Ultrasound Profile of Carotid Plaque: A New Approach Towards Stroke Prediction

14

José Seabra, L.M. Pedro, J. Fernandes e Fernandes, and J. Miguel Sanches

1 Introduction

Carotid artery atherosclerosis is an important cause of death and disability due to stroke. Among patients with carotid plaques, only a few show warning events, whereas the majority present cerebral events associated with previous asymptomatic plaques.

Numerous studies reported the importance of the degree of stenosis as an indicator of stroke in both symptomatic and asymptomatic groups [1, 2]. Indeed, disease severity and selection of patients for surgery is based on previous occurrence of clinical symptoms, such as stroke, transient ischemic attack (TIA), amaurosis fugax (AF), and the degree of stenosis presented by the plaque.

Moreover, it has been shown [1,2] that surgery associated with a degree of stenosis of more than 70 % resulted in an absolute reduction of 17 % in the risk of ipsilateral stroke after 2 years and 11.6 % at 3 years. These observations suggest that not all severe stenotic plaques are harmful; in fact, as reported by Inzitari et al. [3], the majority of asymptomatic high-grade stenotic plaques remain asymptomatic. Moreover, a study performed by Polak et al. [4] reported evidence that atheromatous plaques with relatively low degree of stenosis may produce symptoms as well. In addition, endarterectomy carries a non-negligible risk for the patient as well as significant financial costs for the patient, hospital, and health system in general. Consequently, an optimized characterization and identification of symptomatic lesions must be carried out to objectively select patients which should be treated with surgery among those to whom a correct medication represent a better solution.

While preliminary studies [5, 6] were based on a subjective evaluation of the plaque appearance to interpret the lesion severity, nowadays advanced methods of image processing allow the extraction of a large number of features from B-mode ultrasound (BUS) images of the carotid plaque. Furthermore, specialized techniques can be employed to identify a subset of salient features, which may be used as input to a classification system. The use of feature selection methods significantly simplifies the classification task, which will be faster and use less memory, while usually achieving a high classification performance. The evolution of artificial intelligence methods in conjunction with optimized computer performance has allowed the development of computer-aided diagnosis (CAD) systems [7]. Such systems are expected to help physicians on supporting the evaluation of pathologic findings during the diagnostic procedure.

1.1 Background

Different techniques relying on both qualitative and quantitative assessment of carotid plaque echo-morphology can be found in the literature, although no single technique has emerged as the method of choice.

As suggested by histopathological studies, other factors besides stenosis, including plaque structure and echomorphology (information on plaque gray-scale intensities) have shown to be associated with neurological symptoms [8–10]. Echogenic material reflects strongly the ultrasound signal and comprises fibrous tissue and calcium deposits, whereas echolucent material has less reflecting ability and includes blood and lipids. As previously referred, echolucent plaques are more likely to lead to development of neurologic events than echogenic ones.

El-Barghouty et al. [8] in a study with 94 plaques has provided a characterization of plaques based on the

J. Seabra • J.M. Sanches (✉)
Instituto de Sistemas e Robótica, Instituto Superior Técnico, Torre Norte, 6° Piso, Av. Rovisco Pais, 1049-001 Lisboa, Portugal
e-mail: jseabra@ist.ist.utl.pt; jmrs@ist.utl.pt

L.M. Pedro • J. Fernandes e Fernandes
Cardiovascular Institute and the Lisbon University Medical School, Hospital de Santa Maria, 1649-035 Lisbon, Portugal
e-mail: lmendespedro@clix.pt; fernandes456@hotmail.it

L. Saba et al. (eds.), *Multi-Modality Atherosclerosis Imaging and Diagnosis*,
DOI 10.1007/978-1-4614-7425-8_14, © Springer Science+Business Media, LLC 2014

gray-scale median (GSM), reporting an association between echolucency (GSM ≤ 32) and the incidence of cerebral computed tomography brain infractions. Subsequently, the study conducted by Iannuzzi et al. [11] in 242 stroke and 336 transient ischemic attack patients showed that the features more consistently associated with cerebrovascular events were low echogenicity, thicker plaques, and presence of longitudinal motion.

Then, Wilhjelm et al. [12] in a study with 52 patients scheduled for surgery, presented a comparison between subjective visual classification of the plaque appearance, first and second-order statistical features, and histological analysis of the excised plaques. Some correlation was found between the three types of information, where the best performing feature was found to be the histogram contrast.

From thereon, various studies have emerged, either using exclusively or combining texture information, based on first and second order statistics [13], Fourier power spectrum [13], fractal properties [14], and Law's texture energy [7]. Particularly, Christodoulou et al. used statistical, model-based, and Fourier-based methods as well as a combination of classifiers and found an average diagnostic performance (accuracy) of 73.1 % in identifying symptomatic lesions, using a dataset composed of 230 plaque images.

The comprehensive study conducted by Pedro et al. [15] in 215 carotid plaques, combined quantitative (e.g., the degree of stenosis and histogram features) and qualitative information resulting from visual inspecting the plaques on BUS images for developing an ultrasound score. This score, designated as Activity Index (AI) provides promising results in identifying plaques with a high likelihood of developing symptoms despite the significant number of false positive samples.

Moreover, the work developed by Mougiakakou et al. [7] extended previous studies on the characterization of carotid plaques from BUS images, by systematically studying all available first-order statistics, as well as Law's texture energy features. In that study, a CAD system is used to support diagnosis based on a neural network trained via a combination of back propagation with a genetic algorithm. This study produced promising results in identifying atheromatous lesions at high risk of stroke in a population of 108 plaques, thus suggesting the use of CAD systems as valuable tools in modern clinical practice. A rather interesting study performed by Kyriacou et al. [16] in a population of 137 asymptomatic and 137 symptomatic plaques proposed a multilevel binary and gray-scale morphological analysis method that have strong connections to prior clinical studies on what constitutes an unstable, symptomatic plaque. The multilevel approach is used to decompose the BUS image in its low-image, middle-image, and hi-image parts corresponding to hypoechoic, isoechoic, and hyperechoic image components, as originally proposed by AbuRahma [17]. The power spectra was computed from such images, showing significant differences between the symptomatic and asymptomatic spectra. Moreover, the derived pattern spectra were used as classification features with two different classifiers, the Probabilistic Neural Network (PNN) and the Support Vector Machine (SVM) and, as noted by the authors, the low-images alone provide better results than complicated multifeature classifier systems ($\approx$74 % versus 73 % [13]). An elegant explanation of why the pattern spectra for low-images performed better is that such images capture the (more echolucent) lipid components and there is clinical evidence that unstable plaques have large lipid components.

Furthermore, the importance of speckle in BUS images as well as its statistical modeling for tissue characterization has been previously documented [18]. In this chapter, we investigate the usefulness of a recently proposed de-speckling algorithm [19] which is able to decompose an ultrasound image into its noiseless and speckle components for feature extraction and, consequently, for tissue characterization. It is expected that such echo-morphology and texture parameters obtained from these image sources could contribute to a better analysis of the symptomatic plaque and differentiation from the asymptomatic lesion.

Here, it is argued that an optimal method for identifying vulnerable lesions should attempt to include not only morphological and textural features extracted from pixel intensity information within the plaque but also diagnostic information regarding plaque structure and appearance (e.g., stenosis, evidence of surface disruption, and presence of echogenic cap), interpreted and given by experienced physicians. The combination of this information is expected to produce a more comprehensive description of the profile of an active plaque, potentially providing the identification of lesions that would developed symptoms in the future.

1.2 Chapter Organization

This chapter is composed of two main parts, reflecting two different studies that were performed. The workflow of the current chapter is presented in Fig. 14.1.

Section 2 describes a cross-sectional study for characterizing the symptomatic plaque including a detailed descrip-

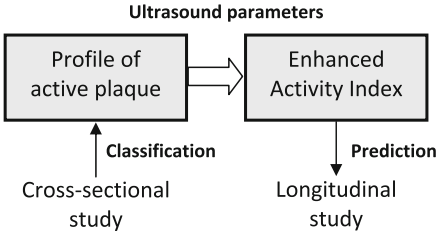

Fig. 14.1 Framework presented in the current chapter

tion of the data (Sect. 2.1) and methods (Sect. 2.2) used. Then, the results and main observations about the first study are drawn in Sect. 2.3.

The second study, presented in Sect. 3, involves the development and testing of a diagnostic measure to quantify plaque activity in a group of asymptomatic subjects. The data used in the longitudinal study is described in Sect. 3.1. The design of the plaque activity measure is described in Sect. 3.2 and experimental results are given in Sect. 3.3. Finally, Sect. 4 concludes the chapter.

2 Cross-Sectional Study

This study introduces a classification framework which enables to distinguish between symptomatic and asymptomatic lesions (Fig. 14.1). This method uses a collection of ultrasound image processing methods for feature extraction and tissue classification. In addition, it provides the identification of the most relevant parameters for plaque classification, consequently yielding an ultrasound profile of the "active," symptomatic plaque.

2.1 Data Management

This study included 221 carotid bifurcation plaques acquired from 99 patients, 75 males and 24 females. Mean age in this group of subjects was 68 years old (41–88). This data set was specifically assigned for training and testing the performance of a classification framework in separating symptomatic from asymptomatic lesions.

Patients were observed through neurological consultation at Instituto Cardiovascular and Hospital de Santa Maria, Lisbon, Portugal. A typical exam included a noninvasive examination with color-flow duplex scan of one or both carotids, performed with ATL-HDI 3000 equipment (Philips Medical Systems, Bothell, WA, USA) using a L12-5 scan probe (5–12 MHz broadband linear-array transducer). A plaque was considered symptomatic when AF or focal transitory, reversible or established neurological symptoms in the ipsilateral carotid territory, was observed in the previous 6 months. From this data set, 70 plaques were symptomatic while the remaining 151 did not reveal symptoms. This study was based on ultrasound images of plaques acquired at a fixed time frame (cross-sectional study).

2.2 Methods: CAD System

The conceptual idea of the cross-sectional study relies on a computer-assisted diagnostic framework (Fig. 14.2) designed with the purpose of distinguishing between symptomatic and asymptomatic lesions and, consequently, providing an accurate description of the vulnerable plaque.

The CAD system is supported by an user-friendly graphical interface, developed in MATLAB (Version R2007b, The Mathworks, Natick, MA, USA). This program provides the physician with several functionalities, including image normalization, definition of plaque(s) contour(s), adding relevant patient information (e.g., age, clinical history, medication, risk factors) and about subjective plaque structural characteristics (e.g., degree of stenosis, evidence of surface disruption, presence of fibrous cap and echolucent areas). Physicians can easily give their clinical input through an application designed for this purpose. In addition, the CAD system incorporates a chain of image processing tasks, such as image normalization as well as estimation of the envelope image and de-speckling. Operations involving envelope RF (ERF) image retrieval and de-speckling are employed to create new sources of information used for plaque characterization. Finally, the designed CAD system supplies the calculation of a measure which indicates the risk of the plaque to developing symptoms formulated in two

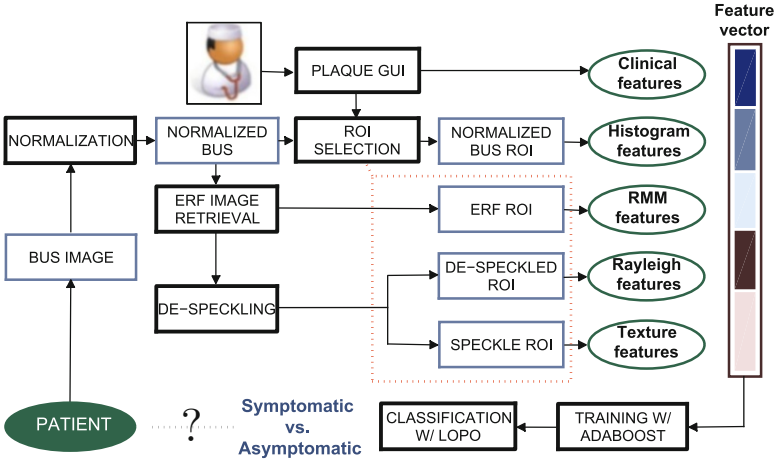

Fig. 14.2 Plaque classification framework

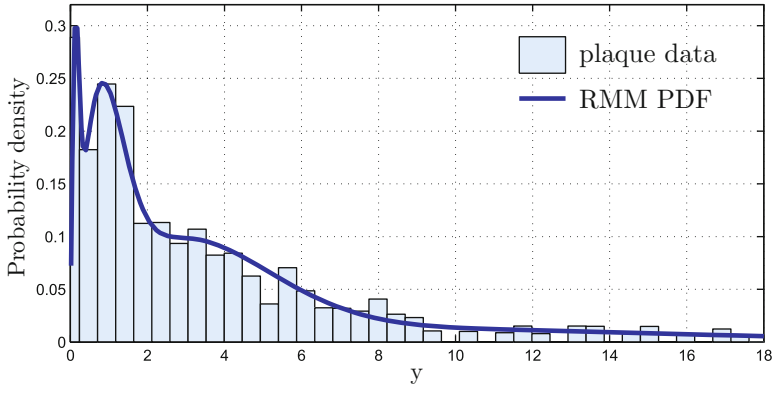

Fig. 14.3 RMM applied to plaque intensities in the ERF image: the RMM PDF is obtained as in (14.1) using the estimated *weights* and *Rayleigh* parameters

different ways: the *Activity Index*, early proposed by Pedro et al. [15] and the *Enhanced Activity Index* (EAI), which is detailed ahead in this chapter.

The main steps of the CAD system, supported by the described application, are explained next.

2.2.1 Image Processing

Image normalization is an important step to guarantee that images acquired under different conditions yield comparable and reproducible features and classification results. Image normalization was achieved as previously reported [9]; in particular, pixel intensities across the image were linearly scaled so that the adventitia and blood intensities would be in the range of 185–195 and 0–5, respectively (Fig. 14.3, top-left). This is an interactive procedure since the user must select two regions in the image, one corresponding to the adventitia (accounting for the most echogenic part) and the other to the blood (corresponding to the less echogenic component).

The normalized image is used to segment existing plaque(s) in the image. Each plaque is delineated by drawing around its structure and the obtained contour is a result of a two-step procedure: (a) contour interpolation according to a maximum distance (10 pixels) allowed between two consecutive points defined manually and (b) contour smoothing using basis functions (sinc functions).

De-speckled and speckle components of the image are required to compute echo-morphological and textural features. It is argued that the speckle-free (noiseless) component of the ultrasound image contains echogenic contents providing important insight on plaque morphology (and surrounding tissues). On the other hand, the speckle component, due to its multiplicative nature which makes it possible to dissociate it from the underlying anatomy, enables to better investigate the spatial relationship among pixels (texture) in the image. In a first step, an estimate of the envelope image (Fig. 14.3, top-right) is obtained from the normalized BUS image through the proposed decompression method. Subsequently, the ERF image is used to compute speckle-free and speckle components, displayed in Fig. 14.3, bottom-left and Fig. 14.3, bottom-right.

2.2.2 Feature Extraction

Features used for training the plaque classifier comprise subjective input given by the physician together with information automatically extracted from the normalized BUS, ERF, de-speckled and speckle images. As a consequence, features used for the purpose of plaque characterization include:

- *BUS morphological features* given by a physician during BUS examination. The 4-element vector of morphological parameters include (a) *evidence of plaque disruption*, defined by an interruption in the echogenic surface of the plaque; (b) *presence of echogenic cap*, identified as an equivalent of a thick fibrous cap and characterized by an echogenic line over the visible structure of the plaque; (c) *the degree of stenosis*, quantified using cross-section area measurement combined with hemodynamic assessment; and (d) *plaque echo-structure appearance*, where uniform plaques are defined as homogeneous while plaques presenting significant areas of echolucency are defined as heterogeneous.
- *Histogram features* extracted from the histogram of normalized pixel intensities inside the plaque. A total of 13 histogram features is estimated, including the *mean gray value*, *median gray value*, *percentage of pixels with grey value lower than 40*, *standard deviation of gray values*, *kurtosis*, *skewness*, *energy*, *entropy*, *10-, 25-, 50-, 75-, and 90-percentiles*.
- *RMM features* consist of the parameters of a mixture of Rayleigh distribution (RMM), early proposed in [20], used to model plaque echo-morphology contents. The RMM method is applied on envelope data, whose pixel intensities approximately follow Rayleigh statistics. Given this, gray-scale intensities within the plaque are considered random variables described by the following mixture of K distributions:

$$p(y_i|\Psi) = \sum_{k=1}^{K} \theta_k \, p(y_i|\sigma_k), \tag{14.1}$$

where $p(y_i|\sigma_k)$ is the Rayleigh PDF. θ_k and σ_k are the weights and Rayleigh parameters of the mixture, respectively, which are estimated using the *Expectation–Maximization* method, $K = 6$ (see Fig. 14.3) and $\Psi = \{\theta_1, \ldots, \theta_k, \ \sigma_1, \ldots, \sigma_k\}$. Hence, a 13-element feature vector is obtained, consisting of 6 *mixture weights*, 6 *Rayleigh parameters*, and the *effective number of RMM components*, determined by the number of mixture components with nonzero weight.

- *Rayleigh features* consist of average theoretical estimators of the Rayleigh distribution, whose parameters are given by the pixels on the de-speckled which contains the plaque. The Rayleigh features include the *mean*, $\mu_\sigma = \overline{\sigma_{i,j}}\sqrt{\pi/2}$, *median*, $\upsilon_\sigma = \overline{\sigma_{i,j}}\sqrt{2\log(2)}$, *variance*, $\sigma_\sigma = \overline{\sigma_{i,j}}\sqrt{(4-\pi)/2}$ of *Rayleigh values* and *percentage of pixels with Rayleigh value lower than 40*, $\sigma_{PP40} = 100 - \exp(-(40^2/2\overline{\sigma_{i,j}{}^2}))$.
- *Texture features* involve the study of the spatial distribution of gray levels inside the plaque region extracted from the speckle image. These features are estimated from gray level cooccurrence matrices (GLCMs), autoregressive (AR) models, and wavelet models. GLCMs are constructed using the relative frequencies $P(i, j, d, \theta)$ with which two neighboring pixels with gray levels i and j at a given distance d and orientation θ occur on the image. The distances used are $d = \{1, 2, 3, 4\}$ pixels and the angles $\theta = \{0, 45, 90, 135\}°$, thus creating 16 different GLCMs. From each computed GLCM different statistics can be derived, namely the *Contrast*, *Correlation*, *Energy*, and *Homogeneity* thus producing a 64-element feature vector. *Contrast* measures the local variations in the GLCM, while *Correlation* gives the joint probability occurrence of the specified pixel pairs. The *Energy* provides the sum of squared elements in the GLCM and, finally, *Homogeneity* measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal. Furthermore, to investigate a possible relation between each pixel and its neighborhood, the AR model is used on the speckle image, $\mathbf{N} = \{\eta_{i,j}\}$. This model assumes $\eta_{i,j}$ to be a 2D random variable where each pixel depends on its causal neighbors according to [21]:

$$\eta_{i,j} = \sum_{n,m}^{p,q} a_{n,m}\eta_{i-m,j-n} + u_{i,j}, \tag{14.2}$$

where $a_{n,m}$ are the AR coefficients to be estimated and $u_{i,j}$ are the residues. Considering a 1st order model such that $(p, q) = (1, 1)$, we estimate 3 *AR coefficients*. Alternatively, plaque texture can be studied using multi-level 2D wavelet decomposition. This technique consists of using low and high pass filters onto the approximation coefficients at level l in order to obtain the approximation at level $l + 1$ and the details in three orientations (horizontal, vertical, and diagonal). Here, decomposition is made along $l = 4$ levels. For each level, the *percentage of energy* for the approximation E_A as well as horizontal E_H, vertical E_V, and diagonal E_D details is computed. Hence, a 13-element wavelet-based feature vector is obtained composed of 4 $(E_\mathrm{H}) + 4\,(E_\mathrm{V}) + 4\,(E_\mathrm{D}) + E_\mathrm{A}$.

Therefore, each plaque is described by a feature vector $\mathbf{x}$ of 4 (Clinical) + 13 (Histogram) + 13 (RMM) + 4 (Rayleigh) + 80 (Texture) = 114 features.

2.2.3 Classification

The aforementioned features which describe each plaque are used to train the AdaBoost (Adaptive Boosting) classifier [22]. The use of such classifier is motivated by the promising results achieved when classifying plaque components on IVUS images [23]. AdaBoost is a binary classifier which consists in designing a *strong* classifier by linearly combining a set of *weak* classifiers. At each round of the boosting algorithm, the classification error in classifying the training data set is minimized by selecting the best discriminative value of one feature in the vector $\mathbf{x}$. The classifier performance is assessed by means of the LOPO cross-validation technique, where the training set is built taking at each time all patients' data, except one, used for testing. Performance results are given in terms of Sensitivity: Sens = TP/(TP + FN), Specificity: Spec = TN/(TN + FP), Precision or PPV (*Positive Predictive Value*): Prec = TP/(TP + FP) and Accuracy: Acc = (TP + TN)/(TP + TN + FP + FN), where TP = True Positive, TN = True Negative, FP = False Positive and FN = False Negative. Hence, a good classifier for diagnostic purposes would present a high sensitivity, meaning that it would be able to detect most of the symptomatic lesions, and a high PPV, which indicates that few asymptomatic lesions were identified as symptomatic.

2.2.4 Feature Analysis

A considerable amount of features was collected after ultrasound image processing. Naturally, not all the features are important to accurately characterize the plaque status, whether it is symptomatic or not. Hence, at this point an attempt is made to identify the most relevant ultrasound parameters for this particular problem. Hypothesis testing is a common method of drawing inferences about one or more populations based on statistical evidences from population samples (features). Here, we want to investigate if the statistical properties of a given feature significantly differ from the symptomatic to the asymptomatic group. Different hypothesis tests, including the z-, t-, Kolmogorov–Smirnov, and Mann–Whitney U-tests, make different assumptions

about the distribution of the random variable (feature value) being sampled in the data. For example, the z-test and the t-test both assume that the data are independently sampled from a normal distribution. In this work, the Mann–Whitney U-test [24] was chosen because it was the one providing the most promising results. This method performs a two-sided rank sum test of the null hypothesis that feature values in symptomatic and asymptomatic populations are independent samples from identical continuous distributions with equal medians, against the alternative that they do not have equal medians. Moreover, the *p-value* is the probability of rejecting the null hypothesis assuming that the null hypothesis is true. Clinically significant features will have a *p-value* which is typically lower than 0.05 or 0.01. In this work, features were considered to be relevant for differentiating between symptomatic and asymptomatic groups when the $p\text{-}value < 0.05$.

2.3 Experimental Results

This section presents three types of results. First, a suitable feature set, which is statistically relevant for the plaque classification problem is investigated and identified. Secondly, AdaBoost is trained with different ultrasound feature sets in order to evaluate which feature source is more effective to distinguish between plaques with and without symptoms. Then, an overall comparison study between state-of-the-art classifiers (degree of stenosis and AI) and the proposed method is provided.

Before implementing AdaBoost, it is of crucial importance to investigate the best feature set to describe and identify symptomatology in carotid plaques. This will allow to draw some conclusions about the different sources of information employed for plaque classification.

The use of a Mann–Whitney (M–W) U hypothesis test, described in Sect. 2.2.4, enables to identify ultrasound parameters with statistical significance. Table 14.1 presents the parameters and corresponding sources and *p-values* of the so-called *best feature set*.

Table 14.1 Optimal ultrasound parameter set

Ultrasound parameter	Type
Degree of stenosis	Clinical
Plaque echo-structure appearance	Clinical
Evidence of plaque disruption	Clinical
Presence of echogenic cap	Clinical
Mean	Histogram
Skewness	Histogram
Percentile 10, 50	Histogram
4th; 5th; 6th Rayleigh parameters	Rayleigh mixture models
5th; 6th mixture components	Rayleigh mixture models
# mixture components	Rayleigh mixture models
Wavelet decomposition energy	Speckle
GLCM homogeneity	Speckle

A closer look at the 16-element feature set allows to verify that both subjective and image-based parameters are useful for plaque classification. In particular, features from different image sources, namely the normalized image, the envelope RF image, and speckle field are considered statistically relevant. This preliminary observation justifies the use of an ultrasound preprocessing set of operations since it enables to estimate useful parameters for plaque classification.

Furthermore it is interesting to study the classifier performance under different conditions. Hence, AdaBoost is trained with five different parameter sets, considering only morphological information ($F.1$), parameters used to estimate the AI ($F.2$), the total feature set ($F.3$), and a feature set composed of the most relevant features ($F.4$), summarized in Table 14.1. Moreover, a last feature set ($F.5$) is also considered, including again the best feature set except that now all parameters were computed from just one image source—the normalized image, thus discarding information contained on de-speckled and speckle images. After training, the diagnostic value of each classifier is tested on the validated database, according to the LOPO technique. Classification performance is shown in Fig. 14.4, while a detailed description is given in Table 14.2.

Several observations can be made: first, morphological features are important markers of symptomatology as justified by the significant accuracy and specificity values

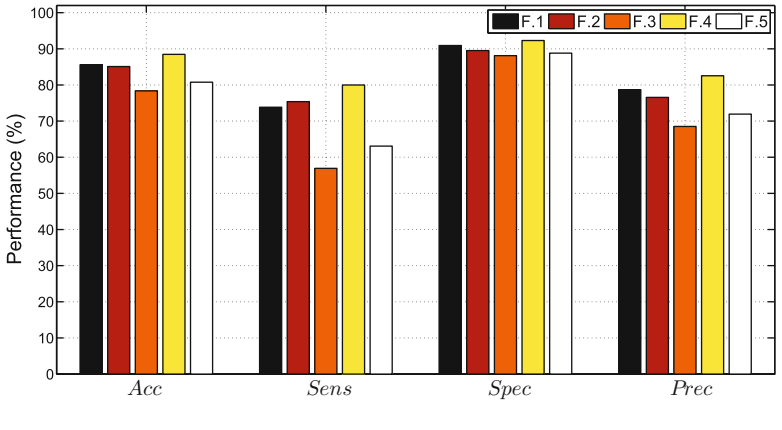

Fig. 14.4 Plaque classification with AdaBoost trained with different feature sets

Table 14.2 Detailed classification of results obtained with AdaBoost (results obtained with best feature set are highlighted)

LOPO (%)	$F.1$	$F.2$	$F.3$	$F.4$	$F.5$
Acc	85.57	85.10	78.37	**88.46**	80.77
Sens	73.85	75.38	56.92	**80.00**	63.08
Spec	90.91	89.51	88.11	**92.31**	88.81
Prec	78.69	76.56	68.52	**82.54**	71.93

obtained with $F.1$. However, it is important to bear in mind that this result is attained exclusively with subjective parameters, quantified and interpreted by physicians, which naturally know the patient clinical status. It is not quantifiable to what extent this a priori knowledge influences the estimation of morphological parameters. Secondly, morphological parameters are combined with histogram parameters, namely the GSM and P40, thus determining the feature set used in the AI method. Results obtained with $F.2$ are similar to $F.1$, except for sensitivity which is higher in $F.1$ meaning that it is able to detect more TP at the expense of getting more FP (lower precision).

Moreover, in order to investigate the usefulness of the proposed feature set we have trained AdaBoost with all the ultrasound parameters proposed in this work. Naturally, $F.3$ results in lower classifier performance and this can be mostly explained by the fact that some undesirable features are cluttering the classifier, which tweaks in favor of those features thus leading to poor classification performance. Finally, the collection of features which were found to be statistically relevant for this particular problem were used to train the studied classifier. Using $F.4$, all the performance criteria are significantly better than the reference classifier ($F.2$) up to 80 % sensitivity (improvement of 5 %) and 88.5 % accuracy (improvement of 3 %).

Hence, it was clearly identified a set of features, including morphology, echogenicity, and texture, which proved to be suitable for identifying symptomatic plaques among plaques presenting no symptoms. As it was previously detailed, such features were extracted after applying a set of processing operations, including envelope RF estimation and de-speckling.

In order to assess the usefulness of the aforementioned sources of information, a comparison is made between classification results when features are computed from different image sources ($F.4$), as proposed along this chapter, and when such features are exclusively obtained from normalized images ($F.5$). Results clearly show that the classification performance is substantially improved from $F.5$ to $F.4$ showing that it is preferable to use the mentioned RMM and textural features when these are extracted from their sources, envelope RF image, and speckle, respectively, rather than computing such features on the normalized image.

A third result, presented in Fig. 14.5 and Table 14.3, is designed to show a general perspective of the classification performances obtained with different approaches. Hence, in this study we have included the gold-standard method, based on the degree of stenosis with a clinical meaningful cut-off of 80 %, together with a recent approach based on the AI score and, finally, the best classifier investigated so far throughout this chapter, that is, the AdaBoost method trained with the so-called *best feature set*.

By comparing the outcomes of each classifier, it is observed that AdaBoost trained with the estimated *best feature set* outperforms the other two approaches which are often referred in literature. In particular, specificity and precision values are significantly higher for AdaBoost with respect to the other classifiers which suggests that the number of detected FP is relatively small. As a consequence, the accuracy obtained in correct plaque classification is also high.

Furthermore, it should be interesting to perform a direct comparison between the effectiveness of the proposed classification method and other related work [13, 16]. Even if this is not possible because the ultrasound data used is not the same, the margin between the accuracy obtained with the AdaBoost method (≈88 %) and these studies (≈74 %) is large enough to argue that the proposed method indeed outperforms other related plaque classification approaches.

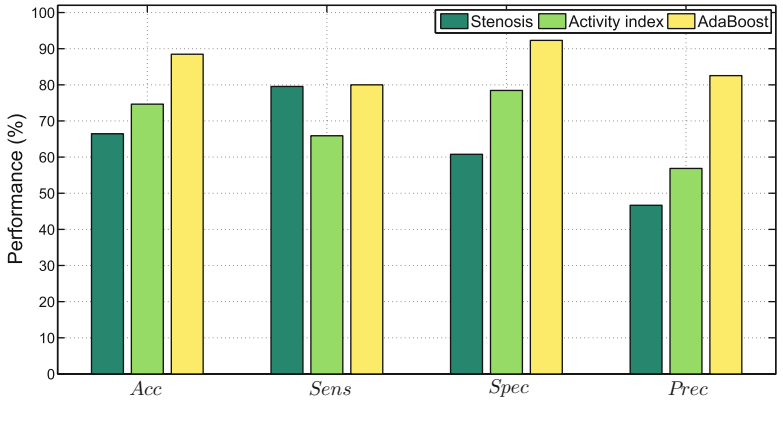

Fig. 14.5 Plaque classification using two state-of-the-art approaches together with proposed method

Table 14.3 Detailed classification results according to different feature sets and performance criteria

LOPO (%)	Degree of stenosis	AI	AdaBoost w/*best feature set*
Acc	66.44	74.66	**88.46**
Sens	79.55	65.91	**80.00**
Spec	60.78	78.43	**92.31**
Prec	46.67	56.86	**82.54**

3 Longitudinal/Natural History Study

A study which aims at classifying plaques with and without symptoms using both morphological- and image-based features has just been presented. However, despite its clinical significance, the CAD system as it has been described so far is not capable of identifying those asymptomatic lesions at high risk of becoming symptomatic. In fact, this kind of information would be more useful for physicians because they would be able to observe an asymptomatic lesion and quantitatively evaluate if such lesion is prone to develop symptoms.

As a consequence, the identification of a subset of "dangerous" or "active" plaques, featuring high neurological risk would help in the indication of treatment. Needless to say, this decision has important clinical and economical consequences for all the parts involved in this process.

As mentioned before, the absolute benefit of surgical intervention based on the degree of stenosis alone as a decision-making criterion is low in the asymptomatic disease and in symptomatic disease with moderate obstruction [1,2]. This clearly motivates the need for developing new strategies for plaque risk prediction.

In this study, a quantitative tool to evaluate plaque activity is proposed, designated by EAI. This method makes use of information gained during the cross-sectional study, particularly the estimated feature set which provides the best discrimination of symptomatic lesions among those that are harmless. Hence, the so-called *best feature set* represents an ultrasound *input* for an algorithm which aims at predicting the occurrence of symptoms in a longitudinal study conducted in a group of asymptomatic subjects (Fig. 14.1).

Again, the diagnostic power of the proposed EAI is compared to other strategies for identifying plaques at high risk, namely the one based on the degree of stenosis and the AI [15].

3.1 Data Management

This study presents a score that correlates with plaque activity and tests its diagnostic power on a group of 112 asymptomatic plaques, acquired from 112 patients. BUS images were collected from the ACSRS (*Asymptomatic Carotid Stenosis and Risk Study*) [25], consisting in a multicentre natural history study of patients with asymptomatic internal carotid diameter stenosis greater than 50 % in relation to the bulb. The degree of stenosis was graded using multiple established ultrasonic duplex criteria. The distribution of plaques according to the degree of stenosis was an average value of ≈75 % (50–99) and no. of plaques with degree of stenosis > 70 % = 80. Patients were followed for possible occurrence of symptoms for a mean time interval of 37.1 weeks. At the end of the study, 13 out of 112 patients (11.6 %) had developed symptoms (3 AF, 6 TIA, 4 stroke).

3.2 Methods: Enhanced Activity Index

A quantitative diagnostic measure—EAI—is developed considering the knowledge gathered in the cross-sectional study. Recall that the first study enabled to identify the "profile" of the "active" plaque by taking into account a set of parameters that are statistically relevant for separating symptomatic and asymptomatic lesions. Hence, the implementation of EAI is performed as follows:

1. The ultrasound "profile" of the "active" plaque is considered, by taking the features which are relevant for plaque classification, using the M–W U statistical test.
2. Reference values are taken for each ultrasound feature, f_i, and group, symptomatic (S), and asymptomatic (A), considering the mean $(\mu_i(S), \mu_i(A))$ and variance $(\sigma_i^2(S), \sigma_i^2(A))$.
3. The EAI*, renamed for convenience, is computed as the Bayes factor given by

$$\mathrm{EAI}^* = \frac{R_S}{R_A}, \tag{14.3}$$

where

$$R_k = \sum_i p(f_i|\omega_k) \approx \mathcal{N}(\mu_i(k), \sigma_i^2(k)), \quad k = \{S, A\} \tag{14.4}$$

are the marginal likelihoods of each group (S or A) and correspond to the sum of the conditional probabilities of each feature belonging to each group, respectively. Such conditional probabilities in (14.4) are computed assuming a normal distribution (Fig. 14.6). In (14.3), R_S and R_A represent the likelihoods of each plaque producing symptoms or stabilize, respectively. Hence, when EAI = 1, the result is inconclusive, while for EAI < 1 the plaque will stay harmless with a significant probability which is higher as EAI decreases. Contrarily, plaques showing an EAI > 1 are prone to produce symptoms, being more "dangerous" when EAI increases.

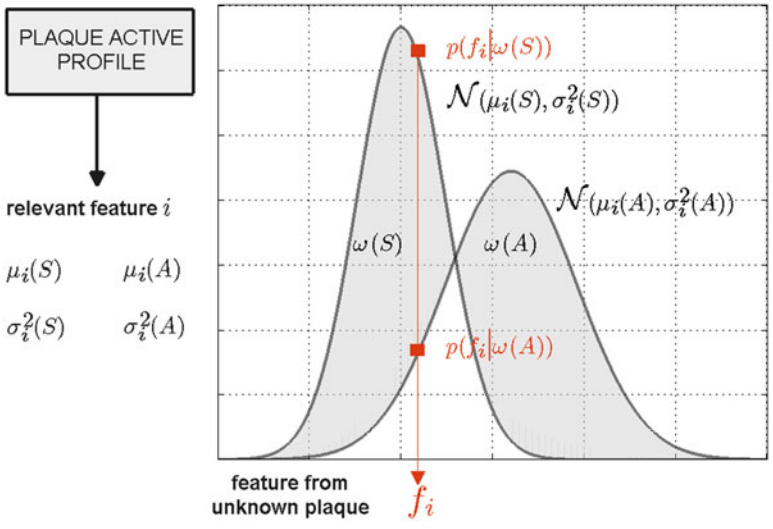

Fig. 14.6 Illustrative concept of conditional probabilities for a particular plaque feature f_i used to compute the EAI

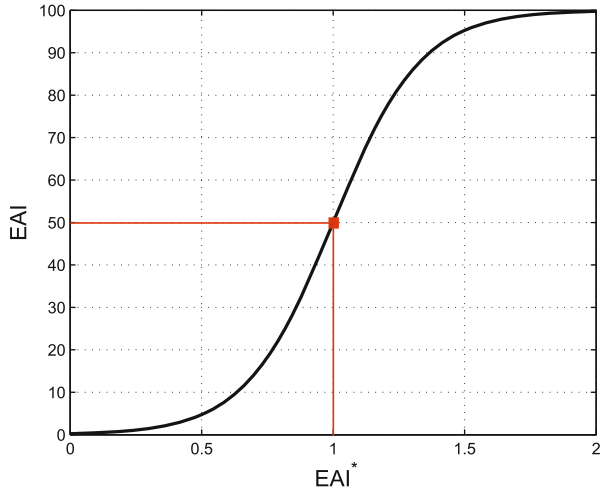

Fig. 14.7 Mapping function for re-scaling the EAI onto a 0–100 scale, where 100 represents maximum risk and 0 accounts for stability

4. The EAI is rescaled using a sigmoid mapping function which places the EAI onto a 0 100 scale. This function, sketched in Fig. 14.7 is defined as

$$\text{EAI} = \frac{100}{1 + \exp -(\text{EAI}^* - 1)}. \tag{14.5}$$

This mapping technique is useful to make the predictive power of the proposed EAI method comparable to AI [15] and degree of stenosis [2].

3.3 Experimental Results

The longitudinal study (Fig. 14.1) investigates the diagnostic power of EAI for identifying plaques at high risk of originating cerebrovascular events. To make this study feasible, the proposed method should be compared with other strategies of plaque risk prediction (e.g. degree of stenosis and AI).

Such comparison is here performed using ROC (*Receiver Operating Characteristic*) curve analysis [26]. In general, when considering the results of a particular test in two populations, one population with a disease, the other population without the disease, one rarely observes a perfect separation between the two groups. For every possible cut-off point or criterion value which one selects to discriminate between the two populations, there will be some cases with the disease correctly classified as positive (TPF = True Positive fraction) but some samples with the disease classified as negative (FNF = False Negative fraction). On the other hand, some cases without the disease will be correctly classified as negative (TNF = True Negative fraction) but some samples without the disease will be classified as positive (FPF = False Positive fraction). In a ROC curve the TPF (Sensitivity) is plotted as function of the FPF (100-Specificity) for different cut-off points, therefore each point on the ROC plot represents a sensitivity/specificity pair corresponding to a particular decision threshold. A test with perfect discrimination (no overlap in the two distributions) has a ROC plot that passes through the upper left corner (100 % TPF and 0 % FPF). Sometimes, the ROC is used to generate a summary statistic.

Moreover, the area under the ROC curve (ROC AUC) statistic is often used in machine learning for model comparison. This measure indicates that a predictive method is more accurate as higher is the ROC AUC. Similarly, it can be interpreted as the probability that when one randomly picks one positive and one negative example, the classifier will assign a higher score to the positive example than to the negative. In engineering, the area between the ROC curve and the no-discrimination line is also used. This area is often simply known as the discrimination. Moreover, the intersection of the ROC curve with the line at 90° to the no-

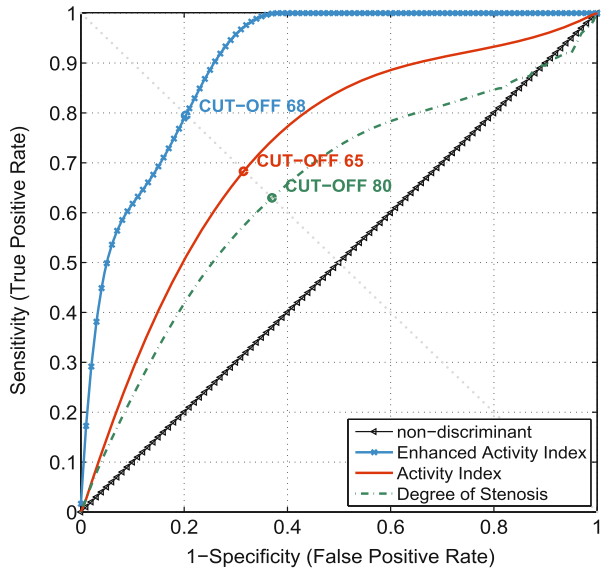

Fig. 14.8 ROC curves for different prediction methods

discrimination line is also considered as an heuristic method to investigate the cut-off providing the best discriminative power of the test (or predictive method).

Figure 14.8 presents the ROC curves obtained with the three studied predictive methods. The ROC AUCs are 0.6496 (64.96 %), 0.7329 (73.29 %), and 0.9057 (90.57 %) for degree of stenosis, AI and EAI, respectively. Naturally, the ROC AUC of the no-discrimination line is 0.5000 (50.00 %). These results markedly show that the EAI method is the most accurate method among the ones tested. Additionally, the AI method is also better than the degree of stenosis, as expected, because the former includes other parameters besides the stenosis degree. When computing the differences between the ROC AUCs and the no-discrimination line, one obtains 14.96 %, 23.29 %, and 40.57 % of discrimination for degree of stenosis, AI and EAI, respectively. This observation gives a clue about the amount of effective diagnostic information that is gained since the no-discrimination line corresponds to random guessing. The results reinforce the idea that the EAI method is the most discriminant among the investigated approaches.

According to the aforementioned heuristic method, the cut-off values providing the best trade-off between TP and FP rates for each predictive method are respectively 80, 65, and 68 for degree of stenosis, AI, and EAI (Fig. 14.8).

Note that the choice of a method to identify the best cut-off value is critical for the performance of the predictive method. The heuristic method that was presented (a diagonal line perpendicular to the nondiscriminant line, intersecting it at 0.5 FPR and 0.5 TPR) assigns equal importance to the detection of TP (TPR) and TN (1-FPR). In practice, we can argue that the relative importance of TPR and TNR should change according to each scenario. Hence, when the decision-making strategy intends to assign more relevance to the TPR, the cut-off line (or curve) should be shifted upper right while the opposite should happen when an increase importance is to be given to the TNR.

Figure 14.9 provides a different viewpoint about the diagnostic power of the proposed EAI prediction method when compared to the other methods. Particularly, this result allows the comparison of FP and FN samples according to different cut-offs applied for each studied method. Hence, the shorter the bars corresponding to the FP and FN are, the better is the cut-off or the predictive method, depending on if one is studying a particular method or comparing the three methods at the same time. To make an equivalence between the results shown in Fig. 14.9 and the ROC curves, it can be said that as the bars of FN and FP get smaller (hence, the TPR increases and the FPR decreases), the predictive method moves up and left, respectively.

It can be clearly observed that the application of the EAI method provides lower FP values when compared to the other methods regardless the cut-off chosen. Also evident is the fact that the prediction method based on stenosis is the one resulting in the highest number of FP, which is a natural observation because it is by far the simplest discriminative test used. Other observation that can be made from comparing Fig. 14.9b, c is that, generally, the number of FP is significantly lower for EAI when compared to AI.

In fact, Fig. 14.9 provides an objective interpretation of the trade-off between FP and FN. However, choosing an optimal cut-off is highly subjective. The reader should bear in mind that a method with a good diagnostic power should be naturally able to identify as much TP samples as possible, while providing a small number of FN and FP. In fact, the cut-off should be chosen according to a justifiable criterion: (a) is it more important to identify and treat all subjects that will develop a neurological complication even though a large number of patients must be operated? (b) should one be worried about sparing as many patients as we can from surgery?, or (c) should one decide on combining low FP and FN rates? If we pick the latter, the most suitable cut-off for degree of stenosis is 70 % while the best cut-offs for AI are 56 and 50, respectively. Additionally, it is worth to note that the application of EAI with the mentioned cut-off is able to identify all TP, in other words, is capable of predicting all plaques that developed symptoms.

Results presented in Fig. 14.9 are detailed in Table 14.4, providing different performance criteria for each method and cut-offs. For instance, note that the EAI method with cut-off 52 shows 100 % sensitivity and 30.95 % positive predictive value. This is, indeed, the most important result to outline since the EAI, with this particular cut-off, is able to identify all the plaques which will develop symptoms while detecting the smallest number of false positives.

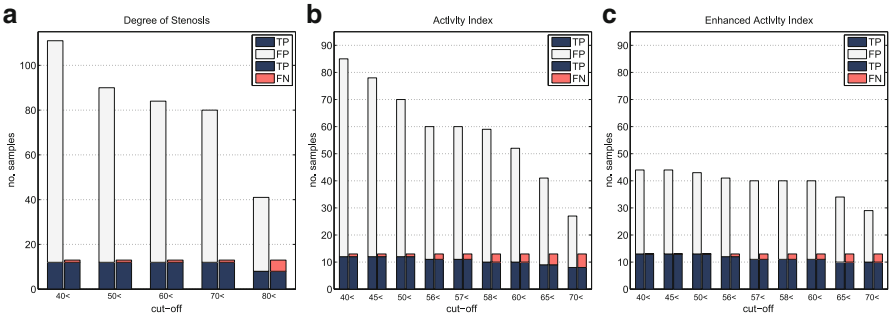

Fig. 14.9 Bar plot with TP, FP and FN values for different cut-offs, according to the stenosis predictor (**a**), AI (**b**) and EAI (**c**)

Table 14.4 Diagnostic (discriminative) power of the investigated methods for different cut-offs, according to performance criteria including sensitivity, specificity, accuracy, PPV and NPV

	Cut-off	Sens (%)	Spec (%)	Acc (%)	PPV (%)	NPV (%)
Stenosis	**>40**	92.31	0	10.71	10.81	0
	>50	92.31	21.21	29.46	13.33	95.46
	>60	92.31	27.27	34.82	14.29	96.43
	>70	***92.31***	***31.31***	***38.39***	***15.00***	***96.87***
	>80	61.54	66.67	66.07	19.51	92.96
AI	**>40**	92.31	26.26	33.93	14.12	96.30
	>45	92.31	33.33	40.18	15.36	97.06
	>50	***92.31***	***41.41***	***47.32***	***17.14***	***97.62***
	>56	84.62	50.50	54.46	18.33	96.15
	>60	76.92	57.58	59.82	19.23	95.00
	>65	69.23	67.68	67.86	21.95	94.37
	>70	61.54	80.81	78.57	29.63	94.12
EAI	**>40**	100	68.69	72.32	29.55	100
	>50	100	69.70	73.21	30.23	100
	>52	***100***	***70.71***	***74.10***	***30.95***	***100***
	>56	92.31	70.71	73.21	29.27	98.59
	>60	84.61	70.71	72.32	27.50	97.22
	>65	76.92	75.76	75.89	29.41	96.15
	>70	76.92	80.81	80.36	34.48	96.39

Results obtained for the best cut-off for each method (combination of best Sens and PPV) are highlighted

In predictive analysis, a table of confusion, also known as a confusion matrix, is a table with two rows and two columns that reports the number of TN, FP, FN, and TP. Table 14.5 summarizes the true predictive value of each method according to the aforementioned cut-offs, selected after comprehensive appreciation of Table 14.4.

As it is observed in Table 14.5 the EAI method was able to identify the 13 patients who had developed symptoms by the end of the follow-up (longitudinal) study, whereas the degree of stenosis and the AI methods were unable to identify, respectively, 1 and 2 patients who developed neurological complications later. Moreover, as far as the false positive number is concerned, the EAI method yields 29 FP against 49 and 68, respectively, for the AI and degree of stenosis. This means that if the decision of surgery for plaque removal was based in the former method, only 29 patients were unnecessarily operated. This number is significantly smaller than the one observed for the other methods, suggesting that the EAI is the most cost-effective method. Thus, the EAI method demonstrates to have the best diagnostic power among the methods investigated because it provides the most accurate selection of a subset of patients potentially at high risk within a population of asymptomatic patients.

Table 14.5 Confusion matrix with prediction outcome of the investigated methods: stenosis (underlined), AI (italics), and EAI (bold)

		Actual value					
		P			N		
Prediction	P′	12	*11*	**13**	68	*49*	**29**
	N′	1	*2*	**0**	31	*50*	**70**

4 Conclusions

Carotid plaques are the commonest source of neurological symptoms due to embolization or flow reduction.

Throughout this chapter it has been motivated the need for defining accurately the ultrasound "profile" of the "active" plaque, that is, the asymptomatic lesion with an increased likelihood of becoming symptomatic. This is of considerable importance because currently the treatment planning, based only on the patient's clinical history and degree of stenosis, is not optimal and cost-effective.

First, a cross-sectional study was performed for training and testing the Adaptive Boosting classifier using the LOPO cross-validation technique. This classifier consists of ultrasound parameters, accounting for morphology, echogenicity, and texture, extracted from different image sources, after application of a set of processing operations, described in previous chapters. A suitable statistical hypothesis test is applied in order to identify a subset of features which are statistically meaningful to discriminate between plaques with and without symptoms. An AdaBoost classifier based on the so-called *best feature set* outperforms other state-of-the-art methods, yielding an accuracy of 88 % and sensitivity of 80 % in identifying symptomatic plaques. Moreover, a comparative study of classifiers performance clearly suggests the usefulness of the preprocessing ultrasound methods, proposed throughout this thesis, as well as the value of mixture model and textural features for plaque classification.

Once a suitable ultrasound profile of the symptomatic or active plaque was established, an EAI that quantifies the degree of plaque activity or likelihood to rupture was proposed. This measure was evaluated on a longitudinal study of asymptomatic plaques and compared to other approaches (degree of stenosis and AI), demonstrating the best diagnostic power. In particular, EAI provides correct identification of all plaques that developed symptoms while giving the smallest number of false positives. This result suggests that the EAI could have a significant impact on stroke prediction and treatment planning.

References

1. North American Symptomatic Carotid Endarterectomy Trial Collaborators (1998) Benefit of carotid endarterectomy in patients with symptomatic moderate or severe stenosis. N Engl J Med 339(20):1445–1453
2. European Carotid Surgery Trialists Collaborative Group (1998) Randomised trial of endarterectomy for recently symptomatic carotid stenosis: final results of the MRC European Carotid Surgery Trial (ECST). Lancet 351:1379–1387
3. Inzitari D et al (2000) The causes and risk of stroke in patients with asymptomatic internal carotid artery stenosis. N Engl J Med 342(23):1693–1700
4. Polak JF et al (1998) Hypoechoic plaque at US of the carotid artery: an independent risk factor for incident stroke in adults aged 65 years or older. Cardiovascular Health Study. Radiology 208(3):649–654
5. Widder B, Paulat K, Hackspacher J, Hamann H, Hutschenreiter S, Kreutzer C, Ott F, Vollmar J (1990) Morphological characterization of carotid artery stenoses by ultrasound duplex scanning. Ultrasound Med Biol 16(4):349–354
6. Belcaro G, Nicolaides AN, Laurora G, Cesarone MR, Sanctis M, Incandela L, Barsotti A (1996) Ultrasound morphology classification of the arterial wall and cardiovascular events in a 6-year follow-up study. Arterioscler Thromb Vasc Biol 16:851–856
7. Mougiakakou S, Golemati S, Gousias I, Nicolaides AN, Nikita KS (2007) Computer-aided diagnosis of carotid atherosclerosis based on ultrasound image statistics, laws' texture and neural networks. Ultrasound Med Biol 33(1):26–36 .
8. El-Barghouty NM, Levine T, Ladva S, Flanagan A, Nicolaides A (1996) Histological verification of computerised carotid plaque characterisation. Eur J Vasc Endovasc Surg 11(4): 414–416
9. Elatrozy T, Nicolaides A, Tegos T, Griffin M (1998) The objective characterization of ultrasonic carotid plaque features. Eur J Vasc Endovasc Surg 16:223–230
10. Tegos TJ, Kalodiki E, Daskalopoulou SS, Nicolaides AN (2000) Stroke: epidemiology, clinical picture, and risk factors—part i of iii. Angiology 51(10):793–808
11. Iannuzzi A, Wilcosky T, Mercuri M, Rubba P, Bryan AF, Bond MG (1995) Ultrasonographic correlates of carotid atherosclerosis in transient ischemic attack and stroke. Stroke 26(4):614–619
12. Wilhjelm JE, Grnholdt ML, Wiebe B, Jespersen SK, Hansen LK, Sillesen H (1998) Quantitative analysis of ultrasound b-mode images of carotid atherosclerotic plaque: correlation with visual classification and histological examination. IEEE Trans Med Imaging 17(6):910–922
13. Christodoulou CI, Pattichis CS, Pantziaris M, Nicolaides A (2003) Texture-based classification of atherosclerotic carotid plaques. IEEE Trans Med Imaging 22(7):902–912
14. Asvestas P, Golemati S, Matsopoulos GK, Nikita KS, Nicolaides AN (2002) Fractal dimension estimation of carotid atherosclerotic plaques from b-mode ultrasound: a pilot study. Ultrasound Med Biol 28(9):1129–1136
15. Pedro LM, Fernandes JF, Pedro MM, Gonalves I, Dias NV (2002) Ultrasonographic risk score of carotid plaques. Eur J Vasc Endovasc Surg 24:492–498
16. Kyriacou E et al (2009) Classification of atherosclerotic carotid plaques using morphological analysis on ultrasound images. Appl Intell 30(1):3–23
17. AbuRahma AF, Thiele SP, Wulu JT Jr (2002) Prospective controlled study of the natural history of asymptomatic 60 % to 69 % carotid stenosis according to ultrasonic plaque morphology. J Vasc Surg 36(3):437–443
18. Thijssen J (2003) Ultrasonic speckle formation, analysis and processing applied to tissue characterization. Pattern Recognit Lett 24(4–5):659–675
19. Seabra J, Sanches J (2010) On estimating de-speckled and speckle components from B-mode ultrasound images. In: Proceedings of IEEE international symposium on biomedical imaging, Rotterdam, April 2010. IEEE Engineering in Medicine and Biology Society, pp 284–287
20. Seabra J, Sanches J, Ciompi F, Radeva P (2010) Ultrasonographic plaque characterization using a rayleigh mixture model. In: Proceedings of IEEE international symposium on biomedical imaging, Rotterdam, The Netherlands, April 2010. IEEE Engineering in Medicine and Biology Society, pp 1–4
21. Wear K, Wagner R, Garra B (1995) A comparison of autoregressive spectral estimation algorithms and order determination methods in ultrasonic tissue characterization. IEEE Trans Ultrason Ferroelectr Freq Control 42(4):709–716

22. Schapire RE (2002) The Boosting Approach to Machine Learning An Overview In MSRI Workshop on Nonlinear Estimation and Classification
23. Ciompi F, Pujol O, Gatta C, Rodriguez O, Mauri F, Radeva P (2009) Fusing in-vitro and in-vivo intravascular ultrasound data for plaque characterization. Int J Cardiovasc Imaging 26:763–779
24. Gibbons J, Chakraborti S (1992) Nonparametric statistical inference. CRC Press, Boca Raton
25. Nicolaides AN, Kakkos SK, Griffin M, Sabetai M, Dhanjil S et al (2005) Severity of asymptomatic carotid stenosis and risk of ipsilateral hemispheric ischaemic events: results from the ACSRS study. Eur J Vasc Endovasc Surg 30(3):275–284
26. Zweig MH, Campbell G (1993) Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clin Chem 39(4):561–577