

A Medical Decision Support System Based on Support Vector Machines and the Genetic Algorithm for the Evaluation of Fetal Well-Being

Hasan Ocak

Received: 9 August 2012 / Accepted: 29 December 2012 / Published online: 16 January 2013
© Springer Science+Business Media New York 2013

Abstract A new scheme was presented in this study for the evaluation of fetal well-being from the cardiocotogram (CTG) recordings using support vector machines (SVM) and the genetic algorithm (GA). CTG recordings consist of fetal heart rate (FHR) and the uterine contraction (UC) signals and are widely used by obstetricians for assessing fetal well-being. Features extracted from normal and pathological FHR and UC signals were used to construct an SVM based classifier. The GA was then used to find the optimal feature subset that maximizes the classification performance of the SVM based normal and pathological CTG classifier. An extensive clinical CTG data, classified by three expert obstetricians, was used to test the performance of the new scheme. It was demonstrated that the new scheme was able to predict the fetal state as normal or pathological with 99.3 % and 100 % accuracy, respectively. The results reveal that, the GA can be used to determine the critical features to be used in evaluating fetal well-being and consequently increase the classification performance. When compared to widely used ANN and ANFIS based methods, the proposed scheme performed considerably better.

Keywords Fetal Heart Rate (FHR) · Cardiocotogram (CTG) · Diagnosis · Support Vector Machines (SVM) · Genetic Algorithm (GA)

Introduction

The cardiocotography (CTG), introduced into clinical practice in the late 1960s, is a non-invasive and a low-price technique for assessing fetal status. The baby's fetal heart rate (FHR) and the mother's uterine contractions (UC) are recorded on a paper trace known as cardiocotograph. This is achieved by using a Doppler ultrasound transducer to monitor FHR and a pressure transducer to monitor UC. Since the fetus is not available for direct observations, CTG is widely used by obstetricians for assessing fetal well-being. It allows for the early identification of a pathological state (i.e. congenital heart defect, fetal distress or hypoxia, etc.) and assists the obstetrician to predict future complications and intervene before there is an irreversible damage to the fetus.

CTG recordings are visually analyzed by experts which makes the interpretations subjective and not reproducible. This is the originating point of the debates on whether the CTG monitoring is effective or not, especially in low-risk pregnancies [1]. In addition, it was shown that over-reliance on the test has led to increased misdiagnosis of fetal distress and hence increased the number of cesarean deliveries [2]. Although guidelines were developed for evaluating the CTG recordings, inter-observer and intra observer variations in the interpretations and the misdiagnosis of fetuses due to different experience level of the experts still constitute a big dilemma [3–5]. To increase the utility and the effectiveness of the CTG monitoring and minimize the inconsistencies in the interpretations, a lot of effort has been devoted by

H. Ocak (✉)
Kocaeli University, Department of Mechatronics Engineering,
Umuttepe Campus,
41380 Kocaeli, Turkey
e-mail: hocak@kocaeli.edu.tr

H. Ocak
e-mail: hxo9@cwru.edu

H. Ocak
Case Western Reserve University, Department of Electrical
Engineering and Computer Science, 10900 Euclid Avenue,
Cleveland, OH 44106, USA

researchers from both medical and engineering backgrounds towards developing computerized systems for the analysis and the classification of CTG recordings.

Various techniques have been proposed in the literature for the prediction and classification of fetal state from CTG recordings. In a study presented in [6], FHR signals recorded between 38 and 40 week gestation were classified using a statistical method. Auto-regressive moving-average (ARMA) model parameters of 3 min FHR segments were used as features for a retrospective classification of the FHR segments, using a linear discriminant function. Compared to visual classification results, the computer/observer classification agreement was found to be 85 %. In another study [7], hidden Markov models (HMMs) were used to identify the fetal state as either hypoxic (low supply of oxygen) or normal. Both time and frequency domain features were used to train two separate HMM models. A maximum overall classification rate of 83 % was reported after testing with various HMM configurations. The authors of [8] used time domain parameters like mean, standard deviation, approximate entropy and frequency domain parameters obtained from the power spectral density (PSD) of the FHR signals along with fuzzy inference systems (FIS) to predict two very common fetal pathological conditions: Intra-Uterine Growth Retardation (IUGR) and type-I Diabetes. Support vector machines (SVM) were used in another study [9] to identify fetuses compromised and suspicious of developing metabolic acidosis. The method proposed in [10] uses nonlinear features like fractal dimension, approximate entropy and Lempel Ziv complexity along with Naive Bayes and SVM classifiers to classify FHR signals as normal or pathological. Results based on 189 recordings revealed an overall sensitivity and specificity of 70 %. Methods based on wavelets [11–13], artificial neural-networks (ANN) [14, 15] and adaptive neuro-fuzzy inference systems (ANFIS) [16] have also been proposed. A comparative study of neural networks and statistical methods is presented in [17].

In the first phase of the study, unlike most of the methods presented in the literature which use features extracted only from FHR signals to classify fetal states, features extracted from both FHR and UC data were used in an SVM based classification scheme. SVM has been one of the most widely used and popular classification methods since its introduction to the literature in the early 1990s. In the second phase of the study, the genetic algorithm (GA) was employed to reduce the number of features and find the optimal feature subset that maximizes the classification performance of the SVM classifier. The GA is an adaptive heuristic search algorithm inspired by Darwinian natural selection and genetics in biological systems. The GA has not only a higher probability of locating the global optimum solution in a complex search space than the traditional techniques, but is also less sensitive to initial conditions. It has been applied

in a wide range of real-world applications, including but not limited to machine learning, data mining, financial forecasting and medical decision support systems [18–22]. The new scheme was tested with a large clinical data set that consists of 1831 CTG recordings, out which 1655 were classified as normal and the remaining 176 were classified as pathological by a consensus of three expert obstetricians. The results demonstrated that the GA based optimal feature selection scheme increased the classification performance of the SVM classifier especially for the pathological CTG while reducing the number of features from 21 to 13.

Technical background

Support Vector Machines (SVM)

Support vector machines (SVM), originally introduced to the literature in 1995 by Vapnik [23], are a relatively new learning method that classifies any data into two classes by finding the optimal hyperplane that separates the data. SVM has become one of the most widely used classification technique since its introduction and has been successfully applied in many medical decision support systems [24–27].

Let's consider N training vectors, where each input \mathbf{x}_i belongs to one of the two classes $y_i = -1$ or $+1$. Thus, it is assumed that the training data is of the form

$$\{\mathbf{x}_i, y_i\}, \quad i = 1 \dots N, \quad \mathbf{x} \in \mathbb{R}^d, \quad y \in \{-1, 1\} \quad (1)$$

In the case of linearly separable data, a hyperplane can be drawn to separate the two classes as illustrated in Fig. 1. This hyperplane can be defined as

$$\mathbf{x} \cdot \mathbf{w} - b = 0 \quad (2)$$

where \mathbf{w} is a vector normal to the hyperplane and b is a scalar. The perpendicular distance between the hyperplane

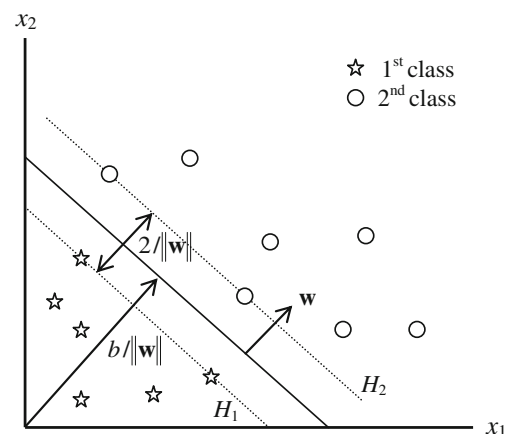


Fig. 1 The optimal hyperplane that separates two classes

and the origin is given by $b/\|\mathbf{w}\|$. One can now define two more hyperplanes, H_1 and H_2 , parallel to the separating hyperplane in a way that they separate the two classes and there are no data points between them.

$$\mathbf{x}_i \cdot \mathbf{w} - b = -1 \text{ for } H_1 \tag{3}$$

$$\mathbf{x}_i \cdot \mathbf{w} - b = +1 \text{ for } H_2 \tag{4}$$

The distance between these two planes are referred to as SVM's margin. By simple geometry, the margin can be computed as $2/\|\mathbf{w}\|$. The goal is to find the separating hyperplane that maximizes this margin while it is equidistant from both H_1 and H_2 . This can be accomplished by minimizing $\|\mathbf{w}\|$. However, in order to prevent the data points from falling into the margin, the following constraints should also be satisfied

$$y_i = +1 \Rightarrow \mathbf{x}_i \cdot \mathbf{w} - b \geq 1 \tag{5}$$

$$y_i = -1 \Rightarrow \mathbf{x}_i \cdot \mathbf{w} - b \leq -1 \tag{6}$$

These two constraints can be rewritten as one

$$y_i(\mathbf{x}_i \cdot \mathbf{w} - b) - 1 \geq 0 \tag{7}$$

Therefore, maximization of the margin leads to the following constrained problem

$$\min \frac{1}{2} \|\mathbf{w}\|^2 \text{ s.t. } y_i(\mathbf{x}_i \cdot \mathbf{w} - b) - 1 \geq 0 \tag{8}$$

which can be solved by using Lagrange multipliers

$$L_p(\mathbf{w}, b, \alpha) = \frac{1}{2} \|\mathbf{w}\|^2 - \sum_{i=1}^N \alpha_i [y_i(\mathbf{x}_i \cdot \mathbf{w} - b) - 1] \tag{9}$$

where α is a vector of Lagrange multipliers. The minimization can now be achieved by setting the partial derivatives of the above equation to zero.

$$\frac{\partial L_p}{\partial \mathbf{w}} = 0 \Rightarrow \mathbf{w} = \sum_{i=1}^N \alpha_i y_i \mathbf{x}_i \tag{10}$$

$$\frac{\partial L_p}{\partial b} = 0 \Rightarrow \sum_{i=1}^N \alpha_i y_i = 0 \tag{11}$$

Substitution of Eqs. 10 and 11 back into Eq. 9 gives a new formulation of the problem

$$L_d = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j \mathbf{x}_i \cdot \mathbf{x}_j \text{ s.t. } \alpha_i \geq 0, \sum_{i=1}^N \alpha_i y_i = 0 \tag{12}$$

If $y_i y_j \mathbf{x}_i \cdot \mathbf{x}_j$ is denoted by H_{ij} , then the optimization problem can be restated as

$$\min_{\alpha} \frac{1}{2} \alpha^T \mathbf{H} \alpha - \frac{1}{2} \sum_{i=1}^N \alpha_i \text{ s.t. } \alpha_i \geq 0, \sum_{i=1}^N \alpha_i y_i = 0 \tag{13}$$

which can now be solved by standard quadratic programming techniques.

In case the data is not linearly separable, a nonlinear classifier can be implemented by applying the kernel trick which transforms the data into high-dimensional space where it can be linearly separated. In this method, the dot product, $\mathbf{x}_i \cdot \mathbf{x}_j$, is replaced by a nonlinear kernel function. In this study, the data was not linearly separable and a Gaussian radial basis function (RBF) was used to realize the nonlinear classifier.

$$k(\mathbf{x}_i, \mathbf{x}_j) = e^{-\|\mathbf{x}_i - \mathbf{x}_j\|^2 / 2\sigma^2} \tag{14}$$

The Gaussian RBF and the σ value were selected by testing with various SVM configurations.

Genetic Algorithms (GAs)

A genetic algorithm is a stochastic optimization technique which is successfully applied to wide variety of applications. Unlike traditional gradient based optimization schemes, genetic algorithms (GAs) can be applied to solve problems that are not well suited for conventional gradient based optimization methods, including the problems in which the objective function is highly nonlinear, discontinuous, nondifferentiable or stochastic. In classical optimization methods, a single point is generated by a deterministic computation at each iteration and the sequence of points approaches to an optimal solution. In GA on the other hand, a population of points are stochastically generated at each iteration and the best point in the population approaches to an optimal solution which is more likely to be the global minimum as compared to the optimal points found by derivative-based deterministic algorithms [28, 29].

The critical step in applying the genetic algorithms to an optimization problem is the representation of a potential solution to the problem as an individual or a gene sequence (also known as chromosome). The most common approach is to encode solutions as binary strings: sequences of 1's and 0's. The typical steps involved in the genetic algorithm are presented in Fig. 2. At the beginning of the process, individuals which are the encoded solutions are randomly created. These individuals are independent initial solution candidates and allow complex and very high dimensional spaces to be searched in parallel decreasing the chances of getting stuck at a local minimum which is the most important

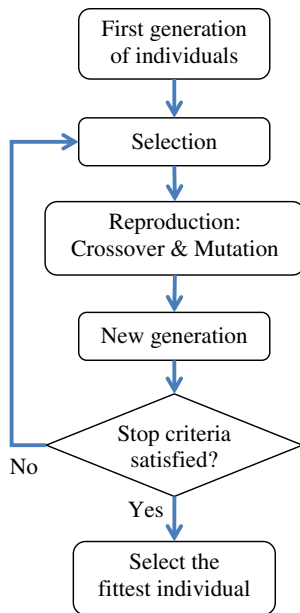


Fig. 2 A typical flow chart of the GA

limitation of gradient based methods. The performance of individuals or chromosomes are measured and assessed by an objective function. Fitness value or scores assigned to individuals by the objective function constitute the basis for applying the survival of the fittest strategy. The ultimate goal is to find the individual or the encoded solution that minimizes this objective function. Based on the fitness scores, the algorithm selects the best two individuals, namely the parents, that will exchange genes to reproduce new individuals, namely the children, for the next generation. Through the gene exchange process, the GA generates better individuals at each evolutionary cycle and eventually finds the best possible individual which is the optimal solution to the objective function. As illustrated in Fig. 3, reproduction is governed by the two genetic operators: crossover and mutation. Crossover is a process in which parents exchange genes whereas mutation

is the random modification of gene(s) in a parent’s chromosome. Both operators are critical and affect the way the search space is explored. Crossover allows new solution regions to be searched while mutation adds random behavior to the search allowing the algorithm to jump to unexplored regions [30]. The evolutionary cycle of the GA resumes for many generations until a termination condition is satisfied. The best gene sequence in the last generation is decoded to obtain the optimal solution for the given problem.

Methods

Clinical data and feature set

The data used in this study was obtained from UCI Machine Learning Repository [31] and originated from a study conducted in University of Porto. The data set consists of fetal heart rate and uterine contraction features extracted from 1831 CTG recordings. The features were obtained using a CTG analysis program SisPorto 2.0 [32]. There are a total of 21 features per CTG recording which are illustrated in Table 1. A consensus classification label was assigned to each of the data as either normal or pathological by three expert obstetricians. Out of the 1831 CTG recordings, 1655 belong to the normal fetal state, and the remaining 176 belong to the pathological state.

SVM based classification

In order to develop an SVM classifier, the available data set which consists of 1831 input codebook vectors, was divided into training and test sets. Out of the 1655 normal CTG data, 827 (50 %) were included in the training and the remaining 828 (50 %) were included in the test data set. Similarly, out of the 176 pathological CTG data, 88 (50 %) were included in the training and the remaining 88 (50 %) were included in

Fig. 3 Genetic operators: crossover and mutation

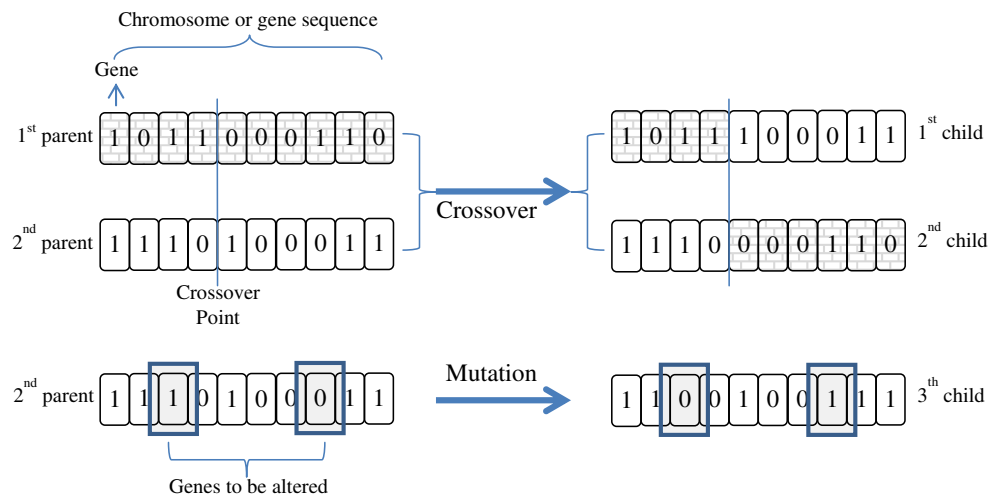


Table 1 Summary of all CTG features used in classification

Feature #	Symbol	Feature info
1	LB	FHR baseline (beats per minute)
2	AC	# of accelerations per second
3	FM	# of fetal movements per second
4	UC	# of uterine contractions per second
5	DL	# of light decelerations per second
6	DS	# of severe decelerations per second
7	DP	# of prolonged decelerations per second
8	ASTV	percentage of time with abnormal short term variability
9	MSTV	mean value of short term variability
10	ALTV	percentage of time with abnormal long term variability
11	MLTV	mean value of long term variability
12	Width	width of FHR histogram
13	Min	minimum of FHR histogram
14	Max	maximum of FHR histogram
15	Nmax	# of histogram peaks
16	Nzeros	# of histogram zeros
17	Mode	histogram mode
18	Mean	histogram mean
19	Median	histogram median
20	Variance	histogram variance
21	Tendency	histogram tendency: -1=left assymmetric; 0=symmetric; 1=right assymmetric

the test data set. The input vectors included in the training and test data sets were selected based on the data indices. The data with even indices were included in the training data set whereas data with odd indices were included in the test data set.

The data used in this study was not fully separable using a linear SVM. Therefore, kernel functions were used to transform the data into a new feature space where a hyperplane could separate the data. By testing with various kernel functions, a Gaussian radial basis function (RBF) was selected as the kernel function. By trial and error, the σ value for the RBF was set to 2.

To account for the misclassified data (points that fall into the SVM's margin) a penalty factor is added to the objective function given in Equation 8. This penalty factor is adjusted by a constant, C , known as the box constraint of the SVM. It is called the box constraint since it keeps the allowable values of Lagrange multipliers in Eq. 13 in a bounded region, a "box", such that $C \geq \alpha_i \geq 0$. C controls the degree of tolerance to the misclassified data. Larger values of C lower the generalization ability of the SVM leading to overfitting of the training data. By trial and error, the boxconstraint for the SVM was set to 0.3. SVM-based classification of the data was implemented in MATLAB.

GA based optimal feature selection

Finding the optimum feature subset that maximizes the classification rates of normal and pathological CTG is a combinatorial problem. The GA, having proven performance in solving such problems, was used to select the features that best correlate with the fetal state. After testing with various GA configurations, the parameters were tuned as given in Table 2. The number of individuals within the population was determined to be 50. The solutions were coded as chromosomes with 21 genes, each having a value of 0 or 1. The 21 genes characterize the 21 features in the feature pool. A gene having a value of 1 implies that the feature subset encoded by the chromosome contains the corresponding feature. The evolutionary process started with the first generation of individuals that were randomly created. The objective function that was used to measure the fitness or the performance of an individual was set as,

$$f = w_1e_n + w_2e_p + w_3N_f \tag{15}$$

where w_1 , w_2 and w_3 are the weights, e_n and e_p are the misclassification rates of the SVM classifier in percentages for the normal and the pathological states, respectively. N_f is the number of 1's in the gene sequence which is equal to the number of features in the encoded feature subset. Only the test data was utilized to evaluate the fitness function for a given individual. As either w_1 or w_2 increased, the penalty for misclassifying an input vector that belongs to the corresponding class would increase as well. Consequently, the GA would give more effort to minimizing the misclassification rate for that class as compared to the other one. Similarly, the penalty for adding an extra feature to the feature subset increases as k_3 is increased. Hence, the GA would attempt to reduce the number of features. In this study, by trial and error w_1 , w_2 and w_3 were set to 1, 1 and 0.1, respectively.

Two individuals with the highest scores, named as elites, in each generation were guaranteed to survive to the next generation. The elite genes were not modified in any way. The reproduction process for generating the children of the new generation was governed by two genetic operators: crossover and mutation. Crossover and mutation rates were set to 80 % and 20 %, respectively. The evolutionary cycle

Table 2 GA configuration

Population size	50
# of genes per individual	21
# of elites	2
crossover percentage	80 %
mutation percentage	20 %
max. # of generations	100
stop criteria	If no improvements for the last 20 generations or for the last 5 min

Table 3 Summary of the classification rates

	Training data set		Test data set	
	Normal CTG	Pathological CTG	Normal CTG	Pathological CTG
Complete feature set used	99.6 %	100 %	97.9 %	97.7 %
GA based optimal feature subset used	99.5 %	100 %	99.3 %	100 %
Improvement	-0.1 %	0 %	1.4 %	2.3 %

was allowed to continue for 100 generations or until the termination condition given in Table 2 was met.

Results and discussion

Results using the complete feature set

As mentioned earlier, the data used in the study was not fully linearly separable. Therefore, Gaussian RBF's were used to transform the data into a new feature space where a hyperplane could separate the data into its two classes. The SVM classifier was implemented in MATLAB. The corresponding classification rates of the SVM are provided in the first row of Table 3. Normal and pathological data in the training set were classified accurately with 99.6 % and 100 %, respectively. The classification rates for the test set were 97.9 % and 97.7 % for the normal and the pathological data, respectively. The results after the GA step are presented in the next section.

Results after selecting the optimal feature set using the GA

The classification performance of the SVM can be enhanced by eliminating irrelevant features and keeping the ones that best correlate with the fetal state. The GA was utilized for that purpose. The parameters of the GA must be set

intelligently since they affect the way the complex optimization surface is explored. New regions are added to the search through the crossover operation. Thus, the number of regions explored in the search space increases as the crossover rate increases. Mutation on the other hand is a critical operator that introduces random behavior to the search. It allows unexplored new solution regions to be discovered preventing premature convergence of the algorithm. By trial and error crossover and mutation rates were set to 80 % and 20 %, respectively.

The stop criterion for the evolutionary cycle was set such that the algorithm will halt when either the maximum number of generations (set as 100) is reached or the best individual in the population does not improve for the last 5 min or twenty generations. Figure 4 illustrates the fitness scores of encoded solutions for a sample run. The figure illustrates 1-) the fitness scores of the fittest individual (bottom bars); 2-) the fitness scores of the worst individual (top bars) and 3-) the population's mean fitness score (line plot in the middle) as a function of the evolutionary cycle. It can be observed that, over successive generations the population evolved toward a set of more optimal solutions. New generations of individuals, on average, contain more good genes than a typical solution in a previous generation. Convergence was achieved at the end of the 32nd generation.

Since the search space is very complex and there are a lot of hills and valleys, the GA does not return the same optimal solution at each run. Besides, there might be more than one

Fig. 4 Best, worst and the mean fitness scores within the population as a function of the generation

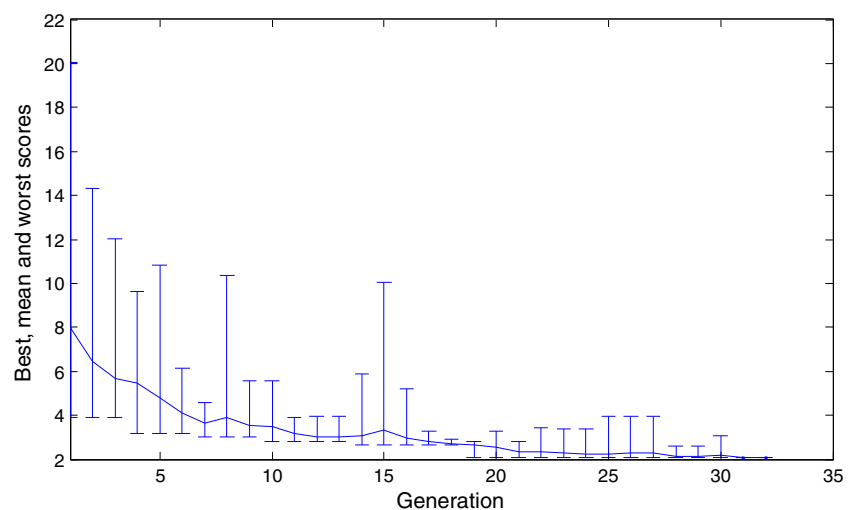
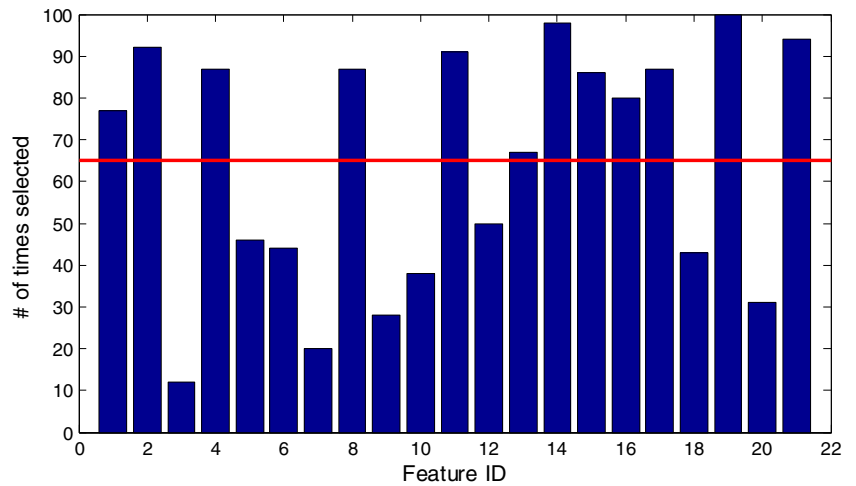


Fig. 5 Number of times features are selected during the 100 different run of the GA



feature subset that would give the best possible classification rate that can be achieved by the algorithm. Therefore, to identify the features that are vital to the classification performance, the GA was ran for 100 different instants. Figure 5 depicts the number of times each feature was selected during these instances. A horizontal red line is drawn in the figure at $y=65$. The 12 features above this red line were selected at least 65 % of 100 different instances. Table 4 includes a list of the feature subset found by the GA for which the best classification rate was achieved. It can be observed the feature subset includes all the 12 features that are above the red line in Fig. 5. In addition to these features, the feature subset also includes the 10th feature which was selected 38 times out of 100 instances.

Table 3 depicts the classification rates achieved by the SVM classifier when the optimal feature subset provided in

Table 3 was used. The normal and pathological data in the training set were classified accurately with 99.5 % and 100 %, respectively. Compared to the results obtained for the complete feature set, the classification percentages are the same except for a 0.1 % drop in the classification accuracy of the normal CTG. The classification rates for the test set were 99.3 % and 100 % for the normal and pathological data, respectively. Compared to the results obtained for the complete feature set, there was a 1.4 % and 2.3 % increase in the classification performances of the normal and the pathological data, respectively. When an obstetrician incorrectly interprets a normal CTG as pathological, he/she would most likely proceed with cesarean delivery. This would just be a wrong decision on the type of delivery that would not put the mother and the baby in extra danger and can be tolerable. On the other hand, when a pathological CTG (implying the baby is under stress) is mistakenly evaluated as normal, a decision of natural delivery might put both the baby’s and the mother’s life in danger. Therefore, misclassification of pathological CTG as normal should be avoided as much as possible. In this study, only about 0.7 % of the normal fetal states were predicted as pathological whereas none of the pathological states were evaluated as normal. The results are promising since the proposed scheme does not make the critical error of misclassifying the pathological CTG as normal. In short, the GA did not only reduce the number of features from 21

Table 4 Features selected by the GA

Feature #	Symbol	Feature info
1	LB	FHR baseline (beats per minute)
2	AC	# of accelerations per second
4	UC	# of uterine contractions per second
8	ASTV	percentage of time with abnormal short term variability
10	ALTV	percentage of time with abnormal long term variability
11	MLTV	mean value of long term variability
13	Min	minimum of FHR histogram
14	Max	maximum of FHR histogram
15	Nmax	# of histogram peaks
16	Nzeros	# of histogram zeros
17	Mode	histogram mode
19	Median	histogram median
21	Tendency	histogram tendency: -1=left asymmetric; 0=symmetric; 1=right asymmetric

Table 5 Comparison of the proposed scheme with the ANN and ANFIS classifiers

	Test data set	
	Normal CTG	Pathological CTG
Proposed scheme	99.3 %	100 %
ANFIS-based classifier	97.2 %	96.6 %
ANN-based classifier	96.0 %	94.3 %

to 13, but also increased the performance of the classifier by 1.4 % for the normal and 2.3 % for the pathological case. Moreover, the GA helped uncover which parameters or features are useful in assessing the fetal state.

Comparison of the proposed scheme with the ANN and ANFIS

The proposed method was compared to widely used ANN and ANFIS based classifiers. A feedforward backpropagation neural network was used for the ANN based classifier. After testing with various configurations, the best performance was achieved with a network having two hidden layers. There were five neurons in the first hidden layer and three neurons in the second hidden layer. Tangent-sigmoid transfer functions were used in the hidden layers whereas a pure linear transfer function was used in the output layer. The network was trained using Levenberg-Marquardt (LM) algorithm. For the ANFIS classifier, a Sugeno-type FIS structure was generated using subtractive clustering. By trial and error, the cluster radius was determined as 0.6. Both classifiers were implemented in MATLAB using the same training and the test data. The classification results of the ANN and ANFIS based schemes along with the proposed method for the test data are presented in Table 5. ANN and ANFIS based classifiers achieved 96.0 % and 97.2 % correct classification rates for the normal CTG, respectively. On the other hand, the corresponding classification rates for the pathological CTG were 94.3 % and 96.6 %, respectively. The test results reveal that, the proposed scheme outperformed the ANN based method by 3.3 % and the ANFIS based method by 2.1 % in terms of predicting the normal state. In addition, it also outperformed the ANN based method by 5.7 % and the ANFIS based method by 3.4 % in terms of predicting the pathological state. In overall, using SVM with the GA selected CTG and UC based features is shown to be an effective method for evaluating the fetal well-being.

ANN and ANFIS based methods were tested using the full feature set. The optimal feature set as presented in the manuscript was obtained for the SVM classifier. The GA selected features are optimal only when used together with an SVM classifier with the provided configuration. Thus, the classification accuracies of the ANN and ANFIS based classifiers get even worse when the optimal feature set was used instead of the full feature set.

Conclusions

CTG recordings are widely used by obstetricians for detecting fetal abnormalities and taking the necessary action

before it is too late. However, interpretations of the recordings after visual analysis by experts are subjective and not reproducible. In this study, a new method was presented for selecting the optimal feature subset that maximizes the classification percentages of an SVM based normal and pathological CTG classifier. Optimal feature selection was accomplished using the genetic algorithm. Test results based on extensive clinical data proved that the genetic algorithm can be used to determine the important features that could be used in assessing the fetal state. The results reveal that the proposed scheme is favorable to ANN and ANFIS based methods. The main shortcoming of the proposed method is that it does not provide a confidence score besides the fetal state returned by the SVM classifier. When modified to give a confidence score, the proposed scheme could be more practical to be realized on a commercial scale.

Conflict of interest The authors declare that they have no conflict of interest.

References

1. MacDonald, D., Grant, A., Sheridan-Pereira, M., Boylan, P., and Chalmers, I., The Dublin randomized controlled trial of intrapartum fetal heart rate monitoring. *Am. J. Obstet. Gynecol.* 152:524–539, 1985.
2. Goddard, R., Electronic fetal monitoring is not necessary for low risk labours. *Brit. Med. J.* 322:1436–1437, 2001.
3. Rooth, G., Huch, A., and Huch, R., Guidelines for the use of fetal monitoring. *Int. J. Gynaecol. Obstet.* 25:159–167, 1987.
4. Bernardes, J., Costa-Pereira, A., Ayres-de-Campos, D., van-Geijn, H. P., and Pereira-Leite, L., Evaluation of interobserver agreement of cardiocotograms. *Int. J. Gynaecol. Obstet.* 57:33–37, 1997.
5. Palomaki, O., Luukkaala, T., Luoto, R., and Tuimala, R., Intrapartum cardiocotography - the dilemma of interpretational variation. *J Perinat Med* 34:298–302, 2006.
6. Jongsma, H. W., and Nijhuis, J. G., Classification of fetal and neonatal heart rate patterns in relation to behavioural states. *Eur. J. Obstet. Gynecol. Reprod. Biol.* 21:293–299, 1986.
7. Georgoulas, G., Stylios, C. D., Nokas, G., and Groumpos, P. P., Classification of fetal heart rate during labour using hidden Markov models. *Proc IEEE Int Joint Conf. Neural Netw.* 3:2471–2474, 2004.
8. Signorini, M. G., de-Angelis, A., Magenes, G., Sassi, R., Arduini, D., and Cerutti, S., Classification of fetal pathologies through fuzzy inference systems based on a multiparametric analysis of fetal heart rate. *Comput. Cardiol.* 27:435–438, 2000.
9. Georgoulas, G., and Stylios, C. D., Predicting the risk of metabolic acidosis for newborns based on fetal heart rate signal classification using support vector machines. *IEEE Trans. Biomed. Eng.* 53:875–884, 2006.
10. Spilka, J., Chudacek, V., Koucky, M., Lhotska, L., Huptych, M., Janku, P., Georgoulas, G., and Stylios, C. D., Using nonlinear features for fetal heart rate classification. *Biomed. Signal Process. Control* 7:350–357, 2012.
11. Georgoulas, G., Stylios, C. D., and Groumpos, P. P., Feature extraction and classification of fetal heart rate using wavelet analysis and support vector machine. *Int. J. Artif. Intell. Tools* 15:411–432, 2005.

12. Vaisman, S., Salem, S. Y., Holcberg, G., and Geva, A. B., Passive fetal monitoring by adaptive wavelet denoising method. *Comput. Biol. Med.* 42:171–179, 2012.
13. Cattani, C., Doubrovina, O., Rogosin, S., Voskresensky, S. L., and Zelianko, E., On the creation of a new diagnostic model for fetal well-being on the base of wavelet analysis of cardiotocograms. *J. Med. Syst.* 30:489–494, 2006.
14. Magenes, G., Signorini, M. G., and Arduini, D., Classification of cardiotocographic records by neural networks. *Proc. IEEE-INNS-ENNS Int. Joint Conf. Neural Netw. (IJCNN'00)* 3:637–641, 2000.
15. Liszka-Hackzell, J. J., Categorization of fetal heart rate patterns using neural networks. *J. Med. Syst.* 25:269–276, 2001.
16. Magenes, G., Signorini, M.G., Sassi, R., Arduini, D., Multiparametric analysis of fetal heart rate: comparison of neural and statistical classifiers. Proceedings of Medicon 2001, Pula, Croatia, 360–363, (2001).
17. Ocak, H., and Ertunc, H. M., Prediction of fetal State from the cardiotocogram recordings using adaptive neuro-fuzzy inference systems. *Neural Comput. Appl.*, 2012. doi:10.1007/s00521-012-1110-3.
18. Shah, S., and Kusiak, A., Cancer gene search with data-mining and genetic algorithms. *Comput. Biol. Med.* 37:251–261, 2007.
19. Hassan, R., Nath, B., and Kirley, M., A fusion model of HMM, ANN and GA for stock market forecasting. *Expert Syst. Appl.* 33:171–180, 2007.
20. Yu, S. N., and Lee, M. Y., Bispectral analysis and genetic algorithm for congestive heart failure recognition based on heart rate variability. *Comput. Biol. Med.* 42:816–825, 2012.
21. Elveren, E., and Yumusak, N., Tuberculosis disease diagnosis using artificial neural network trained with genetic algorithm. *J. Med. Syst.* 35:329–332, 2011.
22. Kocer, S., and Canal, M. R., Classifying epilepsy diseases using artificial neural networks and genetic algorithm. *J. Med. Syst.* 35:489–498, 2011.
23. Vapnik, V., and Cortes, C., Support vector networks. *Mach. Learn.* 20:273–297, 1989.
24. Ubeyli, E. D., Multiclass support vector machines for diagnosis of erythematous-squamous diseases. *Expert Syst. Appl.* 35:1733–1740, 2008.
25. Sahin, D., Ubeyli, E. D., Ilbay, G., Sahin, M., and Yasar, A. B., Diagnosis of airway obstruction or restrictive spirometric patterns by multiclass support vector machines. *J. Med. Syst.* 34:967–973, 2010.
26. Abibullaev, B., and An, J., Decision support algorithm for diagnosis of ADHD using electroencephalograms. *J. Med. Syst.* 36:2675–2688, 2012.
27. Chen, H. L., Yang, B., Wang, G., Wang, S. J., Liu, J., and Liu, D. Y., Support vector machine based diagnostic system for breast cancer using swarm intelligence. *J. Med. Syst.* 36:2505–2519, 2012.
28. Goldberg, D. E., Genetic Algorithms in Search, Optimization and Machine Learning. Addison-Wesley, 1989
29. Melanie, M., An Introduction to Genetic Algorithms. MIT Press, 1996.
30. Holland, J., Genetic algorithms. *Scientific American*, 66–72, 1992.
31. Frank, A., Asuncion, A., UCI Machine Learning Repository. [<http://archive.ics.uci.edu/ml>]. Irvine, CA: University of California, School of Information and Computer Science, 2000.
32. Ayres-de-Campos, D., Bernardes, J., Garrido, A., Marques-de-Sa, J., and Pereira-Leite, L., SisPorto 2.0: a program for automated analysis of cardiotocograms. *J. Matern-Fetal Med* 5:311–318, 2000.