

Multi Feature Texture Analysis for the Classification of Carotid Plaques

C.I. Christodoulou^{1,2}, C. S. Pattichis³, M. Pantziaris², T. Tegos⁴,
A. Nicolaides^{2,4}, T. Elatrozy⁴, M. Sabetai⁴, S. Dhanjil⁴

¹Department of Electronic Engineering, Queen Mary and Westfield College,
University of London, Mile End Road, London E1 4NS, U.K.

²Cyprus Institute of Neurology and Genetics, P.O.Box 3462, 1683 Nicosia, Cyprus

³Department of Computer Science, University of Cyprus, 75 Kallipoleos Str.,
P.O.Box 578, 1678 Nicosia, Cyprus

⁴Irvine Laboratory for Cardiovascular Investigation and Research, St Mary's Hospital
Medical School, Imperial College of Science, Technology and Medicine,
Praed Street, London W2 1PG, U.K.
email: {cschr2, pattichi}@ucy.ac.cy

Abstract

The objective of this work was to develop a computer aided system which will facilitate the automated characterisation of carotid plaques recorded from high resolution ultrasound images for the identification of individuals with asymptomatic carotid stenosis at risk of stroke. The plaques were classified into (i) symptomatic or (ii) asymptomatic. Ten different texture feature sets were extracted from the segmented plaque image using the following algorithms: first order statistics, spatial gray level dependence matrices, gray level difference statistics, neighborhood gray tone difference matrix, statistical feature matrix, Laws texture energy measures, fractal dimension texture analysis, Fourier power spectrum and shape parameters. Although the statistics of all features extracted for the two classes indicated a high degree of overlap, a classification of the plaques was possible using the unsupervised self-organizing feature map (SOFM) classifier and combining techniques. The classification results of the different feature sets were combined using (i) majority voting and (ii) weighted averaging based on a confidence measure derived from the SOFM. Combining the classification results of the ten different feature sets improved significantly the classification results obtained by the individual feature sets, reaching an average diagnostic yield of 75%.

1 Introduction

There is evidence that carotid endarterectomy in patients with asymptomatic carotid stenosis will reduce the incidence of stroke. However, a large number of patients is operated unnecessarily. Therefore it is necessary to identify patients at high risk (>4% stroke

incidence p.a.) which will be considered for carotid endarterectomy, and patients at low risk (>1% p.a.) which will be spared from an unnecessary, expensive and often dangerous operation. There are indications that the morphology of atherosclerotic carotid plaques, obtained by high resolution ultrasound imaging, have prognostic implications [1]. Smooth surface, echogenicity and a homogenous texture are characteristics of stable plaques, whereas irregular surface, echolucency and a heterogenous texture are characteristics of potentially unstable plaques.

The main objective of this work is to develop a computer aided system which will facilitate the automated characterization of carotid plaques recorded from high resolution ultrasound images (duplex scanning and color flow imaging). This work is part of an European Union project (Biomed 2 Program - PL 950629) carried out in centers all over Europe and coordinated by the St. Mary's Hospital, London, U.K. The aim of the project is to evaluate the value of noninvasive investigations in the identification of individuals with Asymptomatic Carotid Stenosis at Risk of Stroke [2]. The computer aided classification of carotid plaques will contribute towards a more standardized and accurate methodology for the assessment of carotid plaques. The developed system should be able, based on extracted texture feature and shape parameters, to automatically classify plaques into one of the following types: (i) Symptomatic because of ipsilateral hemispheric symptoms and (ii) Asymptomatic because they were not connected with ipsilateral hemispheric events. The aim is to identify patients at risk of stroke.

2 Material

A total of 166 carotid plaque images (76 symptomatic + 90 asymptomatic) has been processed. Two sets of data have been selected: (i) for training the system, and (ii) for evaluating its performance. For training the system 58 symptomatic and 58 asymptomatic plaques were used whereas for evaluation of the system the remaining 18 symptomatic and 32 asymptomatic plaques were used. In order to verify the correctness of the classification results a bootstrapping procedure was followed. The system was trained and evaluated using five different bootstrap sets where in each set 116 different plaques were selected at random for training and 50 different plaques for evaluation.

3 Image Preprocessing

Before processing, the images have to be standardised manually by adjusting the image so that the median gray level value of the blood is 15-20 and the median gray level value of the adventitia (artery wall) is 180-200. The scale of the gray level of the images ranges from 0 to 255. This standardisation using blood and adventitia as reference points was necessary in order to extract comparable results when processing images obtained by different operators and different equipment. Following the image standardisation the plaque segments were outlined manually by the expert physician using as guide their corresponding colour blood flow images. Figure 1 illustrates an ultrasound image of the carotid artery with the outline of the carotid plaque.

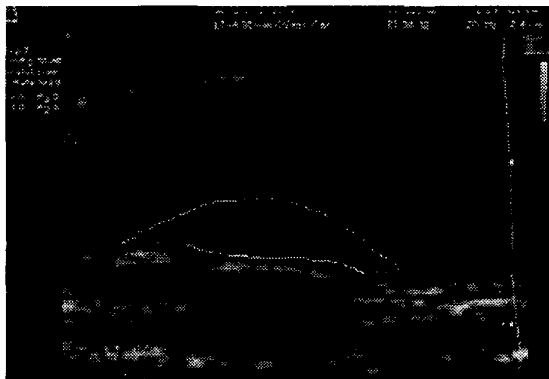


Fig. 1 An ultrasound image of the carotid artery with the outline of the carotid plaque.

4 Feature Extraction

Texture features and shape parameters were extracted from the segmented plaque images in order to be used for the classification of the carotid plaques. Texture contains important information which is used by humans for the interpretation and the analysis of many types of images. Texture refers to the spatial

interrelationships and arrangement of the basic elements of an image [3]. Visually, these spatial interrelationships and arrangements of the image pixels are seen as variations in the intensity patterns or gray tones. Therefore texture features have to be derived from the gray tones of the image. Although it is easy for humans to recognize texture, it is quite a difficult task to be defined, and subsequently to be interpreted by digital computers.

In this study, a total number of 61 texture features and shape parameters were extracted from the plaque segments using the following algorithms:

(i) First Order Statistics (FOS) [4]

- 1) Mean value, 2) Median value, 3) Standard Deviation, 4) Skewness, 5) Kurtosis.

(ii) Spatial Gray Level Dependence Matrices (SGLDM) [5]

- 1) Angular second moment, 2) Contrast, 3) Correlation, 4) Sum of squares: variance, 5) Inverse difference moment, 6) Sum average, 7) Sum variance, 8) Sum entropy, 9) Entropy, 10) Difference variance, 11) Difference entropy, 12), 13) Information measures of correlation.

For each feature the mean values and the range of values were computed, and were used as two different feature sets.

(iii) Gray Level Difference Statistics (GLDS) [6]

- 1) Contrast, 2) Angular second moment, 3) Entropy, 4) Mean.

(iv) Neighborhood Gray Tone Difference Matrix (NGTDM) [3]

- 1) Coarseness, 2) Contrast, 3) Busyness, 4) Complexity, 5) Strength.

(v) Statistical Feature Matrix (SFM) [7]

- 1) Coarseness, 2) Contrast, 3) Periodicity, 4) Roughness.

(vi) Laws Texture Energy Measures (TEM) [8], [9]

- 1) LL - texture energy from LL kernel, 2) EE - texture energy from EE kernel, 3) SS - texture energy from SS kernel, 4) LE - average texture energy from LE and EL kernels, 5) ES - average texture energy from ES and SE kernels, 6) LS - average texture energy from LS and SL kernels.

(vii) Fractal Dimension Texture Analysis (FDTA) [9], [10]

- $H^{(k)}$ parameter (Hurst coefficient) for resolutions $k=1, 2, 3, 4$.

(viii) Fourier Power Spectrum (FPS) [6]

- 1) Radial sum, 2) Angular sum.

(ix) Shape Parameters

- 1) X - coord. maximum length, 2) Y - coord. maximum length, 3) Area, 4) Perimeter, 5) Perimeter²/Area.

5 Plaque Classification

Following the feature extraction, feature classification was implemented based on neural network technology. The self-organising feature map (SOFM) classifier was used for the classification of the carotid plaques into symptomatic or asymptomatic. As input to the classifier were used the ten features sets given in the previous section. All features were normalised by division with their mean values before use.

5.1 SOFM classifier with a confidence measure

The SOFM was chosen because it is an unsupervised learning algorithm where the input patterns are freely distributed over the output node matrix [11]. The weights are adapted without supervision in such a way, so that the density distribution of the input data is preserved and represented on the output nodes. This mapping of similar input patterns to output nodes which are close to each other represents a discretisation of the input space, allowing a visualisation of the distribution of the input data. The output nodes are usually ordered in a two dimensional grid, and at the end of the training phase, the output nodes are labelled with the class of the majority of the input patterns of the training set, assigned to each node. In the evaluation phase, an input pattern is assigned to the output node with the weight vector closest to the input vector, and it is said to belong to the class label of the winning output node where it has been assigned. Figure 2 illustrates the distribution of the 166 plaques on a 12x12 SOFM using as input all the 61 features. The figure illustrates the high degree of overlap between the two classes.

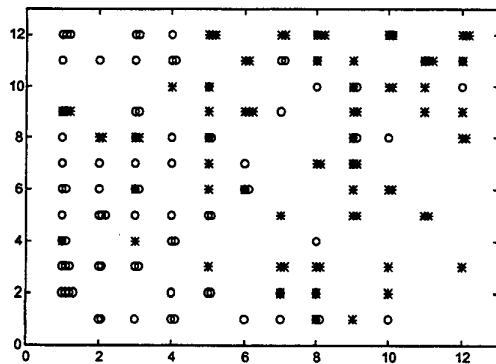


Fig. 2 Distribution of 166 carotid plaques (76 Symptomatic and 90 Asymptomatic) on a 12x12 SOFM using as input all the 61 features (* = Symptomatic, o = Asymptomatic).

Beyond the classification result, a confidence measure was derived from the SOFM classifier characterising how reliable the classification result was. The confidence measure was calculated based on the classes

of the nearest neighbours on the self-organizing map. For this purpose, the output nodes in a neighbourhood window centered at the winning node were considered. The confidence measure was computed for five different window sizes: 1x1, 3x3, 5x5, 7x7, 9x9. For each one of the ten feature sets a different SOFM classifier was trained. The implementation steps for the confidence measure calculation were the following:

Step 1: Train the classifier. An SOFM classifier is trained with the training set, using as input one of the ten feature sets.

Step 2: Label the nodes on the SOFM. Feed again the training set to the SOFM classifier, and label each output node on the SOFM, with the number of the symptomatic or asymptomatic training input patterns assigned to it.

Step 3: Apply the evaluation set. In the evaluation phase a new input pattern is assigned to a winning output node. The number of symptomatic and asymptomatic training input patterns assigned to each node in the given neighbourhood window (e.g. 1x1, ..., 9x9) around the winning node are counted.

Step 4: Compute the confidence measure and classify plaque. Calculate the confidence measure as the percentage of the majority of the training input patterns to the total number of the training input patterns in the given neighbourhood window. More specific in order to set its range from 0 to 1 (0 = low confidence, 1 = high confidence) the confidence measure is calculated as

$$conf = 2(\max\{SN_1, SN_2\} / (SN_1 + SN_2)) - 1 \quad (1)$$

where SN_m is the number of the input patterns in the neighbourhood window for the two classes $m = \{1, 2\}$:

$$SN_m = \sum_{i=1}^L N_{mi} \quad (2)$$

where N_{mi} is the number of the training patterns of the class m assigned to the output node i . L is the number of the output nodes in the $R \times R$ neighbourhood window with $L = R^2$. The evaluation input pattern is classified to the class m of the SN_m with the greatest value, as symptomatic or asymptomatic.

Step 5: Apply a weighting mask. A further enhancement in the calculation of the confidence measure and the plaque classification is to give to the output nodes nearest to the winning output node a greater weight than the ones farther away. So a windowing mask is used and the N_{mi} is multiplied with the value of the mask depending on the square of the distance of the output node from the winning output node. In this case Eq. 2 is modified into

$$sN_{..} = \sum_{i=1}^L N_{..} W_i \quad (3)$$

where W_i is the weight value at the output node i . The winning output node is given a weight value equal to 1.

5.2 Combining classifiers with a confidence measure

In the case of difficult pattern recognition problems, the combination of the outputs of multiple classifiers using for input multiple feature sets extracted from the raw data, can improve the overall classification performance. In the case of noisy or of a limited amount of data, different classifiers often provide different generalisations by realising different decision boundaries. Also, different feature sets provide different representations of the input patterns, containing different classification information. Selecting the best classifier or the best feature set is not necessarily the ideal choice, since potentially valuable information contained in the less successful feature sets or classifiers may not be taken into account. The combination of the results of the different features and the different classifiers increases the probability that the errors of the individual features or classifiers may be compensated by the correct results of the rest [12]. Furthermore the performance of the combiner is never worse than the average of the individual classifiers, but not necessarily better than the best classifier [13]. Also, the error variance of the final result is reduced making the whole system more robust and reliable.

In this work, the usefulness of combining neural network classifiers is investigated in the development of a decision support system for the classification of carotid plaques. For each feature set an SOFM classifier was trained and the ten classification results were combined using: (i) majority voting and (ii) weighted averaging based on a confidence measure derived from the SOFM. In weighted averaging the confidence measure determined the contribution of each feature set to the final classification result. The idea is that some feature sets may be more successful for specific regions of the input population. The confidence measure as calculated in Eq. 3 for the 10 different feature sets, was used as input to the combiner. When an input plaque pattern was classified as symptomatic the confidence measure was multiplied with -1. The final classification result was the average of the 10 confidence measures. If the final result value was negative then the plaque was classified as symptomatic otherwise if it was positive as asymptomatic. Values close to zero mean low confidence of the correctness of the final classification result whereas values close to -1 or 1 indicate a high confidence.

6 Results

A total of 166 (76 symptomatic + 90 asymptomatic) ultrasound images of carotid atherosclerotic plaques were examined. Ten different texture feature sets and shape parameters (a total of 61 features) were extracted from the plaque segments as described in section 4. The statistics of all features extracted for the two classes indicated a high degree of overlap making the classification difficult.

For the classification the unsupervised SOFM classifier was implemented with a 12x12 output node architecture and it was trained for 5000 learning epochs. For training the classifier 58 symptomatic and 58 asymptomatic plaques were used, whereas for evaluation of the system the remaining 18 symptomatic and 32 asymptomatic plaques were used. In order to verify the correctness of the classification results a bootstrapping procedure was followed. The system was trained and evaluated using five different bootstrap sets where in each set 116 different plaques were selected at random for training and 50 different plaques for evaluation. The SOFM classifier yielded a confidence measure on how reliable the classification result is, based on the number of the nearest neighbours on the self-organizing map. Five different neighbourhood windows were tested: 1x1, 3x3, 5x5, 7x7, 9x9. Furthermore the confidence measure was calculated using a weighting mask giving to the output nodes nearest to the winning output node a greater weight than the ones farther away.

Table I tabulates the diagnostic yield obtained for the 10 different feature sets, for the 5 different window sizes. Best window size in average was the 7x7 with 67.8% followed closely by the window sizes 9x9, 5x5 and 3x3. The worst window size was the 1x1 with average diagnostic yield only 39.6%. Best features sets were in average for all windows the GLDS with 65.9%, followed by the SGLDM (mean values) with 65.3% and the FOS with 64.6%. Worst feature set was the shape parameters with diagnostic yield only 51.2%. The best individual diagnostic yield was obtained by the GLDS using a 5x5 neighbourhood window and it was 74.4%, followed by the SGLDM (mean values) with 73.2% and a 7x7 window, the FOS with 72.0% and a 7x7 window and the SFM with also 72.0% using a 9x9 window.

In order to enhance the classification success rate, a multi-feature modular neural network with SOFM classifiers was used which was trained with the ten feature sets. The classification results were combined using majority voting and by averaging the confidence measure as described in section 5. The combination of

the classification results significantly improved the average success rate for the ten feature sets and for all the five window sizes from 61.0% up to 71.2% when combined with majority voting and up to 74.9% when combined with the confidence measure. When combined, from the five different neighbourhood

windows tested the best result was obtained by the 3x3 and was 76.4%. This corresponds to a neighbourhood of 9 output nodes on the 12x12 SOFM. However, also the other window sizes yielded comparable good results.

Table I Mean and standard deviation of the diagnostic yield (DY) for the evaluation set of the modular neural network diagnostic system, after bootstrapping the available data for five different sets of plaques. The diagnostic yield is given for the ten feature sets, their average, their combining using (i) majority voting, and (ii) weighted averaging based on a confidence measure. The DY was computed for five different neighborhood windows.

	Window= Feature set	n ¹	1x1 DY%	3x3 DY%	5x5 DY%	7x7 DY%	9x9 DY%	Average DY%
1	FOS	5	42.0 ± 7.8	68.4 ± 7.1	70.0 ± 9.3	72.0 ± 9.2	70.4 ± 9.1	64.6 ± 8.5
2	SGLDM (mean)	13	43.6 ± 4.3	65.2 ± 6.0	71.2 ± 4.8	73.2 ± 4.1	73.2 ± 5.2	65.3 ± 4.9
3	SGLDM (range)	13	39.2 ± 9.3	62.8 ± 8.5	68.4 ± 7.5	69.2 ± 9.3	70.0 ± 9.1	61.9 ± 8.7
4	GLDS	4	41.2 ± 2.7	68.8 ± 7.1	74.4 ± 5.3	72.8 ± 7.1	72.4 ± 7.1	65.9 ± 5.9
5	NGTDM	5	38.8 ± 5.0	61.2 ± 4.3	64.8 ± 5.7	64.8 ± 6.1	65.6 ± 8.1	59.0 ± 5.8
6	SFM	4	34.4 ± 4.3	63.2 ± 4.1	70.0 ± 6.6	71.6 ± 7.4	72.0 ± 7.3	62.2 ± 5.9
7	TEM	6	38.0 ± 5.5	66.0 ± 4.6	68.8 ± 4.7	67.2 ± 4.7	66.8 ± 3.7	61.4 ± 4.6
8	FDTA	4	48.0 ± 5.9	63.6 ± 7.1	65.2 ± 6.6	67.2 ± 6.4	67.6 ± 7.3	62.3 ± 6.7
9	FPS	2	38.8 ± 4.7	53.6 ± 4.3	60.8 ± 2.0	63.6 ± 5.9	61.6 ± 6.1	55.7 ± 4.6
10	Shape parameters	5	32.0 ± 6.3	51.6 ± 6.7	58.0 ± 7.5	56.8 ± 8.3	57.6 ± 7.9	51.2 ± 7.3
	Average		39.6 ± 5.6	62.4 ± 6.0	67.2 ± 6.0	67.8 ± 6.8	67.7 ± 7.1	61.0 ± 6.3
	Combine with majority voting		70.4 ± 2.6	70.4 ± 7.7	72.4 ± 6.2	71.6 ± 9.9	71.2 ± 9.0	71.2 ± 7.1
	Combine with weighted averaging		72.0 ± 2.5	76.4 ± 6.0	75.2 ± 8.6	76.0 ± 8.0	74.8 ± 8.8	74.9 ± 6.7

¹ Feature set vector size.

7 Discussion

In this study a computer aided system using texture features and neural network technology is proposed for the classification of carotid plaques. Such a system will greatly help in enhancing the significance of non-invasive cerebrovascular tests in the identification of asymptomatic carotid stenosis at risk of stroke. A total number of 61 texture features and shape parameters were extracted from the carotid plaque images. The statistics for all the texture features extracted indicate a high degree of overlap between the symptomatic and asymptomatic groups. This makes the separation and the classification of the two groups difficult.

The SOFM classifier was used for the classification of the carotid plaques. The SOFM was chosen because it is an unsupervised learning algorithm where the input patterns are freely distributed over the output node matrix, allowing an efficient mapping of the input data with no need to create exact classification boundaries. The supervised classifiers back propagation (BP) and

radial basis function (RBF) were tested and failed to converge because of the high degree of overlap between the two classes.

Ten different texture and shape feature sets were extracted from the plaque images and used for training multiple SOFM classifiers. Best feature sets were the GLDS, followed by the SGLDM (mean values) and the FOS whereas worst feature set was the shape parameters. In addition, a confidence measure on how reliable the classification result was computed from the SOFM classifier, based on the number of the nearest neighbours on the self-organizing map. For this purpose different neighbourhood window sizes were tested and evaluated.

In order to enhance the classification success rate, a multi-feature modular neural network with SOFM classifiers was used. The classification results of the ten feature sets were combined by averaging the confidence measure derived from the SOFM. Combining the

classification results of the ten different feature sets improved significantly the classification results obtained by the individual feature sets, reaching an average diagnostic yield of about 75%. The benefits of combining are more obvious in the case where no dominant best feature set or best classifier are available, as this was the case with the features extracted from the carotid plaque images. Combining with the use of a confidence measure, by weighted averaging of the individual classification results instead of using majority voting, can further improve the overall classification performance. The use of the confidence measure improved the final diagnostic yield to 74.9%, compared to 71.2% when using majority voting.

In previous work [14-16] a relationship between plaque morphology and risk of stroke was reported. However in most the above studies, the characteristics of the plaques were usually subjectively defined or using simple statistical measures, and the association with symptoms was established through simple statistical analysis. In this work, a large number of texture features were extracted directly from the plaque ultrasound images and were analysed using multi-feature multi-classifier modular neural network technology.

In conclusion, the results in this work show that it is possible to identify a group of patients at risk of stroke based on texture features extracted from high resolution ultrasound images of carotid plaques. This group of patients will benefit from a carotid endarterectomy whereas other patients will be spared from an unnecessary operation. Because of the difficulty of the problem and the high degree of overlap of the symptomatic and asymptomatic classes, the above results should be verified with more images from more patients. Furthermore other source of information may be used for classification, for example information obtained by a 3-dimensional reconstruction of the carotid plaque, which may lead to a better diagnostic yield.

Acknowledgements

This work was supported by the Cyprus Planning Bureau, the Cyprus Institute of Neurology and Genetics and the University of Cyprus.

References

- [1] JEMU: Consensus concerning the morphology and the risk of carotid plaques, Vol. 176, Paris 14 December 1996.
- [2] ACSRS, The value of noninvasive investigations in the identification of individuals with Asymptomatic Carotid Stenosis at Risk of Stroke, *Protocol of Biomed 2 Program* - PL 950629, 1995.
- [3] Amadasun M., King. R., 1989, "Textural Features Corresponding to Textural Properties", *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 19, No 5, Sept./Oct. 1989.
- [4] Press W.H., Flannery B.P., Teukolsky S.A., Vetterling W.T., 1987, *Numerical Recipes, The Art of Scientific Computing*, Cambridge University Press, Cambridge.
- [5] Haralick R.M., Shanmugam K., Dinstein I., 1973, "Texture Features for Image Classification", *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. SMC-3, pp. 610-621, Nov. 1973.
- [6] Weszka J.S., Dyer C.R., Rosenfield A., 1976, "A Comparative Study of Texture Measures for Terrain Classification", *IEEE Transactions on Systems, Man. & Cybernetics*, Vol. SMC-6, April 1976.
- [7] Wu Chung-Ming, Chen Yung-Chang, 1992, "Statistical Feature Matrix for Texture Analysis", *CVGIP: Graphical Models and Image Processing*, Vol. 54, No 5, pp. 407-419, Sept. 1992.
- [8] Laws K.I., 1980, "Rapid Texture Identification", *SPIE*, 1980, Vol. 238, pp. 376-380.
- [9] Wu Chung-Ming, Chen Yung-Chang, Hsieh Kai-Sheng, 1992, "Texture Features for Classification of Ultrasonic Liver Images", *IEEE Transactions on Medical Imaging*, Vol. 11, No 3, June 1992.
- [10] Mandelbrot B.B., 1982, *The Fractal Geometry of Nature*, San Francisco, CA, Freeman.
- [11] Kohonen T., 1990, "The Self-Organizing Map", *Proceedings of the IEEE*, Vol. 78, No. 9, pp. 1464-1480, Sept. 1990.
- [12] Turner K., Ghosh J., "Theoretical foundations of linear and order statistics combiners for neural pattern classifiers", *Dept. of Electrical and Computer Engineering, University of Texas, Austin*, 1996.
- [13] Perrone M.P., "Averaging/Modular Techniques for Neural Networks", *The Handbook of Brain Theory and Neural Networks*, ed. by M.A. Arbib, MIT Press, Cambridge, Massachusetts, pp. 126-129, 1995.
- [14] Geroulakos G., Domjan J., Nicolaidis A., Stevens J., Labropoulos N., Ramaswami G., Belcaro G., Belcaro G., "Ultrasonic carotid artery plaque structure and the risk of cerebral infraction on computed tomography", *Journal of Vascular Surgery*, Vol. 20, No 2, pp. 263-266, August 1994.
- [15] Iannuzzi A., Wilcosky T., Mercury M., Rubba P., Bryan F., Bond G., "Ultrasonographic Correlates of Carotid Atherosclerosis in Transient Ischemic Attack and Stroke", *Stroke*, Vol. 26, No 4, pp. 614-619, April 1995.
- [16] Polak J., Shemanski L., O'Leary D., Lefkowitz D., Price T., Savage P., Brand W., Reld C., "Hypochoic Plaque at US of the Carotid Artery: An Independent Risk Factor for Incident Stroke in Adults Aged 65 Years or Older", *Radiology*, Vol. 208, Number 3, pp. 649-654, Sept. 1998.