



ELSEVIER

Computerized Medical Imaging and Graphics 28 (2004) 247–255

**Computerized
Medical Imaging
and Graphics**

www.elsevier.com/locate/compmimag

Development of the cubic least squares mapping linear-kernel support vector machine classifier for improving the characterization of breast lesions on ultrasound

N. Piliouras^a, I. Kalatzis^a, N. Dimitropoulos^b, D. Cavouras^{a,*}

^aDepartment of Medical Instrumentation Technology, Technological Educational Institution of Athens, Ag. Spyridonos Street, Egaleo GR-122 10 Athens, Greece

^bMedical Imaging Department, EUROMEDICA Medical Center, 2 Mesogeion Avenue, Athens, Greece

Received 29 July 2003; revised 12 April 2004; accepted 12 April 2004

Abstract

An efficient classification algorithm is proposed for characterizing breast lesions. The algorithm is based on the cubic least squares mapping and the linear-kernel support vector machine (SVM^{L_{SM}}) classifier. Ultrasound images of 154 confirmed lesions (59 benign and 52 malignant solid masses, 7 simple cysts, and 32 complicated cysts) were manually segmented by a physician using a custom developed software. Texture and outline features and the SVM^{L_{SM}} algorithm were used to design a hierarchical tree classification system. Classification accuracy was 98.7%, misdiagnosing 1 malignant and 1 benign solid lesions only. This system may be used as a second opinion tool to the radiologists.

© 2004 Elsevier Ltd. All rights reserved.

Keywords: Breast ultrasound; Breast lesion discrimination; Support vector machine; Classification

1. Introduction

Breast cancer is the second most common cancer type in women and early diagnosis helps cancer to be treated more effectively and eliminates patients' discomfort. Screening mammography, for detecting early breast cancers in asymptomatic women, increases the likelihood for cure and long-term survival. However, in cases of indeterminate mammographic findings, breast biopsy may be required. Avoiding unnecessary biopsies is important due to the discomfort, cost, and probable breast scars inflicted upon the patients, which may complicate future radiographic examinations. Ultrasonography (US) has been suggested as an adjunct method to X-ray mammography for the diagnosis of breast cancer [1,2], because of its sensitivity in discriminating between breast cysts and solid masses. However, the effectiveness of US with respect to the diagnosis of breast cancer is still questionable [3] and the role of US controversial. One of the most important reasons

of that controversy is the considerable overlap of benign and malignant findings in US images, the interpretation of which is subjective and depends on the operator's experience. Therefore, it is evident that a quantitative assessment of the breast US image may be of value in order to avoid unnecessary invasive operations. Previous work on computer-aided characterization of breast lesions from US images [4–11] involves estimation of parameters (or features), related to the breast-lesion's texture and/or contour-shape, and application of texture analysis methods. The latter include either the employment of classical statistical analysis techniques [4,6,7,9,11], the use of statistical classifiers [9], or the training of high precision neural networks classifiers [5,10]. Although US has been shown to reliably (96–100%) characterize benign breast cysts [12], the precision of benign–malignant breast-lesion discrimination varies amongst researchers between 85 and 95%, depending on the database, features, and classification algorithms employed. It is obvious that research has to be continued by testing the proposed methods on different US-image databases and at the same time by developing new features and more powerful and robust classifiers.

* Corresponding author. Tel.: +30-210-5385-375; fax: +30-210-5910-975.

E-mail address: cavouras@teih.gr (D. Cavouras).

In the present study we have developed a high performance and accuracy classification technique, based on the linear-kernel support vector machine (SVM) [13] classification algorithm and the least squares mapping technique, and we have compared its performance with that of the radial basis function kernel SVM classifier on the same database of US images. The latter comprised images of simple cysts, complicated cysts, benign and malignant lesions. Features evaluated were related to lesion-texture and lesion-outline parameters.

2. Material and methods

The study comprised 150 ultrasound breast images of 150 patients with histologically (pathologically) confirmed lesions and/or clinically proved cysts; 59 solid lesions were diagnosed as benign and 52 as malignant, 7 cysts were characterized as simple cysts, and 32 as complicated cysts. Table 1 gives a detailed account of the type of breast lesions involved in the present study.

2.1. Data acquisition

All ultrasound examinations were performed on an HDI-3000 ATL digital ultrasound system with a wide

Table 1
Breast lesion distribution and categorization into four classes

Class description	Description	No. of patients	Class total
I. Cyst	Simple cyst	7	7
II. Benign	Fibroadenoma	34	59
	Calcified fibroadenoma	2	
	Lymph node	1	
	Lipoma	6	
	Giant fibroadenoma	3	
	Degenerated fibroadenoma	1	
	Adenolipoma	1	
	Cystosarcoma phylloides	1	
	Papilloma	5	
	Hemangioma	1	
	Intramammary lymph node	4	
	III. Malignant	Infiltrating ductal carcinoma	
Infiltrating lobular carcinoma		5	
In situ ductal carcinoma		1	
Male breast carcinoma		1	
Cyst wall carcinoma		3	
Papillomatous carcinoma		2	
Myeloid carcinoma		1	
IV. Cystic masses (complicated cysts)	Infected cyst	9	32
	Cyst with debris	3	
	Hematoma	6	
	Cyst with infected wall	2	
	Old infected cyst	4	
	Abcess	3	
	Lymphocele	1	
	Oil cyst	1	
	Sebaceous cyst	3	

band(5–12 MHz) probe, by the same experienced physician (N.D.) using similar scanning conditions to provide for quality and uniformity in patient management and data acquisition. In each examination only one image through the largest diameter of the area of interest (mass or cyst) was obtained. Each breast image was digitized by connecting the video output of the ultrasound scanner to a Screen Machine II frame grabber using $512 \times 512 \times 8$ image resolution. The region of interest was outlined by the physician by means of a custom designed software in C++ (see Fig. 1) for the purposes of the present study. This software saved the selected ultrasound sub-image in a file and automatically calculated the lesion-texture and lesion-outline features for the design of the classification software subsystem.

2.2. Feature generation

2.2.1. Lesion-texture features

From each ultrasound sub-image, features evaluating the gray-level distribution were extracted by means of first- and second-order statistics. The first-order statistics were derived from the ultrasound sub-image gray-level distribution (histogram) and they comprised the mean, standard deviation, skewness, and kurtosis. Regarding the second-order statistics, 13 features were computed from the co-occurrence matrix [14] using one pixel step length and five from the run-length matrix [15]. Each second-order statistic feature was represented by two values, the mean and range over the 0, 45, 90 and 135° co-occurrence and run-length matrices, to ensure rotation independence. Features calculated from the co-occurrence matrix evaluate properties of image gray-level distribution, such as homogeneity, contrast, gray-level local variation, linear-dependencies, lack of order, etc. Similarly, run-length features evaluate the distribution of small (short runs) or large (long runs) structures within the area of interest. In this way, a total of 40 textural features were calculated for each image-sample, 4 from the image-sample histogram, 26 from the co-occurrence and 10 from the run-length matrices (in fact 13 features and 10 features, respectively, each being represented by two values, the mean and range).

2.2.2. Lesion-outline features

Eight features evaluating the shape of lesion's outline were created [16]. The shape was evaluated by means of features computed from the radial distances of the lesion's centroid from its outline edges. The significance of these features is that the breast-lesion shape has been related to the lesion's probability for malignancy. The reason is that a malignant tumor often infiltrates the surrounding tissue in a random manner thus producing irregular lesion shapes. Lesion shapes are often grouped into five classes: round, oval, lobulated, nodular, and stellate, with increasing probability of malignancy from round to stellate.

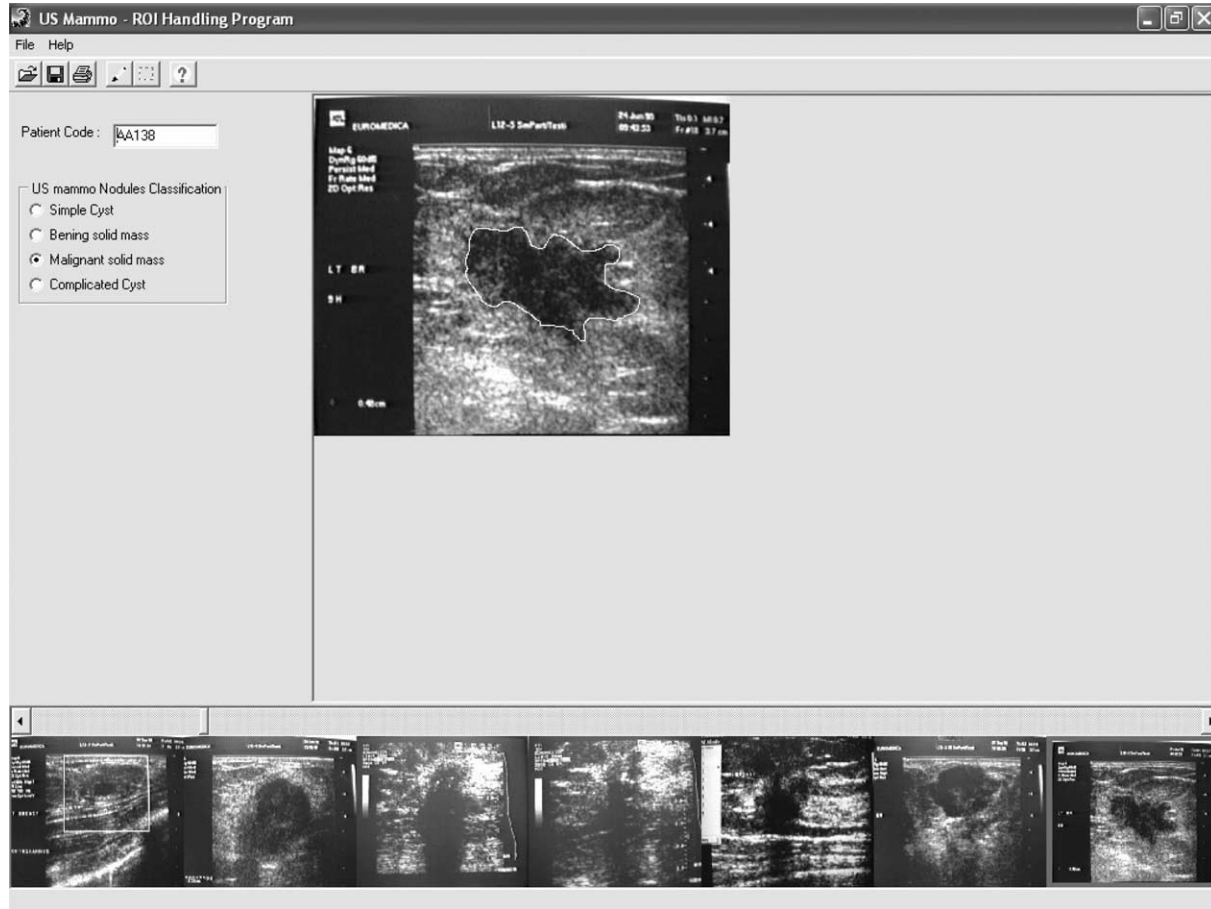


Fig. 1. Interface of the custom-made software system designed to read US images, segment by manual interaction breast lesions, and extract specific features.

2.3. Classification algorithms

2.3.1. The cubic least squares mapping—linear support vector machine classifier (SVM^{LSM})

The idea behind training an SVM classifier, in case of linearly separable classes, is to find two hyperplanes in the feature space that satisfy the following conditions: (i) maximize the area between the hyperplanes (also called the *margin*) and (ii) minimize the number of patterns that lie between the hyperplanes. A third hyperplane through the middle of the margin may then be defined, which is the decision boundary of the two classes. The discriminant function of the SVM classifier may thus be defined [13] as in (1)

$$f(\mathbf{x}) = \text{sign} \left(\sum_{i=1}^{N_S} a_i y_i (\mathbf{x}^T \cdot \mathbf{x}_i) + b \right) \quad (1)$$

where a_i are weight elements, b is a threshold parameter, \mathbf{x}_i are the support vectors (i.e. the training pattern vectors that have their corresponding weights $a_i \neq 0$), N_S is the number of the support vectors, and $y_i \in \{-1, +1\}$ depending on the class. The weights and threshold are determined so as to fulfill the above-mentioned conditions.

In the case of non-linearly separable classes, the SVM classifier first maps the patterns to a higher-dimension feature space using a transformation of the form

$$\mathbf{x}^T \cdot \mathbf{x}_i \mapsto \Phi^T(\mathbf{x}) \cdot \Phi(\mathbf{x}_i) \quad (2)$$

so that the classes become linearly separable.

The above transformation is usually achieved by means of a kernel function $k(\mathbf{x}_i, \mathbf{x})$, which is equivalent to the dot product $\Phi^T(\mathbf{x}) \cdot \Phi(\mathbf{x}_i)$ in (2) and it satisfies Mercer's conditions [13]. The discriminant function of the SVM classifier may thus be described by relation (3):

$$f(\mathbf{x}) = \text{sign} \left(\sum_{i=1}^{N_S} a_i y_i k(\mathbf{x}, \mathbf{x}_i) + b \right) \quad (3)$$

A number of kernels satisfying Mercer's conditions have been proposed [13] such as the linear, the polynomial, the radial basis, and the sigmoidal. Although powerful in its application, an important drawback of the SVM classifier is that its training phase is sometimes very demanding in computational time, especially in cases of high feature space dimensionality.

In the present study, our aim was to significantly reduce feature dimensionality and, at the same time, increase class

separation. This was made possible by suitably transforming the training patterns by means of a least squares mapping technique following cubic vector-augmentation (LSM) and then employing the linear-kernel SVM algorithm.

The LSM procedure has two steps, cubic-augmentation and least squares mapping, which are described below.

In the ‘cubic-augmentation’ step, the pattern vectors were augmented with second and third degree elements. Let $\mathbf{x} = [x_1, x_2, \dots, x_d]$ be a pattern vector, where d is the input space dimensionality. The latter was augmented by the following elements: 1/Quadratic terms: $x_i^2, x_i x_j$, where $i, j = 1, 2, \dots, d$ and $i \neq j$ and 2/Cubic terms: $x_i^3, x_i^2 x_j, x_i x_j^2, x_i x_j x_k$, where $i, j, k = 1, 2, \dots, d$ and $i \neq j \neq k$. By adding the quadratic and cubic terms, as well a constant term (of zeroth degree), the augmented pattern vectors dimensionality [17] is equal to:

$$\hat{d} = \frac{(d+3)!}{3!d!} \tag{4}$$

Of course, higher degree elements could have been defined, however, cubic we found to be a trade-off between the limited discriminating capabilities of first and second degree hyper-surfaces (hyper-planes, -ellipsoids, -paraboloids, -hyperboloids) and the computational requirements [13,17]. Another significant reason was that high-degree surfaces often lead to over-fitting conditions.

In the ‘least squares mapping’ step, the augmented patterns are transformed from the augmented feature space to the *decision* space, where the members of each class are clustered around arbitrarily pre-selected class-representative points, such that the mapping error is minimized [18]. In this way, class separation increases and dimensionality reduces to the number of existing classes. The transformation matrix \mathbf{A} , which is used for the mapping in to the decision space, is calculated so as to minimize the total mean-square error between the training set and the arbitrary pre-selected points in the decision space, as follows

$$\nabla_{\mathbf{A}} \sum_{i=1}^K \sum_{j=1}^{N_i+1} \frac{P_i}{N_i} \|\mathbf{A}\hat{\mathbf{x}}_{ij} - \mathbf{v}_i\|^2 = 0 \tag{5}$$

where $\hat{\mathbf{x}}_{ij}^T$ is the j th augmented pattern vector of class i ; \mathbf{v}_i is the decision space vector around which the patterns of class i are to be clustered; K is the number of classes; N_i is the number of patterns of class i ; and P_i is the a priori probability of class i .

The solution of Eq. (5) leads to relation (6):

$$\mathbf{A} = \left[\sum_{i=1}^K \sum_{j=1}^{N_i+1} \frac{P_i}{N_i} \mathbf{v}_i \hat{\mathbf{x}}_{ij}^T \right] \left[\sum_{i=1}^K \sum_{j=1}^{N_i+1} \frac{P_i}{N_i} \hat{\mathbf{x}}_{ij} \hat{\mathbf{x}}_{ij}^T \right]^{-1} \tag{6}$$

The LSM transformation is then applied according to relation (7):

$$\mathbf{z}_{ij} = \mathbf{A}\hat{\mathbf{x}}_{ij} \tag{7}$$

The transformed data ($\mathbf{A}\hat{\mathbf{x}}_{ij}$) were used by the linear-kernel SVM classifier, resulting in the following discriminant function:

$$f(\mathbf{x}) = \text{sign} \left(\sum_{i=1}^{N_s} a_i y_i [(\mathbf{A}\hat{\mathbf{x}})^T \cdot (\mathbf{A}\hat{\mathbf{x}}_i)] + b \right) \tag{8}$$

2.3.2. Radial basis function kernel support vector machine classifier (LSM^{RBF})

For comparison reasons, an SVM classifier [13] fed with the original (unmapped) data was also employed, using as kernel the widely used Gaussian radial basis function (RBF)

$$k(\mathbf{x}, \mathbf{x}_i) = \exp \left(\frac{-\|\mathbf{x}_i - \mathbf{x}\|^2}{2\sigma^2} \right) \tag{9}$$

resulting in the following discriminant function for the LSM^{RBF}

$$f(\mathbf{x}) = \text{sign} \left(\sum_{i=1}^{N_s} a_i y_i e^{-\|\mathbf{x}_i - \mathbf{x}\|^2 / 2\sigma^2} + b \right) \tag{10}$$

where σ is the standard deviation.

2.3.3. Third degree polynomial classifier (POL)

In the feature selection step, a simple and fast, but efficient enough, third degree polynomial classifier [17], equipped with a minimum Euclidean-distance rule, was used. The discriminant function of PC-classifier is given by the following relation

$$\begin{aligned} f_c(\mathbf{X}) = & \sum_{i=1}^d a_{cii} x_i^3 + \sum_{i=1}^{d-1} \sum_{j=i+1}^d a_{cij} x_i^2 x_j + \sum_{i=1}^{d-1} \sum_{j=i+1}^d a_{ciji} x_i x_j^2 \\ & + \sum_{i=1}^{d-2} \sum_{j=i+1}^{d-1} \sum_{k=j+1}^d a_{cijk} x_i x_j x_k + \sum_{i=1}^d a_{cii} x_i^2 \\ & + \sum_{i=1}^{d-1} \sum_{j=i+1}^d a_{cij} x_i x_j + \sum_{i=1}^d a_{ci} x_i - b_c \end{aligned} \tag{11}$$

where d is the number of features, a_c are weight coefficients, b_c is a threshold parameter, and x_j are the pattern vector elements.

2.4. System design

In this phase it is required that the overall classification system is designed for optimum operation. Accordingly, system architecture, best feature-combinations, and best classification algorithms are determined so that highest classification accuracy in correctly characterizing breast lesions is achieved.

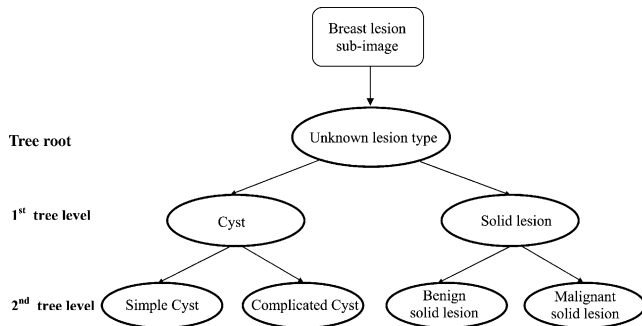


Fig. 2. The hierarchical tree structure used for classifying breast lesions, employing the SVM^{L_{SM}} classifier at each decision-node.

2.4.1. System architecture

The computer software for classifying the breast lesions into four classes (simple cysts, complicated cysts, benign solid masses, and malignant solid masses) was designed as a two-level hierarchical decision tree (Fig. 2). At each level two-class discrimination tests were performed. At the first level data were grouped into cysts (comprising simple cysts and complicated cysts) and non-cysts (consisting of benign and malignant solid breast lesions). At the second level two discrimination tests were performed concerning cysts–complicated cysts and malignant–benign breast solid lesions. At each node of the hierarchical tree the highest performance classifier was chosen and it was designed to work optimally (highest classification accuracy with smallest number of features). This was achieved after testing the performance of each classifier (POL, SVM^{RBF}, SVM^{L_{SM}}) separately at each tree-node.

2.4.2. Best feature selection

Features are essential in the design of the classifier. Ideally, all features at hand (45) should be employed in the design of the classifier, but since a number of them may be redundant due to mutual correlations [17], an optimum number of them has to be selected to achieve highest classification accuracy. Choosing the best feature-combination that will maximize the performance of the classifier at each node of the hierarchical tree is a necessary but time-consuming and computationally demanding procedure. One method (exhaustive search [17]) involves designing the classifier by means of every possible feature-combination (i.e. 2, 3, 4, etc. feature-combinations), each time testing the classifier’s performance and finally selecting that

feature-combination that demonstrates the highest classification accuracy with the smallest number of features. Considering the complexity of the present application, the exhaustive method was time prohibiting, thus the following compromise had to be adopted. At each level of the hierarchical tree-node the discriminatory ability of each one of the 45 features was tested employing the Student’s *t*-test; only the highest discriminating features ($p < 0.001$) were selected and were employed in the design of the classifier. At each tree-node, the performance of the fast PC-classifier was tested for all possible combinations of the selected high discriminating features, in order to determine the highest classification accuracy with the minimum number of features. In this stage, the PC-classifier performance was evaluated by classifying all data involved in its design (self-consistency test). Finally, the best feature-combination, giving highest classification accuracy with smaller number of features, was selected at each hierarchical tree decision-node.

2.4.3. Best classifier selection

Classifier performance was evaluated by the leave-one-out method (LOO) [17]. According to the latter, the classifier is designed by all but one sub-image, which is then classified. The process is repeated, each time leaving a different sub-image out, until all data have been processed. In this way, the classifier is evaluated by data that are not involved in its design. It is evident, however, that the classifier has to be re-designed each time a sub-image is left out. The performances of all classifiers (POL, SVM^{RBF}, SVM^{L_{SM}}) were evaluated at each decision-node of the hierarchical tree, employing the node’s best feature-combination. This huge workload was split into two and it was performed on two workstations (Pentium IV, working at 2.4 MHz), requiring several hours of processing time. In the end the best classifier was chosen to work at each tree decision-node.

2.5. System performance evaluation

The accuracy of the system was performed by presenting each breast-lesion sub-image to the input of the hierarchical tree structure. The latter evaluated the pre-selected best feature-combinations at each decision-node and assigned the sub-image to one of four classes (simple cyst,

Table 2
Classification results achieved by the three classifiers at each decision-node of the hierarchical tree

Hierarchical tree decision-node	Best feature-combination ^a	POL, LOO (S-C)	SVM ^{RBF} , LOO (S-C) N_{sv}	SVM ^{L_{SM}} , LOO (S-C) N_{sv}
Cysts–non-cysts	ASM, ENT, SEN, LRE, AR, ZC	88.0% (100%)	89.3% (100%) 71	100% (100%) 2
Simple cysts–complicated cysts	KU, RDm, RDs, Rde	100% (100%)	82.1% (100%) 36	100% (100%) 2
Benign lesions–malignant lesions	SVA, SEN, DEN, DVA, SRE	73.0% (99.1%)	62.2% (100%) 95	98.2% (98.2%) 4

POL, polynomial classifier; SVM^{RBF}, SVM classifier with RBF kernel; SVM^{L_{SM}}, SVM classifier using cubic least squares mapping and linear kernel; LOO, leave-one-out method; S-C, self-consistency test; N_{sv} , number of support vectors.

^a See Table 3.

Table 3
Textural and outline features and abbreviations

Name	Abbreviation	Type
Kurtosis	KU	First order
Angular second moment	ASM	Co-occurrence
Entropy	ENT	Co-occurrence
Sum variance	SVA	Co-occurrence
Sum entropy	SEN	Co-occurrence
Difference entropy	DEN	Co-occurrence
Difference variance	DVA	Co-occurrence
Short run-length emphasis	SRE	Run-length
Long run-length emphasis	LRE	Run-length
Mean of radial distance	RDM	Outline
Standard deviation of radial distance	RDs	Outline
Entropy of radial distance	RDe	Outline
Area ratio parameter	AR	Outline
Zero crossing count	ZC	Outline

complicated cyst, benign solid lesion, malignant solid lesion). The whole procedure was performed by employing the LOO method and results were presented in a 4 × 4 truth table.

3. Results and discussion

Table 2 presents the results obtained by the three different classifiers at each decision-node of the hierarchical

tree structure. Results were obtained employing the LOO and the self-consistency tests. At the first level, highest classification accuracy for cysts–non-cysts discrimination employing the LOO method was achieved by the SVM^{LSM} while the other classifiers scored lower. Best feature-combination comprised six features (see Tables 3 and 4) textural features (three from the co-occurrence and one from the run-length matrices) and two lesion-outline features. It is interesting to note that the SVM^{LSM} managed to discriminate correctly all the cysts, using a very small number of two support vectors. This implies that data had been well separated by the least squares mapping process and thus the linear-kernel SVM could work fast and accurately. This can be seen in Fig. 3a, where the six-dimensional feature space is transformed into a two-dimensional decision space (dimensionality is equal to the number of classes) by the least squares mapping process. It is an easy task for the linear-kernel SVM to define its support vectors and draw its decision line. In contrast, the non-linear SVM^{RBF} had to employ 71 support vectors (Table 2), causing over-fitting in achieving 100% self-consistency. That result evidently collapsed to 89.3% when the LOO method was applied, implying that the confidence put to the SVM^{RBF} classifier’s accuracy is lower when presented with new data.

At the second level of the tree structure, the SVM^{LSM} managed to separate successfully all simple cysts from complicated cysts, while scoring a high of 98.2% by both

Table 4
Truth table presenting the results of the hierarchical tree classification system in discriminating breast lesions, employing the SVM^{LSM} classifier and the leave-one-out method at each tree decision-node

		Classification				Accuracy (%)
		Simple cysts	Complicated cysts	Malignant masses	Benign masses	
Biopsy	Simple cysts	7	0	0	0	100
	Complicated cysts	0	32	0	0	100
	Malignant masses	0	0	51	1	98.1
	Benign masses	0	0	1	58	98.3
Overall accuracy						98.7

SVM^{LSM}, SVM classifier using cubic least squares mapping and linear kernel.

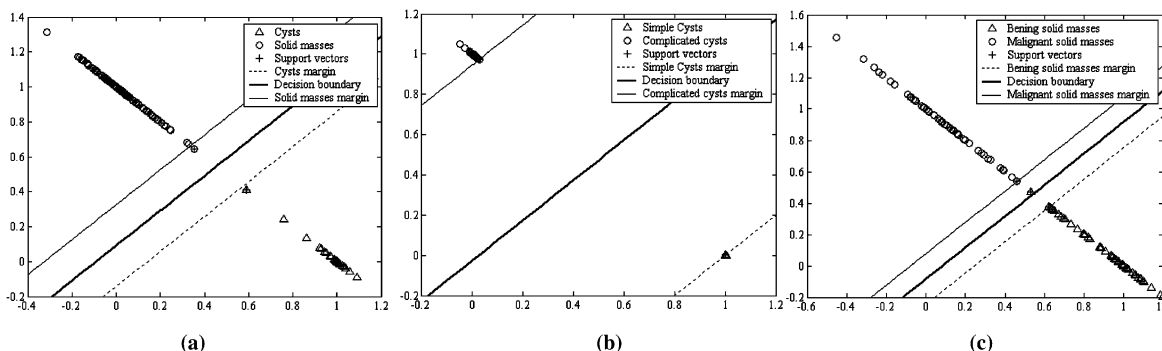


Fig. 3. Scatter diagrams of the SVM^{LSM} data mapping and decision boundaries regarding (a) cysts vs. non-cysts, (b) simple cysts vs. complicated cysts, and (c) malignant lesions vs. benign lesions.

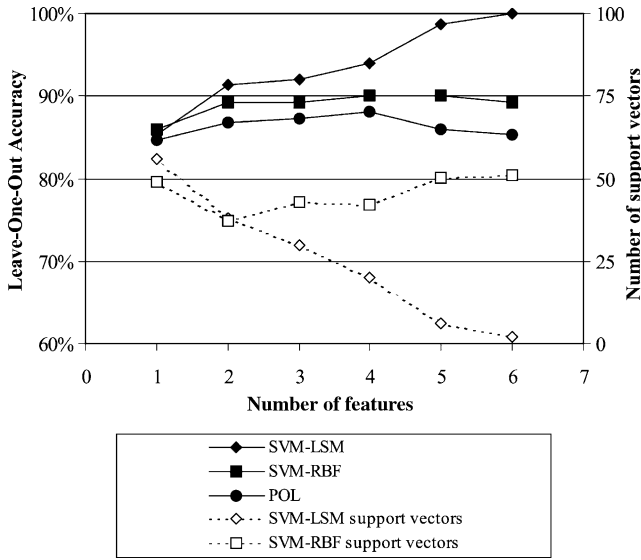


Fig. 4. Variation of classification accuracies and support vectors involved vs. the number of features for the POL, SVM^{RBF} and the SVM^{LSM} classifiers, employing the LOO method.

LOO and self-consistency tests in discriminating between malignant and benign solid lesions, employing only four support vectors. This discrimination has been recognized by previous researchers [3] to being a difficult task.

As it can be seen in Table 2, there were only two solid lesions misdiagnosed by the SVM^{LSM}, one from each class. Fig. 3b and c demonstrate the high effectiveness of the algorithm in discriminating between simple cysts from

complicated cysts and benign from malignant solid lesions, respectively. Results obtained by previous workers [5,6,9, 10] ranged between 85 and 95%, employing other classification algorithms and using different to ours US-image databases. In a previous report [5] using the MLP and auto-correlation features, a 95% high precision was reached in discriminating malignant from benign solid breast-nodules. Regarding the polynomial and SVM^{RBF} algorithms of the present work, they gave a low 73 and 62.2% using the LOO and 99.1 and 100% employing the self-consistency tests, respectively. Another finding of the present work is that, in discriminating between benign and malignant lesions, the best feature-combination included lesion-texture features only (see Table 3), implying that the lesion shape may be indicative but not decisive of the type of solid lesion.

A comparative evaluation of the three classifiers' performances in discriminating between cysts and non-cysts, may be obtained by employing the LOO method and a varying number of feature-combinations, as shown in Fig. 4. Also depicted, are the variations of the support vectors employed by the SVM^{RBF} and SVM^{LSM} classifiers. It is observed that although the classifiers start with approximately equal performances, their precision differs as the number of features increases. The classification accuracy of the SVM^{LSM} attains highest value (100%) at six features while the performance of the SVM^{RBF} and the POL seem to peak (90 and 88%, respectively) at four feature-combinations. It is worth noting that the number of support vectors employed by both SVM classifiers

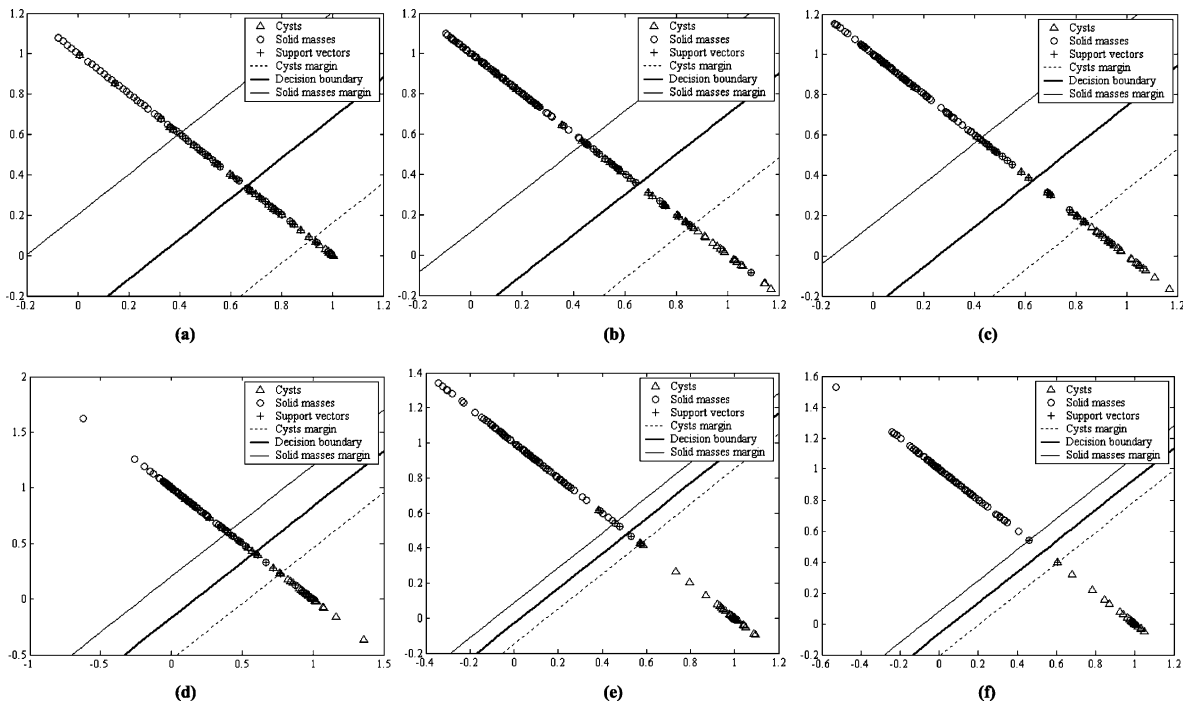


Fig. 5. Scatter diagrams displaying class separation and SVM^{LSM} classifier operation as the number of features increases from two features (a) to six features (f).

decreases with increasing classification accuracy, implying that when data is clustered into separable classes then fewer support vectors are required by the SVM classification algorithm. The strong point of the proposed SVM^{LSM} classifier is that if classes may be dragged further apart in a higher-dimensional feature space by increasing the number of features, then the transformation from a high-dimensional feature space by least squares mapping into a two-dimensional decision space (equal to the number of classes) will result in separable classes so that the linear-kernel SVM can work with high precision. This gradual transformation from non-separable into separable classes by increasing the number of features from two to six may be clearly observed in Fig. 5a–f, respectively.

Table 4 shows the classification accuracy of the hierarchical tree structure system when presented with all the sub-images of the present study. Overall 98.7% classification accuracy using LOO was found, which is an indication of the system's diagnostic capability. However, final judgment on the system's reliability may be passed only after the system is presented with a large number of new breast lesions in a clinical environment. Such process is currently underway.

In conclusion, an efficient classification method is proposed, based on the Least Squares Mapping linear-kernel SVM algorithm, which was found to outperform the polynomial and the SVM^{RBF} classifiers in discriminating breast lesions. Additionally, a high precision classification system was designed capable of characterizing breast lesions as simple cysts, complicated cysts, benign lesions, and malignant lesions from their ultrasound images. These results are promising, considering the radiologists' accuracy (86.67%), in discriminating benign from malignant lesions on ultrasound [19]. This system could be a useful diagnostic tool by providing a second opinion to the physician.

4. Summary

In the present study we have developed an efficient classification algorithm (SVM^{LSM}) for characterizing breast lesions. In the proposed algorithm, the pattern vectors first undergo a least-squares error minimization mapping, and second, a linear-kernel support vector machine classifier is applied to the mapped data.

A large number of confirmed lesions were manually segmented from an equal number of breast ultrasound images by an experienced radiologist. For each lesion, textural and outline features were calculated and they were fed to the SVM^{LSM} algorithm. A two-level hierarchical tree classification structure was designed to characterize, at the 1st level, breast lesions as cysts or solid masses, and, at the second level, to discriminate simple cysts

from complicated cysts and benign from malignant solid masses.

Highest classification accuracy was achieved in discriminating cysts from solid masses and simple cysts from complicated cysts, and 98.2% in differentiating benign from malignant solid lesions. The proposed SVM^{LSM} based pattern recognition system may be employed to assist the radiologist in characterizing breast lesions on ultrasound.

References

- [1] Dennis MA, Parker SH, Klaus AJ, Stavros AT, Kaske TI, Clark SB. Breast biopsy avoidance: the value of normal mammograms and normal sonograms in the setting of a palpable lump. *Radiology* 2001; 219:186–91.
- [2] Baker JA, Soo MS, Rosen EL. Artifacts and pitfalls in sonographic imaging of the breast. *AJR* 2001;176:1261–6.
- [3] Zonderland HM, Coerkamp EG, Hermans J, van de Vijver MJ, van Voortuisen AE. Diagnosis of breast cancer: contribution of US as an adjunct to mammography. *Radiology* 1999;213:413–22.
- [4] Rahbar G, Sie AC, Hansen GC, Prince JS, Melany ML, Reynolds HE, Jackson VP, Sayre JW, Basset LW. Benign versus malignant solid breast masses: US differentiation. *Radiology* 1999;213:889–94.
- [5] Chen DR, Chang RF, Huang YL. Computer-aided diagnosis applied to US of solid breast nodules by using neural networks. *Radiology* 1999;213:407–12.
- [6] Sivaramakrishna R, Powell KA, Lieber ML, Chilcote WA, Shekhar R. Texture analysis of lesions in breast ultrasound images. *Comp Med Im Graph* 2002;26:303–7.
- [7] Shankar PM, Reid JM, Ortega H, Piccoli CW, Goldberg BB. Use of non-Rayleigh statistics for the identification of tumors in ultrasonic B-scans of the breast. *IEEE Trans Med Imaging* 1993;12(4): 687–92.
- [8] Gara BS, Krasner BH, Horii SC, Ascher S, Mun SK, Zeman RK. Improving the distinction between benign and malignant breast lesions: the value of sonographic texture analysis. *Ultrason Imaging* 1993;15:267–85.
- [9] Stavros AT, Thickman D, Rapp CL, Dennis MA, Parker SH, Sisney GA. Solid breast nodules: use of sonography to distinguish between benign and malignant lesions. *Radiology* 1995;199:123–34.
- [10] Chen DR, Chang RF, Huang YL. Breast cancer diagnosis using self-organizing map for sonography. *Ultrasound Med Biol* 2000;26(3): 405–11.
- [11] Giger ML, Al-Hallaq H, Huo Z, Moran C, Wolverton DE, Chan CW, Zhong W. Computerized analysis of lesions in US images of the breast. *Acad Radiol* 1999;6:665–74.
- [12] Jackson VP. The role of US in breast imaging. *Radiology* 1990;177: 305–11.
- [13] Keeman V. Learning and soft computing, support vector machines, neural networks, and fuzzy logic models. Cambridge, MA: MIT Press; 2001. p.121–91.
- [14] Haralick RM, Shanmugam K, Dinstein I. Textural features for image analysis. *IEEE Trans Syst Man Cybern* 1973;SMC-3:610–21.
- [15] Galloway MM. Texture analysis using gray level run lengths. *Comput Graphics Image Process* 1975;4:172–9.
- [16] Bruce R, Adhami RR. Classifying mammographic mass shapes using the wavelet transform modulus-maxima method. *IEEE Trans Med Imaging* 1999;18:1170–7.
- [17] Theodoridis S, Koutroumbas K. Pattern recognition. San Diego: Academic Press; 1998.
- [18] Ahmed N, Rao KR. Orthogonal transforms for digital signal processing. Berlin: Springer; 1975. p. 225–58.

- [19] Chang RF, Kuo WJ, Chen DR, Huang YL, Lee JH, Chou YH. Computer-aided diagnosis for surgical office-based breast ultrasound. *Arch Surg* 2000;135:696–9.

Nikolaos Piliouras completed his Bachelor in Electronics and Telecommunication Engineering from the Hellenic Air Force Academy, Greece in 1989. He then received his PhD from the Department of Geology of the University of Patras in 2003, at the field of Earth Tomography. He is currently in cooperation with the Medical Image and Signal Processing Laboratory of the Medical Instrument Technology Department at the Technological Educational Institution of Athens, involved in several research projects and in the educational activities of the Institution. Dr Piliouras's research interests are image and signal processing as well as biomedical engineering.

Ioannis Kalatzis completed his BSc in Physics from the University of Athens, Greece in 1987 and his PhD in Medical Physics from the Medical School of the University of Athens in 2000. During his PhD thesis he was involved in tomographic image processing in the Nuclear Medicine Department of the Hippokrateio General Hospital of Athens, as well in research projects concerning image processing and pattern recognition applications in medicine. He is currently research fellow at the Laboratory of Medical Image and Signal Processing in the Department of Medical Instruments Technology of the Technological Educational Institution of Athens. Dr Kalatzis's research interests include medical image and signal processing and analysis, with focus in the field of pattern recognition.

Nikolaos Dimitropoulos MD, is Director of the Department of Radiology, Mammography and US in 'Euromedica' Medical Center, Athens, Greece. He received his PhD in Radiology from Medical School, Athens University, Greece. He has specialized in mammography and ultrasonography at Guys and Middlesex Hospitals, London, UK, and Albert—Ludwigs Universität, Freiburg, Germany. He has worked as clinical radiologist and he was head of radiological departments in several Greek hospitals and clinics. His research interests include diagnostic imaging and medical image processing and analysis in X-rays and ultrasound.

Dionisis Cavouras was born in Kalamata, Greece, in 1951. He received his BSc (1974) in Electronic Engineering, and the MSc (1976) and PhD (1981) in Systems Engineering from the City University, London, England. He was a research assistant at the Department of Nuclear Medicine, Guy's Hospital, London, England between 1976 and 1981, he worked as a Research Fellow at the Department of Computed Tomography, Hellenic Air-force Hospital, Athens, Greece, between 1994 and 1991 and since then he is a professor of medical imaging processing at the Department of Medical Instruments Technology in Technological Educational Institution of Athens, and Director of the Laboratory of Medical Image and Signal Processing. He is also an adjunct lecturer on medical image processing at the Department of Informatics, University of Athens, Greece, and the Department of Medical Physics, University of Patras, Greece. His research interests include medical image processing, image analysis, pattern recognition, medical statistics, and medical physics. He has published numerous technical and medical papers as well as conference proceedings.