Biomedical Signal Processing and Control 45 (2018) 98–108

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Open-access software for analysis of fetal heart rate signals

Zafer Cömert [a,*], Adnan Fatih Kocamaz [b]

[a] Bitlis Eren University, Department of Computer Engineering, Bitlis, Turkey
[b] İnönü University, Department of Computer Engineering, Malatya, Turkey

ARTICLE INFO

Article history:
Received 27 November 2017
Received in revised form 3 May 2018
Accepted 27 May 2018

Keywords:
Biomedical signal processing
Decision support system
Cardiotocography
Software
Image-based time-frequency features

ABSTRACT

Cardiotocography (CTG) comprises fetal heart rate (FHR) and uterine contraction (UC) signals that are simultaneously recorded. In clinical practice, a visual examination is subjectively performed by observers depending on the guidelines to evaluate CTG traces. Owing to this visual assessment, the variability in the interpretation of CTG between inter- and even intra-observers is considerably high. In addition, traditional clinical practice involves different human factors that distort the quantitative quality of the evaluation. Automated CTG analysis is the most promising way to tackle the main shortcomings of CTG by providing reproducibility of the evaluation as well as the quantitative results. In this study, open-access software (CTG-OAS) developed with MATLAB® is introduced for the analysis of FHR signals. The software contains important processes of the automated CTG analysis, from accessing the database to conducting model evaluations. In addition to traditionally used morphological, linear, nonlinear, and time–frequency features, the developed software introduces an innovative approach called image-based time–frequency features to characterize FHR signals. All functions of the software are well documented, and it is distributed freely for research purposes. In addition, an experimental study on the publicly accessible CTU-UHB database was performed using CTG-OAS to test the reliability of the software. The experimental study obtained results that included an accuracy of 77.81%, sensitivity of 76.83%, specificity of 78.27%, and geometric mean of 77.29%. These fairly promising results indicate that the software can be a valuable tool for the analysis of CTG signals. In addition, the results obtained using CTG-OAS can be easily compared to different algorithms. Moreover, different experimental setups can be designed using the tools provided by the software. Thus, the software can contribute to the development of new algorithms.

1. Introduction

Cardiotocography (CTG) is the most widely used monitoring technique for determining the fetal state during the antenatal period. CTG is composed of fetal heart rate (FHR) and uterine contraction (UC) signals that are simultaneously recorded by electronic fetal monitoring (EFM) devices [1]. In clinical practice, the EFM device produces a paper strip (sometimes called a CTG trace or CTG strip) during the test. After the test is completed, observers interpret the paper strip with the naked eye depending on guidelines such as those laid down by the International Federation of Gynecology and Obstetrics (FIGO) [2].

In the context of these guidelines, the routinely judged FHR components, often called FIGO-based or morphological features, are the baseline heart rate, variability, accelerations, and decelerations. These components are the most robust indicators to ascertain fetal well-being [3]. For this reason, in almost all studies that address automated CTG analysis, these basic morphological features are confirmed as an indispensable part of the analyses. The reproducibility of the visual assessment of CTG strips has poor value because this is associated with the expertise level of the observers [4]. Therefore, disagreement regarding the interpretation of CTG has been reported as considerably high in various studies [5,6]. In other words, there are numerous factors influencing the fetal heart rhythm that are based on complex physical mechanisms and maturational changes, and the interpretation of these factors using only a paper strip is not an easy task. Thus, the major drawbacks of FHR monitoring stem from the reading and interpretation of CTG traces rather than technical aspects [7].

The initial studies on automated CTG analysis focused on the detection of morphological features that clinicians examine with the naked eye [8,9]. Nevertheless, this is not a simple task owing to a lack of standards on how computers estimate the diagnos-

* Corresponding author.
E-mail addresses: zcomert@beu.edu.tr (Z. Cömert), fatih.kocamaz@inonu.edu.tr (A.F. Kocamaz).

https://doi.org/10.1016/j.bspc.2018.05.016
1746-8094/

tic indices. In general, the morphological features are enlarged with other diagnostic indices obtained from linear [10,11], nonlinear [12–14], discrete wavelet transform (DWT) [15–17], empirical mode decomposition (EMD) [18,19], time–frequency [20,21] and image-based time–frequency (IBTF) [22,23] domains to identify FHR signals.

In summary, automated CTG analysis is the most promising way to overcome CTG's drawbacks. This approach aims to reduce the subjective nature of fetal state evaluation and to ensure a more objective and quantitative assessment [24]. Furthermore, computerized systems can be beneficial means for achieving, storing, transmitting, and assessing the signals [24]. Numerous algorithms and software have tried to develop a model that enables a more quantitative, objective, and consistent implementation of CTG analysis, and some of these attempts resulted in commercial products.

SYSTEM 8000 was developed for antenatal FHR analysis [25], and the improved version of this system, *sonicadFetalCare*, is commercially available nowadays. In addition, 2CTG2 was developed to analyze antepartum recordings using a set of standard parameters obtained from different domains [26]. A graphical tool for the automatic and objective analysis of CTG tracings, CTG Analyzer was generated using MATLAB® [27]. In addition, a medical expert system, NST-EXPERT, was designed to evaluate the fetal condition [28]. The Computer Aided Foetal Evaluator (CAFE) is a successor to this system [29].

SisPorto, which closely followed the FIGO guidelines, was introduced in 1990. Today, the fourth version of the SisPorto program has been distributed [30]. In addition, *K2 Medical Systems* presented an intelligent system consisting of central and local units to collect information such as CTG and results of blood sample analyses from a patient's bed [31]. Nowadays, the commercial version of this system is known as INFANT® [32]. This topic is still under discussion since automated CTG systems have not yet gained enough clinical acceptance.

In this study, we introduce a prototype of an open-access software for CTG analysis (CTG-OAS). This software contains important processes such as accessing the database, preprocessing, feature transform, and classification in terms of an automated CTG analysis. In addition, the applicability of texture features such as contrast, correlation, energy, and homogeneity is explored for detecting fetal hypoxia. A comparison of the three different classifier performances is also examined for this particular purpose. An experimental study is performed on the publicly accessible CTU-UHB database using only CTG-OAS, and fairly promising results are achieved. This software, which was developed mainly for research purposes, has great potential for the automated analysis of antenatal CTG. In addition, while the software cannot be used directly as a diagnostic tool, it can be used to ensure technical support for observers.

The rest of this paper is organized as follows: The database and the components of the software are summarized in Section 2. Experimental results and a discussion are presented in Section 3. Finally, concluding remarks are given in Section 4.

2. Material and methods

2.1. CTU-UHB intrapartum CTG database

A publicly accessible intrapartum CTG database called CTU-UHB [33] was used in this study to test the performance of the developed software. A total of 9164 intrapartum recordings were collected between 2010 and 2012 through STAN S21/S31 and Avalon FM40/FM50 EFM devices in the obstetrics ward of the University Hospital in Brno, Czech Republic. Then, 552 instances were carefully selected from these recordings, considering several technical and clinical criteria. All signals sampled at 4 Hz were stored in electronic form in the OB TraceVue® system. Each recording was divided into four parts, and each part was interpreted by nine experienced obstetricians. Furthermore, an annotated file providing biochemical markers, clinical features, and details of the evaluation was presented. Please refer to [33] for more detailed information about the database.

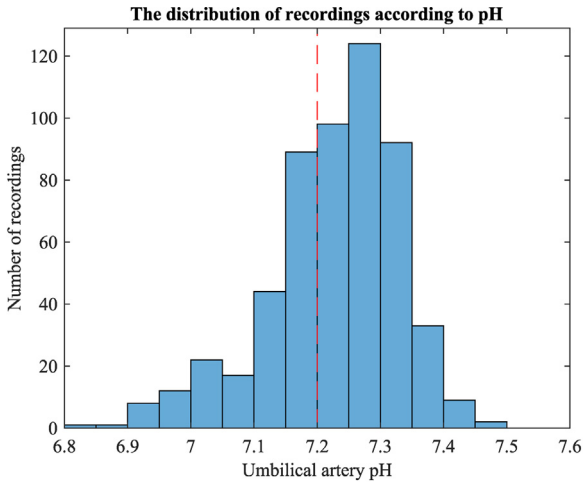

Fig. 1. Histogram of total number of CTG recordings used in the study, ordered by umbilical artery pH. There is a total of 552 recordings ranging from 6.85 to 7.47.

In a supervised learning approach, labeling of the data is a mandatory step that provides the network training to be performed. Either the views of clinicians [34] or fetal outcomes [35] can be chosen to label the signals. The fetal outcomes (pH values, the base deficit of newborn umbilical artery blood measured, etc.) are admitted as the objective annotation, whereas expert annotations (visual inspection) or evaluations of newborns (Apgar score) in the delivery room are agreed to be subjective in perinatal terminology [35]. In experimental studies, we prefer to use pH values to provide an objective evaluation of fetal hypoxia [36]. However, no specific umbilical artery pH value has been determined for separating the FHR signals as normal and abnormal. Furthermore, different values were used in previously reported studies [37,38]. These various works demonstrate that a cord artery pH of less than 7.20 points to fetal distress; otherwise, the pH value refers to fetal well-being [39].

The adjusted umbilical artery pH value emphasized with a red dashed line in Fig. 1 is used as a borderline for separating the recordings. A total of 552 recordings are taken into consideration, and the numbers of normal and hypoxic recordings are 375 and 177, respectively. In addition, we focus on the last 15 min of the signals, representing the second stage of labor in this study.

2.2. Basic components of CTG-OAS

CTG-OAS equipped with advanced signal processing and machine learning tools is an open-access software package developed for FHR signal analysis. The software is freely distributed for research purposes [40]. CTG-OAS was developed with the MATLAB® graphical user interface development environment (GUIDE). A block diagram of CTG-OAS is illustrated in Fig. 2. The software ensures the necessary processing steps are carried out: data preparation, feature transform, and classification regarding automated CTG analysis. CTG-OAS ensures a basic preprocessing scheme and also has the ability to provide a comprehensive feature set extracted from morphological, linear, nonlinear, time–frequency, DWT, and statistical domains to represent FHR signals.

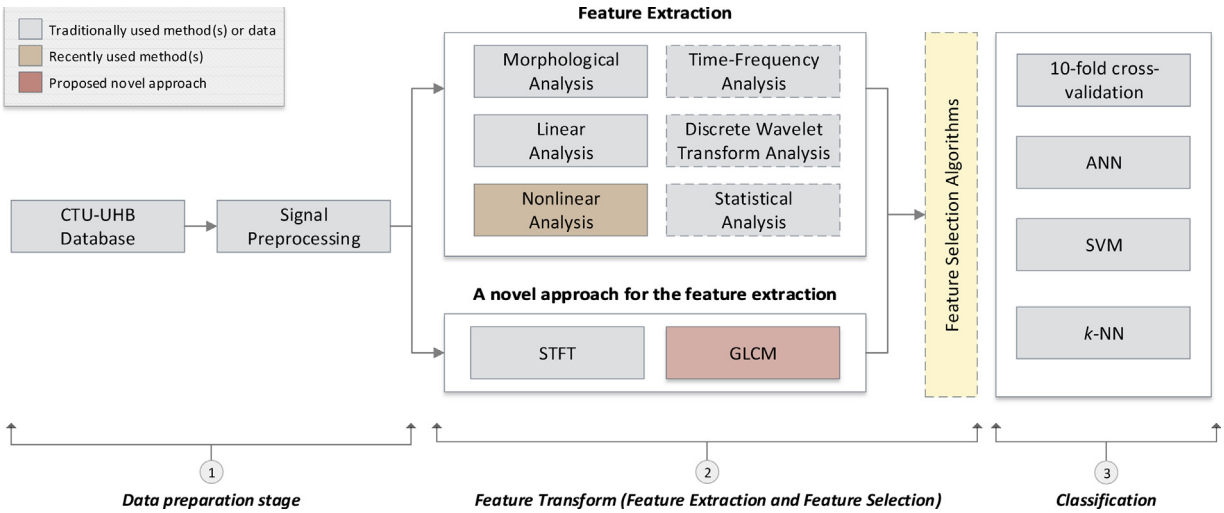

Fig. 2. Block diagram of CTG-OAS components. Blocks with dashed lines represent unused components of software in experimental study.

In addition, an innovative method consisting of a combination of a Short Time Fourier Transform (STFT) and Gray Level Co-Occurrence Matrix (GLCM) is used in the feature extraction stage by the software [22]. Thus, the software can offer a quite comprehensive feature set. In addition to the fairly comprehensive feature extraction stage, several feature selection algorithms are embedded in CTG-OAS. Furthermore, popular machine learning techniques such as the Artificial Neural Network (ANN), Support Vector Machine (SVM), and k Nearest Neighbors (k-NN) are also embedded in the software to separate normal and hypoxic fetuses [41]. We designed a web platform to share developments in the software and to distribute it freely (cf. www.ctganalysis.com).

2.3. Preprocessing

Preprocessing is an essential step for almost all biomedical signal processing applications. This process affects not only the values of extracted features but also the classification performance [24,42]. From this point of view, owing to the nature of FHR, FHR signals include several types of noise that occur owing to the movements of both mother and fetus, displacements of the transducer, and other factors such as labor-related stress [13]. As a result, the FHR signal can be contaminated by these factors.

Segment selection, outlier detection, and interpolation constitute the main preprocessing procedures in our case study [43]. The segments, which last 15 min and consist of 3600 samples owing to a 4-Hz sampling frequency, are used in the experimental work. It should be noted that the sampling frequency rate can affect the FHR variability [44]. A basic artifact rejection scheme is employed: extreme values ($\geq$200 bpm and $\leq$50 bpm) are interpolated as in [45]. The cubic Hermite spline interpolation technique [46] provided by MATLAB® is utilized to fill the missing beats in order to produce robust and accurate values for spectral analysis [47]. Furthermore, long gaps (>15 s) are not included in the subsequent feature extraction process. In the last step of the preprocessing, FHR signals are detrended by using second-order polynomials owing to the use of nonlinear signal processing techniques.

2.4. Feature extraction

A mixture of features is utilized in this study. First, we focus on morphological features that describe the shapes of and changes in the FHR signals. The baseline, number of accelerations, and deceleration patterns are calculated as described in [2,48,49].

The feature set is enlarged with linear and nonlinear features to ensure more effective signal identification. To this end, the mean (μ), standard deviation (σ), long-term irregularity (*LTI*), delta (Δ), short-term variability (*STV*), interval index (*II*), mean absolute deviation (*meanAD*), and median absolute deviation (*medianAD*) of the FHR signals are considered as time-domain features. A detailed explanation of these features is not within the scope of this paper and can be found in [10,11,14,17,50]. In addition, it should be emphasized that the FHR variability (FHRV) in short and long terms is of high clinical importance and is rather significant with regard to clinical settings [51].

Nevertheless, there is no a clear standardization in the definition and evaluation of FHR changes in clinical applications [52]. The computation of *STV* varies according to the signal acquisition approach, which is either external or internal [53]. Real beat-to-beat variability can be estimated for internal recordings, whereas a correlation-based technique is frequently employed for external recordings. In addition, several ways for computing this important index such as Arduini, Dalton, Organ, Sonicaid 8000, Van Geijn, Yeh, and Zugaib have been proposed in the literature [54].

To compute *STV* indices in this study, the FHR time-series were divided into 2.5-s blocks. Each of these blocks was represented by the average value of 10 samples (owing to a 4-Hz sampled frequency), which constitute the block. *STV* indices were computed by dividing the sum of the differences between two consecutive blocks by the number of minutes [48]. The long-term variability indices (also called the delta in this study) estimated the difference between the minimum and maximum values in a 60-s block-averaged over the duration of the signal [55].

Recently, nonlinear time-series analysis has gained increased popularity in FHR signal analysis [56–58]. In particular, as new diagnostic indexes, Approximation Entropy (ApEn), Sample Entropy (SampEn), and Lempel Ziv Complexity from the nonlinear domain have been found to be rather remarkable [12,13]. These features are utilized in the context of the experimental study. The tolerance and embedding dimension parameters of both ApEn and SampEn are set to 0.20 and 2, respectively.

As previously stated, it must be emphasized that the mathematical forms of all methods are not presented here because the focus of our research is mostly on embedding these methods in the software. As an exception, we explain how to extract new IBTF features in order to emphasize one of the innovative aspects of this study and to show how to develop a new feature extraction approach with CTG-OAS.

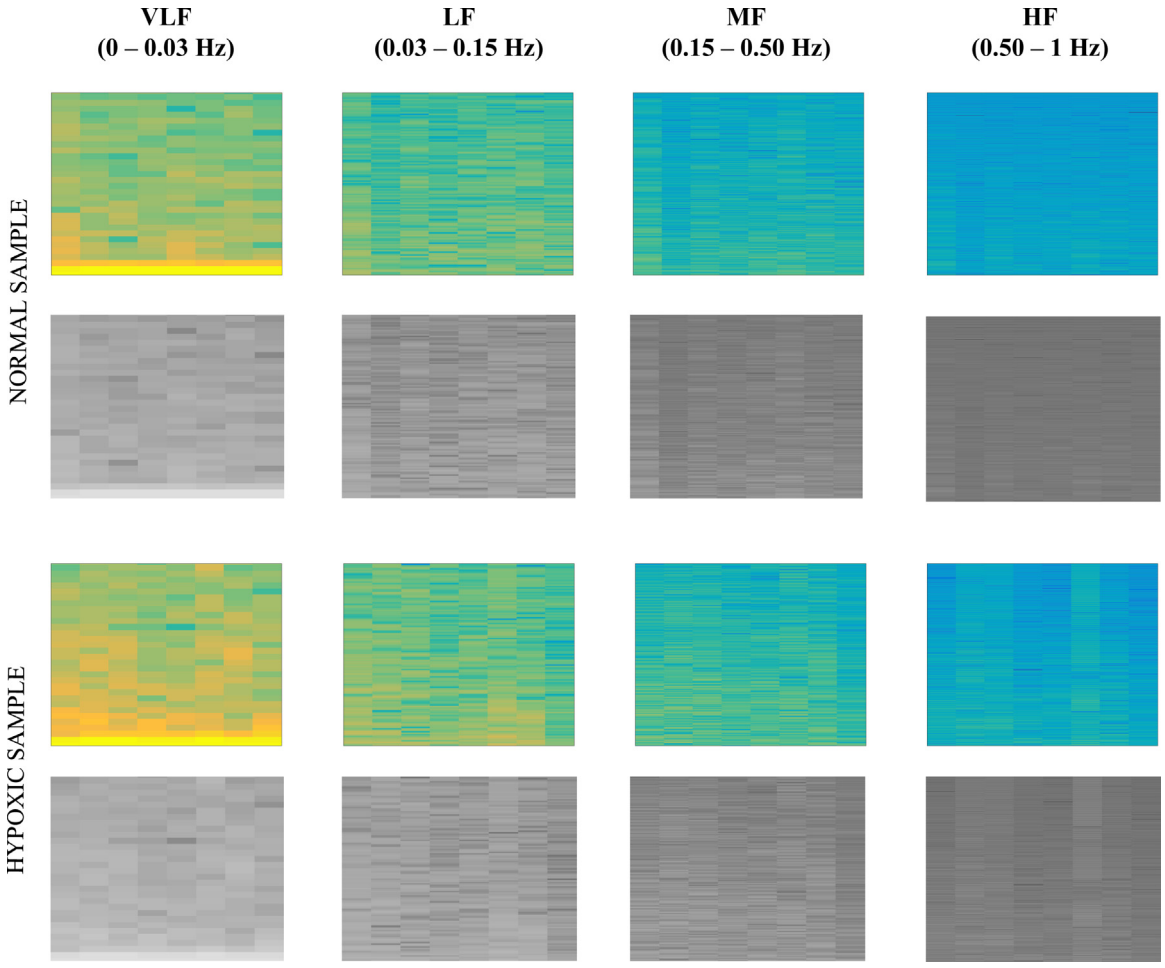

Fig. 3. Spectrograms and gray-level images for normal and hypoxic sample recordings.

IBTF features, often called as texture features, have been used successfully in electroencephalography [59] and electromyogram [60] signal analysis. However, the applicability of texture features has not been explored in FHR signal analysis. The basic information source of IBTF analysis is spectrogram images that are obtained using standard STFT. In the next step, the spectrogram images are converted into gray-level images that are used as input to a GLCM, as shown in Fig. 3. Given an image I of size $X \times Y$, the gray-tone spatial-dependence matrix M can be defined as follows:

$$M_{\delta=1,\theta=0}(p,q) = \sum_{i=1}^{X}\sum_{j=1}^{Y} \begin{cases} 1, & if I(i,j) = p \text{ and } I(i,j+\delta) = q \\ 0, & otherwise \end{cases} \quad (1)$$

where p and q represent the grayscales. X and Y define the sizes of the image. The GLCM calculates how often a pixel with gray-level value i occurs depending on the distance (δ) and angle (θ) parameters to adjacent pixels with value j. The GLCM can be used to acquire valuable information, which represents the image characteristics, by extracting the contrast, correlation, energy, and homogeneity features [61]. The mathematical forms of these features are described as follows. For a more thorough description of the GLCM, [61,62] can be reviewed.

$$Contrast = \sum_{i,j} |i-j|^2 \bar{G}_{\delta,\theta}(i,j) \quad (2)$$

$$Correlation = \sum_{i,j} \frac{(i-\mu_x)(j-\mu_y)}{\sigma_x \sigma_y} \bar{G}_{\delta,\theta}(i,j) \quad (3)$$

$$Energy = \sum_{i,j} \left(\bar{G}_{\delta,\theta}(i,j)\right)^2 \quad (4)$$

$$Homogeneity = \sum_{i,j} \frac{1}{1+|i-j|} \bar{G}_{\delta,\theta}(i,j) \quad (5)$$

Herein, μ and σ indicate the mean and standard deviation, respectively, and the marginal distribution is expressed as $\bar{G}_{\delta,\theta}(i,j)$ Contrast recognizes the number of local intensity variations. Correlation reveals the gray-level linear dependencies in the image. Energy explains the uniformity of the image. Homogeneity indicates the closeness of the distribution of elements in the co-occurrence matrix to the co-occurrence matrix diagonal.

Last, IBTF features were taken into account with distance ($\delta = 1$) and angle ($\theta = 90^\circ$) parameters in this study. The values of these parameters can be changed. In addition, IBTF features can be extracted from different empirically determined bands. In this way, a large number of IBTF features can be ensured for signal characterization. Owing to performance concerns, we focused only on the spectrograms of very low-frequency bands (VLF, 0–0.03 Hz) considering a great number of experimental experiences. Thus, we determine only four IBTF features: contrast, correlation, energy and homogeneity.

Consequently, CTG-OAS ensures that a total of 18 informative features, which are summarized in Table 1, are extracted from different domains to represent FHR signals. The output of a sample recording produced by CTG-OAS is shown in Fig. 4. The figure consists of squares and rectangles and represents a modeled CTG trace.

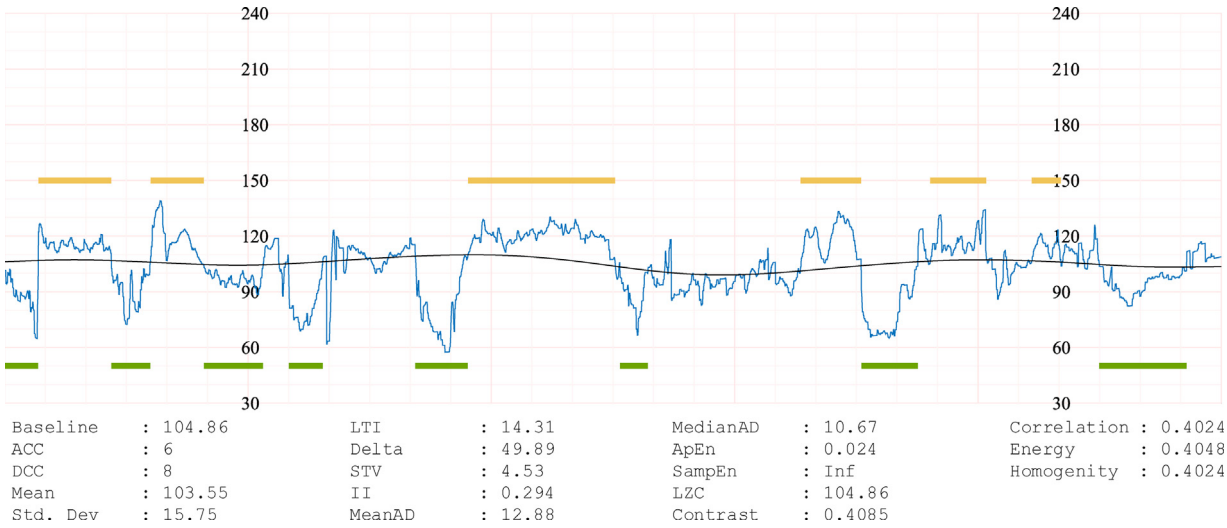

Fig. 4. Analysis report of trace 1061 (internal number of CTU-UHB database) produced by CTG-OAS.

Table 1
The features used in the experimental study.

	Features	Parameters settings
Morphological features [2,48]	Baseline, The number of accelerations, The number of decelerations.	After the virtual baseline is estimated, the parameters representing upper and lower limit are adjusted to 8 so as to remove accelerations and decelerations [75].
Time domain (linear) features [10,11,14,17]	Mean (μ), Standard deviation (σ), Long-term irregularity, (LTI), Delta (Δ), Short-term variability (STV), Interval index (II), Mean absolute deviation (MeanAD), Median absolute deviation (MedianAD).	
Nonlinear features [12,13]	Approximation Entropy (ApEn), Sample Entropy (SampEn), Lempel Ziv Complexity (LZC).	Embedding dimension ($m = 2$) Tolerance ($r = 0.20$)
Image-based time-frequency features [22]	Contrast, Correlation, Energy, Homogeneity.	Distance ($\delta = 1$) Angle ($\theta = 90^{\circ}$)

The squares correspond to 30 s on the horizontal and 10 bpm on the vertical axis, whereas the rectangles correspond to 3 min on the horizontal and 30 bpm on the vertical axis. The black solid line indicates the estimated baseline, and the orange and green lines on the figure represent the detected acceleration and deceleration patterns, respectively.

In addition, the computed values of the diagnostic features are given at the bottom of Fig. 4. Although CTG-OAS can ensure a larger feature set and automatic feature selection algorithms, we select the features used in this study manually considering the usage frequency of the features as well as our experience. On the other hand, as the number of features used to identify the signals increases, it is important to be aware that the clinical interpretation difficulty, system complexity, and computational costs also increase [63].

2.5. Classifiers

2.5.1. k-Nearest neighbors (k-NN)

A *k*-NN algorithm that stores all available cases and classes is a simple and efficient machine learning tool and is used for regression/classification tasks [64]. *k*-NN performs predictions based on a similarity measure. Predictions for a new instance are performed by searching the entire training set considering *k* neighbors. For this particular purpose, a distance measure such as Euclidean is generally used [65]. It should be noted that the best distance metric can be changed depending on the properties of the data, the computational complexity of *k*-NN increases with the size of the training dataset, and *k*-NN produces more efficient results when the dimension of the inputs is small. Furthermore, normalization of the features, elimination of the missing data problem, and lower dimensionality will lead to achieving better results for the *k*-NN. Although there is a theoretical weakness, choosing the *k*value by running an algorithm many times with different *k* values and choosing the one with the best performance is frequently preferred [66].

2.5.2. Artificial neural network (ANN)

ANNs are flexible computing structures inspired by the human nervous system. They are composed of interconnected processing components called neurons. Interconnected neurons of various weights provide the transmission of information across networks [67]. The networks are trained by adjusting the weights of the interconnections using the input data. Training algorithms follow various methods to calculate the weights [68]. It should be remembered a lack or extreme number of neurons in a hidden layer may lead to drawbacks with regard to generalization and overfitting.

2.5.3. Support vector machine (SVM)

SVM was introduced in 1995 by Vapnik [69] as a new machine learning tool. As a machine learning tool, SVM is designed for a wide range of applications in numerous areas that require regression, pattern recognition, and classification because of its superior generalization ability. SVM is frequently used for two-group classification problems by finding optimal hyperplanes to divide multidimensional data into two classes. Thus, two sets of linearly separable data can be classified easily. However, data collected in real life are generally nonlinear. SVM solves this challenge by using a kernel-induced feature space that maps the data to a very-high-dimension feature space [70]. A linear decision border is searched in this fea-

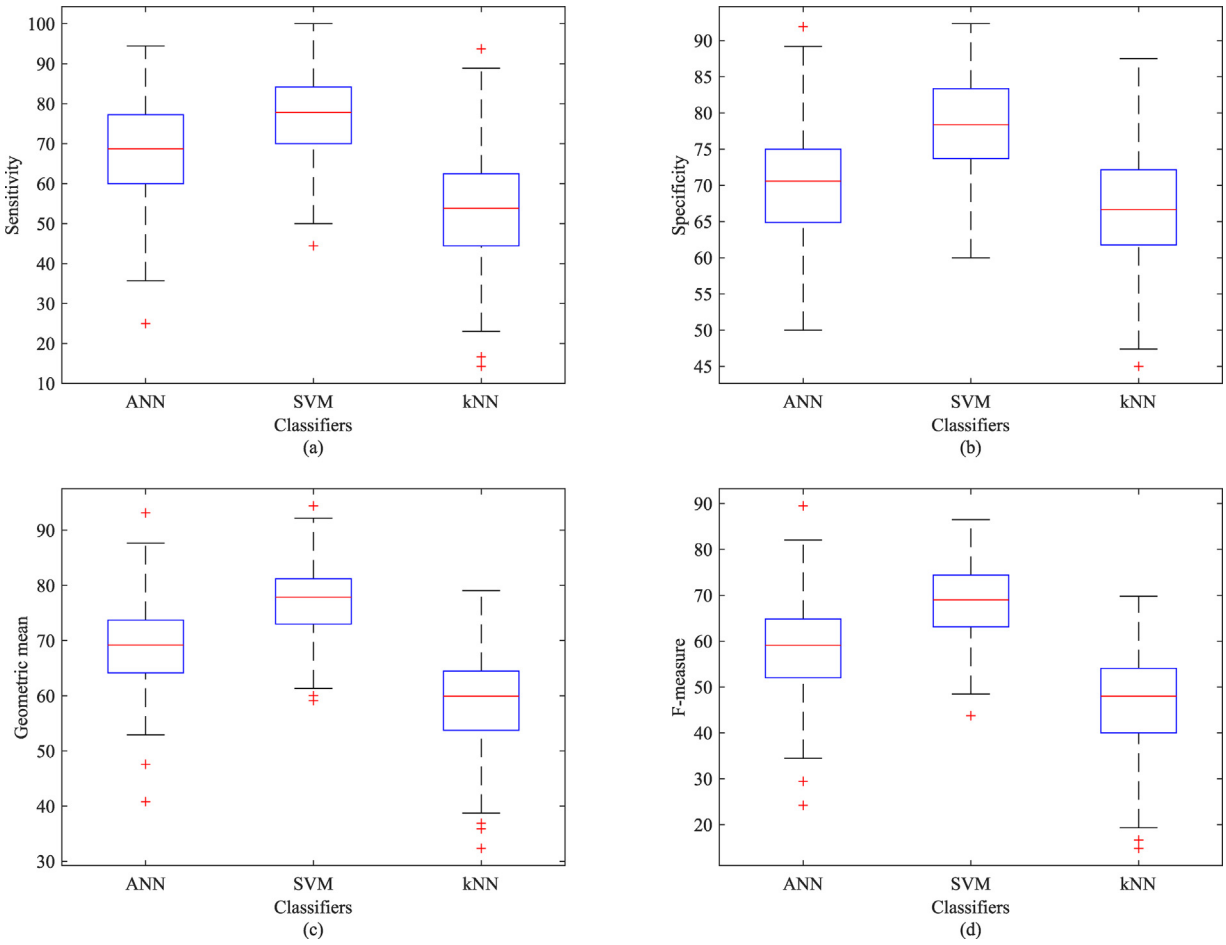

Fig. 5. Performance measures: (**a**) Se, (**b**) Sp, (**c**) GM, and (**d**) F-measure of three different classifiers.

ture space, and special parameters of the decision border ensure high generalization ability.

3. Experimental results and discussion

The main objective of this study is to ensure a software package that is equipped with advanced and modern tools for analyzing FHR signals. To this end, an experimental study from achieving the signals to model evaluation was conducted using CTG-OAS. The processes with indicated blocks in Fig. 1 were followed in the experiment, and a total of 18 specified features were used to feed the classifiers. ANN, SVM, and *k*-NN classifiers were employed to identify hypoxic fetuses in this experiment. In the tuning process of SVM, a hyperplane, which is used to transform data into a new space that can be nonlinearly separable, was adjusted with a Gaussian radial basis function (RBF).

The σ value for the RBF was tested in the range from 1 to 10 and was set to 10. The penalty factor that is used to control the rate of misclassified data is adjusted by a constant *C*, known as the box constraint of the SVM. By trial and error, the penalty factor for the SVM was set to 0.05. A pattern recognition network with a Levenberg-Marquardt backpropagation (LM) training algorithm was utilized for the ANN. The network topology was configured as {18, [10,6,2], 2}. This means that the network has an input layer with 18 nodes, three hidden layers consisting of 10, 6, and 2 nodes, respectively, and one output layer with 2 nodes that are used to represent the two classes.

The tangent-sigmoid and softmax transfer functions were used in the hidden layers and output layer, respectively. *Cross Entropy* was selected to calculate the network performance. The rest of the parameters were used with their default values for the ANN. As for *k*-NN, the *k* value expressing the number of neighbors was evaluated in the range from 1 to 10 and was adjusted to 3, and Euclidean distance measurement was used.

Table 2
Confusion matrix for binary problem.

	Actual Positive	Actual Negative
Predicted as positive	True Positive (TP)	False Positive (FP)
Predicted as negative	False Negative (FN)	True Negative (TN)

Table 3
The performance metrics with their formulas and short descriptions.

Performance metric	Formula	Description
Accuracy (Acc)	$\frac{TP+TN}{TP+FP+FN+TN}$	The overall performance of the model.
Sensitivity (Se)	$\frac{TP}{TP+FN}$	The positive test result of the model.
Specificity (Sp)	$\frac{TN}{TN+FP}$	The negative test result of the model.
Geometric (Gm)	$\sqrt{Se.Sp}$	The geometric mean of Se and Sp.
F-Measure	$\frac{2TP}{2TP+FP+FN}$	The harmonic mean of precision and recall.

The elements of a confusion matrix that are indicated in Table 2 were used to measure the performances of the models. True Positive (TP) and True Negative (TN) indicate the number of hypoxic and normal fetuses identified correctly, whereas False Positive (FP) and False Negative (FN) indicate the number of hypoxic and normal fetuses identified incorrectly.

The commonly used performance metrics are given in Table 3 with their formulas and short descriptions. Accuracy (Acc) gives

Table 4
The results of FHR signals classification without IBTF features.

	Acc (%)	Se (%)	Sp (%)	GM (%)	F-mea. (%)	AUC
ANN	$67.31 \pm 6.1^{(87.2)}_{(50.0)}$	$64.99 \pm 10.6^{(94.1)}_{(37.5)}$	$68.40 \pm 7.4^{(94.4)}_{(50.0)}$	$66.34 \pm 6.5^{(83.4)}_{(47.4)}$	$55.51 \pm 8.6^{(80.0)}_{(31.1)}$	$0.71 \pm 0.07^{(0.91)}_{(0.52)}$
SVM	$73.62 \pm 5.9^{(89.2)}_{(52.7)}$	$71.52 \pm 10.7^{(95.4)}_{(35.7)}$	$74.72 \pm 7.1^{(96.7)}_{(52.3)}$	$72.81 \pm 6.5^{(89.9)}_{(50.2)}$	$63.03 \pm 8.6^{(87.5)}_{(31.5)}$	$0.79 \pm 0.06^{(0.95)}_{(0.57)}$
k-NN	$61.17 \pm 6.4^{(81.8)}_{(47.2)}$	$59.45 \pm 11.4^{(86.6)}_{(27.2)}$	$63.49 \pm 8.1^{(86.4)}_{(41.0)}$	$60.98 \pm 7.1^{(81.9)}_{(39.5)}$	$49.69 \pm 9.1^{(72.3)}_{(17.1)}$	$0.66 \pm 0.07^{(0.89)}_{(0.43)}$

Mean ± standard deviation $^{(\max)}_{(\min)}$.

Table 5
The results of FHR signal classification with the full feature set.

	Acc (%)	Se (%)	Sp (%)	GM (%)	F-mea. (%)	AUC
ANN	$69.70 \pm 6.4^{(92.7)}_{(54.5)}$	$68.52 \pm 11.9^{(94.4)}_{(25.0)}$	$70.29 \pm 7.7^{(91.8)}_{(50.0)}$	$69.02 \pm 7.3^{(93.1)}_{(40.0)}$	$58.62 \pm 9.7^{(89.4)}_{(24.2)}$	$0.76 \pm 0.07^{(0.97)}_{(0.51)}$
SVM	$77.81 \pm 5.6^{(90.9)}_{(60.0)}$	$76.83 \pm 10.4^{(100)}_{(44.4)}$	$78.27 \pm 6.6^{(92.3)}_{(60.0)}$	$77.29 \pm 6.3^{(94.4)}_{(59.1)}$	$68.48 \pm 8.1^{(86.4)}_{(43.7)}$	$0.84 \pm 0.05^{(0.98)}_{(0.68)}$
k-NN	$62.59 \pm 6.4^{(80.0)}_{(43.6)}$	$53.28 \pm 12.4^{(93.7)}_{(14.2)}$	$66.80 \pm 7.9^{(87.5)}_{(45.0)}$	$59.08 \pm 7.8^{(79.1)}_{(32.3)}$	$47.13 \pm 10.1^{(69.7)}_{(14.8)}$	$0.64 \pm 0.08^{(0.87)}_{(0.41)}$

Mean ± standard deviation $^{(\max)}_{(\min)}$.

the overall performance of the model, whereas sensitivity (Se) and specificity (Sp) are the most commonly used metrics in biomedical applications. By the way, the geometric mean (GM) and F-measure are useful tools, especially when the distribution of recordings among the classes is not equal. This situation is expressed as an imbalance problem [71]. In the case of imbalance, classifiers may be more successful in distinguishing the majority class.

Two different feature sets were used in the experimental study (with and without IBTF features) in order to understand the effect of IBTF features on classification performance. Experiments with the parameters mentioned above were repeated 30 times, and a 10-fold cross-validation procedure was utilized in each repetition.

The performance results of the classifiers depending on the feature set isolated from the IBTF features are reported in Table 4, whereas the results obtained by using all features are given in Table 5. The confusion matrices and numbers of their elements are given in Table 6 for the classification results reported in Table 5. It is clear that the IBTF features lead to an increase in the performance of all classifiers except for *k*-NN. A combination of traditionally used and proposed IBTF features provided the best classification results (GM = 77.29% and F-measure = 68.48%). The GM values of the ANN and SVM classifiers increased from 66.34% and 72.81% to 69.02% and 77.29%, respectively.

Similar to other ANN applications, a slight superiority in specificity was observed when comparing the sensitivity owing to an imbalanced data distribution. We noticed a decrease in almost all performance metrics for *k*-NN when using IBTF features. In addition, the best results were provided by SVM, and it significantly improved the quality of classification in comparison to the ANN and *k*-NN. The achieved performance values of the metrics are illustrated in Fig. 5. As a result, it is shown that texture features such as contrast, correlation, energy, and homogeneity can be used as predictors of fetal hypoxia.

In this section of our investigation, as another significant tool for model evaluation with two classes, a receiver operating characteristic (ROC) curve and the corresponding area under this curve (AUC) were utilized [72]. ROC curves of the classifiers for prediction of fetal hypoxia with a full feature set are illustrated in Fig. 6. The highest AUC (close to 1) shows the highest certainty of the fetal state assessment considering the analyzed feature set. Adequate predictive capabilities were obtained by using ANN and SVM. The AUC values of ANN and SVM were 0.76 and 0.84, respectively. *k*-NN could not yield the expected performance, and the results of the classifier can be considered mediocre or unsatisfactory (AUC = 0.64).

The software introduced in this study contributes to automated CTG analysis by ensuring modern and advanced signal processing, feature transform, and classification techniques. In addition, numerous experimental studies can be performed via CTG-OAS. For this reason, we believe that CTG-OAS can be a valuable tool for the analysis of CTG signals as well as the development of new algorithms. By the way, a comparison between this study and related studies was realized considering several significant parameters such as the database, division criteria, and performance metrics in terms of the automated CTG analysis. However, it should be remembered that making a one-to-one comparison is not feasible because of differences in the experimental studies.

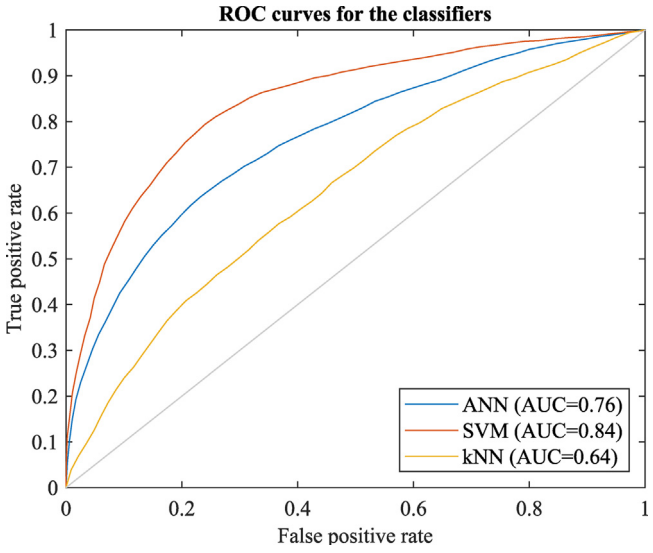

Fig. 6. ROC curves of classifiers for prediction of fetal hypoxia with full feature set.

Table 7 presents a summary of the comparison. As seen in Table 7, private and ad-hoc databases were frequently used in previous studies. Furthermore, different division criteria such as umbilical artery pH values with different threshold values and visual inspections were utilized. Compared to [45], although the same database was used, a different threshold value (pH ≤ 7.05) for umbilical artery pH was taken into account. A completely different procedure that covers Relevance in Estimating Features (RELIEF) for automatic feature selection and the Synthetic Minority Oversampling Technique (SMOTE) for fixing imbalanced data distribution was applied. The GM was achieved as 65.19%.

Regarding [73], the pH value was considered as 7.05. The sensitivity and specificity were reported as 68.50% and 77.70%, respectively. Thus, the GM was computed as 72.96%. These effective results were achieved by using only three features.

Table 6
The aggregated confusion matrix of ANN, SVM and *k*-NN classifiers for the full feature set.

ANN		SVM		*k*-NN	
3639[a]	3346[b]	4072[a]	2436[b]	2843[a]	3728[b]
1671[c]	7904[d]	1238[c]	8814[d]	2467[c]	7522[d]

[a] TP – True Positive.
[b] FP – False Negative.
[c] FN – False Negative.
[d] TP – True Positive.

The same open database was used in [45] and [73] with a different umbilical artery pH value ($pH \leq 7.05$). We carried out the experimental study again by taking into account the specified umbilical artery pH value. As a result, we achieved a rather imbalanced dataset that consisted of 44 hypoxic and 508 normal recordings. The best results were obtained with SVM having a sensitivity of 60.95% and specificity of 73.97%. Consequently, a remarkable performance loss was observed.

The results reported in [13] can be assessed as rather promising. FIGO-based, heart rate variability (HRV)-based, and nonlinear features were employed to identify a total of 217 private FHR signals. The feature selection and classification procedures were conducted by using WEKA software. Although Naïve Bayes, SVM, and C4.5 tree classifiers were employed, SVM was superior to the others. Balanced sensitivity and specificity values (Se = 73.40% and Sp = 76.30%) were indicated, and the GM was calculated as 74.83%.

A combination of DWT features and SVM classifiers were examined in [16] using four different feature sets and segments that lasted for 5 or 10 min. Experiments were performed on 80 signals

Table 7
Comparison of previously reported studies.

Reference and methods	Database	Criterion	Se (%)	Sp (%)	GM (%)	F-measure (%)
Spilka et al. [45], morphological, frequency domain, nonlinear features, RELIEF, SMOTE, 44 fold (stratified) cross validation, Adaboost	CTU-UHB	$pH \leq 7.05$	65.10[a]	64.09[a]	65.19[a]	64.64[a]
Stylios et al. [73], morphological, time domain, frequency domain, nonlinear domain features, 44 fold (stratified) cross validation, LS-SVM	CTU-UHB	$pH \leq 7.05$	68.50	77.70	72.96[a]	–
Spilka et al. [13], FIGO-based, HRV-based, and nonlinear features, Information Gain, Principal Component Analysis, Group of adaptive models evolution neural network, Naïve Bayes, SVM, C4.5 tree	Private	$pH < 7.05$	73.40	76.30	74.83[a]	71.90
Georgoulas et al. [16], DWT, 10-fold cross validation, Nearest Neighbors, Linear, Quadratic, SVM classifiers	Private	$pH < 7.10 pH > 7.20$	93.33	75.00	83.67	–
Krupa et al. [19], EMD, 5-fold cross validation, SVM	Private	Visual inspection	95.00	70.00	81.54[a]	–
Dash et al. [74], NICHD, system identification, and HRV features, 10-fold cross validation, generative model and Bayesian theory	Private	$pH \leq 7.15$	60.90	81.70	70.53[a]	–
This paper, morphological, linear, nonlinear, IBTF features, 10-fold cross validation, 30 repetition, ANN	CTU-UHB	$pH < 7.20$	68.52	70.29	69.02	58.62
This paper, morphological, linear, nonlinear, IBTF features, 10-fold cross validation, 30 repetition, SVM	CTU-UHB	$pH < 7.20$	76.83	78.27	77.29	68.48

– Information not available.
[a] Computed from the confusion matrix or information presented in the studies.

that belonged to the final stage of labor and consisted of 20 risky and 60 healthy signals. The highest performance (GM = 83.67%) was reported in the study.

An analysis depending on EMD features and SVM classifiers was implemented in [19] on a private dataset consisting of 90 randomly selected recordings of 20-min durations. A visual inspection was performed by three obstetricians, and labeling of the signals was executed according to the majority votes of these experts. A quite high sensitivity (Se = 95.00%) was reported; however, a specificity that was as high as the sensitivity (Sp = 70.00%) could not be achieved. Thus, the GM was computed as 81.54%.

Looking into the results presented in [74], where the sensitivity and specificity were 60.90% and 81.70%, respectively, the GM was computed as 70.53%. Generative models were applied to a total of 83 private recordings, and 20 of them were evaluated as not healthy. As expected, compared to specificity, a low sensitivity value was achieved.

The classification performance varies depending on several parameters such as the period of data collection, division criteria (visual inspection or an objective annotation such as pH), and the structure of the database (number of recordings, duration of recordings, distribution of recordings among the classes, etc.). According to the comparison in Table 7, the best sensitivity was reported by Krupa et al. [19]. When the study is examined, it is obviously seen that an ad-hoc and private database was used in the study. The recordings belong to the antepartum period and were divided by three experienced obstetricians into normal and abnormal categories via visual inspection.

In addition, only a total of 90 recordings (consisting of 60 normal and 30 abnormal) were processed. In this manner, the authors reported a rather high sensitivity (95%). As for our case, the experimental study was realized using publicly accessible intrapartum CTG recordings. In the intrapartum period, it is difficult to classify CTG recordings automatically since the recordings exhibit extremely high irregularities. Moreover, an objective biochemical marker, the umbilical artery pH, was considered for separating the recordings. For this reason, the reported sensitivity (76.83%) can be accepted as encouraging. In addition, the results of the experimental study are in agreement with related studies given in Table 7.

From a clinical point of view, automated CTG analysis systems support clinicians in their decision-making processes by ensuring diagnostic indices that cannot be detected using the naked eye [75]. Nevertheless, clinicians note that the visual detection of macroscopic and temporary changes such as acceleration and deceleration patterns in FHR time-series are an integral part of a clinical examination. This situation indicates that automated CTG analysis systems have not yet been fully adopted by clinicians [36].

In addition, clinicians expressed the opinion that the automatic detection and notification of specific patterns such as decelerations, low FHRV, bradycardia, and tachycardia are very beneficial in clinical applications. CTG-AOS was designed to build different experimental studies and was developed as a retrospective evaluation tool for the numerical analysis of FHR signals. Thus, IBTF features may contribute to providing more robust decision-making tools by ensuring new statistical indices for fetal hypoxia detection. Finally, although we achieved rather promising results, we believe that more research is needed using larger databases to determine whether IBTF features are useful for the detection of fetal hypoxia.

4. Conclusion

The main drawbacks of CTG surveillance techniques stem from the reading and interpretation of CTG traces owing to a lack of practical standards and high inter- and even intra-observer variability. For this reason, automated CTG analysis is the most promising way to overcome these drawbacks. The overall goal of this study was to verify an open-access software package called CTG-OAS, which was equipped with advanced and modern signal processing techniques for the analysis of FHR signals. CTG-OAS can be used to realize numerous experimental studies and to develop new algorithms.

A sample experimental study on the usability of IBTF features such as contrast, correlation, energy, and homogeneity was conducted using CTG-OAS for the detection of fetal hypoxia. We examined the effects of the novel feature extraction approach, which is based on a combination of STFT and GLCM, on classification performance. To this end, three popular machine learning techniques (ANN, SVM, and *k*-NN) were employed. The performances of the models were examined by using confusion matrices and were visualized using ROC curves. We achieved quite promising results (Se = 76.83%, Sp = 78.27%, and AUC = 0.84). SVM outperformed the other classifiers.

Finally, a comparison was conducted with related studies considering several significant parameters with regard to automated CTG analysis. Based on the results of the experimental study and the detailed comparison, we believe that CTG-OAS can be a valuable tool for the analysis of FHR signals as well as in developing new algorithms.

The details of developments in CTG-OAS can be followed at www.ctganalysis.com, and the software can be downloaded from this website. In future work, we will try to enhance the software performance using different machine learning techniques and several heuristic automatic feature selection algorithms to build stronger decision support tools for the interpretation of CTG signals.

Conflicts of interest

The authors declare that there is no conflict of interest related to this paper.

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

Acknowledgments

The authors would like to thank Chudáček et al. for providing open-access CTU-UHB Intrapartum Cardiotocography Database. The database is available at https://physionet.org/physiobank/database/ctu-uhb-ctgdb/.

References

[1] R.M. Grivell, Z. Alfirevic, G.M. Gyte, D. Devane, Antenatal cardiotocography for fetal assessment, Cochrane Database Syst. Rev. (2010) 1–48, http://dx.doi.org/10.1002/14651858.CD007863.pub2.

[2] D. Ayres-de-Campos, C.Y. Spong, E. Chandraharan, FIGO consensus guidelines on intrapartum fetal monitoring: cardiotocography, Int. J. Gynecol. Obstet. 131 (2015) 13–24, http://dx.doi.org/10.1016/j.ijgo.2015.06.020.

[3] C. Kouskouti, K. Regner, J. Knabl, F. Kainer, Cardiotocography and the evolution into computerised cardiotocography in the management of intrauterine growth restriction, Arch. Gynecol. Obstet. 295 (2017) 811–816, http://dx.doi.org/10.1007/s00404-016-4282-8.

[4] D. Ayres-de-campos, J. Bernardes, A. Garrido, J. Marques-de-sá, L. Pereira-leite, SisPorto 2.0: a program for automated analysis of cardiotocograms, J. Matern. Fetal. Med. 9 (2000) 311–318, http://dx.doi.org/10.3109/14767050009053454.

[5] J. Bernardes, A. Costa-Pereira, D. Ayres-De-Campos, H.P. Van Geijn, L. Pereira-Leite, Evaluation of interobserver agreement of cardiotocograms, Int. J. Gynecol. Obstet. 57 (1997) 33–37, http://dx.doi.org/10.1016/S0020-7292(97)02846-4.

[6] S. Rhose, A.M.F. Heinis, F. Vandenbussche, J. van Drongelen, J. van Dillen, Inter- and intra-observer agreement of non-reassuring cardiotocography analysis and subsequent clinical management, Acta Obstet. Gynecol. Scand. 93 (2014) 596–602, http://dx.doi.org/10.1111/aogs.12371.

[7] H.P. van Geijn, 2 developments in CTG analysis, Baillieres. Clin. Obstet. Gynaecol. 10 (1996) 185–209, http://dx.doi.org/10.1016/S0950-3552(96)80033-2.
[8] M. Kazandi, F. Sendag, F. Akercan, M.C. Terek, G. Gundem, Different types of variable decelerations and their effects to neonatal outcome, Singap. Med. J. 44 (2003) 243–247.
[9] R. Mantel, H.P. van Geijn, F.J.M. Caron, J.M. Swartjes, E.E. van Woerden, H.W. Jongswa, Computer analysis of antepartum fetal heart rate: 2. Detection of accelerations and decelerations, Int. J. Biomed. Comput. 25 (1990) 273–286, http://dx.doi.org/10.1016/0020-7101(90)90031-O.
[10] H. Gonçalves, A.P. Rocha, D. Ayres-de-Campos, J. Bernardes, Linear and nonlinear fetal heart rate analysis of normal and acidemic fetuses in the minutes preceding delivery, Med. Biol. Eng. Comput. 44 (2006) 847, http://dx.doi.org/10.1007/s11517-006-0105-6.
[11] H. Gonçalves, J. Bernardes, A. Paula Rocha, D. Ayres-de-Campos, Linear and nonlinear analysis of heart rate patterns associated with fetal behavioral states in the antepartum period, Early Hum. Dev. 83 (2007) 585–591, http://dx.doi.org/10.1016/j.earlhumdev.2006.12.006.
[12] Z. Cömert, A.F. Kocamaz, Evaluation of fetal distress diagnosis during delivery stages based on linear and nonlinear features of fetal heart rate for neural network community, Int. J. Comput. Appl. 156 (2016) 26–31, http://dx.doi.org/10.5120/ijca2016912417.
[13] J. Spilka, V. Chudáček, M. Koucký, L. Lhotská, M. Huptych, P. Janků, G. Georgoulas, C. Stylios, Using nonlinear features for fetal heart rate classification, Biomed. Signal Process. Control. 7 (2012) 350–357, http://dx.doi.org/10.1016/j.bspc.2011.06.008.
[14] M.G. Signorini, G. Magenes, S. Cerutti, D. Arduini, Linear and nonlinear parameters for the analysis of fetal heart rate signal from cardiotocographic recordings, IEEE Trans. Biomed. Eng. 50 (2003) 365–374, http://dx.doi.org/10.1109/TBME.2003.808824.
[15] E. Salamalekis, P. Thomopoulos, D. Giannaris, I. Salloum, G. Vasios, A. Prentza, D. Koutsouris, Computerised intrapartum diagnosis of fetal hypoxia based on fetal heart rate monitoring and fetal pulse oximetry recordings utilising wavelet analysis and neural networks, BJOG Int. J. Obstet. Gynaecol. 109 (2002) 1137–1142, http://dx.doi.org/10.1111/j.1471-0528.2002.01388.x.
[16] G. Georgoulas, C. Stylios, P. Groumpos, Feature extraction and classification of fetal heart rate using wavelet analysis and support vector machines, Int. J. Artif. Intell. Tools. 15 (2006) 411–432, http://dx.doi.org/10.1142/S0218213006002746.
[17] Z. Cömert, A.F. Kocamaz, Using wavelet transform for cardiotocography signals classification, 25th Signal Process. Commun. Appl. Conf., IEEE, Antalya (2017) 1–4, http://dx.doi.org/10.1109/SIU.2017.7960152.
[18] M.R. Ortiz, E.R. Bojorges, S.D. Aguilar, J.C. Echeverria, R. Gonzalez-Camarena, S. Carrasco, M.J. Gaitan, A. Martinez, Analysis of high frequency fetal heart rate variability using empirical mode decomposition, Comput. Cardiol. (2005) 675–678, http://dx.doi.org/10.1109/CIC.2005.1588192.
[19] N. Krupa, M.A. Mohd Ali, E. Zahedi, S. Ahmed, F.M. Hassan, Antepartum fetal heart rate feature extraction and classification using empirical mode decomposition and support vector machine, Biomed. Eng. Online 10 (2011) 1–15, http://dx.doi.org/10.1186/1475-925x-10-6.
[20] J.Y. Kwon, I.Y. Park, J.C. Shin, J. Song, R. Tafreshi, J. Lim, Specific change in spectral power of fetal heart rate variability related to fetal acidemia during labor: comparison between preterm and term fetuses, Early Hum. Dev. 88 (2012) 203–207, http://dx.doi.org/10.1016/j.earlhumdev.2011.08.007.
[21] M. Romano, L. Iuppariello, A.M. Ponsiglione, G. Improta, P. Bifulco, M. Cesarelli, Frequency and time domain analysis of foetal heart rate variability with traditional indexes: a critical survey, Comput. Math Methods Med. (2016) 1–12, http://dx.doi.org/10.1155/2016/9585431.
[22] Z. Cömert, A.F. Kocamaz, A study based on gray level Co-occurrence matrix and neural network community for determination of hypoxic fetuses, Int. Artif. Intell. Data Process. Symp. TR (2016) 569–573, http://dx.doi.org/10.13140/RG.2.2.23901.00489.
[23] Z. Cömert, A.F. Kocamaz, Cardiotocography analysis based on segmentation-based fractal texture decomposition and extreme learning machine, 25th Signal Process. Commun. Appl. Conf. (2017) 1–4, http://dx.doi.org/10.1109/SIU.2017.7960397.
[24] M. Romano, P. Bifulco, M. Ruffo, G. Improta, F. Clemente, M. Cesarelli, Software for computerised analysis of cardiotocographic traces, Comput. Methods Programs Biomed. 124 (2016) 121–137, http://dx.doi.org/10.1016/j.cmpb.2015.10.008.
[25] G.S. Dawes, M. Moulden, C.W. Redman, System 8000: computerized antenatal FHR analysis, J. Perinat. Med. 19 (1991) 47–51.
[26] G. Magenes, M.G. Signorini, M. Ferrario, F. Lunghi, 2CTG2: A new system for the antepartum analysis of fetal heart rate, in: T. Jarm, P. Kramar, A. Zupanic (Eds.), 11th Mediterr. Conf. Med. Biomed. Eng. Comput., Springer Berlin Heidelberg, Berlin, Heidelberg, 2007, pp. 781–784, http://dx.doi.org/10.1007/978-3-540-73044-6_203.
[27] A. Sbrollini, A. Agostinelli, L. Burattini, M. Morettini, F. DiNardo, S. Fioretti, L. Burattini, CTG analyzer: a graphical user interface for cardiotocography, 2017 39th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (2017) 2606–2609, http://dx.doi.org/10.1109/EMBC.2017.8037391.
[28] A. Alonso-Betanzos, B. Guijarro-Berdinas, V. Moret-Bonillo, S. Lopez-Gonzalez, The NST-EXPERT project: the need to evolve, Artif. Intell. Med. 7 (1995) 297–313.
[29] B. Guijarro-Berdiñas, A. Alonso-Betanzos, O. Fontenla-Romero, Intelligent analysis and pattern recognition in cardiotocographic signals using a tightly coupled hybrid system, Artif. Intell. 136 (2002) 1–27, http://dx.doi.org/10.1016/S0004-3702(01)00163-1.
[30] D. Ayres-de-Campos, M. Rei, I. Nunes, P. Sousa, J. Bernardes, SisPorto 4.0 – computer analysis followingthe 2015 FIGO Guidelines for intrapartum fetal monitoring, J. Mater. Neonatal. Med. 30 (2017) 62–67, http://dx.doi.org/10.3109/14767058.2016.1161750.
[31] R.D. Keith, K.R. Greene, Development, evaluation and validation of an intelligent system for the management of labour, Baillieres, Clin. Obstet. Gynaecol. 8 (1994) 583–605.
[32] I. Nunes, D. Ayres-de-Campos, C. Figueiredo, J. Bernardes, An overview of central fetal monitoring systems in labour, J. Perinat. Med. 41 (2013) 93–99, http://dx.doi.org/10.1515/jpm-2012-0067.
[33] V. Chudáček, J. Spilka, M. Burša, P. Janků, L. Hruban, M. Huptych, L. Lhotská, Open access intrapartum CTG database, BMC Pregnancy Childbirth 14 (2014) 16, http://dx.doi.org/10.1186/1471-2393-14-16.
[34] J. Spilka, V. Chudáček, P. Janků, L. Hruban, M. Burša, M. Huptych, L. Zach, L. Lhotská, Analysis of obstetricians' decision making on CTG recordings, J. Biomed. Inform. 51 (2014) 72–79, http://dx.doi.org/10.1016/j.jbi.2014.04.010.
[35] V. Chudacek, J. Spilka, P. Janku, M. Koucky, L. Lhotska, M. Huptych, Automatic evaluation of intrapartum fetal heart rate recordings: a comprehensive analysis of useful features, Physiol. Meas. 32 (2011) 1347–1360, http://dx.doi.org/10.1088/0967-3334/32/8/022.
[36] G. Georgoulas, P. Karvelis, J. Spilka, V. Chudáček, C.D. Stylios, L. Lhotská, Investigating pH based evaluation of fetal heart rate (FHR) recordings, Health Technol. (Berl.) 7 (2017) 241–254, http://dx.doi.org/10.1007/s12553-017-0201-7.
[37] H. Cao, D.E. Lake, J.E. Ferguson 2nd, C.A. Chisholm, M.P. Griffin, J.R. Moorman, Toward quantitative fetal heart rate monitoring, IEEE Trans. Biomed. Eng. 53 (2006) 111–118, http://dx.doi.org/10.1109/TBME.2005.859807.
[38] G.G. Georgoulas, C.D. Stylios, G. Nokas, P.P. Groumpos, Classification of fetal heart rate during labour using hidden Markov models, IEEE Int. Jt. Conf. Neural Netw. (2004) 2471–2475, http://dx.doi.org/10.1109/IJCNN.2004.1381017.
[39] E. Salamalekis, E. Hintipas, I. Salloum, G. Vasios, C. Loghis, N. Vitoratos, C. Chrelias, G. Creatsas, Computerized analysis of fetal heart rate variability using the matching pursuit technique as an indicator of fetal hypoxia during labor, J. Mater. Fetal. Neonatal. Med. 19 (2006) 165–169, http://dx.doi.org/10.1080/14767050500233290.
[40] Z. Cömert, A.F. Kocamaz, Novel software for comprehensive analysis of cardiotocography signals CTG-OAS, in: A. Karci (Ed.), Int. Conf. Artif. Intell. Data Process., IEEE, Malatya, 2017, pp. 1–6, http://dx.doi.org/10.1109/IDAP.2017.8090210.
[41] Z. Cömert, A.F. Kocamaz, Comparison of machine learning techniques for fetal heart rate classification, Acta Phys. Pol. A 132 (2017) 451–454, http://dx.doi.org/10.12693/APhysPolA.131.451.
[42] M. Cesarelli, M. Romano, P. Bifulco, F. Fedele, M. Bracale, An algorithm for the recovery of fetal heart rate series from CTG data, Comput. Biol. Med. 37 (2007) 663–669, http://dx.doi.org/10.1016/j.compbiomed.2006.06.003.
[43] M. Romano, G. Faiella, P. Bifulco, G. D'Addio, F. Clemente, M. Cesarelli, Outliers detection and processing in CTG monitoring, in: L.M. Roa Romero (Ed.), XIII Mediterr. Conf. Med. Biol. Eng. Comput. 2013 MEDICON, Sept. 2013, Seville, Spain, Springer International Publishing, Cham, 2013, pp. 651–654, http://dx.doi.org/10.1007/978-3-319-00846-2_161 (25–28).
[44] H. Gonçalves, A. Costa, D. Ayres-de-Campos, C. Costa-Santos, A.P. Rocha, J. Bernardes, Comparison of real beat-to-beat signals with commercially available 4 Hz sampling on the evaluation of foetal heart rate variability, Med. Biol. Eng. Comput. 51 (2013) 665–676, http://dx.doi.org/10.1007/s11517-013-1036-7.
[45] J. Spilka, G. Georgoulas, P. Karvelis, V.P. Oikonomou, V. Chudáček, C. Stylios, L. Lhotská, P. Jankru, Automatic evaluation of FHR recordings from CTU-UHB CTG database, in: M. Bursa, S. Khuri, M.E. Renda (Eds.), Inf. Technol. Bio- Med. Informatics 4th Int. Conf. ITBAM 2013, Prague, Czech Republic, August 28, 2013. Proc., Springer Berlin Heidelberg, Berlin, Heidelberg, 2013, pp. 47–61, http://dx.doi.org/10.1007/978-3-642-40093-3_4.
[46] F.N. Fritsch, R.E. Carlson, Monotone piecewise cubic interpolation, SIAM J. Numer. Anal. 17 (1980) 238–246, http://dx.doi.org/10.1137/0717021.
[47] M. Cesarelli, M. Romano, M. Ruffo, P. Bifulco, G. Pasquariello, A. Fratini, PSD modifications of FHRV due to interpolation and CTG storage rate, Biomed. Signal Process. Control. 6 (2011) 225–230, http://dx.doi.org/10.1016/j.bspc.2010.10.002.
[48] P. Fergus, A. Hussain, D. Al-Jumeily, D.-S. Huang, N. Bouguila, Classification of caesarean section and normal vaginal deliveries using foetal heart rate signals and advanced machine learning algorithms, Biomed. Eng. Online 16 (2017) 89, http://dx.doi.org/10.1186/s12938-017-0378-z.
[49] S.N. Al-yousif, M.A.M. Ali, Cardiotocography trace pattern evaluation using MATLAB program, in: Int. Conf. Biomed. Eng. Technol., IACSIT Press, Singapore, 2011, pp. 153–158.
[50] H.P. van Geijn, H.W. Jongsma, J. de Haan, T.K.A.B. Eskes, Analysis of heart rate and beat-to-beat variability: interval difference index, Am. J. Obstet. Gynecol. 138 (1980) 246–252, http://dx.doi.org/10.1016/0002-9378(80)90242-2.
[51] M. Romano, P. Bifulco, A.M. Ponsiglione, G.D. Gargiulo, F. Amato, M. Cesarelli, Evaluation of floatingline and foetal heart rate variability, Biomed. Signal Process. Control. 39 (2018) 185–196, http://dx.doi.org/10.1016/j.bspc.2017.07.018.
[52] M. Malik, J.T. Bigger, A.J. Camm, R.E. Kleiger, A. Malliani, A.J. Moss, P.J. Schwartz, Heart rate variability Standards of measurement, physiological

interpretation, and clinical use, Eur. Heart J. 17 (1996) 354–381 http://eurheartj.oxfordjournals.org/content/17/3/354. abstract.
[53] W.J. Parer, J.T. Parer, R.H. Holbrook, B.S.B. Block, Validity of mathematical methods of quantitating fetal heart rate variability, Am. J. Obstet. Gynecol. 153 (1985) 402–409, http://dx.doi.org/10.1016/0002-9378(85)90078-X.
[54] M. Cesarelli, M. Romano, P. Bifulco, Comparison of short term variability indexes in cardiotocographic foetal monitoring, Comput. Biol. Med. 39 (2009) 106–118, http://dx.doi.org/10.1016/j.compbiomed.2008.11.010.
[55] P. Fergus, M. Selvaraj, C. Chalmers, Machine learning ensemble modelling to classify caesarean section and vaginal delivery types using Cardiotocography traces, Comput. Biol. Med. 93 (2018) 7–16, http://dx.doi.org/10.1016/j.compbiomed.2017.12.002.
[56] P. Van Leeuwen, S. Lange, H. Bettermann, D. Grönemeyer, W. Hatzmann, Fetal heart rate variability and complexity in the course of pregnancy, Early Hum. Dev. 54 (1999) 259–269, http://dx.doi.org/10.1016/S0378-3782(98)00102-9.
[57] M. Cesarelli, M. Romano, P. Bifulco, G. Improta, G. D'Addio, An application of symbolic dynamics for FHRV assessment, Stud. Health Technol. Inform. 180 (2012) 123–127.
[58] D.K. Lake, J.R. Moorman, C. Hanqing, Sample entropy estimation using sampen, PhysioNet (2012) https://www.physionet.org/physiotools/sampen/.
[59] A. Şengür, Y. Guo, Y. Akbulut, Time-frequency texture descriptors of EEG signals for efficient detection of epileptic seizure, Brain Inf. 3 (2016) 101–108, http://dx.doi.org/10.1007/s40708-015-0029-8.
[60] A. Sengur, M. Gedikpinar, Y. Akbulut, E. Deniz, V. Bajaj, Y. Guo, DeepEMGNet: an application for efficient discrimination of ALS and normal EMG signals, in: T. Bvrezina, R. Jabłoński (Eds.), Mechatronics 2017 Recent Technol. Sci. Adv., Springer International Publishing, Cham, 2018, pp. 619–625, http://dx.doi.org/10.1007/978-3-319-65960-2_77.
[61] R.M. Haralick, Statistical and structural approaches to texture, Proc. IEEE 67 (1979) 786–804, http://dx.doi.org/10.1109/PROC.1979.11328.
[62] R.M. Haralick, K. Shanmugam, I. Dinstein, Textural features for image classification, IEEE Trans. Syst. Man. Cybern. SMC-3 (1973) 610–621, http://dx.doi.org/10.1109/TSMC.1973.4309314.
[63] J. Spilka, J. Frecon, R. Leonarduzzi, N. Pustelnik, P. Abry, M. Doret, Sparse support vector machine for intrapartum fetal heart rate classification, IEEE J. Biomed. Heal. Inf. 1 (2016), http://dx.doi.org/10.1109/JBHI.2016.2546312.
[64] D.W. Aha, D. Kibler, M.K. Albert, Instance-based learning algorithms, Mach. Learn. 6 (1991) 37–66, http://dx.doi.org/10.1007/BF00153759.
[65] K.Q. Weinberger, L.K. Saul, Distance metric learning for large margin nearest neighbor classification, J. Mach. Learn. Res. 10 (2009) 207–244 https://www.scopus.com/inward/record.uri?eid=2-s2.0-61749090884&partnerID=40&md5=ee1807133eef2ca55697144ca4961c9f.
[66] H. Wang, I. Duntsch, Nearest neighbours without k: a classification formalism based on probability, Brock Univ. (2003) 1–10.
[67] M.T. Hagan, H.B. Demuth, M.H. Beale, O. DeJesús, O. DeJesús, Neural Network Design, Martin Hagan, 2014 https://books.google.com.tr/books?id=4EW9oQEACAAJ.
[68] N. Karayiannis, A.N. Venetsanopoulos, Artificial Neural Networks: Learning Algorithms, Performance Evaluation, and Applications, Springer, US, 1992 https://books.google.com.tr/books?id=n0RYqCELd-IC.
[69] C. Cortes, V. Vapnik, Support-vector networks, Mach. Learn. 20 (1995) 273–297, http://dx.doi.org/10.1007/BF00994018.
[70] V.N. Vapnik, The Nature of Statistical Learning Theory, Springer Science & Business Media, 2013 https://books.google.com.tr/books?id=EoDSBwAAQBAJ.
[71] Y. Sun, M.S. Kamel, A.K.C. Wong, Y. Wang, Cost-sensitive boosting for classification of imbalanced data, Pattern Recognit. 40 (2007) 3358–3378, http://dx.doi.org/10.1016/j.patcog.2007.04.009.
[72] T.C.W. Landgrebe, R.P.W. Duin, Efficient multiclass ROC approximation by decomposition via confusion matrix perturbation analysis, IEEE Trans. Pattern Anal. Mach. Intell. 30 (2008) 810–822, http://dx.doi.org/10.1109/TPAMI.2007.70740.
[73] C.D. Stylios, G. Georgoulas, P. Karvelis, J. Spilka, V. Chudáček, L. Lhotska, Least squares support vector machines for FHR classification and assessing the pH based categorization, in: E. Kyriacou, S. Christofides, C.S. Pattichis (Eds.), XIV Mediterr. Conf. Med. Biol. Eng. Comput. 2016 MEDICON 2016, March 31st-April 2nd 2016, Paphos, Cyprus, Springer International Publishing, Cham, 2016, pp. 1211–1215, http://dx.doi.org/10.1007/978-3-319-32703-7_234.
[74] S. Dash, J.G. Quirk, P.M. Djuric, Fetal heart rate classification using generative models, IEEE Trans. Biomed. Eng. 61 (2014) 2796–2805, http://dx.doi.org/10.1109/TBME.2014.2330556.
[75] A. Agostinelli, E. Braccili, E. Marchegiani, R. Rosati, A. Sbrollini, L. Burattini, M. Morettini, F. DiNardo, S. Fioretti, L. Burattini, Statistical baseline assessment in cardiotocography, 2017 39th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (2017) 3166–3169, http://dx.doi.org/10.1109/EMBC.2017.8037529.