", Nouv Presse Med.5(26):1623-7, 1976.

[11] Le Gal, M., Chavanne, G., Pellier, D., "Diagnostic value of clustered microcalcifications discovered by mammography (apropos of 227 cases with histological verification and without a palpable breast tumor)", Bull Cancer;71(1):57-64, 1984.

[12] Fogel DB, Wasson EC 3rd, Boughton EM, Porto VW, "Evolving artificial neural networks for screening features from mammograms", Artif Intell Med, 14(3):317-26, 1998.

[13] Chan H P, Sahiner B, Petrick N, Helvie M, Lam K L, Adler D and Goodsitt M, "Computerized classification of malignant and benign microcalcifications on mammograms: texture analysis using an artificial neural network", Phys.Med.Biol. 42, 549-67, 1997.

[14] Jiang Y, Nishikawa RM, Wolverton DE, Metz CE, Giger ML, Schmidt RA, Vyborny CJ, Doi K, "Malignant and benign clustered microcalcifications: automated feature analysis and classification", Radiology,198(3):671-8, 1996.

[15] Wu YC, Freedman MT, Hasegawa A, Zuurbier RA, Lo SC, Mun SK, "Classification of microcalcifications in radiographs of pathologic specimens for the diagnosis of breast cancer", Acad Radiol 2(3):199-204, 1995.

[16] Nishikawa R M, Giger M L, Doi K, Vyborny C J, Schmidt R A, "Computer-aided detection of clustered microcalcification on digital mammograms", Med. & Biol. Eng. & Comput., 33,174-78, 1995.

[17] Karssemeijer N, "Adaptive noise equalization and recognition of microcalcification clusters in mammograms", Int.J.Patt.Rec.&Im.Analysis 7, 1993.

[18] Zang W, Doi K, Giger M L and Nishikawa R M, "Computerized detection of clustered microcalcifications in digital mammograms: applications of artificial neural networks", Med.Phys., 19, 555-60, 1992.

[19] Wu Y, Doi K, Giger ML, Nishikawa RM, "Computerized detection of clustered microcalcifications in digital mammograms: applications of artificial neural networks", Med. Phys., 19(3):555-60, 192.

[20] G. Spyrou, M. Nikolaou, M. Koussaris, A. Tsibanis, S. Vassilaros, and P. Ligomenides, "A system for Computer Aided Early Diagnosis of breast cancer based on microcalcifications analysis", Res-Systemica, vol. N°2, Special Issue, 2002.

[21] G. Spyrou, K. Koufopoulos, S. Vassilaros and P. Ligomenides, "Computer aided image analysis and classification schemes for the early diagnosis of breast cancer," Hermis International Journal of Computer Mathematics and its Applications, vol. 4, pp.175-181, 2003.

[22] G. Spyrou, S. Kapsimalakou A. Frigas K. Koufopoulos S. Vassilaros P. Ligomenides ""Hippocrates-mst": A prototype for Computer-Aided microcalcification analysis and risk assessment for breast cancer", Medical & Biological Engineering & Computing, submitted for publication, vol. 4, pp.1007-1015, 2006.

[23] A. Frigas, S. Kapsimalakou, G. Spyrou, K. Koufopoulos, S. Vassilaros, A. Chatzimichael, J. Mantas, P. Ligomenides, "Evaluation of a Breast Cancer Computer Aided Diagnosis System", Stud Health Technol Inform, vol. 124, pp.631-6,2006.

[24] I.Andreadis, G.Spyrou, A.Antaraki, G.Zografos, D. Koulocheri, G. Giannakopoulou, K.Nikita, P. Ligomenides, "Combining SVM nad Rule-Based classifiers for optimal classification in breast cancer diagnosis", HERCMA 2007.

[25] American College of Radiology (ACR), Illustrated breast imaging reporting and data system (BI-RADS), 4th edn. American College of Radiology, Reston, VA, 2003.

[26] C. J. C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition," Knowledge Discovery Data Mining, 2:1-43, 1998.

[27] N. Cristianini, J. Shawa-Taylor, "An introduction to Support Vector Machines and other kernel based learning methods," Cambridge: UK: Cambridge University Press, 2000.