



Available at

www.ElsevierComputerScience.com

POWERED BY SCIENCE @ DIRECT®

Pattern Recognition Letters 25 (2004) 249–258

Pattern Recognition  
Letters

www.elsevier.com/locate/patrec

# Development of the probabilistic neural network–cubic least squares mapping (PNN–LSM<sup>3</sup>) classifier to assess carotid plaque's risk

N. Piliouras<sup>a</sup>, I. Kalatzis<sup>a</sup>, P. Theocharakis<sup>a</sup>, N. Dimitropoulos<sup>b</sup>,  
D. Cavouras<sup>a,\*</sup>

<sup>a</sup> Department of Medical Instrumentation Technology, Technological Educational Institution of Athens, Ag. Spyridonos Street, Egaleo, GR-122 10 Athens, Greece

<sup>b</sup> Department of Medical Imaging, EUROMEDICA Medical Center, 2 Mesogeion Avenue, GR-11527 Athens, Greece

Received 18 July 2003; received in revised form 29 September 2003

## Abstract

An efficient classification algorithm based on the cubic least squares mapping (LSM<sup>3</sup>) and the probabilistic neural network (PNN) classifier is proposed for assessing the carotid plaque's risk of causing brain infarcts. Ultrasound images of 24 high-risk and 32 low-risk carotid plaques were manually segmented by an experienced physician using a custom developed software. Three textural features, related to the plaque's internal composition, the PNN, and the PNN–LSM<sup>3</sup> classification algorithms were used to design a classification system. PNN classification accuracy was 92.9%, misdiagnosing one high-risk and three low-risk plaques while the PNN–LSM<sup>3</sup> managed to classify all plaques correctly. The proposed system may be of value to patient management as a second opinion tool, after it is tested on more data in a clinical environment.

© 2003 Elsevier B.V. All rights reserved.

*Keywords:* Carotid plaque; Probabilistic neural network; Least squares mapping; Ultrasound; Classification

## 1. Introduction

Atherosclerotic carotid plaque composition, structure, and stenosis are important factors in assessing the risk factor for stroke in both symp-

tomatic and asymptomatic patients (El-Barghouty et al., 1996). The method mostly used in non-invasively examining the carotid plaque is ultrasound (US), which, however, involves the subjective evaluation of US images by the physician and, thus, it depends upon the experience of the operator (Hatsukami et al., 1994; Rakebrandt et al., 2000).

Objective methods for estimating the risk associated with carotid plaque mainly concern evaluation of the gray-scale median value (GSM) from

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

E-mail address: [cavouras@teiath.gr](mailto:cavouras@teiath.gr) (D. Cavouras).

the US images (Biasi et al., 1999; Tegos et al., 2001a,b; Cheng et al., 2002; Denzel et al., 2003). GSM is a parameter that evaluates carotid plaque echogenicity. Low GSM refers to echolucent type of carotid plaques, due to their low echo levels, and high GSM concerns echogenic carotid plaques, due to higher echo levels. Echolucent plaques consist mainly of high lipid or haemorrhagic content while echodense contain fibrous and calcified tissue. It has been found that patients with low GSM (<50) have a higher risk of suffering a brain stroke (Biasi et al., 1999) than patients with higher GSM. Recent studies have employed computer processing to calculate textural features from the carotid plaque's US image and have applied statistical methods (Hatsukami et al., 1994; Noritomi et al., 1997; Montauban van Swijndregt et al., 1998; Wilhjelm et al., 1998; Zhang et al., 1998; Arnold et al., 2001) for assessing the carotid risk factor. Textural features comprised GSM and features emanating from the run-length and/or co-occurrence matrices (Wilhjelm et al., 1998; Arnold et al., 2001). Other features used were (i) the degree of stenosis, (ii) the echogenicity area index, and (iii) the plaque's surface structure (Pedro et al., 2002). However, to our knowledge, more advanced computer-aided methods, such as pattern recognition, have not been applied to the solution of this problem.

The contribution of the present study was

- (1) to design and implement a computer-based image analysis system, using pattern recognition methods, for selecting, processing, and classifying US carotid plaque images as high- or low-risk of causing brain damage. Classification was based on textural features derived from the carotid plaque's US image;
- (2) to develop a new efficient classification algorithm, by exploiting the power and speed of the probabilistic neural networks (PNN) and the capabilities for dimensionality reduction and class separation of the cubic least squares mapping (LSM<sup>3</sup>) technique. The power of the proposed algorithm (PNN–LSM<sup>3</sup>) was compared with that of the classical PNN.

## 2. Material and methods

### 2.1. Data acquisition

Fifty-six patients with various degrees of carotid stenosis were examined on an HDI-3000 ATL digital ultrasound system by the same experienced physician (ND). The scanning protocol was kept fixed (El-Barghouty et al., 1996) to provide for uniformity in data acquisition. All patients were examined in the supine position employing anterior, lateral, and posterior longitudinal scans. Thirty-eight patients were asymptomatic and the rest 18 were symptomatic (6 strokes, 2 global symptoms, 10 transient ischaemic attacks). Brain infarctions were found on CT in six symptomatic and in eight non-symptomatic patients. Employing a custom developed software system, designed in C++ for the purposes of the present study, each US image was digitized by connecting the US scanner's video output to a frame grabber using  $512 \times 512 \times 8$  image resolution. The same software was used by the physician for outlining the boundary of each carotid plaque (see Fig. 1), for the automatic computation of all carotid plaque features, and for the classification of the carotid plaque type. Data processing was performed on a Pentium IV, 2.4 GHz computer.

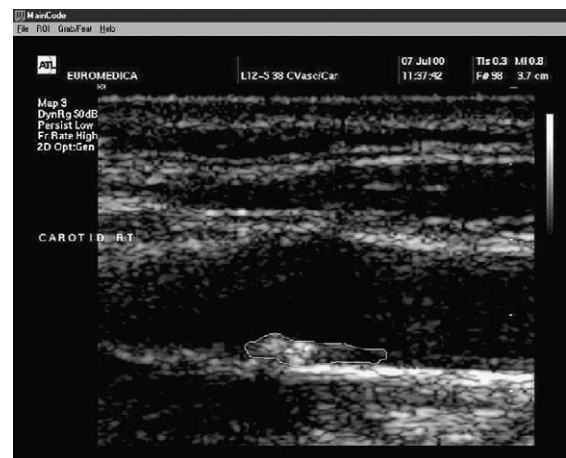


Fig. 1. Interface of the custom-made software system designed to read US images, segment by manual interaction carotid plaque-lesions, and extract specific features.

In determining a golden standard for subsequent processing, the plaques were separated into two groups: 32 were characterized as low-risk and 24 as high-risk of causing brain infarcts. Factors considered for this categorization were (1) the presence of previous symptoms, (2) the degree of stenosis as assessed on US, (3) the gray-scale median ( $GSM > 50$  for echogenic plaques and  $GSM \leq 50$  for echolucent plaques following an US scanning protocol as close as possible to that suggested by Biasi et al. (1999), (4) the physician's subjective assessment during US scanning regarding the type of plaque surface (smooth or ulcerated) and the plaque's internal structure; plaques predominantly echolucent with a thin echogenic cap or intermediate echolucent lesions with small areas of echogenicity were classified as echolucent and of high-risk and plaques with intermediate echogenic lesions with small areas of echolucency or uniformly echogenic lesions were classified as echogenic and of low-risk, and (5) the final clinical outcome.

## 2.2. Feature generation

From each ultrasound carotid plaque 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 image gray-level distribution (histogram) and they comprised the mean, standard deviation, skewness, and kurtosis. Regarding the second-order statistics, 12 features were computed from the co-occurrence matrix (Haralick et al., 1973), using one pixel step length, and five from the run-length matrix (Galloway, 1975). Each second-order statistic feature was represented by two values, the mean and range over the  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  and  $135^\circ$  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, 38 features were calculated for each carotid

plaque image, 4 from the image histogram, 24 from the co-occurrence and 10 from the run-length matrices (in fact 12 features and 5 features, respectively) each being represented by two values, the mean and range (Haralick et al., 1973; Galloway, 1975). An additional feature was also calculated from each plaque image, the ratio of echodense to echolucent (D/L) plaque content. Each plaque image was processed by characterizing each pixel as echodense or echolucent depending on its gray-level proximity to preset values of 113 and 40, respectively. These values were the gray-level averages of a large sample of dense and lucent US carotid plaque lesions, which had been interactively selected by the physician.

All 39 features were normalized to zero mean and unit standard deviation (Theodoridis and Koutroumbas, 1999), according to relation

$$\tilde{x}_i = \frac{x_i - \mu}{\sigma}, \quad (1)$$

where  $x_i$  and  $\tilde{x}_i$  are the  $i$ th feature values of pattern vector  $\mathbf{x}$  before and after the normalization, respectively, and  $\mu$  and  $\sigma$  are the mean value and standard deviation, respectively of feature  $i$  over all plaques (both high- and low-risk).

## 2.3. Design of the cubic least squares mapping-probabilistic neural network (PNN-LSM<sup>3</sup>) classifier

Fig. 2 shows the structure of the PNN-LSM<sup>3</sup>. From each carotid plaque-image a number of optimum features (e.g. 4 features) are computed and are presented into the input of the classifier. The cubic least squares mapping (LSM<sup>3</sup>) transforms the feature space into the decision space, reducing dimensionality to the number of classes. The LSM<sup>3</sup> procedure has two steps.

In the "cubic" step, the pattern vectors are 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 is augmented by the following elements: (1) *Quadratic terms*:  $x_i^2$ ,  $x_i x_j$ , where  $i, j = 1, 2, \dots, d$  and  $i \neq j$ . (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

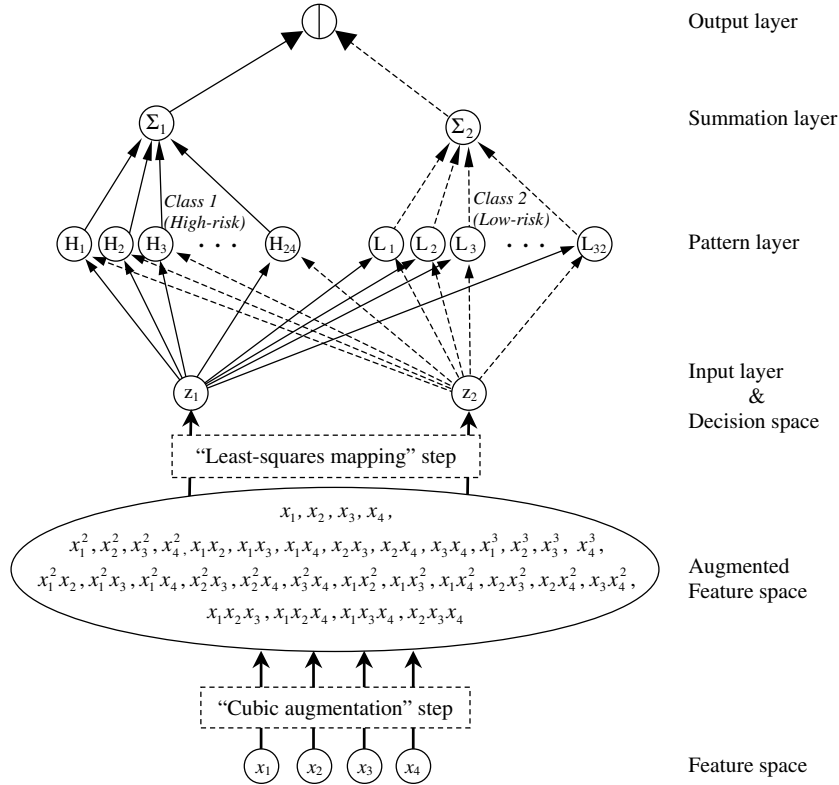


Fig. 2. Schematic diagram of the PNN-LSM<sup>3</sup> classifier.

term (of zeroth degree), the augmented pattern vectors dimensionality (Theodoridis and Koutroumbas, 1999) is equal to

$$\hat{d} = \frac{(d + 3)!}{3!d!}. \tag{2}$$

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 (Kecman, 2001; Theodoridis and Koutroumbas, 1999). Another significant reason was that high-degree surfaces often lead to over-fitting conditions.

Following the cubic step, the LSM step (Ahmed and Rao, 1975) follows. In LSM, the augmented patterns are transformed from the augmented feature space to the *decision* space, where the members of each class are clustered around se-

lected points, such that the mapping error is minimized, making the classes more separable. The transformation matrix  $M$ , which is used for the mapping in to the decision space, is calculated to minimize the total mean-square error between the data and the arbitrary pre-selected class-points in the decision space.

The decision space error for a single augmented pattern vector  $i$  of class  $k$  is given by

$$e_i = M\hat{x}_{ki} - C_k, \tag{3}$$

where

$$M = \begin{bmatrix} m_{11} & m_{12} & \cdots & m_1 \hat{d} \\ m_{21} & m_{22} & \cdots & m_2 \hat{d} \\ \vdots & \vdots & \ddots & \vdots \\ m_{N_c 1} & m_{N_c 2} & \cdots & m_{N_c} \hat{d} \end{bmatrix}$$

is the mapping matrix,  $N_c$  is the number of classes,  $\hat{d}$  is the augmented feature space dimensionality,

$\hat{\mathbf{x}}_{ki}$  is the  $i$ th augmented pattern vector of class  $k$ , and  $\mathbf{C}_k$  is the decision space point where the patterns of class  $k$  are to be clustered.

Considering equal a priori probabilities for each class  $k$ , the total mean-square error for the whole training set is given by

$$e = \sum_{k=1}^{N_c} \left( \frac{1}{N_k} \sum_{i=1}^{N_K} (\mathbf{M}\hat{\mathbf{x}}_{ki} - \mathbf{C}_k)^T (\mathbf{M}\hat{\mathbf{x}}_{ki} - \mathbf{C}_k) \right), \quad (4)$$

where  $N_k$  is the number of patterns of class  $k$ .

To minimize the mean-square error  $e$ , the mapping matrix  $\mathbf{M}$  can be calculated solving the equation

$$\nabla_{\mathbf{M}} e \equiv \begin{bmatrix} \frac{\partial e}{\partial m_{11}} & \frac{\partial e}{\partial m_{12}} & \cdots & \frac{\partial e}{\partial m_{1 \hat{d}}} \\ \frac{\partial e}{\partial m_{21}} & \frac{\partial e}{\partial m_{22}} & \cdots & \frac{\partial e}{\partial m_{2 \hat{d}}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial e}{\partial m_{N_c 1}} & \frac{\partial e}{\partial m_{N_c 2}} & \cdots & \frac{\partial e}{\partial m_{N_c \hat{d}}} \end{bmatrix} = 0. \quad (5)$$

By substituting Eq. (4) to Eq. (5)

$$\nabla_{\mathbf{M}} \left[ \sum_{k=1}^{N_c} \left( \frac{1}{N_k} \sum_{i=1}^{N_K} (\mathbf{M}\hat{\mathbf{x}}_{ki} - \mathbf{C}_k)^T (\mathbf{M}\hat{\mathbf{x}}_{ki} - \mathbf{C}_k) \right) \right] = 0 \quad (6)$$

or

$$\sum_{k=1}^{N_c} \left( \frac{1}{N_k} \sum_{i=1}^{N_K} \left[ \nabla_{\mathbf{M}} (\hat{\mathbf{x}}_{ki}^T \mathbf{M}^T \mathbf{M} \hat{\mathbf{x}}_{ki}) - 2 \nabla_{\mathbf{M}} (\hat{\mathbf{x}}_{ki}^T \mathbf{M}^T \mathbf{C}_k) + \nabla_{\mathbf{M}} (\mathbf{C}_k^T \mathbf{C}_k) \right] \right) = 0. \quad (7)$$

Basic matrix operations give

$$\nabla_{\mathbf{M}} (\hat{\mathbf{x}}_{ki}^T \mathbf{M}^T \mathbf{M} \hat{\mathbf{x}}_{ki}) = 2 \mathbf{M} (\hat{\mathbf{x}}_{ki} \hat{\mathbf{x}}_{ki}^T), \quad (8)$$

$$\nabla_{\mathbf{M}} (\hat{\mathbf{x}}_{ki}^T \mathbf{M}^T \mathbf{C}_k) = \mathbf{C}_k \hat{\mathbf{x}}_{ki}^T, \quad (9)$$

$$\nabla_{\mathbf{M}} (\mathbf{C}_k^T \mathbf{C}_k) = 0. \quad (10)$$

By substituting Eqs. (8)–(10) to Eq. (7)

$$\sum_{k=1}^{N_c} \left( \frac{1}{N_k} \sum_{i=1}^{N_K} [2 \mathbf{M} (\hat{\mathbf{x}}_{ki} \hat{\mathbf{x}}_{ki}^T) - 2 \mathbf{C}_k \hat{\mathbf{x}}_{ki}^T] \right) = 0 \quad (11)$$

or

$$\mathbf{M} \sum_{k=1}^{N_c} \left( \frac{1}{N_k} \sum_{i=1}^{N_K} \hat{\mathbf{x}}_{ki} \hat{\mathbf{x}}_{ki}^T \right) - \sum_{k=1}^{N_c} \left( \frac{1}{N_k} \sum_{i=1}^{N_K} \mathbf{C}_k \hat{\mathbf{x}}_{ki}^T \right) = 0. \quad (12)$$

Solving Eq. (12) matrix  $\mathbf{M}$  is obtained:

$$\mathbf{M} = \left[ \sum_{k=1}^{N_c} \left( \frac{1}{N_k} \sum_{i=1}^{N_K} \mathbf{C}_k \hat{\mathbf{x}}_{ki}^T \right) \right] \left[ \sum_{k=1}^{N_c} \left( \frac{1}{N_k} \sum_{i=1}^{N_K} \hat{\mathbf{x}}_{ki} \hat{\mathbf{x}}_{ki}^T \right) \right]^{-1}. \quad (13)$$

The dimensions of mapping matrix  $M$  are  $N_c \times \hat{d}$ .

Pattern vectors  $\hat{\mathbf{x}}_{ki}$  may now be transformed from the augmented feature space (of dimensionality  $\hat{d}$ ) to the decision space (of dimensionality equal to the number of classes,  $N_c$ ) by the transformation

$$\hat{\mathbf{x}}_{ki} \mapsto \mathbf{M} \hat{\mathbf{x}}_{ki}. \quad (14)$$

Transformations for increasing classification accuracy are also employed by other recently developed algorithms, such the support vector machines classifier (Kecman, 2001). In the present work, the transformed patterns in (14) are then fed into the PNN classifier.

The PNN (Specht, 1990) are implemented by a feed-forward and one-pass structure and encapsulate the Bayes' decision rule together with the use of Parzen estimators of data's probability distribution function. The discriminant function of a PNN classifier for class  $k$  is given by the following equation:

$$g_k(\mathbf{x}) = \frac{1}{(2\pi)^{d/2} \sigma^d N_k} \times \sum_{i=1}^{N_K} \exp \left( - \frac{(\mathbf{x} - \mathbf{x}_{ki})^T (\mathbf{x} - \mathbf{x}_{ki})}{2\sigma^2} \right), \quad (15)$$

where  $\sigma$  is a smoothing parameter.

Following the LSM<sup>3</sup> step, in Eq. (15) the pattern vectors  $\mathbf{x}$  and  $\mathbf{x}_{ki}$  must be substituted by the cubic-augmented and least-squares-mapped vectors  $\mathbf{M}\hat{\mathbf{x}}$  and  $\mathbf{M}\hat{\mathbf{x}}_{ki}$  correspondingly (transformation (14)). In addition, because the dimensionality of LSM<sup>3</sup> vectors is equal to the number of classes

( $N_c$ ),  $d$  must be substituted by  $N_c$ . Thus, the discriminant function of the PNN–LSM<sup>3</sup> classifier for class  $k$  is given by

$$g_k(\mathbf{x}) = \frac{1}{(2\pi)^{N_c/2} \sigma^{N_c} N_k} \times \sum_{i=1}^{N_k} \exp\left(-\frac{(\mathbf{M}\hat{\mathbf{x}} - \mathbf{M}\hat{\mathbf{x}}_{ki})^T (\mathbf{M}\hat{\mathbf{x}} - \mathbf{M}\hat{\mathbf{x}}_{ki})}{2\sigma^2}\right). \quad (16)$$

The test pattern  $\mathbf{x}$  is classified to the class with the larger discriminant function value.

#### 2.4. System design

System design requires the determination of the best features-combination for achieving highest classification accuracy. Ideally, all textural features should be employed in the design of the classifiers, but, since a number of them may be redundant due to high mutual correlation (Theodoridis and Koutroumbas, 1999) and in order to reduce computational complexity, an optimum number of them had to be selected so as to achieve highest classification accuracy with the least number of features. One method is the exhaustive (Theodoridis and Koutroumbas, 1999) which involves designing each 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. However, considering the large number of possible combinations of the 39 textural features, the exhaustive method proved too time consuming to be adopted, thus a compromise between accuracy and computational speed had to be reached. Each one of the 39 features was tested by employing the Student's  $t$ -test; only 14 features with the highest class discrimination power ( $p < 0.001$ ) were selected and by applying the exhaustive search method on them, optimum design parameters of the classifiers were determined (highest precision with smallest number of features). These optimal features were then employed in the final design of the classifiers.

#### 2.5. System performance evaluation

System evaluation was performed by means of the leave-one-out (LOO) method (Theodoridis and Koutroumbas, 1999). Each classifier was designed by all but one carotid plaque US image, which was then classified. The process was repeated, each time leaving a different plaque-image out, until all data have been processed. In this way, each classifier was evaluated by data that were not involved in its design. It is evident, however, that the classifiers had to be re-designed each time a plaque-image was left out (i.e. application of the exhaustive method to search for optimum features). This huge workload required few hours of processing time. In the end, the best operating parameters were determined for each classifier.

### 3. Results and discussion

Table 1 shows the classification accuracy achieved by the PNN classifier in discriminating high-risk from low-risk plaques, using three textural features and employing the LOO method. The classifier distinguished correctly 23 high-risk plaques and incorrectly assigned one plaque to the low-risk group, thus giving a 95.8% precision in characterizing plaques with high-risk of causing embolization. In the case of the low-risk carotid plaques, 29 were assigned to the correct class while only three was wrongly classified to the high-risk group, scoring a 90.6% group discrimination accuracy. Overall, the PNN achieved 92.9% precision in distinguishing correctly the low-risk from

Table 1

Best classification accuracy achieved by the PNN classifier employing the ratio dense/lucent pixels, the correlation, and the short run-length emphasis best feature combination

PNN	High-risk (echolucent)	Low-risk (echodense)	Accuracy (%)
High-risk (echolucent)	23	1	95.8
Low-risk (echodense)	3	29	90.6
Overall accuracy			92.9

the high-risk carotid plaques. This is the highest classification achieved by the PNN with the minimum number of textural features and with the best tuning of parameters  $\sigma = 0.25$  (see relation (16)), determined by trial and error. Fig. 3 is a three dimensional scatter diagram of the three textural features and the decision surface drawn by the PNN classifier. The non-linear nature of the decision boundary could not separate all data since three echogenic “low-risk” plaques and one echolucent “high-risk” plaque were situated well within the opposite regions. As shown in Fig. 4, the classification accuracy of the PNN varied with the number of features scoring a low of 83.9% for one feature, 91.1% for two features, and reaching a maximum of 92.9% for three features. Thereafter, the classification accuracy of the PNN stabilised for up to 5-feature combinations.

The performance of the PNN–LSM<sup>3</sup> classifier is shown in Table 2 and in Figs. 4 and 5. As it can be observed from Fig. 4, the classification accuracy of the PNN–LSM<sup>3</sup> employing LOO increases with the number of features, reaching highest precision of 94.6%, 96.4%, and 100% using one (short run-length emphasis), two (short run-length emphasis and D/L), and three features (short run-length emphasis, D/L, and correlation), respectively. Thereafter, the classification accuracy of the classifier stabilised to 100% for higher number

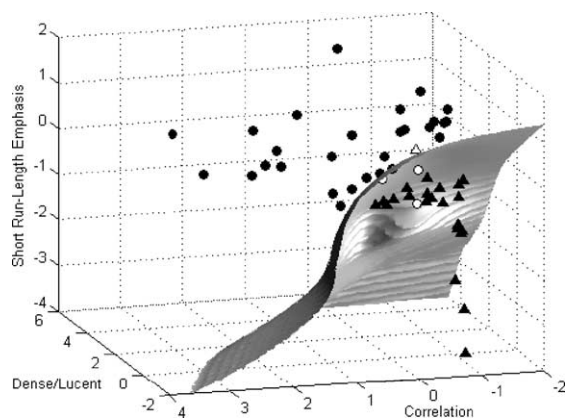


Fig. 3. Three dimensional scatter diagram and decision surface of the PNN classifier employing optimum 3-feature combination for discriminating high-risk (▲) from low-risk (●) carotid plaque lesions (misclassified patterns: △ and ○).

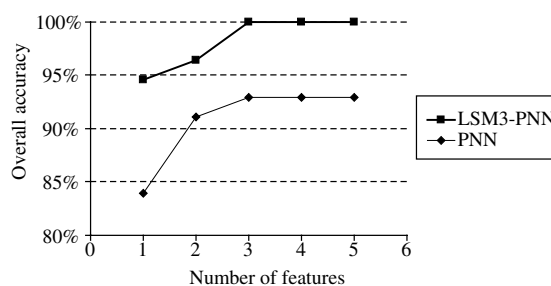


Fig. 4. Classification accuracy variation with increasing number of features employed for both the PNN and the PNN–LSM<sup>3</sup> classifiers.

Table 2  
Best classification accuracy achieved by the PNN–LSM<sup>3</sup> classifier employing the ratio dense/lucent pixels, the correlation, and the short run-length emphasis best feature combination

PNN–LSM <sup>3</sup>	High-risk (echolucent)	Low-risk (echodense)	Accuracy (%)
High-risk (echolucent)	24	0	100
Low-risk (echodense)	0	32	100
Overall accuracy			100

of features. A better understanding in the potentiality of the classifier may be assessed from Fig. 5a–d, where the scatter diagrams of best classifier performances using 1-, 2-, 3-, and 4-feature combinations (D/L, correlation, entropy, and short run-length emphasis) are displayed. The first thing worth noting is that the dimensions of the scatter diagrams are independent of the number of features involved and depend solely on the number of classes. This is achieved by means of the LSM process. Accordingly, data are gathered in a least-squares manner around two arbitrary pre-selected points corresponding to the two classes ( $C_1 = [1 \ 0]^T$  for echolucent and  $C_2 = [0 \ 1]^T$  for echogenic plaques, see vectors  $C_k$  in relation (3)). Another important observation is that as the classes are dragged further apart in the multidimensional features space by increasing the number of features involved, that separation is also clearly depicted on the two dimensional decision space. This gradual transformation from

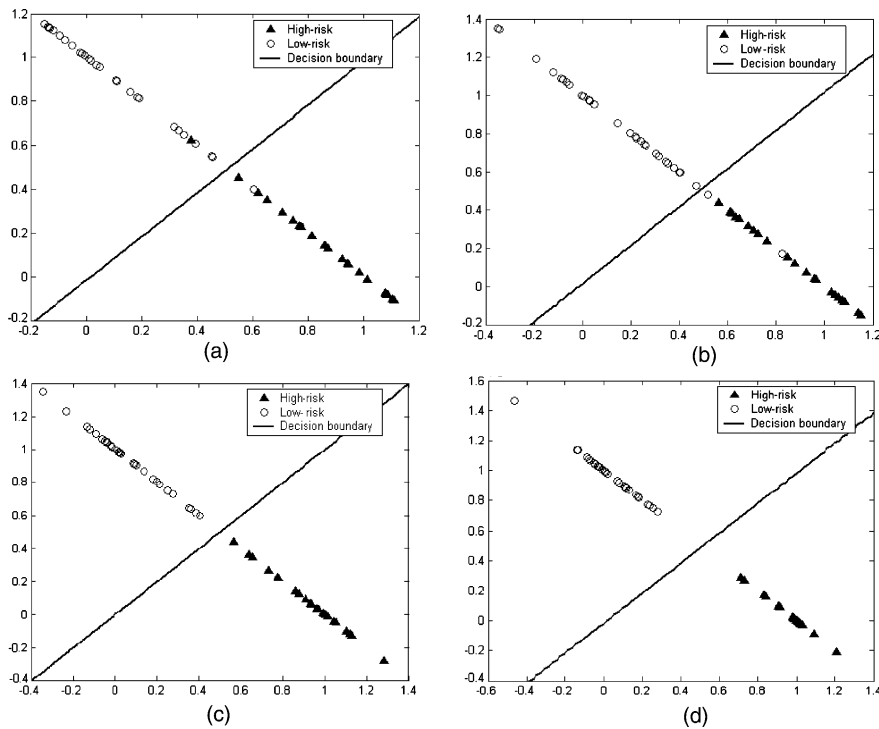


Fig. 5. Scatter diagrams displaying class separation and PNN–LSM<sup>3</sup> classifier operation as the number of features increases from one (a) to four features (d).

non-separable into separable classes by increasing the number of features from 1 to 4 may be clearly observed in Fig. 5a–d, respectively. It is therefore evident that in this decision space of separable classes the classifier can operate with high precision.

Another important factor concerning the performance of the PNN–LSM<sup>3</sup> classifier is the degree of the LSM process. Fig. 6 shows scatter diagrams employing the best 4-feature combination (D/L, correlation, entropy, and short run-length emphasis). In Fig. 6a, the LSM process includes the first degree terms only ( $x_i, i = 1, 2, 3, 4$ ). It is evident that there exists class overlapping and the classifier gives an 83.9% precision. In Fig. 6b the LSM process includes the first and second degree terms ( $x_i, x_i^2, x_i x_j, i, j = 1, 2, 3, 4, i \neq j$ ), class overlapping is less, and classification accuracy increases to 92.2%. In Fig. 6c, the LSM process comprises the first, second, and third degree terms ( $x_i, x_i^2, x_i x_j, x_i^3, x_i^2 x_j, x_i x_j x_k, i, j, k = 1, 2, 3, 4, i \neq j \neq k$ ),

resulting in complete class separation and thus rendering highest classifier performance (100%) possible.

For comparison reasons, the proposed PNN–LSM<sup>3</sup> algorithm was tested against the statistical Bayesian classifier. The latter achieved an overall highest accuracy of 85.7%, employing the LOO method and a 4-feature combination.

Regarding the best feature combination giving highest precision, it is interesting to note that the ratio D/L, the correlation, and the short run-length emphasis are meaningful textural parameters related to plaque structure. The ratio D/L shows the prevalence of dense or lucent pixels within the plaque, correlation is a measure of gray-tone linear-dependencies in the plaque image, and short run-length emphasis indicates the existence of small structures in the plaque. In a previous work (Arnold et al., 2001), textural features from the co-occurrence matrix were employed on carotid plaque US images in analyzing

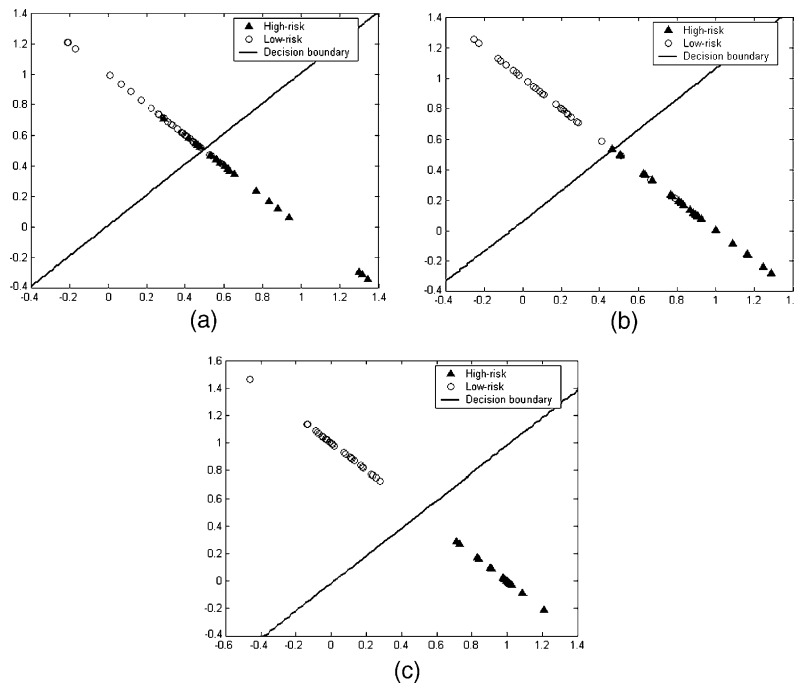


Fig. 6. Scatter diagrams demonstrating class separation and PNN-LSM classifier performance improvement for the first (PNN-LSM<sup>1</sup>), second (PNN-LSM<sup>2</sup>), and third (PNN-LSM<sup>3</sup>) degree LSM process, employing best 4-feature combination (D/L, correlation, entropy, and short run-length emphasis).

plaque composition by multiple discriminant analysis of variance with classification accuracy ranging between 68% for haemorrhage and 100% for calcium. In comparison, our work showed high classification accuracy in distinguishing high-risk (high lipid and/or haemorrhagic content) from low-risk plaques (mostly calcium and/or fibrous tissue). This was achieved by using image processing and powerful pattern recognition methods.

In conclusion, the risk embolization associated with carotid plaques has been assessed on US images by means of image analysis and pattern recognition methods. Textural features related to the plaque's structure have been determined, providing highest classification accuracy. An efficient classification method was proposed, based on the cubic LSM and on the PNN classifier, which was found to outperform the PNN classifier. Finally, the proposed system may be of value to patient management as a second opinion tool, but it

should be tested on more data in a clinical environment.

## References

- Ahmed, N., Rao, R., 1975. Orthogonal Transforms for Digital Signal Processing. Springer-Verlag, Berlin.
- Arnold, A., Taylor, P., Poston, R., Modaresi, K., Padayachee, S., 2001. An objective method for grading ultrasound images of carotid artery plaques. *Ultrasound Med. Biol.* 27, 1041–1047.
- Biasi, G.M., Sampaolo, A., Mingazzini, P., De Amicis, P., El-Barghouty, N.M., Nicolaides, A., 1999. Computer analysis of ultrasonic plaque echolucency in identifying high risk carotid bifurcation lesions. *Eur. J. Vasc. Endovasc. Surg.* 17, 476–479.
- Cheng, S.W.K., Ting, A.C.W., Wu, L.L.H., 2002. Ultrasonic analysis of plaque characteristics and intimal-medial thickness in radiation-induced atherosclerotic carotid arteries. *Eur. J. Vasc. Endovasc. Surg.* 24, 499–504.
- Denzel, C., Fellner, F., Wutke, R., Bazler, K., Müller, K.M., Lang, W., 2003. Ultrasonographic analysis of arteriosclerotic

- plaques in the internal carotid artery. *Eur. J. Ultrasound* 16, 161–167.
- El-Barghouty, N., Levine, T., Ladva, S., Flanagan, A., Nicolaides, A., 1996. Histological verification of computerised carotid plaque characterisation. *Eur. J. Vasc. Endovasc. Surg.* 11, 414–416.
- Galloway, M.M., 1975. Texture analysis using gray level run lengths. *Comput. Graphics Image Process.* 4, 172–179.
- Haralick, R.M., Shanmugam, K., Dinstein, I., 1973. Textural features for image analysis. *IEEE Trans. Syst. Man. Cybern.* 3, 610–621.
- Hatsukami, T.S., Thackray, B.D., Primozich, J.F., Ferguson, M.S., Burns, D.H., Beach, K.W., Detmer, P.R., Alpers, C., Gordon, D., Strandness Jr., D.E., 1994. Echolucent regions in the carotid artery: preliminary analysis comparing three dimensional histologic reconstructions to sonographic findings. *Ultrasound Med. Biol.* 20, 743–749.
- Keeman, V., 2001. *Learning and Soft Computing, Support Vector Machines, Neural Networks, and Fuzzy Logic Models*. MIT Press, Cambridge, MA.
- Montauban van Swijndregt, A.D., Elbers, H.R.J., Moll, F.L., De Letter, J., Ackerstaff, R.G.A., 1998. Ultrasonographic characterization of carotid plaques. *Ultrasound Med. Biol.* 24, 489–493.
- Noritomi, T., Sigel, B., Swami, V., Justin, J., Gahtan, V., Chen, X., Feleppa, E.J., Roberts, A.B., Shirouzu, K., 1997. Carotid plaque typing by multiple-parameter ultrasonic tissue characterization. *Ultrasound Med. Biol.* 23, 643–650.
- Pedro, L.M., Fernandes e Fernandes, J., Pedro, M.M., Gonçalves, I., Dias, N.V., Fernandes e Fernandes, R., Carneiro, T.F., Balsinha, C., 2002. Ultrasonographic risk score of carotid plaques. *Eur. J. Vasc. Endovasc. Surg.* 24, 492–498.
- Rakebrandt, F., Crawford, D.C., Havard, D., Coleman, D., Woodcock, J.P., 2000. Relationship between ultrasound texture classification images and histology of atherosclerotic plaque. *Ultrasound Med. Biol.* 26, 1393–1402.
- Specht, D.F., 1990. Probabilistic neural networks. *Neural Networks* 3, 109–118.
- Tegos, T.J., Sabetai, M.M., Nicolaides, A.N., Elatrozy, T.S., Dhanjil, S., Stevens, J.M., 2001a. Patterns of brain computed tomography infarction and carotid plaque echogenicity. *J. Vasc. Surg.* 33, 334–339.
- Tegos, T.J., Stavropoulos, P., Sabetai, M.M., Khodabakhsh, P., Sassano, A., Nicolaides, A.N., 2001b. Determinants of carotid plaque instability: echoicity versus heterogeneity. *Eur. J. Vasc. Endovasc. Surg.* 22, 22–30.
- Theodoridis, S., Koutroumbas, K., 1999. *Pattern Recognition*. Academic Press, New York.
- Wilhjelm, J.E., Grønholdt, M.-L.M., Wiebe, B., Jespersen, S.K., Hansen, L.K., Sillesen, H., 1998. Quantitative analysis of ultrasound B-mode images of carotid atherosclerotic plaque: correlation with visual classification and histological examination. *IEEE Trans. Med. Imag.* 17, 910–922.
- Zhang, X., McKay, C.R., Sonka, M., 1998. Tissue characterization in intravascular ultrasound images. *IEEE Trans. Med. Imag.* 17, 889–899.