A one-class approach to cardiotocogram assessment

Dimitris Gavrilis, George Nikolakopoulos, *Member, IEEE,* and George Georgoulas

Abstract— **Cardiotocogram (CTG) is the most widely means for the assessment of fetal condition. CTG consists of two traces one depicting the Fetal Heart Rate (FHR), and the other the Uterine Contractions (UC) activity. Many automatic methods have been proposed for the interpretation of the CTG. Most of them rely either on a binary classification approach or on a multiclass approach to come up with a decision about the class that the tracing belongs to. This work investigates the use of a one-class approach to the assessment of cardiotocograms building a model for the healthy data. The preliminary results are promising indicating that normal traces could be used as part of an automatic system.**

I. INTRODUCTION

Since its introduction, cardiotocography is the prevalent tool for the assessment of the condition of the fetus. The assessment is based on an eye inspection of the fetal heart rate (FHR) and the uterine contractions (UC), the two signals that constitute the Cardiotocogram (CTG) (Figure 1), by an experienced clinician following internationally issued guidelines [1], [2]. However that assessment is far from being characterized as trivial and a considerable amount of expertise is needed. To make things even more complicated, it has been proven that the assessment is not consistent either between different clinicians (inter-observer variability) or between the same clinician (intra-observer variability) [3], [4].

To overcome this subjectivity in the interpretation of CTG a number of computerized approaches have been pursued. Most of them follow a similar path: a set of features is extracted and then a classifier is trained to discriminate between two, the most common case, or more classes/categories. The class formation is based either on expert annotation or on biochemical quantities such as the pH of umbilical cord blood sampled just after delivery. For the feature extraction stage a plethora of signal processing techniques has been proposed, ranging from "simple" morphological features that quantify the FIGO guidelines [5][6], to features coming from wavelets analysis [7], nonlinear domain analysis [8] and more [9],[10]. For the classification part, also a variety of methods have been employed such as expert systems [11], neural networks [9], support vector machines [7], neuro-fuzzy systems [12], ordinal classifiers [13] etc.

In this work a one-class classification scheme is tested, to the best of the author's knowledge, for the first time. Starting

Dimitris Gavrilis is with the Digital Curation Unit – IMIS, Athena Research Center, Maroussi, 15125, Greece (e-mail: d.gavrilis@dcu.gr).

George Nikolakopoulos is with the Control Engineering Group at the Department of Computer Science, Electrical and Space Engineering at Luleå University of Technology, Luleå, Sweden (e-mail:geonik@ltu.se).

George Georgoulas is with the Laboratory of Knowledge and Intelligent Computing (KIC), Department of Computer Engineering, Technical Institute of Epirus, 47100, Greece (e-mail: author@kic.teiep.gr).

from the fact that a normal CTG tracing is a reassuring factor a one-class classifier is built using only normal data. The trained classifier is tested using both (unseen) healthy data and data that were characterized either as suspicious or pathological by three experts. The preliminary results indicate that the approach could be promising for signaling an alarm in case of an anomalous situation.
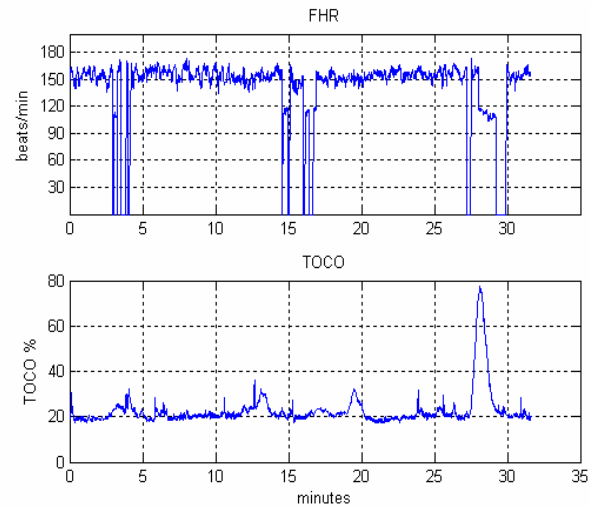

Figure 1. The two signals that constitute the CTG: on top the FHR and on the bottom the UC

The rest of this paper is structured as follows: section II presents in brief the two classifiers that are used in this study along with two dimensionality reduction techniques that are utilized to shed some light onto why the proposed approach can be useful as well as to highlight some potential drawbacks. Section III summarizes the experimental setting along with the results. Finally section IV concludes the paper.

II. METHODS

A. One-class classification

One-class classifiers assume that only information from one class is available even though it is known that instances belonging to other class will be encountered [14]. The notion of one-class classification is closely related to novelty detection [15]-[17] which has found many applications both in biomedical [18][19] and electro-mechanical systems [20][21].

This approach is usually preferred when there is a lack of representative data from all possible classes, making the building of reliable models difficult or not impossible. This could be considered the case in the CTG analysis where most of the collected data belong to healthy fetuses. In this work two simple one-class classifiers are tested: a nearest neighbor (NN) classifier and a Parzen window classifier.

Nearest Neighbor one class classifier

NN methods are based on the general principal that "nature does not make jumps" ("Natura non facit saltus") [22]. This means that data that belong to the "target" class will appear in dense clusters while data that come from other unknown classes will be located at a distance quite far apart from those clusters with the distance in most case being the Euclidian. To put it more formally a new data point $\mathbf{x}_{new}$ it is assigned to the target class or to the class of "outliers" based on the distance to its nearest neighbor $\mathbf{y}$ after normalizing it by the distance of $\mathbf{y}$ to its nearest neighbor $\mathbf{z}$

$$D = \frac{\|\mathbf{x}_{new} - \mathbf{y}\|}{\|\mathbf{y} - \mathbf{z}\|} \tag{1}$$

NN one class classifier belongs to the boundary methods which means that it builds a boundary that surrounds the "target" class.

Parzen density estimator - one class classifier

As it names implies this classifier belongs to the density based methods which estimate the density, $p(\cdot)$, of the data points that belong to the "target" class and then set a threshold, θ, on this density:

$$\begin{cases} p(\mathbf{x}) > \theta & \mathbf{x} \in \text{"target"} \\ p(\mathbf{x}) \leq \theta & \mathbf{x} \notin \text{"target"} \end{cases} \tag{2}$$

In the case of the Parzen estimator, the density is given by:

$$p(\mathbf{x}) = \sum_{i=1}^{N} \exp\left(-\left(\mathbf{x} - \mathbf{x}_i\right)^T h^{-2} \left(\mathbf{x} - \mathbf{x}_i\right)\right) \tag{3}$$

where N is the number of data points of the "target" class and h a free parameter that can be optimized by maximizing the likelihood on the training data using leave-one-out [23].

A. Dimensionality Reduction

Dimensionality reduction is the process of transforming a dataset $\mathbf{X}$ of dimension $N \times D$ where N is the number of data points which have a dimensionality of D, to a new data set $\mathbf{Y}$ of dimension $N \times d$ where $d < D$ (usually $d = D$).

The dimensionality reduction can be used as part of the classification pipeline as a means to alleviate the curse of dimensionality, which often leads to superior generalization performance, or as part of a visualization process to help the user to get a better understanding of its data. Many algorithms have been proposed over the years [24]. In this work the oldest dimensionality reduction method, Principal Component Analysis (PCA) is used both as part of the one class classification process as well as for visualization, while ISOMAP is used to confirm the visualization results of the PCA.

Principal Component Analysis

PCA is a linear technique that achieves dimensionality reduction but projecting the data using a set of orthogonal vectors. The steps of PCA are summarized as follows:

- Compute the mean value for each of the D variables of the original data set

- Subtract the respective mean from the original variable

- Calculate the covariance matrix of the new zero mean data matrix

- Calculate the eigenvalues and its corresponding eigenvectors of the covariance matrix calculated in the previous step.

- Project the zero mean data matrix to the eigenvectors corresponding to the d largest eigenvalues.

ISOMAP

Isomap is part manifold learning family. In Isomap, the distance between two points is measured over a manifold using graph distances [25]. The procedure is summarized as follows:

- Construct a neighborhood graph by considering each data point and its k nearest neighbors.

- Calculate for each data pair the shortest distance along the graph and set it as the distance between that pair (estimate of the actual geodesic distance).

- Use the classical Multidimensional scaling (MDS) algorithm [24] along with the distances calculated in step 2 to create the low dimensional embedding.

Note: The MDS algorithm minimizes the following quantity

$$\sum_i \sum_j \left(\|\mathbf{x}_i - \mathbf{x}_j\| - \|\mathbf{y}_i - \mathbf{y}_j\|\right)^2 . \tag{6}$$

where $\mathbf{x}_i$ lie on the original space and $\mathbf{y}_i$ lie on the reduced dimension space.

III. EXPERIMENTS – RESULTS

A. Data set

For this study the Cardiotocogram data set [26] from UCI repository is used [27]. The data set was created using the SisPorto 2.0 software [5]. 21 features were extracted, which are summarized in Table I. The tracings were assigned into one out of three classes, namely, normal, suspect or pathological following a consensus procedure among three expert clinicians. The total number of cases for each class is summarized in Table II.

B. Experimental procedure

Before training the one class classifiers, the data are normalized to have zero mean and standard deviation equal to one because they exhibit quite large differences in their respective ranges.

After data normalization PCA is applied to reduce the original dimensionality. The number of the retained principal components is empirically determined to account for 75% of the explained variance.

For evaluating the performance of the two one class algorithms a 10 fold cross validation (CV) procedure is repeated 10 times (10x10 CV). In other words the normal data are divided into 10 folds and each time 9/10 of them is used for training the classifier and the rest 1/10 along with the data belonging to the suspect and pathological classes (pooled together to form the "non-normal" class) is used for testing.

TABLE I.   EXTRACTED FEATURES

Feature Abbreviation	Feature description
LB	baseline value
AC	accelerations
FM	foetal movement
UC	uterine contractions
ASTV	percentage of time with abnormal short term variability
mSTV	mean value of short term variability
ALTV	percentage of time with abnormal long term variability
mLTV	mean value of long term variability
DL	light decelerations
DS	severe decelerations
DP	prolongued decelerations
Width	histogram width
Min	low freq. of the histogram
Max	high freq. of the histogram
Nmax	number of histogram peaks
Nzeros	number of histogram zeros
Mode	histogram mode
Mean	histogram mean
Median	histogram median
Variance	histogram variance
Tendency	histogram tendency

TABLE II.   EXPERIMENTAL DATA

Class	Number of cases
Normal	1655
Suspect	295
Pathological	176

C. Results

The results of the aforementioned procedures are summarized in the following Table III and IV for the case of the NN and the Parzen classifier respectively. These matrices are the aggregated confusion matrices from the 10x10 CV procedure.

From these matrices we can observe that the two classifiers exhibit quite a distinct behavior. The NN is better in detecting the normal cases, while the Parzen detector is better in detecting the non-Normal cases. This is summarized in Table V which tabulates the accuracy rates for Normal and Non-Normal classes for the two classifiers.

TABLE III.   NN

Confusion Matrix		Estimated class	
		Normal	Non-Normal
True class	Normal	12413	4137
	Non-Normal	9202	37898

TABLE IV.   PARZEN

Detection/Confusion Matrix		Estimated class	
		Normal	Non-Normal
True class	Normal	8753	7797
	Non-Normal	6417	40683

Classification rates	NN	Parzen
Normal	75%	52.89%
Non-Normal	80.46%	86.38%

IV. CONCLUSIONS-DISCUSSION

This work presented a one-class classification approach for the categorization of CTGs. Two different methods were tested, a NN classifier and a Parzen density estimator, with the NN achieving high classification rates for both classes.

The effectiveness of the one class classifier could be partially explained by observing the two dimensional projection of the original space. Figures 2 and 3 depict the projection of the original space into two dimensions using PCA and ISOMAP respectively. The two methods are quite different (the first one is a linear method which preserves the global characteristics of the data set, while the second one is a nonlinear method which tries to preserve the local characteristics of the data set) but the resulting images are quite similar: There is a dense area in the center, occupied by normal data and two areas on the "sides" occupied mostly by the suspect and pathological cases. Therefore a one class classifier can capture some of the true structure of the normal class.
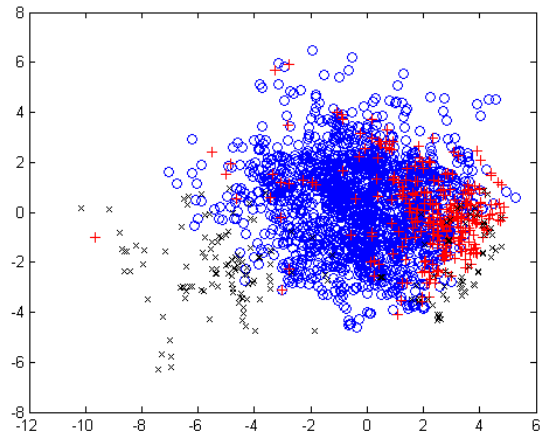

Figure 2.   Projection of the original data set to two dimensions using PCA. Blue circles correspond to normal data, red crosses correspond to suspect data, while pathological cases are marked with blac x

On the other hand it is obvious that there is quite some overlap, especially between the suspect class and the normal class. Therefore there is always going to be a trade of between sensitivity and specificity. What is interesting is that some of the pathological cases lie close or inside the "suspect" area, while some others occupy an area in the complete opposite direction being closer to the normal area rather than the "suspect" area. This observation even though confirmed by two different algorithms, requires further investigation.

In this work we only user pre-extracted features provided by [5]. In the future other features will be tested [8] combined with more advanced classifiers in order to verify the validity of the proposed approach
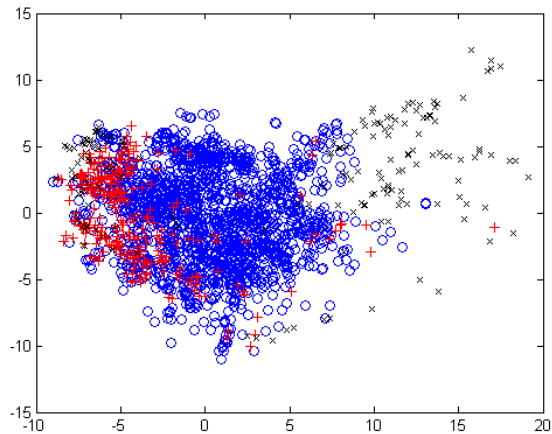

Figure 3.   Projection of the original data set to two dimensions using Isomap. Blue circles correspond to normal data, red crosses correspond to suspect data, while pathological cases are marked with blac x

REFERENCES

[1]   FIGO, "Guidelines for the Use of Fetal Monitoring," Int. J. Gynaecol. Obstet., vol. 25, pp. 159–167, 1986.

[2]   ACOG: American College of Obstetricians and Gynecologists Practice BulletinNo. 106: Intrapartum fetal heart rate monitoring: nomenclature, interpretation, and general management principles. Obstet. Gynecol. 114(1), 192-202 (Jul 2009).

[3]   J. Bernardes, A. Costa-Pereira, D. Ayres-de-Campos, H. P. van Geijn, and L. Pereira-Leite, "Evaluation of interobserver agreement of cardiotocograms," Int. J. Gynecol. Obset., vol. 57, no. 1, pp. 33–37, 1997.

[4]   S. C. Blackwell, W. A. Grobman, L. Antoniewicz, M. Hutchinson, and C. G. Bannerman, "Interobserver andintraobserver reliability of the NICHD 3-Tier Fetal Heart Rate Interpretation System," Am. J. Obstet. Gynecol., vol. 205, no. 4, pp. 378.e1–378.e5, Oct. 2011.

[5]   D. Ayres-de-Campos, J. Bernardes, A. Garrido, J. P. M. de Sa, and L. Pereira-Leite, "SisPorto 2.0: a program for automated analysis of cardiotocograms," J. Matern.. Fetal Med., vol. 9, pp. 311–318, 2000.

[6]   G. Georgoulas, C. Stylios, and P. Groumpos. "Predicting the risk of metabolic acidosis for newborns based on fetal heart rate signal classification using support vector machines." Biomedical Engineering, IEEE Transactions on 53, no. 5 (2006): 875-884.

[7]   Georgoulas, G., Stylios, C., & Groumpos, P. (2006). Feature extraction and classification of fetal heart rate using wavelet analysis and support vector machines. International Journal on Artificial Intelligence Tools, 15(03), 411-432.

[8]   J. Spilka, V. Chudáček, M. Koucký, L. Lhotská, M. Huptych, P. Janků, G. Georgoulas, and C. Stylios, "Using nonlinear features for fetal heart rate classification," Biomedical Signal Processing and Control, vol. 7, no. 4, pp. 350–357, 2012.

[9]   Georgoulas, G., Gavrilis, D., Tsoulos, I. G., Stylios, C., Bernardes, J., & Groumpos, P. P. (2007). Novel approach for fetal heart rate classification introducing grammatical evolution. Biomedical Signal Processing and Control, 2(2), 69-79.

[10]   Chudáček V Andén J Mallat S Abry P and Doret M 2014 Scattering Transform for Intrapartum Fetal Heart Rate Variability Fractal Analysis: a Case Study IEEE Transactions on Biomedical Engineering 61 1100 – 1108

[11]   A. Alonso-Betanzos, B. Guijarro-Berdinas, V. Moret-Bonillo, and S. Lopez-Gonzalez, "The NST-EXPERT project: the need to evolve," Artif. Intell. Med., vol. 7, no. 4, pp. 297–313, 1995

[12]   Ocak, H., & Ertunc, H. M. (2013). Prediction of fetal state from the cardiotocogram recordings using adaptive neuro-fuzzy inference systems. Neural Computing and Applications, 23(6), 1583-1589.

[13]   P Karvelis, J Spilka, G Georgoulas, V Chudáček, C D Stylios and L Lhotská Combining Latent Class Analysis Labeling with Multiclass Approach for Fetal Heart Rate Categorization, Physiological Measurements, in press.

[14]   D. M .J. Tax, "One-class classiffication; concept-learning in the absence of counter-examples, " Ph.D. thesis, Delft University of Technology, 2001.

[15]   M. Markou, and S. Singh, "Novelty detection: a review-part 1: statistical approaches," Signal Processing, vol. 83, no. 12, 2003, pp. 2481-2497.

[16]   M. Markou, S. Singh, "Novelty detection: a review-part 2: neural network based approaches, " Signal Processing, vol. 83, no. 12, 2003, pp. 2499-2521.

[17]   Pimentel, M. A., Clifton, D. A., Clifton, L., & Tarassenko, L. (2014). A review of novelty detection. Signal Processing, 99, 215-249.

[18]   Clifton, L., Clifton, D. A., Pimentel, M. A., Watkinson, P. J., & Tarassenko, L. (2013). Gaussian processes for personalized e-health monitoring with wearable sensors. Biomedical Engineering, IEEE Transactions on, 60(1), 193-197.

[19]   Fairley, J., Georgoulas, G., Stylios, C., & Rye, D. (2010). A hybrid approach for artifact detection in EEG data. In Artificial Neural Networks–ICANN 2010 (pp. 436-441). Springer Berlin Heidelberg.

[20]   G Georgoulas, G Nikolakopoulos and M. Mustafa, A Data Fusion Approach to Bearing Fault Detection and Diagnosis, in Proceedings, POWERENG 2015

[21]   G. Georgoulas, T. Loutas, C.D. Stylios, and V. Kostopoulos, "Bearing fault detection based on hybrid ensemble detector and empirical mode decomposition," Mechanical Systems and Signal Processing, vol. 41, no. 1, 2013, pp. 510-525.

[22]   R. Battiti and M. Brunato, "The LION way. Machine Learning Plus Intelligent Optimization," 2013.

[23]   The Data Description Toolbox, <http://homepage.tudelft.nl/n9d04/dd_tools.html>.

[24]   L.J. van der Maaten, E.O. Postma, H.J. van den Herik, Dimensionality reduction: A comparative review, Journal of Machine Learning Research 10 (2009) 66-71.

[25]   Tenenbaum, J. B., De Silva, V., & Langford, J. C. (2000). A global geometric framework for nonlinear dimensionality reduction. Science, 290(5500), 2319-2323.

[26]   https://archive.ics.uci.edu/ml/datasets/Cardiotocography#

[27]   Frank A, Asuncion A (2010) UCI machine learning repository. University of California, School of Information and Computer Science, Irvine. http://archive.ics.uci.edu/ml