Classification of Fetal Heart Rate Signals Based on Features Selected Using the Binary Particle Swarm Algorithm

G. Georgoulas[1], C. Stylios[1] , V. Chudacek[2], M. Macas[2], J. Bernardes[3] and L. Lhotska[2]

[1] Dept. of Communications, Informatics and Management, TEI of Epirus, Kostakioi, Artas, Greece
[2] Czech Technical University, Gerstner Laboratory, Prague, Czech Republic
[3]Department of Gynecology and Obstetrics, Porto Faculty of Medicine, Porto, Portugal

Abstract—**The objective of this study is the investigation of a novel feature selection method based on the implementation of the binary Particle Swarm Algorithm and its application to the demanding problem of the classification of Fetal Heart Rate signals during the intrapartum period. The results are promising paving the way for more research towards the improvement of the proposed algorithm.**

Keywords—**binary Particle Swarm Optimization, Fetal Heart Rate, K-nn, Support Vector Machines.**

I. INTRODUCTION

Electronic fetal monitoring is an essential tool for fetal surveillance both during the antepartum and the intrapartum period. It is mainly based on the monitoring and evaluation of the Fetal Heart Rate, (FHR) usually by eye inspection by the physician in charge following specific guidelines [1],[2]. Evaluation and interpretation of FHR gives an indication of the fetal health condition [3]. Many researches have tried to toggle with the problem of FHR analysis and interpretation using an algorithmic manner, a particularly difficult problem in the intrapartum period [4]. While the availability of commercialized systems for the intrapartum period is still in the very beginning [10], the development of powerful and low cost personal computers have led to proposing of new automatic or semiautomatic systems [5],[15], which are however still in an experimental or evaluation phase.

Due to the very complicated nature of the problem at hand many different categories of features extracted from the FHR signal have been proposed [14],[15] and tested for characterization of the fetal condition. Feature selection is a complicated process [16] and in difficult problems we tend to extract many features from different domains and then we proceed to a refinement stage in an attempt to find the most representative ones. By doing so, we can reduce the redundant information, alleviating simultaneously the well known problem of the "curse of dimensionality".

In this work we propose the use of the binary version of the particle swarm optimization for the selection of those features that convey the most relevant information for the classification of FHR signals. During the selection phase we also experimented to select the appropriate model for the k-nn classifier [17] employed for the classification (i.e. selection of the appropriate number k of nearest neighbors). This particular classifier proved to be inefficient to deal with the complexity of this specific problem and as a result we had to replace it during the final classification stage with a Support Vector Machine (SVM) [18] classifier, which yielded better results.

The remaining of the paper is structured as follows: Section II gives a brief introduction to continuous and binary Particle Swarm Optimization (PSO) algorithm. Section III presents all the stages comprising the proposed method for FHR classification and the paper concludes with Section IV where the results are summarized and discussed along with some proposals for future work.

II. PARTICLE SWARM OPTIMIZATION

The Particle Swarm Optimization (PSO) is a population based stochastic optimization technique developed by Kennedy and Eberhart [19],[20] inspired by the social behavior of animals such as flocking or fish schooling. Since its introduction, PSO has found many applications in solving optimization problems in real number spaces [21],[20].

A potential solution to a minimization (or maximization) problem is represented by a particle having coordinates x_{id} and rate of change v_{id} in *D*-dimensional space. In its original formulation, the updates of the particles are accomplished according to the following equations

$$v_{id}(t+1) = v_{id}(t) + \eta_1 r\left[p_{id} - x_{id}(t)\right] + \eta_2 R\left[p_{nd} - x_{id}(t)\right] \quad (1)$$

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t+1) \quad (2)$$

where $v_{id}(t)$ is the current velocity of the i^{th} particle, $x_{id}(t)$ is the current position of the i^{th} particle, p_{id} is the particle's locations at which the best fitness has been achieved so far and p_{nd} is the best particle among the neighbors at which the best fitness has been achieved so far. *R* and *r* are two independently generated random numbers uniformly distributed in [0, 1] and η_1, η_2 are learning factors. Since its introduction variations of the above formulation have been proposed [21],[22] improving the performance of the continuous version of the PSO.

As already mentioned the above formulation has found many applications in optimization. However apart from the well known continuous PSO there is also a discrete binary version of the PS algorithm proposed by Kennedy and Eberhart [23], [21]. In the binary version the formula for the velocity remains unchanged except from the fact that the particle positions p_{nd} , p_{id} , $x_{id}(t)$ can only take integer values. The parameter $v_{id}(t)$ determines the probability threshold for the particle's predisposition of deciding 1 or 0. Since probabilities should be constrained in the interval [0,1] the logistic function is used. The position of each particle is determined by the following relation

$$\text{if } \rho_{id} < S\left(v_{id}(t)\right) \text{ then } x_{id}(t) = 1, \text{ else } x_{id}(t) = 0 \qquad (3)$$

where $S\left(v_{id}(t)\right) = \dfrac{1}{1+\exp\left(-v_{id}(t)\right)}$ and ρ_{id} is a random vector drawn from a uniform distribution between 0.0 and 1.0. Moreover a constant parameter $V_{\max}$ can be used in order to limit the velocity term, allowing always a possibility for a change.

Apart from the above binary version another one has been proposed by Agrafiotis and Cedeno [24], in order to be used for feature selection applied to regression problems. In the present work we employed the binary algorithm proposed by Kennedy and Eberhart with $V_{\max} = \pm 4$ [19][20].

III. Proposed FHR analysis procedure

The proposed procedure consists of four stages. The first stage pre-processes the FHR signals before we proceed to the second stage that performs the feature extraction. The third stage employs the bPSO in order to select an appropriate set of features and also select the "optimum" number of neighbours for the k-nn algorithm. The fourth stage which performs the classification was originally comprised by the k-nn model selected during the previous stage but had to be modified, replacing the k-nn classifier with an SVM.

A. Preprocessing stage

Before the extraction of the features from the FHR signals an artifact removal process has to take place in order to remove the "noise" contained in almost all FHR recordings. The procedure for artifact removal, which was first introduced in [25], detects a stable FHR segment, which is defined as a segment where the difference (in beats/min) between five adjacent samples is less than 10 beats/min. Whenever a difference between adjacent beats higher than 25 beats/min is found, a linear interpolation is applied between the first of those two signals and the first signal of a new stable FHR segment. After the removal of artifacts segments of equal size (20 minutes) were cropped starting from the end of the recordings (with the exclusion of 1-2 final minutes) for further analysis.

B. Feature Extraction stage

The FHR signal is a time series and as a result it is natural to try to extract features from the complementary domains of time and frequency. Furthermore the more of 30 years of use of this signal by gynecologists and obstetrics led as to extract a set of features more closely related to the physician's approach to FHR interpretation directly associated with the "morphology" of the FHR signal. Therefore we can divide the extracted features into three classes.

1) Time domain features:

In time domain we extracted the following 7 features {mean FHR, Standard deviation of FHR, Delta value, STV value, II value, LTI value, Delta total value}. A detailed description of each one of the above parameters can be found in [14],[15].

2) Frequency domain features:

In order to extract features from the frequency domain we adopted a partitioning of the frequency band proposed by Magenes et al [4] since it seems to give a more adequate characterization of the fetal condition [15]. Therefore we partitioned the frequency range into 4 bands and calculated the corresponding energies:

- the Very Low Frequency (*VLF*) 0-0.03 Hz
- the Low Frequency (*LF*) 0.03-0.15
- the Movement Frequency (*MF*) 0.15-0.5 Hz,
- the High Frequency (*HF*) 0.5-1 Hz
- As a fifth feature for this feature set we used the ratio:
- $LF/(HF+MF)$

3) "Morphological" features

As mentioned in the introduction, everyday FHR interpretation is based upon certain morphological characteristics, according to predefined guidelines [2]. In this work we extracted the following seven morphological parameters:

1. Baseline
2. Number of accelerations
3. Number of Small accelerations
4. Number of Mild decelerations
5. Number of Prolonged decelerations
6. Number of Severe decelerations
7. The percentage of the decelerations time

A detailed description of each one of the above parameters can be found in [15]. Therefore we had a total of 19 features which will be tested using the feature selection stage in order to eliminate those that are not appropriate for the specific classification task.

C. Feature selection using bPSO

The main purpose of feature selection is to reduce the number of features without compromising the classification performance. Moreover sometimes the elimination of less discriminatory features can lead to an increased generalization performance of the classification method.

In this work the selection was performed using bPSO. Each particle represented a subset of features. The length of each particle was equal to the number of features augmented by 2 bits used to code the number of neighbors of the K-nearest neighbor classifier employed for the evaluation of the fitness of each individual. A zero value at the d^{th} position declares absence of the corresponding feature from the subset of features while a one declares the inclusion of the corresponding feature in the subset. Figure 1 illustrates schematically the information coded in each particle.

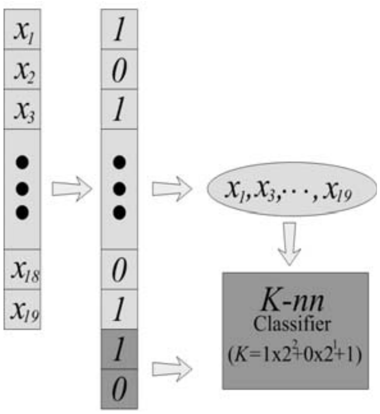

Fig. 1 Feature selection scheme using bPSO

The fitness value of each particle was the negative geometric mean [26]:

$$-\sqrt{a^{+} \cdot a^{-}} \tag{4}$$

where a^{+} is the accuracy, which is observed separately on positive examples (cases "at risk"/abnormal) and a^{-} is the accuracy observed separately on negative examples (normal/healthy cases). The minus was used in order to have a minimization problem while the geometric mean was selected in order to penalize classifiers with imbalanced performance [26].

D. Classification stage

The final stage is responsible for the classification of the FHR signal. Originally the classifier selected was the k-nn where *k* was part of the selection stage. To be more specific the last two bits of each particle were used to code the value of k. Therefore the value of *k* was given by

$$k = b_{end-1} \cdot 2^4 + b_{end} \cdot 2^2 + 1 \tag{5}$$

where the one at the end of the relation was added in order to always have an odd value for *k*, allowing *k* to take one of the following values $k=\{1,3,5,7\}$. However even though, during the training phase the k-nn classifier gave very good results this was not the case during the testing phase. As a result we had to change the classifier and use an SVM with RBF kernels which resulted in a better performance [18].

IV. Results and conclusions

In this research work, the data set consisted of 160 recordings. 130 of the recordings belonged to fetuses with umbilical artery pH>7.2 and consisted the normal group while the rest 30 having pH<7.1 consisted the "at risk" group. In order to alleviate the problem arising from the imbalance segregation of examples we used the Synthetic Minority Oversampling Technique [27] with 600% oversampling of the minority class.

The performance of the proposed procedure was estimated using the 10 fold stratified cross validation method [28]. The results achieved by the k-nn and the SVM classifier are summarized in Table 1. As it can be seen even though the k-nn achieves greater overall accuracy the SVM classifier manages to achieve higher sensitivity.

Table 1 Classification Performance

	k-nn	SVM
"at risk"	63.3%	70%
normal	88.5%	79.2%
overall	83.8%	77.5%

These results, pertaining to the analysis of fetuses born with arterial pH > 7.20 or lower than 7.10, are promising and are comparable with the best results reported in the literature regarding the analysis of FHR tracings with other methods, pertaining to fetuses representing the whole arterial pH range [29]. Further investigation should be given towards the use of bPSO for feature selection. In future work we will test the extraction of other FHR features, namely in the morphological domain, looking at other approaches to the FIGO guidelines for fetal monitoring [2], [29].Moreover, we will test the use of a hybrid version of

PSO in order to simultaneously selected the "best" features and also select the best model for the SVM classifier, regarding the analysis of the whole range of fetuses, born with arterial pH > 7.20, 7.20-7.10 and <7.10..

ACKNOWLEDGMENT

This research work has partially supported by the joint research project "Intelligent System for Medical Data Analysis and Evaluation Based On Computational Intelligence Techniques" under the Greek-Czech bilateral cooperation agreement, funded by Greek GSRT and Czech MEYS and partially by the research program No. MSM 6840770012 "Transdisciplinary Research in the Area of Biomedical Engineering II" of the CTU in Prague, sponsored by the MEYS of the Czech Republic

The data for this study were collected within the Research Project POSI/CPS/40153/2001, from Fundação para a Ciência e Tecnologia, Portugal.

REFERENCES

1. Parer J T (1997) Handbook of fetal heart rate monitoring. 2nd ed. Philadelphia, PA: Saunders
2. Rooth G, Huch A, Huch R (1987) Guidelines for the use of fetal monitoring. Int J Gynaecol Obstet 25: 159-67
3. Geijn H P V (1996) Developments in CTG analysis. Bailliere's Clin Obstet Gynaecol 10: 2: 185-209
4. Dawes G, Moulden M, Sheil O, Redman C W G (1992) Aproximate entropy, a statistic of regularity, applied to fetal heart rate data before and during labour. Obstet Gynecol 80:5:763-8
5. Arduini D, Rizzo G, Piana G, Bonalumi A, Brambilla P, Romanini C (1993) Computerized Analysis of Fetal Heart Rate: I. Description of the System (2CTG). J Matern Fetal Invest 3: 159-163
6. Mantel R, Geijn H P V, Caron F J M, Swartjes J M, Woerden E E V, Jongsma H W (1990) Computer analysis of antepartum fetal heart rate: 1. Baseline determination. Int J Biomed Comput 25: 4: 261-272
7. Dawes G S, Moulden M, Redman C W (1995) Computerized analysis of antepartum fetal heart rate. Am J Obstet Gynecol 173:. 4: 1353-1354
8. Jezewski J, Wrobel J (1993) Foetal monitoring with automated analysis of cardiotocogram: The KOMPOR system. In Proc. 15th Ann. Card. IEEE/EMBS, San Diego, CA, 1993, pp. 638-639
9. Maeda K (1990) Computerized analysis of cardiotocograms and fetal movements. Bailliere's Clin Obstet Gynaecol 4: 4:1797-813
10. Campos D A, Bernardes J, Garrido A, Sa J P M, Pereira-Leite L (2000) SisPorto 2.0: a program for automated analysis of cardiotocograms. J Matern Fetal Med 9: 311-318
11. Salamalekis E, Thomopoulos P, Giannaris D, Salloum I, Vasios G, Prentza A, Koutsouris D (2002)Computerised intrapartum diagnosis of fetal hypoxia based on fetal heart rate monitoring and fetal pulse oximetry recordings utilising wavelet analysis and neural networks. Br J Obstet Gynaecol 109:10:1137-1142
12. Skinner J F, Garibaldi J M, Ifeachor E C (1999) A Fuzzy System for Fetal Heart Rate Assessment. In Proc. 6th Fuzzy Days on Computational Intelligence, Dortmund, Germany, 1999, pp. 20-29
13. Guijarro-Berdinas B, Alonso-Betanzos A, Fontenla-Romero O (2002) Intelligent analysis and pattern recognition in cardiotocographic signals using a tightly coupled hybrid system. Artificial Intelligence 136:1-27
14. Magenes G, Signorini M G, Arduini D (2000) Classification of cardiotocographic records by neural networks Neural Networks. Proc. IEEE-INNS-ENNS Int. Joint Conf. on Neural Networks (IJCNN'00), 2000, 3: 637-641
15. Georgoulas G, Stylios C, Groumpos P (2006) A Novel methodology for Fetal Heart Rate Signal Classification During the Intrapartum Period, IEEE Trans. Biomed Engineering 53:5:875-884
16. Sa J P M (2001) Pattern Recognition: Concepts, Methods and Applications, Springer
17. Duda R O, Hart P E, Stork D G (2001) Pattern Classification. 2nd ed. John Wiley & Sons
18. Burges C J C (1998) A Tutorial on Support Vector Machines for Pattern Recognition. Data Mining and Knowledge Discovery 2:2:121-167
19. Eberhart R C, Kennedy J (1995) A new optimizer using particle swarm theory. Proc. 6th International Symposium on Micromachine and Human Science, Nagoya, Japan, 1995, pp 39-43
20. Kennedy J, Eberhart R C (1995) Particle Swarm Optimization. Proc. IEEE International Conference on Neural Networks, Piscataway, NJ, 1995, pp 1942-1948
21. Parsopoulos K E, Vrahatis M N, (2002) Recent approaches to global optimization problems through Particle Swarm Optimization. Natural Computing 1:235-306
22. Kennedy J, Eberhart R C (2001) Swarm Intelligence. Morgan Kaufman, San Francisco
23. Kennedy J, Eberhart R C (1997) A discrete binary version of the particle swarm algorithm. Proc. of the 1997 Conference on Systems, Man, and Cybernetics, Piscataway, NJ, 1997, pp 4104-4109
24. Agrafiotis D K, and Cedeno W (2002) Feature Selection for Structure-Activity Correlation Using Binary Particle Swarms. J Med Chem 45: 1098-1107
25. Bernardes J, Moura C, Sa J P M, L. Pereira-Leite L (1991) The Porto system for automated cardiotocographic signal analysis. J Perinat Med 19:61-65
26. Provost F, Fawcett T (1997) Analysis and Visualization of Classifier Performance: Comparison under Imprecise Class and Cost Distribution. In Proc. 3rd Int. Conf. on Knowledge Discovery and Data Mining, 1997, pp. 43-48
27. Chawla N V, Hall L., Kegelmeyer W (2002) SMOTE: Synthetic Minority Over-sampling Technique. Journal of Artificial Intelligence Research 16:341–378
28. Breiman L, Friedman J H, R. A. Olshen R A, Stone C J (1984) Classification and Regression Trees. Belmont: Wadsworth Int. Group
29. Bernardes J, Ayres-de-Campos D, Costa-Pereira A, Pereira-Leite L, Garrido A (1998) Objective computerized FHR analysis. Int J Gynecol Obstet 62:141-7.

Address of the corresponding author:

Author:	Chrysostomos Stylios
Institute:	Technological Educational Institute of Epirus
Street:	Kostakioi
City:	Artas
Country:	Greece
Email:	stylios@teleinfom.teiep.gr