

Neuro-fuzzy Classification of Transcranial Doppler Signals with Chaotic Measures and Spectral Parameters

Ali Ozturk

Computer Engineering Department
KTO Karatay University
KONYA, TURKEY
ali.ozturk@karatay.edu.tr

Ahmet Arslan

Computer Engineering Department
Selcuk University
KONYA, TURKEY
ahmetarslan@selcuk.edu.tr

Abstract—Transcranial Doppler (TCD) is a non-invasive diagnosis method which is used in diagnosis of various brain diseases by measuring the blood flow velocities in brain arteries. In this study, chaos analysis of the TCD signals recorded from the middle arteries of the temporal region of brain of the 82 patients and 23 healthy people was investigated. Among 82 patients, 20 of them had cerebral aneurism, 10 had brain hemorrhage, 22 had cerebral oedema and the remaining 30 had brain tumor. Maximum Lyapunov exponent which is the strongest quantitative indicator of chaos was found to be positive for all TCD signals. The correlation dimension was found as greater than 2 and as fractional value for all TCD signals. These two features were used for training a NEFCLASS model. The NEFCLASS model had two input nodes for D2 and maximum Lyapunov exponent values and five output nodes representing the subject group to which the inputs belonged. In order to make k-fold cross-validation, the data set was randomly divided into 5 subsets of equal size. In an iterated manner, 4 of these subsets were used for training and the remaining 1 subset was used for testing. This operation was repeated for 3 times. The average accuracy for train and test set was found as %81 and %79, respectively. The performance of the NEFCLASS model was also assessed in the same manner with spectral parameters (i.e. resistivity index and pulsatility index) which were obtained from Doppler sonograms. The average accuracy was found as %67 and %63 for train and test set, respectively.

Keywords—Transcranial Doppler signals; nonlinear analysis; chaotic measures; neuro-fuzzy classification; NEFCLASS

I. INTRODUCTION

TCD study of the adult intracerebral circulation is used to evaluate intracranial stenoses, cerebral arteriovenous malformations, cerebral vasospasm and cerebral hemodynamics in general [1]. To make intra-brain blood flow measurements require sophisticated imaging methods. The Doppler principle utilized in medicine is different from the classical Doppler method. Since the targets in the human body don't emit radiation, it is necessary to transmit a signal into the body and then observe the frequency changes between that transmitted signal and the received signals reflecting from the targets. Under these conditions, there is a shift in the ultrasound frequency given by

$$f_d = f_t - f_r = 2f_t v \cos \theta / c \quad (1)$$

Where f_t and f_r are the transmitted and received ultrasound frequencies, v is the velocity of the target, c is the velocity of the sound in the medium, and θ is the angle between ultrasound beam and the direction of motion of the target.

In the literature, the features extracted from Doppler signals via spectral analysis methods were widely used for automatic medical diagnosis [2, 3, 4, 5]. In [3], the pulsatility index (PI) and resistivity index (RI) are used for the classification of carotid arterial Doppler signals. In [5], two different artificial neural networks were compared in classifying the TCD signals using the spectral parameters. The chaotic invariant measures were used for the classification of TCD signals in [6, 7]. However, there is no study which investigates the comparison of the classification performance of a neuro-fuzzy classifier on the spectral parameters and chaotic measures.

The TCD signals may be regarded as a series of discrete values varying in time. These kinds of series are called time series. Chaos theory (Nonlinear time series analysis) provides some necessary tools to quantitatively analyze an experimental time series such as a TCD signal. The values such as maximum Lyapunov exponent and correlation dimension are some of the quantitative measures of chaos. For example, if the maximum Lyapunov exponent is found to be positive, then we say that there is an underlying chaotic behavior in the system that generates the nonlinear time series. This means that the system have a picture of chaotic attractor embedded in the phase space which includes nearby orbits that diverges from each other in time. According to the Takens theorem [8], if the real dimension of the attractor is D_A , then choosing the embedding dimension D_E as $D_E > 2D_A$ prevents the orbits from crossing each other resulting in a non-periodic flow. The chaotic attractors generally have a correlation dimension which is a fractional value greater than 2.

Neuro-fuzzy classification approaches aim at creating fuzzy classification rules from data. NEFCLASS is able to learn fuzzy rules and fuzzy sets by simple heuristics. The main idea of NEFCLASS is to quickly create interpretable fuzzy classifiers. The learning algorithm of NEFCLASS has two stages: rule (structure) learning and parameter learning. Rule learning is achieved via a variation of the approach by Wang and Mendel [9]. Each input space corresponding to a feature is partitioned by a given number of initial fuzzy sets. Then, by

processing the training data only once, the antecedents for the rules are created by checking which areas of the input space contain data. An evaluation procedure then creates a rule base by assigning appropriate consequents (class labels) to the discovered antecedents and selects only a certain number of rules with good performance [10,11,12].

The two most strong chaos indicators (i.e. the maximal Lyapunov exponent and the correlation dimension) were calculated for each signal. In order to make k-fold cross-validation, the data set was randomly divided into 5 subsets of equal size. In an iterated manner, 4 of these subsets were used for training and the remaining 1 subset was used for testing. This operation was repeated for 3 times. The NEFCLASS-J, which is a Java enabled version of NEFCLASS, was used for this purpose. The NEFCLASS model had two input nodes corresponding to the chaotic invariant measures and five output nodes which represent the four different brain diseases and the healthiness. The NEFCLASS model with 2 input and 5 output nodes was also used in the same manner as described above to assess the classification performance with spectral parameters which were obtained from Doppler sonograms (i.e. resistivity index (RI) and pulsatility index (PI)).

II. MATERIALS AND METHOD

A. Hardware

The hardware of the system used for this study involves a 2 Mhz ultrasound transducer, analog Doppler unit (Multi Doppler Transducer XX, DWL Gmb, Uberlingen, Germany), analog/digital interface board (Sound Blaster Pro-16), and PIII 600 Mhz microprocessor PC with printer. The Doppler unit is also equipped with imaging software that makes it possible to focus the sample volume at a desired location in the temporal region.

The analog Doppler unit can work in both continuous and pulse wave modes. The sample volume sizes can be adjusted according to vessel diameter for recording different blood flow rates. The signal obtained from the blood vessel is transferred to a PC via a 16-bit sound card on an analog/digital interface board.

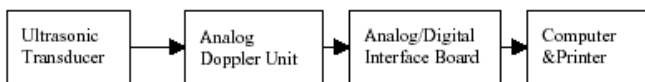


Fig.1. Block diagram of measurement system

B. Spectral Parameters

The spectral values obtained from Doppler sonograms may differ due to the factors such as observation angle and probe frequency when they are not applied at an optimal level. However, since these factors will spread out all over the spectral wave, the rates devoted to the full wave will not be affected. The values such as maximum systolic height (S), end diastolic height (D) and mean height (M) are extracted from the Doppler sonograms which are obtained via spectral analysis methods. The two waveform indices which are RI (resistivity index) and PI (pulsatility index) are calculated using these values as in the following [13].

$$RI = \frac{(S - D)}{S} \quad (2)$$

$$PI = \frac{(S - D)}{M} \quad (3)$$

Both of these indices reflect the resistance to flow downstream from the point of insonation. They are influenced by many factors such as proximal stenosis, distal stenosis and peripheral resistance. To obtain the spectral parameters mentioned above, the sonograms produced by Auto Regressive Burg method were used [14].

C. Correlation Dimension

Various dimension definitions are given to indicate the complexity of dynamical systems. The most commonly used one in the literature is the correlation dimension. If an attractor is chaotic, the neighboring trajectories that are initially close will diverge exponentially in time and the points on these trajectories will be dynamically uncorrelated. Since these points are located on the attractor they are correlated in phase space, although they appear randomly scattered points. The idea behind the dimension quantifiers is that the weight $p(\epsilon)$ of a typical ϵ -ball covering part of the invariant set scales with its diameter like $p(\epsilon) \approx \epsilon^D$, where the value for D depends also on the weight definition. Using the square of the probability p_i to find a point of the set inside the ball, the dimension is called the correlation dimension, which is computed most efficiently by the correlation sum [15]:

$$C(m, \epsilon) = \frac{1}{N_{pairs}} \sum_{j=m}^N \sum_{k=j-w}^N \Theta(\epsilon - |s_j - s_k|) \quad , \quad (4)$$

Where s_i are m -dimensional delay vectors, $N_{pairs}=(N-m+1)(N-m-w+1)/2$ the number of pairs of points covered by the sums, Θ is the Heaviside step function and w is the Theiler window [16]. Since successive elements of a time series are not usually independent, Theiler suggested exclusion of these temporally correlated points from the pair counting operation. This is achieved by simply ignoring all pairs of points whose time indices differ by less than w . On sufficiently small length scales and when the embedding dimension m exceeds the box-dimension of the attractor [17],

$$C(m, \epsilon) \propto \epsilon^{D_2} \quad (5)$$

Since the box-dimension ϵ is not known a priori, the convergence of the estimated D_2 values in m must be checked.

In order to estimate the correlation dimension, we have to provide the time delay information between embedding vectors. Fraser and Swinney [18] suggested a method called mutual information to obtain a reasonable delay. Unlike the autocorrelation function, this method takes into account also nonlinear correlations. To find the optimum delay, we have to compute the following mutual information S for different τ delay values,

$$S = - \sum_{ij} p_{ij}(\tau) \ln \frac{p_{ij}(\tau)}{p_i p_j} \quad (6)$$

Where p_i is the probability to find a time series value in the i -th interval, and $p_{ij}(\tau)$ is the joint probability that an observation falls into the i -th interval and the observation time τ later falls into the j -th interval. The time delay τ where the mutual information S takes the first minimum value, is chosen as the optimum delay.

D. The Maximal Lyapunov Exponent

The maximal Lyapunov exponent can be determined without the explicit construction of a model for the time series. Consider the representation of the time series data as a trajectory in the embedding space, and assume that you observe a very close return s'_n to a previously visited point s_n . Then one can consider the distance $\Delta_0 = s_n - s'_n$ as a small perturbation, which should grow exponentially in time. For the time series: $\Delta_t = s_{n+t} - s'_{n+t}$, if one finds that $|\Delta_t| \approx \Delta_0 \varepsilon^{\lambda t}$ then λ is the maximal Lyapunov exponent.

According to Rosenstein et. al. [19], the following S value is computed for different N values.

$$S(t) = \frac{1}{N} \sum_{n=0}^{N-1} \left(\ln \frac{1}{|g_n|} \sum |s_{n+t} - s'_{n+t}| \right) \quad (7)$$

If S exhibits a linear increase with an identical slope, then this slope can be taken as an estimate of the maximal Lyapunov exponent.

E. The NEFCLASS Model

A NEFCLASS system has a 3-layer feedforward architecture that is derived from a generic fuzzy perceptron [20]. The first layer contains the input nodes representing the pattern features. The activation of a node is equal to its external input. The hidden layer holds the rule nodes representing the fuzzy rules and third layer consists of the output nodes, one for each class. The activation of a rule node $a_R^{(p)}$ and the activation of an output unit $a_c^{(p)}$ when a pattern is applied are computed by

$$a_R^{(p)} = \min_{x \in U_1} \{W(x, R)(a_x^{(p)})\}, \quad (8)$$

$$a_c^{(p)} = \sum_{R \in U_2} \frac{W(R, c) \cdot a_R^{(p)}}{W(R, c)} \text{ or } a_c^{(p)} = \max_{R \in U_2} \{a_R^{(p)}\} \quad (9)$$

Where $W(x, R)$ is the fuzzy weight on the connection from input node x to rule node R . $W(R, c)$ is the weight on the connection from rule node R to output node c . For semantical reasons all these weights are fixed at 1 [21].

The fuzzy sets and the linguistic rules which perform the approximation and define the resulting NEFCLASS system are obtained from the training data set by learning. A NEFCLASS system that is created from scratch starts with no hidden nodes at all. They are created during a first run through the learning task. A rule is created by finding for a given input pattern p the combination of fuzzy sets, where each yields the highest

degree of membership for the respective input feature. If this combination is not identical to the antecedents of an already existing rule, and if the maximum number of hidden layer nodes is not reached, a new rule node is created. Another way would be to use a scoring system, not restricting the number of rules at first, and then later select those rules having the highest scores.

After the rule base is created, the learning algorithm adapts the membership function of the antecedents. The learning procedure for the fuzzy sets is a simple heuristic. It results in shifting the membership functions and in making their supports larger or smaller. In this paper, we defined triangular membership function with 4 fuzzy sets (e.g., very small (vsm), small (sm), large (lg) and very large (vlg)) for each input node and also used *min* as the t-norm to determine the activation of a rule node.

Automatic pruning techniques used in the NEFCLASS implementation try to reduce the number of rules and variables to increase the comprehensibility of the induced model. The size of a fuzzy rule base is automatically determined by continuously selecting rules until all training patterns are covered by at least one rule. After training the membership functions to improve the performance of the rule base, a pruning mechanism tries to remove rules and variables in order to increase the comprehensibility of the classifier.

III. EXPERIMENTAL RESULTS AND DISCUSSION

The general profile of the subjects from which the TCD signals are obtained is given in the following table.

TABLE I. THE GENERAL PROFILE OF THE SUBJECTS USED IN THIS STUDY

	Males	Females	Age Range	Avg. Age
Cerebral aneurysm	12	8	55-65	59.5±0.5
Brain Hemorrhage	4	6	21-36	27.0±0.5
Cerebral Oedema	11	11	3-40	25.0±0.5
Brain Tumor	12	18	12-41	29.5±0.5
Healthy	15	8	23-65	31.5±0.5
Summary	54	51	3-65	35.0±0.5

In order to find the correlation dimension of TCD time series, the *d2* program of TISEAN package [22] was used. This program gets the time series file as input and produces 4 different files with different extensions the file having the extension *.d2* contains the necessary information to estimate the correlation dimension. The local slopes of the correlation sums for different embedding dimension were written in this file. To obtain the correlation dimension, we plotted the contents of the file in x-y coordinates where the ε values correspond to x coordinate and the local slopes of the correlation sums correspond to y coordinate.

For correlation dimension estimation we needed time delay information which was obtained by mutual information method. According to this method, the first delay value τ at

which time delayed mutual information takes the minimum is a good candidate for a reasonable time delay. For all the nonlinear time series investigated in this study the time delay τ was found to be 1 according to the mutual information method.

In order to find out the maximal Lyapunov exponent for the TCD time series, the *lyap_r* program of TISEAN package [22] was used. This program implements the method of Rosenstein et al. [19] which is known to be reliable in small data sets. The *lyap_r* program produces a file with two columns, first of which shows the iteration number and the second one shows the logarithm of the stretching factor. Maximal Lyapunov exponent was obtained by plotting these values using linear scales for both of the axes. The first slope of the resulting curve gives the maximal Lyapunov exponent.

The statistical summary for the estimated values of these two chaotic invariant measures is given in the following table.

TABLE II. THE STATISTICAL SUMMARY OF CHAOTIC INVARIANT MEASURES FOR EACH PATIENT GROUP

	Avg. D_2	Min. D_2	Max. D_2	Avg. λ_1	Min. λ_1	Max. λ_1
Cerebral aneurysm	3.58	3.23	3.84	3.28	2.96	3.38
Brain Hemorrhage	2.67	2.41	2.92	3.12	2.71	3.26
Cerebral Oedema	2.46	2.23	2.62	1.84	1.74	1.98
Brain Tumor	2.12	2.08	2.32	1.47	1.42	1.58
Healthy	3.71	3.42	3.92	3.68	3.60	3.83

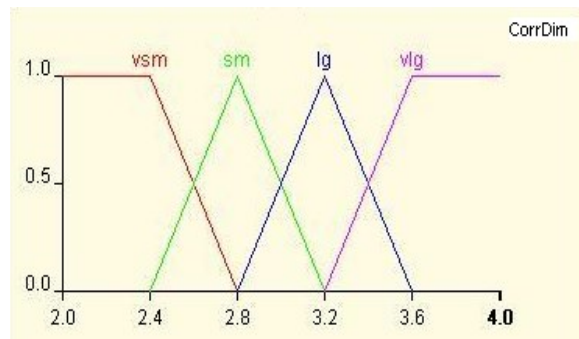
In order to make k-fold cross-validation for NEFCLASS, the data set was randomly divided into 5 subsets of equal size. In an iterated manner, 4 of these subsets were used for training the classification algorithms and the remaining 1 subset was used for testing. This operation was repeated for 3 times. The rule bases were created with the proposed NEFCLASS learning algorithm by applying the samples in the training subset for 1000 epochs. Automatic pruning was applied to each NEFCLASS system. The learning rate was given as 0.01.

The rule set of the NEFCLASS system created after 1000 epochs with application of automatic pruning consisted the following 6 rules.

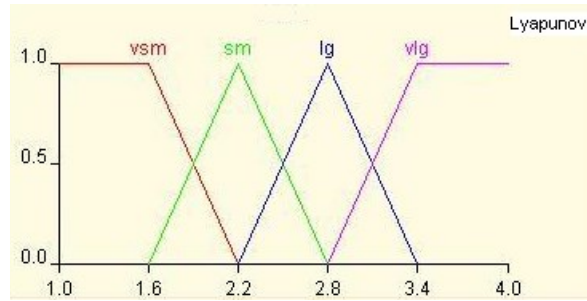
The final fuzzy sets of the two input nodes after the application 1000 epochs can be seen in Figure 3.

TABLE III. FUZZY RULE SET OBTAINED AFTER 1000 EPOCHS WITH AUTOMATIC PRUNING

1. if CorrDim is very large and Lyap is very large then Class 1
2. if CorrDim large and Lyap is very small then Class 2
3. if CorrDim is very small and Lyap is very small then Class 3
4. if CorrDim is small and Lyap is very large then Class 4
5. if CorrDim is large and Lyap is large then Class 5
6. if CorrDim is very large and Lyap is large then Class 5

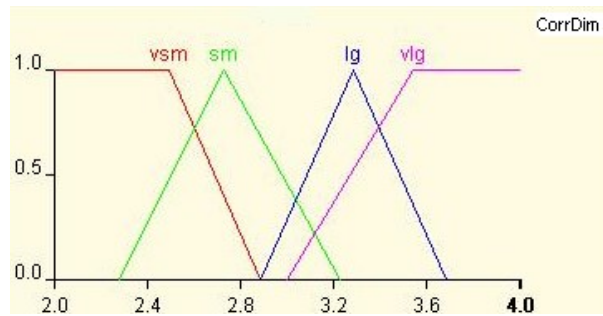


(a)

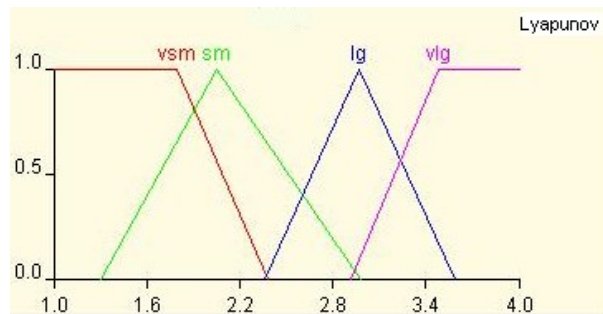


(b)

Fig.2. Initial fuzzy sets for two input nodes of the NEFCLASS system (a) Correlation dimension (b) Lyapunov exponent



(a)



(b)

Fig.3. Final fuzzy sets for two input nodes of the NEFCLASS system of 1000 epochs (a) Correlation dimension (b) Lyapunov exponent

The classification accuracy and performance of the NEFCLASS system was determined using the following cost functions, i.e. MAE (Mean Absolute Error), RMSE (Root Mean Squared Error) and RAE (Relative Absolute Error).

$$MAE = \frac{1}{n} \sum_{j=1}^n |O_j - T_j| \quad (10)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{j=1}^n (O_j - T_j)^2} \quad (11)$$

$$RAE = \frac{\sum_{j=1}^n |O_j - T_j|}{\sum_{j=1}^n |T_j - \bar{T}|} \quad (12)$$

Where O_j is the j^{th} observed output value, T_j is the j^{th} expected output value and \bar{T} is the average of the expected output values.

The performance of the NEFCLASS system for train and test sets after 5-fold cross validation in 3 iterations is given in the following tables for the chaotic measures.

TABLE IV. THE PERFORMANCE OF NEFCLASS SYSTEM FOR TRAIN AND TEST SETS FOR EACH ITERATION OF 5-FOLD CROSS VALIDATION (CHAOTIC MEASURES)

Data sets	Train set		Test set	
	False classification	Accuracy	False. classification	Accuracy
1st step	12	%86	4	%81
2nd step	19	%77	4	%81
3rd step	16	%81	5	%76
Average	16	%81	4	%79

TABLE V. THE MAE, RMSE AND RAE VALUES FOR TRAIN AND TEST SETS FOR EACH ITERATION OF 5-FOLD CROSS VALIDATION (CHAOTIC MEASURES)

Data sets	Train set			Test set		
	MAE	RMSE	RAE	MAE	RMSE	RAE
1st step	0,07	0,14	0,22	0,07	0,15	0,23
2nd step	0,10	0,19	0,31	0,10	0,19	0,31
3rd step	0,10	0,19	0,30	0,10	0,20	0,32
Average	0,09	0,17	0,28	0,09	0,18	0,29

According to Table II, the average accuracy is %81 for the training set and %79 for the test set.

In Figure 4, the NEFCLASS system which is obtained after 1000 training epochs is given. The system includes two input nodes for correlation dimension (D_2) and maximum Lyapunov exponent (λ_1). The five output nodes correspond to 4 patient classes and 1 healthy class. The fuzzy rule set given in Table I is mapped from the input nodes to the internal rule nodes.

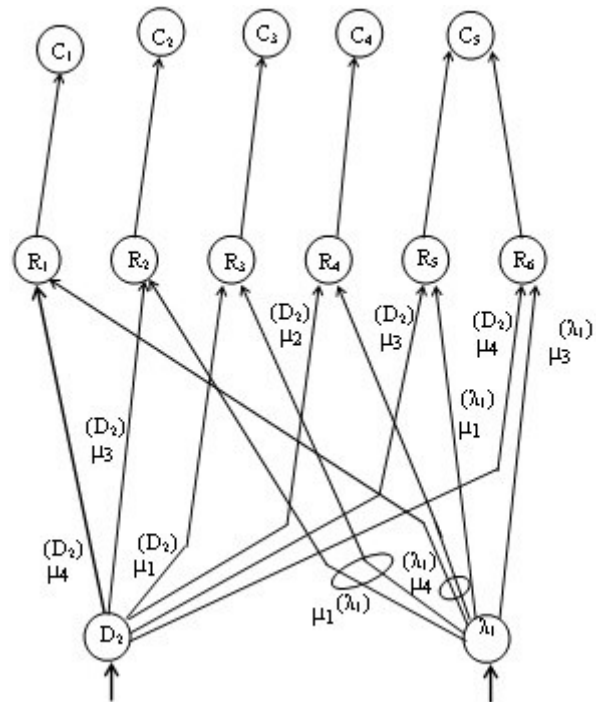


Fig.4. The NEFCLASS system for two input nodes and five output nodes with fuzzy rules after 1000 epochs

In order to compare the performance of the NEFCLASS model with spectral parameters, the NEFCLASS was also trained with resistivity index (RI) and pulsatility index (PI) values obtained from Doppler sonograms of the TCD signals. In the following figure, Doppler sonogram of a patient with cerebral hemorrhage was given as a sample.

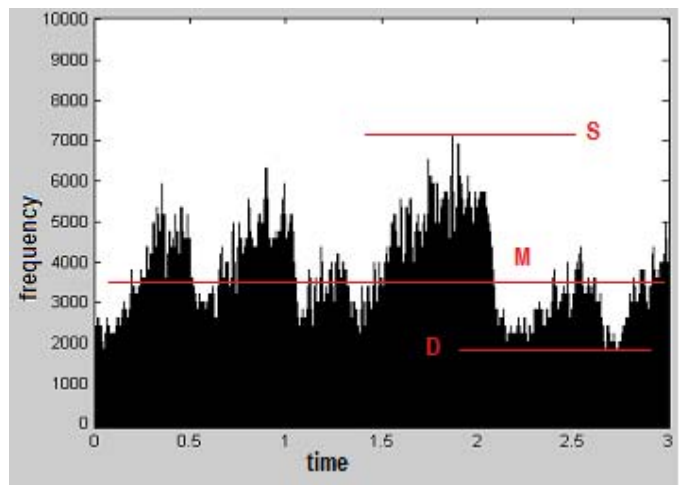


Fig.5. Doppler sonogram of a patient with Brain Hemorrhage

The RI and PI values were calculated from S, D and M values according to the equations 3 and 4. The performance of the NEFCLASS system for train and test sets after 5-fold cross validation in 3 iterations is given in the following tables for the spectral parameters.

TABLE VI. THE PERFORMANCE OF NEFCLASS SYSTEM FOR TRAIN AND TEST SETS FOR EACH ITERATION OF 5-FOLD CROSS VALIDATION (SPECTRAL PARAMETERS)

Data sets	Train set		Test set	
	False classification	Accuracy	False. classification	Accuracy
1st step	25	%70	8	%61
2nd step	29	%65	8	%62
3rd step	29	%65	7	%67
Average	28	%67	8	%63

TABLE VII. THE MAE, RMSE AND RAE VALUES FOR TRAIN AND TEST SETS FOR EACH ITERATION OF 5-FOLD CROSS VALIDATION (SPECTRAL PARAMETERS)

Data sets	Train set			Test set		
	MAE	RMSE	RAE	MAE	RMSE	RAE
1st step	0,13	0,22	0,41	0,15	0,25	0,46
2nd step	0,14	0,24	0,43	0,15	0,27	0,48
3rd step	0,14	0,24	0,45	0,14	0,24	0,42
Average	0,14	0,23	0,43	0,15	0,25	0,45

According to the results, the performance of the NEFCLASS model was better on both train and test sets when chaotic measures were used. The accuracy level was %81 and %79, respectively. When the NEFCLASS model was trained and tested with spectral parameters, its accuracy decreased to %67 and %63 on train and test sets, respectively.

The proposed algorithm for the classification of the TCD signals after training the NEFCLASS system is given in the following.

Algorithm Classification of the TCD signal

Input The TCD signal

Output The category of the TCD signal

- 1) Compute the time delay (τ) value
- 2) Compute the embedding dimension (m)
- 3) Reconstruct the attractor
- 4) Compute the maximal Lyapunov exponent (λ_1)
- 5) Compute the correlation dimension (D_2)
- 6) Apply the λ_1 and D_2 values to the trained NEFCLASS system
- 7) Return the output of the NEFCLASS system

REFERENCES

[1] D.H. Evans, W.N. McDicken, R. Skidmore, J.P. Woodcock, Doppler Ultrasound: Physics, Instrumentation and Clinical Applications, Wiley, Chichester, 1989.

[2] U. Ergun, S. Serhatlioglu, F. Hardalac, I. Guler, Classification of carotid artery stenosis of patients with diabetes by neural network and logistic regression, *Comp. Biol. Med.*, 24, 389–405, 2004.

[3] E.D. Ubeyli, I. Guler, Adaptive neuro-fuzzy inference systems for analysis of internal carotid arterial Doppler signals, *Comp. Biol. Med.*, 35, 687–702, 2005.

[4] I. Guler, F. Hardalac, N. Barisci, Application of FFT analyzed cardiac Doppler signals to fuzzy algorithm, *Comp. Biol. Med.* 32, 435–444, 2002.

[5] S. Serhatlioglu, F. Hardalac, I. Guler, Classification of transcranial Doppler signals using artificial neural network, *J. Med. Syst.* 27, 205–214, 2003.

[6] Ozturk A., Arslan A., Classification of transcranial Doppler signals using their chaotic invariant measures, *Computer Methods and Programs in Biomedicine*, 86(2): 171-180, 2007.

[7] Ozturk A., Arslan A., Hardalac F., Comparison of neuro-fuzzy systems for classification of transcranial Doppler signals with their chaotic invariant measures, *Expert Systems with Applications*, 34(2): 1044-1055, 2008.

[8] F. Takens, Detecting Strange Attractors in Turbulence, *Lecture Notes in Math.*, 898, 226-381, Springer, New York, 1981.

[9] L.X. Wang and J.M. Mendel, Generating fuzzy rules by learning from examples, *IEEE Transactions on System, Man and Cybernetics*, 22(6):1414-1427, 1992.

[10] D. Nauck, U. Nauck and R. Kruse, NEFCLASS for Java – new learning algorithms, In *Proc. IEEE 18th Int. Conf. of the North American Fuzzy Information Processing Society (NAFIPS99)*, 472-476, New York, NY, 1999.

[11] D. Nauck and R. Kruse, A neuro-fuzzy method to learn fuzzy classification rules from data, *Fuzzy Sets and Systems*, 89: 277-288, 1997.

[12] D. Nauck, U. Nauck and R. Kruse, Generating classification rules with the neuro-fuzzy system NEFCLASS, In *Proc. Biennial Conference of the North American Fuzzy Info. Proc. Society NAFIPS'96*, Berkeley, CA, June 1996.

[13] S. Serhatlioglu, Z. Bozgeyik, Y. Ozkan, F. Hardalac, I. Guler. 2003. Neurofuzzy Classification of the Effect of Diabetes Mellitus on Carotid Artery. *Journal of Medical Systems*. Vol. 27, No.5. October 2003.

[14] Kay, S. M. *Modern Spectral Estimation: Theory and Application*. Englewood Cliffs, NJ: Prentice-Hall, 1988.

[15] P. Grassberger and I. Procaccia, Characterization of Strange Attractors, *Physical Review Letters*, 50, 1983.

[16] J. Theiler, Estimating fractal dimension, *Journal of Optical Society of America*, A 7, 1055, 1990.

[17] T. Sauer and J. Yorke, How many delay coordinates do you need?, *Int. J. Bifurcation and Chaos* 3, 737, 1993.

[18] A. M. Fraser and H. L. Swinney, Independent coordinates for strange attractors from mutual information, *Phys. Rev. A* 33, 1134, 1986.

[19] M. T. Rosenstein, J. J. Collins, C. J. De Luca, A practical method for calculating largest Lyapunov exponents from small data sets, *Physica D* 65, 117, 1993.

[20] D. Nauck and R. Kruse, Designing neuro-fuzzy systems through backpropagation, *Fuzzy Modelling: Paradigms and Practice*, Kluwer, 1996.

[21] D. Nauck and R. Kruse, NEFCLASS – a neuro-fuzzy approach for the classification of data, *Proc. of the 1995 ACM Symposium on Applied Computing*, Nashville, Feb. 26-28, Pages 461-465, ACM Press, New York, February 1995.

[22] R. Hegger, H.Kantz, and T.Schreiber, Practical implementation of nonlinear time series methods: The TISEAN package, *Chaos*, 9, 413, 1999.