

Full Length Research Paper

New technique based on uterine electromyography nonlinearity for preterm delivery detection

Safaa M. Naeem*, Ahmed F. Seddik and Mohamed A. Eldosoky

Department of Biomedical Engineering, Helwan University, Helwan, Egypt.

Received 3 October, 2013; Accepted 29 September, 2014

Detecting uterine electromyography (EMG) signals can yield a promising approach to determine and take actions to prevent preterm deliveries. This paper objective is to predict this risk using such uterine signals. Previous classification studies have used only linear signal processing which depends on the spectral characteristics of the uterine EMG signals that did not give clinically acceptable results. On the other hand some studies have made linear and non-linear analysis for the signals and have found that the non-linear parameters can distinguish the preterm delivery in better way than the linear parameters. In this research, two methods will be taken; the first method is to take some linear parameters to a suitable neural network and the second one is to take some non-linear parameters to the same network. Then, the two results are compared by calculating parameters False Positive Rate, False Negative Rate, True Positive Rate, True Negative Rate and Accuracy to evaluate the classification performance. Besides, a linear parameter, discrete cosine transform, which depends on the spectral characteristics of the signals, is taken as an additional feature to the same network so the research will have a third method to illustrate the difference between the traditional previous classification method and the proposed ones. Applying the second method gives better results than the first and the third methods. The paper can propose a method depends on the uterine EMG nonlinearity which gives best results to detect preterm delivery compared with those used in previous studies.

Key words: Uterine electromyography (EMG) signals, term-preterm deliveries prediction, neural network performance evaluation, discrete cosine transform.

INTRODUCTION

A most urgent challenge in healthcare currently is the phenomenon of preterm labor, or labor prior to 37 completed weeks of gestation. Preterm labor leaves serious impacts on economy and society as a whole. The complications of preterm birth include significant neurological, mental, behavioral and pulmonary problems in later child's life (Diab et al., 2010). So any promising

technique that could improve the chances of preterm birth prediction is required. Analysis of uterine electromyography (EMG), termed as electrohysterogram (EHG), records is one such technique.

Uterine EMG has been the subject of research for many years from 1950. The uterine EMG has been proved to be of interest for pregnancy and parturition

*Corresponding author. E-mail: safaa_naim@h-eng.helwan.edu.eg

Author(s) agree that this article remain permanently open access under the terms of the [Creative Commons Attribution License 4.0 International License](http://creativecommons.org/licenses/by/4.0/)

monitoring (Diab et al., 2007). Uterine EMG classification, in previous researches, depends on the spectral characteristics of EMG activity. Wavelet transform is a tool that has been used to describe the uterine EMG activity and Power Spectral Density (Diab et al., 2007) as well as the Wavelet Packet Transform (Moslem et al., 2012). Most of the used signal-processing techniques were linear which rely on the changes in the frequency power spectrum of the uterine activity and included the following: the peak frequency of the power spectrum (Garfield et al., 2005); the burst energy levels (Maul et al., 2004); the mean power frequency (Hassan et al., 2010), the use of the peak frequency, the duration and number of bursts, the means and deviations of the frequency spectrum (Maner and Garfield, 2007); and the approaches of analyzing contractions using multiple techniques such as the kurtosis and skewness coefficient. Other approaches included calculating the root mean square of the signals and the median frequency of the power spectrum and the autocorrelation zero-crossing (Fele-Žorž et al., 2008).

It is known that the underlying physiological mechanisms of biological systems are non-linear processes (Akay, 2001). As the uterus is composed of billions of intricately interconnected cells whose responses are non-linear, it may be regarded as a complex, non-linear dynamic system. To analyze the outputs of such a system, non-linear signal processing techniques are applicable. Therefore, one can hypothesize that non-linear signal processing techniques may yield better results in analysis of the EHG than linear ones (Naeem et al., 2013). These techniques included time reversibility and approximate entropy (Hassan et al., 2010), another research estimated Max-Lyapunov exponent, correlation dimension and sample entropy (Fele-Žorž et al., 2008).

In this research, a combination between most of these previous techniques is done as some of the linear techniques and some of non-linear techniques are used in order to estimate their ability to recognize uterine EMG records of term and preterm deliveries using artificial neural networks ANNs.

The following linear features are chosen: the mean power frequency, the root mean square, the peak and median frequencies of the power spectrum and the autocorrelation zero-crossing and taking them into a suitable ANN to calculate the classifier parameters; while the non-linear features are: time reversibility, approximate entropy, the Max-Lyapunov exponent, the correlation dimension, phase space reconstruction based on the derivatives approach and singular spectrum analysis, adjusted amplitude Fourier transform and the sample entropy of the signal and also taking them into the same ANN to calculate the classifier parameters. Finally, compare the results to estimate the ability of linear and non-linear techniques to differentiate uterine EMG records of term and preterm deliveries. It is expectable

that the classification method depends on the nonlinear uterine EMG features gives better results.

In addition that, using a linear parameter discrete cosine transform (DCT) which depends on the spectral characteristics and the frequency contents of the uterine EMG signals and is taken as an additional linear feature to the same ANN to illustrate the difference between the traditional previous classification method and the proposed ones. The results show that this method is better than the linear method but the non-linear one is still the best.

MATERIALS AND METHODS

Database description

This research uses Term Preterm ElectroHysteroGram Data Base (TPEHG DB). The records were obtained during regular check-ups either around the 22nd week or around the 32nd week of gestation at the University Medical Centre Ljubljana, Department of Obstetrics and Gynecology, Slovenia (PhysioBank database Website [Online], 2011) and used for studies by Ivan Verdenik (Fele-Žorž et al., 2008). The DB used contains 300 uterine EMG records of which:

- (i) 262 records were obtained where delivery was on term
- (ii) 143 before the 26th week of gestation.
- (iii) 119 during-after the 26th week of gestation.
- (iv) -38 records were obtained during pregnancies which ended prematurely
- (v) 19 before the 26th week of gestation.
- (vi) 19 during-after the 26th week of gestation.

Each record is composed of three channels, recorded from four electrodes as shown in Figure 1. The differences in the electrical potentials of the electrodes produced three channels: S1 = E2-E1, S2 = E2-E3 and S3 = E4-E3.

In this paper, used records were digitally filtered using band pass filter (0.3 to 3 Hz) neglecting either these records were taken after or before the 26th week of gestation but the research uses them generally to make a classification into two classes, term and preterm signals.

Feature extraction

Linear features

Mean power frequency: The mean power frequency (MPF) is the frequency at which the average power within the epoch is reached (Frequency signal analysis-BIOPAC Systems Inc. [Online] (2012) and computed from the power spectral density (PSD) of the signal obtained by Welch's averaged periodogram method (Hassan et al., 2010).

Peak frequency: The peak frequency is the frequency at which the maximum power occurs during the epoch (Maner and Garfield, 2007). For each signal, $x(t)$, the peak frequency, f_{max} , is calculated as following (Akay, 2001):

$$f_{max} = \arg \left(\frac{f_s}{N} \max_{i=0}^{N-1} P(i) \right) \quad (1)$$

Where f_s and N denotes the sampling frequency and the number of samples, respectively. P is the frequency-power spectrum.

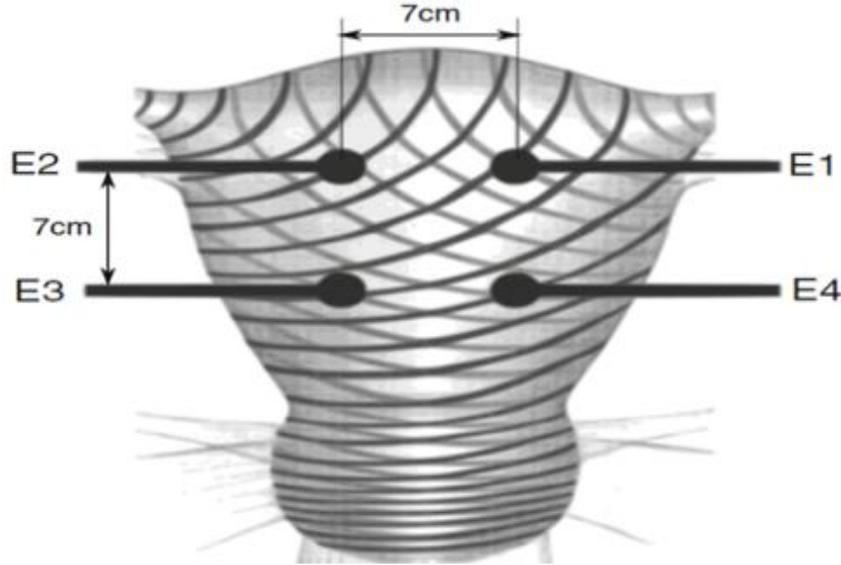


Figure 1. The placement of the electrodes on the abdomen, above the uterine surface.

Root mean square: The root mean square value (RMS) of a signal, $x(i)$, with length N is the root of the mean of the squares of all samples in a signal (Fele-Žorž et al., 2008):

$$RMS = \sqrt{\frac{1}{N} \sum_{i=0}^{N-1} x(i)^2} \quad (2)$$

Median frequency: The median frequency was defined as the frequency just above where the sums of the parts above and below in the frequency-power spectrum, P , are the same (Fele-Žorž et al., 2008) or it is the frequency at which 50% of the total power within the epoch is reached (Frequency signal analysis-BIOPAC Systems Inc. [Online] (2012).

Autocorrelation zero-crossing: The autocorrelation zero-crossing, $\tau_{R_{xx}}$, is defined as the first zero-crossing starting at the peak in the autocorrelation, $R_{xx}(\tau)$, of the signal $x(t)$ (Fele-Žorž et al., 2008):

$$R_{xx}(\tau_{R_{xx}}) = 0, R_{xx}(\tau) = \sum_{i=0}^{N-1} x(i)x(\tau+i) \quad (3)$$

Discrete cosine transform (DCT): DCT generates real spectrum of a real signal and thereby avoids redundant data and computation. The DCT of a real sequence, $x(n)$, with length N is defined as:

$$y(k) = w(k) \sum_{n=1}^N x(n) \cos \frac{\pi(2n-1)(k-1)}{2N}, k = 1, \dots, N \quad (4)$$

$$w(k) = \begin{cases} 1/\sqrt{N} & , k = 1 \\ \sqrt{2/N} & , 2 \leq k \leq N \end{cases} \quad (5)$$

Non-linear features

Approximate entropy: As mentioned by Pincus (1991) the approximate entropy, ApEn, is defined as a measure that quantifies the regularity and predictability of the signals. The ApEn value is low for regular time series and high for complex, irregular ones (Hassan et al., 2010). This paper uses the method applied in [6] to compute the ApEn.

$$ApEn(m, r, \tau, N) = \Phi^m(r) - \Phi^{m+1}(r) \quad (6)$$

$$\Phi^m(r) = \frac{1}{N - (m-1)\tau} \sum_{i=1}^{N-(m-1)\tau} \log C_i^m(r) \quad (7)$$

Where r , the filter parameter value, is $r=0.2 \cdot SD$, SD is the standard deviation of the signal and $C_i^m(r)$ is the correlation sum.

Sample entropy: The sample entropy, SampEn, is a measure of complexity that can be easily applied to any type of time series data. It is conceptually similar to approximate entropy (ApEn), but SampEn does not depend on the data size as much as ApEn does (Lee, 2010).

Phase space reconstruction

Reconstruction based on derivative approach: The phase space dimension or reconstruction dimension, usually symbolized by letter d or E , is defined as the number of states that can be displayed in phase space. Phase space in d - dimensions will display a number of points $\{x(n)\}$ of the system, where each point is given by:

$$\vec{x}(n) = [x(n), x(n+T), \dots, x(n+(d-1)T)] \quad (8)$$

Where n is a moment in time of a system variable, and T is a period between two consecutive measurements of the variable. There is a problem with the phase space graphical presentation, if it has more

than three dimensions (Jovic and Bogunovic, 2007). Phase space reconstruction is a standard procedure when analyzing chaotic systems. It shows the trajectory of the system in time. Here the phase space reconstruction is obtained by a method based on derivatives approach (Packard et al., 1980) that is, by taking $\vec{x}^T(n)$, $\vec{x}''(n)$,...etc.

Reconstruction based on the singular spectrum approach (SSA): SSA is a method of decomposition of time-series into the sum of a small number of independent components. The basic SSA algorithm has two stages: decomposition and reconstruction. The decomposition stage requires embedding and singular value decomposition (SVD). Embedding decomposes the original time series into the trajectory matrix; SVD turns the trajectory matrix into the decomposed trajectory matrices which will turn into the trend, seasonal, monthly components, and white noises according to their singular values. The reconstruction stage demands the grouping to make subgroups of the decomposed trajectory matrices and diagonal averaging to reconstruct the new time series from the subgroups that is, the concept of SSA consists of four steps: embedding, SVD, grouping, and diagonal averaging and all these steps mentioned in details (Yung, 2009).

Amplitude adjusted Fourier transform (AAFT)

The AAFT algorithm generates surrogate data set and this paper takes the same steps mentioned in Garfield et al. (2005) to create the AAFT for uterine EMG data. The idea is to first rescale the value in the original time series so they are Gaussian. Then the FT algorithm can be used to make surrogate time series which have the same Fourier spectrum as the rescaled data. Finally, the Gaussian surrogate is then rescaled back to have the amplitude distribution as the original time series.

Time reversibility

A time series is said to be reversible only if its probabilistic properties are invariant with respect to time reversal (Hassan et al., 2010). In this research a simple equation described in Hassan et al. (2010) is used to calculate the time reversibility T_R for a signal X :

$$T_R(\tau) = \frac{1}{N-\tau} \sum_{n=\tau+1}^N (X_n - X_{n-\tau})^2 \quad (9)$$

Where N is the signal length and in this paper we used $\tau=1$. Time irreversibility can be taken as a strong signature of nonlinearity.

Maximal Lyapunov exponent and correlation dimension

To calculate both parameters a practical method in Fele-Žorž et al. (2008) is used which is based on input data, represented in a phase space. The phase space is a construct which demonstrates or visualizes the changes of the dynamical variables of a system. For any time series, the phase space which is the same as original phase space of the system is reconstructed by using time-delayed samples as the coordinates of the new system.

The maximal Lyapunov exponent estimates the amount of chaos in a system and represents the maximal velocity with which different, almost identical states of the system, diverge (Fele-Žorž et al., 2008). Then the Lyapunov exponent can be calculated as the following equation as the maximum Lyapunov exponent, λ , is a measure of how fast a trajectory converges from a given point into some other trajectory:

$$\lambda = \lim_{t \rightarrow \infty} \lim_{\|\Delta y_0\| \rightarrow 0} \frac{1}{t} \log \frac{\|\Delta y_t\|}{\|\Delta y_0\|} \quad (10)$$

Where $\|\Delta y_0\|$ represents the Euclidean distance between two states of the system at some arbitrary time t_0 and $\|\Delta y_t\|$ represents the Euclidean distance between the two states of the system at some later time t .

In chaos theory, the correlation dimension, D_{corr} , is a measure of the dimensionality of the space occupied by a set of random points, often referred to as a type of fractal dimension. It is proportional to the probability of the distance between two points on a trajectory being less than some r [8]:

$$D_{corr} = \lim_{r \rightarrow 0} \frac{\log(C(r))}{\log(r)} \quad (11)$$

Where $C(r)$ is the correlation integral.

Principal component analysis

PCA is an orthogonal linear transformation that transforms the original time series by projecting it to a new set of coordinates in order of decreasing variance. The transformation is by definition an optimum transformation in the least squares sense. This method reduces the dimension of the representation space to keep only the most important information represented in fewer dimension space domains (The iPredict website [Online], 2012).

The PCA is applied for the parameters (AAFT, derivative phase space reconstruction, SSA and DCT) to obtain only ten features from 24001, 24000, 23999 and 24001 features respectively. To explain why only 10 features are chosen, two notes must be taken in consideration. First, the principal component coefficients for a $M \times N$ matrix - where M is the number of patterns and N is the number of features- are a $N \times N$ matrix. Second, the condition for classifier pattern matrix construction is that the term and preterm signals must have the same number of features. An explanation example, a 150×24001 matrix is the PCA input obtained from applying the AAFT on the training term signals and a 19×24001 matrix from the training preterm ones. After applying the PCA, a 24001×24001 matrix is obtained from the term AAFT spectrum and a 24001×24001 matrix from the preterm one. Then we can take only the first ten most principal components that have the highest variance into the ANN pattern matrix.

Artificial neural network

An important step is the classification step and actually, in this research, three types of artificial neural network ANNs are used to reach best results. One of them is unsupervised learning method (Kohonen self-organizing network) (Goyal and Goyal, 2011) and the others are supervised learning methods (feed-forward back propagation network and trainable cascade-forward back propagation network) (Goyal and Goyal, 2011). Each one of the previous networks is used for linear and non-linear features separately and gives its own parameters to compare. In the research a training data of 150 term signals and 19 preterm ones are used in addition to testing data of 111 term signals and 19 preterm ones.

RESULTS AND DISCUSSION

For each classifier, some parameters can be calculated to evaluate its performance. These parameters are:

Table 1. Resulting values of the used classifiers (linear).

Used classifier	Resulting values			
	<i>FP</i>	<i>FN</i>	<i>TP</i>	<i>TN</i>
Kohonen	12	52	7	59
Feed-forward	19	0	0	111
Cascade-forward	19	0	0	111

Table 2. Comparison of calculated parameters for the three used classifiers (linear).

Used classifier	Calculated parameter				
	<i>FPR</i>	<i>FNR</i>	<i>TPR</i>	<i>TNR</i>	<i>ACC (%)</i>
Kohonen	0.17	0.88	0.12	0.83	50.7
Cascade-forward	0.15	0.00	0.00	0.85	85.3
feed-forward	0.15	0.00	0.00	0.85	85.3

$$FPR = \frac{FP}{FP + TN}, \quad FNR = \frac{FN}{FN + TP} \quad (12)$$

$$TPR = \frac{TP}{TP + FN}, \quad TNR = \frac{TN}{TN + FP} \quad (13)$$

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (14)$$

Where *TP*, *TN*, *FP* and *FN* stand respectively for True Positive, True Negative, False Positive and False Negative values. The values of *FPR*, *FNR*, *TPR*, *TNR* and *ACC* stand respectively for False Positive Rate, False Negative Rate, True Positive Rate (Sensitivity), True Negative Rate (Specificity) and Accuracy.

In the linear method as shown in Table 1 the Kohonen network can recognize 59 signals from 111 and seven signals from 19 for term and preterm uterine EMG respectively while the feed-forward and the cascade-forward networks cannot recognize any preterm uterine EMG signal. From these values Table 2 can be created and it indicates that the Kohonen network has a low sensitivity (0.12) which is higher than the feed-forward and cascade-forward networks' sensitivities (0.00) and it also has high FNR (0.88) while it is the opposite for the other classifiers with no FNR (0.00). From these values, we can observe that no one of the three used classifiers can separate between term and preterm deliveries in a perfect way where the Kohonen network can recognize some of the preterm records but also it classifies many term records as preterm ones. On the other hand the feed-forward and the cascade-forward networks cannot recognize any preterm record although they classify all the term records correctly. Figure 2 shows the representation of the results on the ROC graph.

In the non-linear method as shown in Table 3, the Kohonen network can recognize 80 signals from 111 and

nine signals from 19 for term and preterm uterine EMG respectively and the feed-forward network can recognize 107 signals from 111 and seven signals from 19 for term and preterm uterine EMG respectively while the cascade-forward networks can recognize 110 from 111 term signals and ten from 19 preterm ones. From these values, Table 4 can be created and it indicates that the Kohonen network has a low sensitivity (0.22) and high specificity (0.89) but it also has high FNR (0.78) while the feed-forward network has moderate sensitivity (0.64), high specificity (0.90) and low FNR (0.36). The cascade-forward network has high sensitivity (0.91) and high specificity (0.92) which are higher than the other used classifiers and it also has low FNR (0.09) and low FPR (0.08) which are lower than the others. From these values, we can observe that the cascade-forward network gives the best results where the Kohonen network can recognize some of the preterm records but also it classifies many term records as preterm ones. On the other hand the feed-forward network recognizes term and preterm records with lower errors than the Kohonen network but higher than these for the cascade-forward network. Figure 3 shows the representation of the results on the ROC graph.

In the additional linear method as shown in Table 5, the Kohonen network can recognize 77 signals from 111 and seven signals from 19 for term and preterm uterine EMG respectively and the feed-forward network can recognize 108 signals from 111 and five signals from 19 for term and preterm uterine EMG respectively while the cascade-forward networks can recognize 109 from 111 term signals and eight from 19 preterm ones. From the above values, Table 6 can be created and it indicates that the Kohonen network has a low sensitivity (0.17) and high specificity (0.87) but it also has high FNR (0.83) while the feed-forward network has moderate sensitivity (0.63), high specificity (0.89) and moderate FNR (0.37).

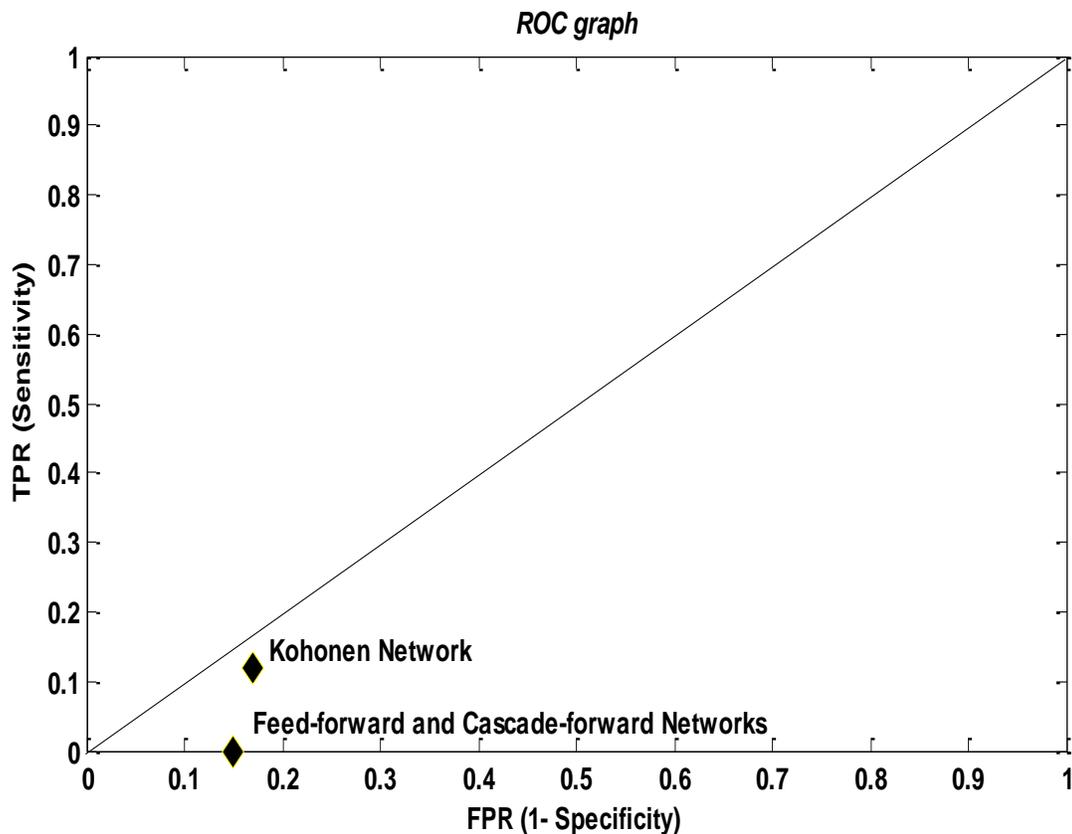


Figure 2. Representation of the results on ROC (Linear).

Table 3. Resulting values of the used classifiers (non-linear).

Used classifier	Resulting values			
	FP	FN	TP	TN
Kohonen	10	31	9	80
Feed-forward	12	4	7	107
Cascade-forward	9	1	10	110

Table 4. Comparison of calculated parameters for the three used classifiers (non-linear).

Used classifier	Calculated parameter				
	FPR	FNR	TPR	TNR	ACC (%)
Kohonen	0.11	0.78	0.22	0.89	68.5
Cascade-forward	0.10	0.36	0.64	0.90	87.7
Feed-forward	0.08	0.09	0.91	0.92	92.3

For the cascade-forward network, a sensitivity of (0.80) is obtained which is better than that in the previous linear method as it can recognize eight preterm signals from 19 ones and also it has low FNR (0.20) which is higher than

that shown in Table 2 as it cannot recognize two term signals from 111 ones. In spite of that, the non-linear method shown in Table 4 is still the best. Figure 4 shows the representation of the results on the ROC graph.

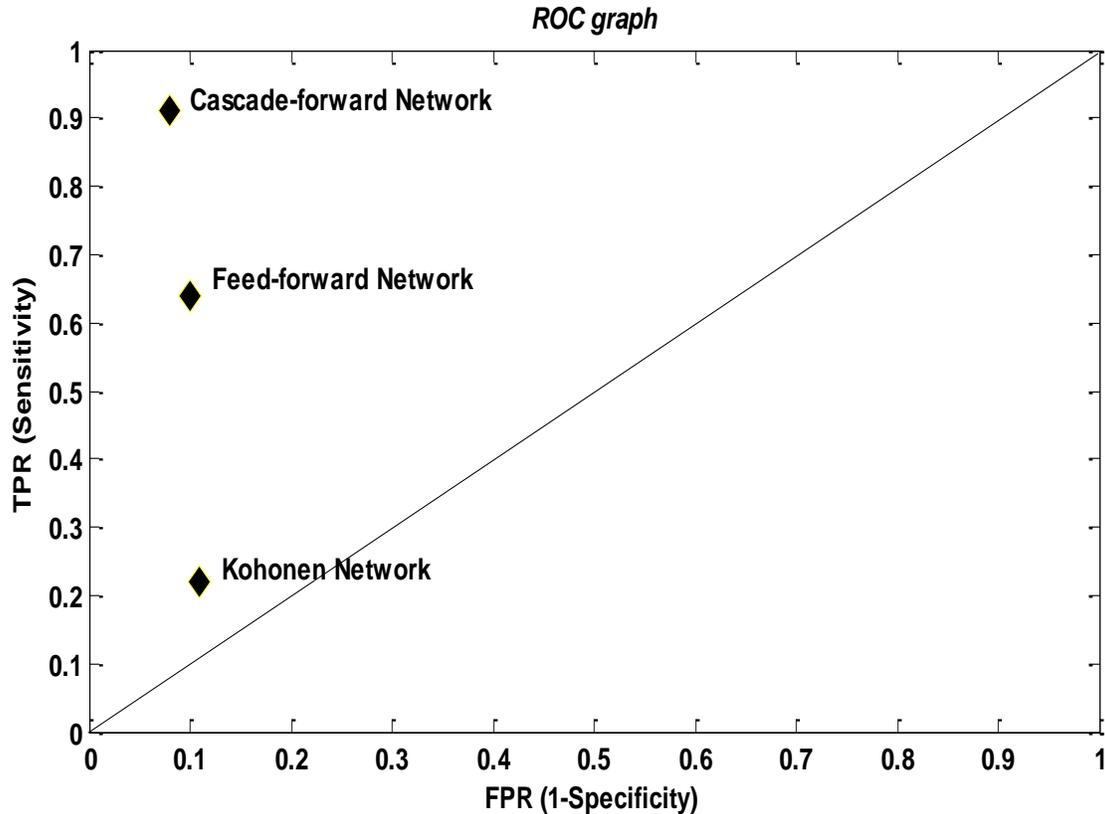


Figure 3. Representation of the results on ROC (Non- Linear).

Table 5. Resulting values of the used classifiers (linear+ DCT feature).

Used classifier	Resulting values			
	FP	FN	TP	TN
Kohonen	12	34	7	77
Feed-forward	14	3	5	108
Cascade-forward	11	2	8	109

Table 6. Comparison of calculated parameters for the three used classifiers (linear+ DCT feature).

Used classifier	Calculated parameter				
	FPR	FNR	TPR	TNR	ACC (%)
Kohonen	0.13	0.83	0.17	0.87	64.6
Cascade-forward	0.11	0.37	0.63	0.89	86.9
feed-forward	0.09	0.20	0.80	0.91	90

Conclusion

From the above results presented in this research some observations can be inferred. Firstly, using non-linear parameters of uterine EMG signals as ANN features can

separate between term and preterm uterine EMG signals with results which are better than these for linear ones even if a spectral characteristic linear parameter (DCT) is used. Secondly, to get best classification accuracy with minimum error, you should use the trainable cascade-

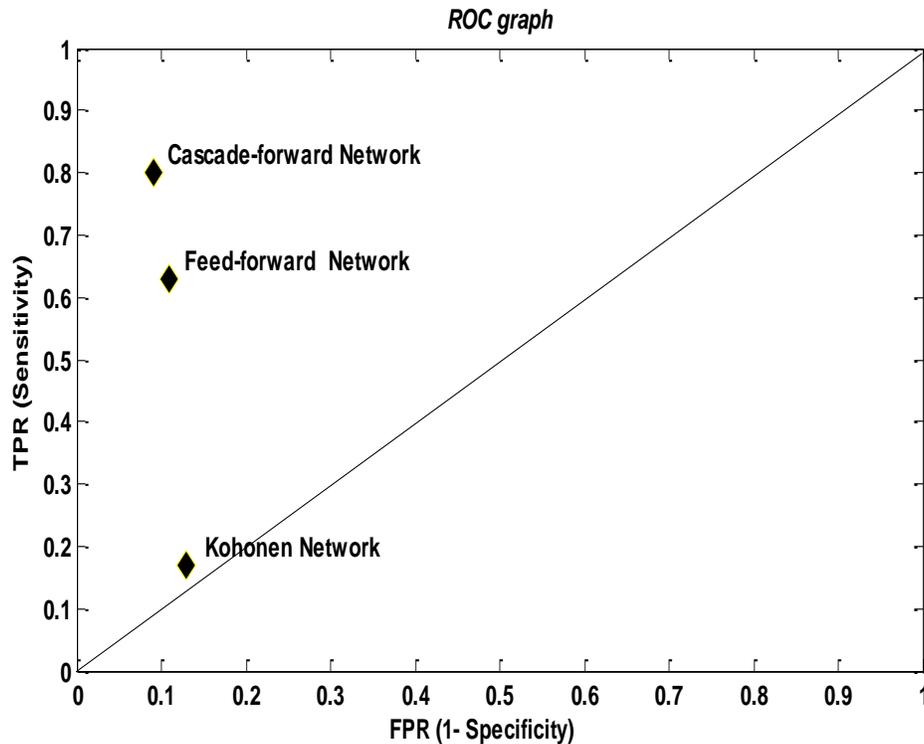


Figure 4. Representation of the results on ROC (Linear+ DCT feature).

forward back propagation network. Finally, the Kohonen network gives worse results in using both linear and non-linear parameters.

Conflict of Interest

The authors have not declared any conflict of interest.

REFERENCES

- Diab MO, El-Merhie A, El-Halabi N, Khoder L (2010). "Classification of uterine EMG signals using supervised classification method," *JbiSE*, 3:837-842.
- Diab MO, Marque C, Khalil MA (2007). "Classification for Uterine EMG Signals: Comparison between AR Model and Statistical Classification Method," *IJCC*.
- Moslem B, Diab MO, Khalil MA, Marque C (2012). "Combining data fusion with multiresolution analysis for improving the classification accuracy of uterine EMG signals," *EURASIP Journal on Advances in Signal Processing*.
- Garfield RE, Maner WL, MacKay LB, Schlembach D, Saade GR (2005). "Comparing uterine electromyography activity of antepartum patients versus term labor patients," *Am. J. ObstetGynecol*, 193:23-29.
- Maul H, Maner WL, Olson G, Saade GR, Garfield RE (2004). "Non-invasive transabdominal uterine electromyography correlates with the strength of intrauterine pressure and is predictive of labor and delivery," *J Matern Fetal Neonatal Med*. 15:297-301.
- Hassan M, Terrien J, Alexandersson A, Marque C, Karlsson B (2010). "Nonlinearity of EHG signals used to distinguish active labor from normal pregnancy contractions," In Proceedings of the 32nd Annual International Conference of the IEEE EMBS: 31 August- 4 September 2010; Buenos Aires, Argentina.
- Maner WL, Garfield RE (2007). "Identification of human term and preterm labor using artificial neural networks on uterine electromyography data," *Ann. Biomed. Eng.* 35:465-473.
- Fele-Žorž G, Kavšek G, Novak-Antolič Z, Jager F (2008). "A comparison of various linear and non-linear signal processing techniques to separate uterine EMG records of term and pre-term delivery groups," *Med. Biol. Eng. Comput.* 46:911-922.
- Akay M (2001). *Nonlinear biomedical signal processing. In dynamic analysis and modeling*, IEEE Inc. New York. 2.
- Naeem SM, Ali AF, Eldosoky MA (2013). "Comparison between Using Linear and Non-linear Features to Classify Uterine Electromyography Signals of Term and Preterm Deliveries," In Proceedings of the 30th National Radio Science Conference NRSC: 16-18 April 2013; Cairo, Egypt. pp. 488-498.
- Pincus SM (1991). "Approximate entropy as a measure of system complexity," *Proc. Natl. Acad. Sci. U S A*, 88:2297-301.
- Lee K (2010). File exchange – Mathworks. [Online]. Available: <http://www.mathworks.com/matlabcentral/fileexchange/35784-sample-entropy>.
- Jovic A, Bogunovic N (2007). "Feature extraction for ECG time-series mining based on chaos theory," In Proceedings of the ITI 29th Int. Conf. on Information Technology Interfaces: 25-28June 2007; Cavtat, Croatia.
- Packard NH, Cruchfield JP, Farmer JD, Shaw RS (1980). "Geometry from a Time Series," *Physical Rev. Lett*, 45:712-715.
- Yung NK (2009). *Singular Spectrum Analysis*, Master's Thesis. University of California, Los Angeles.
- Nielsen F (2001). *Neural Networks – algorithms and applications*, Niels Brock Business College.
- Goyal S, Goyal GK (2011). Cascade and Feed-forward Backpropagation Artificial Neural Network Models for Prediction of Sensory Quality of Instant Coffee Flavoured Sterilized Drink," *Can. J. Artificial Intelligence, Mach. Learn. Pattern recognition*.