

# UNSUPERVISED CLUSTERING AND ANALYSIS OF CONTRACTION-DEPENDENT FETAL HEART RATE SEGMENTS

Liu Yang,<sup>†</sup> Cassandra Heiselman,<sup>◇</sup> J. Gerald Quirk,<sup>◇</sup> and Petar M. Djurić,<sup>†</sup>

<sup>†</sup>Department of Electrical and Computer Engineering,

<sup>◇</sup> Department of Obstetrics, Gynecology and Reproductive Medicine  
Stony Brook University, Stony Brook, NY, USA 11794-2350

## ABSTRACT

The computer-aided interpretation of fetal heart rate (FHR) and uterine contraction (UC) has not been developed well enough for wide use in delivery rooms. The main challenges still lie in the lack of unclear and nonstandard labels for cardiotocography (CTG) recordings, and the timely prediction of fetal state during monitoring. Rather than traditional supervised approaches to FHR classification, this paper demonstrates a way to understand the UC-dependent FHR responses in an unsupervised manner. In this work, we provide a complete method for FHR-UC segment clustering and analysis via the Gaussian process latent variable model, and density-based spatial clustering. We map the UC-dependent FHR segments into a space with a visual dimension and propose a trajectory-based FHR interpretation method. Three metrics of FHR trajectory are defined and an open-access CTG database is used for testing the proposed method.

**Index Terms**— cardiotocography, trajectory of contraction-dependent fetal heart rate segments, unsupervised clustering

## 1. INTRODUCTION

Electrical fetal monitoring (EFM) provides continuous surveillance of fetal heart rate (FHR) and uterine activity (UA) in delivery rooms. This is also referred as cardiotocography (CTG). In the past decades, EFM techniques have been developed well and clinically used in many countries [1]. The aim of CTG monitoring is to alert obstetricians and gynecologists (OB-GYNs) of fetal hypoxia and acidosis during childbirth, allowing them to make timely decisions for surgical intervention. Traditionally, the patterns of CTG are visually examined by experienced OB-GYNs based on published guidelines [2, 3, 4]. A consistent interpretation of CTG and an accurate assessment of fetal wellbeing are not yet possible, owing to the noisy nature of measurements, different standards of reading resolution, as well as the subjective opinions of clinicians [5]. Hence, computer-aided FHR analysis has become a fast-growing research area in recent years [6].

To achieve the goal of automatically and intelligently predicting the risk of fetal distress, data-driven machine learning (ML) techniques have been continuously studied. With the assistance from data and computation, the latent patterns and information that are hidden from human eyes can be explored. The representative works of FHR tracing interpretation mainly follow two directions: feature-based and neural network-based approaches. Among them, the FHR analysis with features are relatively plentiful. A comprehensive study in [7] reaches encouraging results by considering time-domain, frequency-domain and nonlinear features. In addition, nonparametric Bayesian

approaches are also promising. For example, [8] makes use of fourteen FHR features in the hierarchical Dirichlet process mixture model for classification. Further, deep Gaussian process (GP) methods have also been investigated [9] for both supervised and unsupervised fetal status learning. Except for studying with features, some research leverages the power of neural network. In [10, 11], convolutional neural networks produce good performance by automatically conducting feature extraction and selection.

Even though many cutting-edge ML tools have been tested for FHR analysis, a reliable, interpretable, and applicable solution for CTG monitoring and prediction has not been found and/or accepted for a widespread use. The main challenges include the following two aspects. First, according to [3, 4], clinically, the FHR tracings are classified into three categories: normal, atypical or indeterminate, and abnormal. While there is a broad consensus on the normal and abnormal classes, a large portion of FHR recordings, e.g. 50%-84%, are grouped to be uncertain. This fact leads to a problem that the agreement on the standards of CTG labeling are hard to achieve. While umbilical blood pH value is a common index of fetal status, experts have not settled on the threshold that can be used for classifying the fetuses. Also, such indicator cannot reflect the state of the fetus long before delivery. Second, most of the existing works of FHR assessment only consider information from FHR tracings, whereas little attention is put on other data resources, such as UA and clinical information. The inclusion of knowledge from UA signals could lead to improvements in evaluating fetal wellbeing. For example, [12] demonstrates a causal relationship from UA to FHR. In [9] and [13, 14], the features computed according to uterine contractions (UCs) show significant influence on performance enhancement.

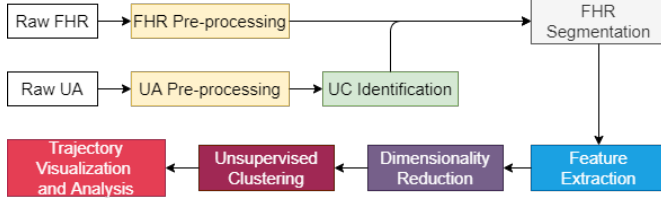
In this paper, we propose to analyze segments of FHR data around the UCs and cluster them in an unsupervised way. The main tools of this work include signal filtering and interpolation, the Gaussian process latent variable model (GPLVM) for reducing dimensionality, and density-based spatial clustering of applications with noise (DBSCAN) as an unsupervised clustering method. Additionally, in this paper, we use a method described in our previous work [15] for identifying contractions from UA recordings. We demonstrate the performance of the proposed method on real data.

## 2. PROBLEM AND BACKGROUND

In this paper, we intend to take another approach to the problem of FHR monitoring. First of all, we do not consider the use of supervised learning for fetal status classification or prediction, but only use the experts' labels as a reference for analysis. Second, we do not take the entire FHR recordings or long segments as objects, but focus on the

---

This work has been supported by NIH under Award 1R01HD097188-01.



**Fig. 1:** Flowchart of the proposed method.

state of the fetus and its changes before and after contractions. We believe that the well-being of a fetus is well reflected through fetal responses to every UCs. The advantages of our approach are 1) we remove all the difficulties associated with pH values, and 2) we start a time-dependent, also called event/contraction-dependent analysis, which will allow for online FHR monitoring.

Some previous research supports our idea. To our knowledge, [16] is the first paper that mentioned using unsupervised technique for FHR clustering. Unfortunately, due to the undesirable results, this method was questioned by the authors. Years later, [9] considered using a five-layer deep GP to project the raw FHR recordings to a low-dimensional space. As uncertainty declines over layers, the normal and abnormal FHR samples are separated. Another recent work, [17], applied the phase space reconstruction to recover the latent manifolds of six consecutive FHR segments in the last thirty minutes before delivery. By comparing the manifolds of healthy and unhealthy fetal recordings and computing the Euclidean distances between them, the probabilities that they are anomalies, i.e., abnormal samples, can be obtained. These results show that, as birth approaches, the statistical difference of outlier probabilities between data of healthy and unhealthy fetuses increases until the last five minutes.

### 3. MODEL AND METHOD

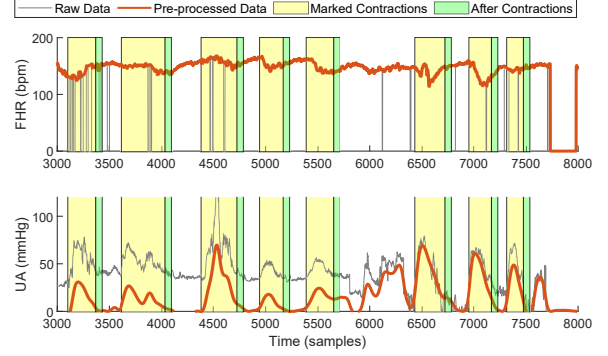
Our idea for contraction-dependent FHR segment clustering and analysis is elaborated in this section. We describe the data pre-processing, the access and use of UC information, feature extraction, dimensionality reduction, unsupervised clustering, result visualization and analysis. The flowchart of the entire solution is shown in Fig. 1.

#### 3.1. Data Pre-processing

Given a CTG database, a critical first step is data pre-processing. The details about UA pre-processing can be found in [15]. For raw FHR, any zero-value sample means missing data, and samples outside of the range 50 bpm ~ 200 bpm are artifacts. Moreover, any successive samples with absolute differences greater than 25 bpm are regarded as abrupt increase or decrease, and if the differences are in 10 bpm ~ 25 bpm, they are unstable samples. The pre-processing of FHR includes two steps: 1) interpolating the missing samples and artifacts whose duration is less than 15 seconds, and 2) interpolating the unstable segments from an abrupt change to the first sample of the following stable segment, that is defined as any successive five samples with differences smaller than 10 bpm. Figure 2 compares the raw data and pre-processed data on a CTG example.

#### 3.2. UC Identification and FHR Segmentation

We applied the GP-based UC identification method proposed in [15] for accessing the contraction information, such as onset, offset and peak time. With these data, we can mark the corresponding FHR



**Fig. 2:** Data pre-processing and segmentation.

segments during contractions and after contractions, as shown in Fig. 2. The 15-second window length is used after contractions. Note that, for any contraction event, if the FHR segment still contains more than 50% missing samples after pre-processing, this event is ignored. Besides, if the interval between two consecutive events is smaller than 15s, the first one is discounted.

#### 3.3. Feature Extraction

Considering the state of a fetus to be latent in its responses to the UC pressure from the mother, three categories of features are extracted from the FHR contraction segments and the FHR post-contraction segments. Table 1 gives a list of 18 features that are used in this study and suggested from previous works [7, 9]. In addition, we also compute the differences of these features from the FHR segments during and after contraction and add them as another group of features. This is because the timing of response also reflects how well the fetus recovers from a big contraction pressure. To this end, for each contraction event, a total of 54 features are calculated.

#### 3.4. Dimensionality Reduction using GPLVM

Dimensionality reduction is usually an essential step of analyzing high-dimensional data. A GPLVM is a generalization of probabilistic principal component analysis in which the assumption of linear mapping is replaced by a nonlinear kernel function [18]. Suppose a high-dimensional observation is  $\mathbf{Y} \in \mathbb{R}^{N \times D}$  where each column  $\mathbf{y}_d$  represents the samples in the  $d$ th dimension,  $d = 1, 2, \dots, D$ . For our FHR study, each row of  $\mathbf{Y}$  matches the feature vector extracted from a single contraction event. Let  $\mathbf{X} \in \mathbb{R}^{N \times Q}$  collect latent variables in a low-dimensional space, i.e.,  $Q \ll D$ , that can be the corresponding coordinates of features from every FHR segments. The GPLVM defines  $D$  independent nonlinear mappings from the latent space to the observation space:

$$\mathbf{y}_d = \mathbf{f}_d(\mathbf{X}) + \mathbf{e}_d, \quad (1)$$

where each function  $\mathbf{f}_d(\cdot)$  is drawn from a Gaussian-process, and the noises  $\mathbf{e}_d$  are independent for different  $d$ . Then, the marginal distribution of  $\mathbf{y}_d$  is

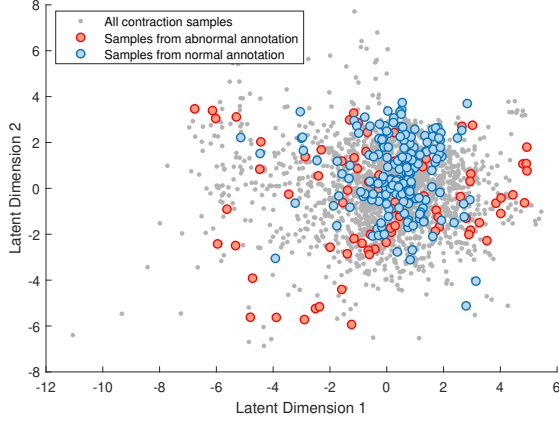
$$p(\mathbf{y}_d|\mathbf{X}) = \mathcal{N}(\mathbf{y}_d|\mathbf{0}, \Sigma_d), \quad (2)$$

where the covariance matrix  $\Sigma_d$  consists of the uncertainty from the input  $\mathbf{X}$  and the observation error, or

$$\Sigma_d = \mathbf{K}_d + \sigma^2 \mathbf{I}. \quad (3)$$

| Category  | Features  |
|-----------|---|
| Time      | Mean, Median, Standard deviation, Minimum, Maximum, Difference of median and minimum, Difference of maximum and median, Short-term variability, Short-term irregularity |
| Frequency | Energies in 0-0.06Hz (VLF), 0.06-0.3Hz (LF), 0.3-1Hz (MF) and 1-2Hz (HF), Ratio of energies: LF/(MF+HF)   |
| Nonlinear | Poincaré SD1, Poincaré SD2, Approximate entropy, Sample entropy   |

**Table 1:** Features of FHR Segments



**Fig. 3:** An example of sample distributions of normal and abnormal classes.

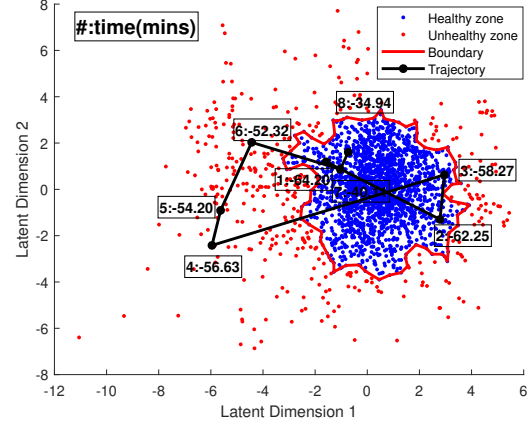
Here, the matrix  $\mathbf{K}_d \in \mathbf{R}^{N \times N}$  is obtained by evaluating the kernel function of the  $d$ th GP on all pairs of input data points. In general, in order to determine the importance of different dimensions, the automatic relevance determination (ARD) technique is used for defining the covariance functions. For example, the ARD version of the squared exponential kernel is

$$k_{\text{RBF-ARD}}(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \exp \left( -\frac{1}{2} \sum_{q=1}^Q \frac{(x_q - x'_q)^2}{l_q^2} \right). \quad (4)$$

A straightforward solution of GPLVM training is optimizing the log likelihood  $\log p(\mathbf{Y}|\mathbf{X}) = \sum_{d=1}^D \log p(\mathbf{y}_d|\mathbf{X})$  over the latent variable  $\mathbf{X}$  and all the hyperparameters in the kernel functions [18]. In [19], a variational inference framework of GPLVM was proposed where a low bound of marginal likelihood is obtained and maximized by introducing auxiliary variables and variational integration.

### 3.5. Unsupervised Clustering

We are not going to use any clinical measurements for labeling FHR segments, and therefore a good choice of unsupervised data clustering algorithm is important. For this purpose, we plotted the sample points of contraction-dependent FHR segments in a 2D space (see Fig. 3), which are mapped from the feature space by using a GPLVM. The details about the data set are explained in Section 4.1. In this example figure, the grey dots represent all contraction events detected from all FHR recordings used in this study. For comparing the sample distributions from normal and abnormal FHR tracings, we selected and emphasized the samples, that are from the recordings with more than 90% expert consensus about their classes, by blue and red colors. From the figure, we can observe that the FHR segments



**Fig. 4:** An example of a trajectory of an FHR recording.

from normal recordings are clustered but the ones from abnormal recordings are scattered. Based on these observations, we propose to build a “Healthy Zone” where most of the FHR contraction segments from normal tracings are located. We note that the states of healthy fetuses are largely confined to this healthy zone. In contrast, the states of non-healthy fetuses may frequently jump out of or stay out of the healthy zone from contraction to contraction.

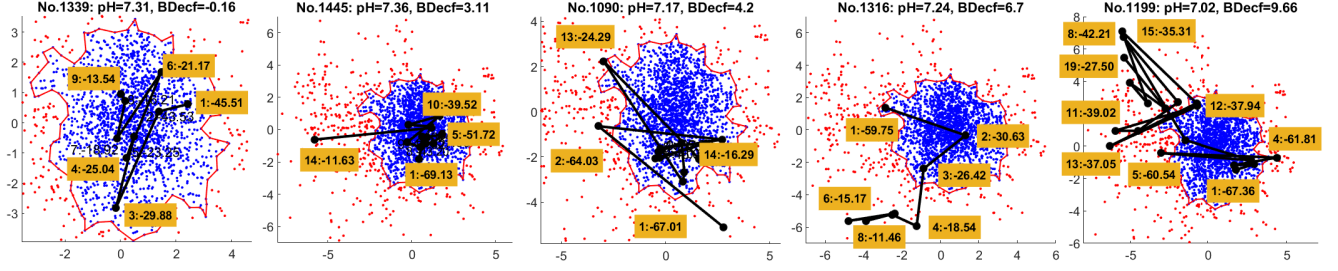
Accordingly, we applied a density-based spatial clustering technique, called DBSCAN, since it is able to identify a main cluster with a high density. There are two parameters in DBSCAN that need to be chosen. One is the maximum distance range between neighbors, and the other is a minimum sample size considered for a cluster. Note that, in this work, we treat the main cluster as the *healthy zone*, and all other clusters and outliers are in the *unhealthy zone*.

### 3.6. Trajectory Analysis and Metrics

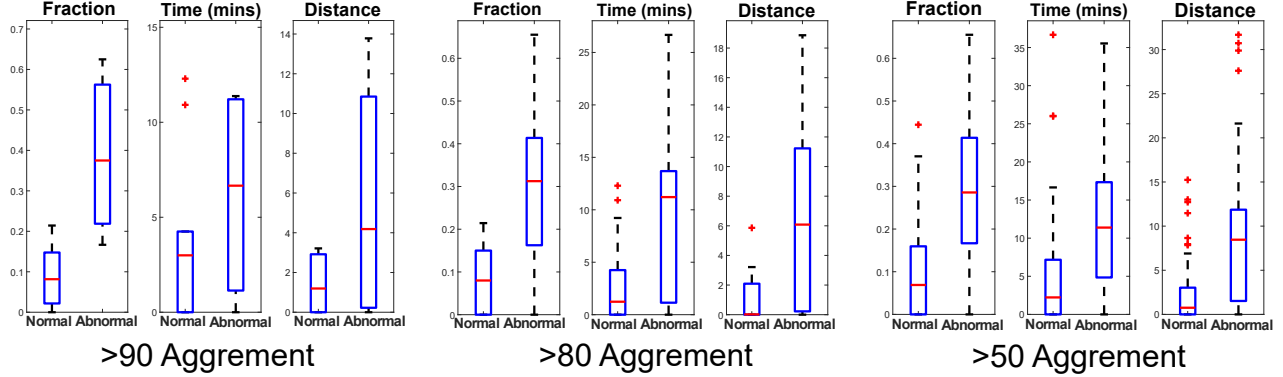
With the binary labels of sample points after clustering, a boundary of the “healthy” cluster can be created by generating an alpha shape of the finite set of points being assigned to the healthy zone. The function boundary in MATLAB R2021b is used. In the same space, for each FHR recording, we can also draw a trajectory of its contraction-dependent segments by connecting their correlated sample points with lines. In this way, the changes of the fetal status at every contractions can be visually checked if the dimensionality of latent space is smaller than three. Figure 4 shows an example of FHR data trajectory where 8 valid contractions were detected. The box near each black dot indicates the order of the contraction and its peak time. The time 0 minute is assigned to birth, resulting in negative time indices.

In order to quantitatively analyze the trajectories, we propose the following three metrics:

$$\text{Fraction} = n_{\text{unhealthy}} / n_{\text{all}} \quad (5)$$



**Fig. 5:** Five representative trajectories of FHR recordings. The left three are in normal class and the right two are in abnormal class.



**Fig. 6:** Comparison of three metrics of trajectories from normal and abnormal FHR recordings.

$$\text{Time} = \sum t_{\text{unhealthy}} \quad (6)$$

$$\text{Distance} = \sum d(\mathbf{x}_{\text{unhealthy}}, \text{Boundary}) . \quad (7)$$

The “Fraction” is the ratio of the number of samples in unhealthy zone and the total number of contractions, representing how often the fetal state travels out of the healthy region. The “Time” is a sum of contraction durations when the samples are in unhealthy zone and the rest periods between them. The “Distance” means the sum of the lengths from points outside the healthy zone to the boundary.

## 4. EXPERIMENTS

### 4.1. CTU-UHB Intrapartum CTG Database

An open access intrapartum CTG database [20], collected from April 2010 to August 2012 at the obstetrics ward of the University Hospital in Brno, Czech Republic, is popularly used for FHR studies. There are 552 available CTG recordings with corresponding clinical information. Each CTG signal includes a raw FHR and a raw UA tracings. They are both sampled at 4Hz. An inter- and intra-observer agreement on evaluating this dataset by 9 experienced obstetricians was published in [5]. By analyzing the annotation results, we selected 83 normal recordings and 50 abnormal recordings that all have more than 50% confidence about their classes. Among these recordings, 2237 contractions were detected. In addition, we also marked the recordings with more than 90% or 80% confidence on their classes.

### 4.2. Results and Discussion

We chose a 2D latent space as an example where we plotted the trajectories of all the selected recordings. Figure 5 presents five

representative trajectories including (1) always in healthy zone, (2) jumping into the unhealthy zone as the fetus is nearing birth, (3) occasional excursion from the healthy zone but close to the border, (4) gradual movement from the healthy zone to the unhealthy zone and (5) large and irregular jumps out of the healthy zone. Intuitively, the fetus might be well in the situations of (1), (2) and (3), because 1) with strong pressure, even the FHR from a healthy baby may show abnormal patterns as delivery approaches, and 2) occasional moves between healthy and unhealthy regions may imply that the fetus still has the ability to recover by itself. However, for the situation (4) and (5), the FHR contraction samples start to move to or stay in the unhealthy zone frequently, which should be an alarm for fetal distress and should alert OB-GYNs for further examination. In Fig. 6, we compare the distributions of three metrics over two classes by using different amount of data. We can see that, although more outliers appear when more data are used, the difference between normal and abnormal classes is still significant. An experiment in a 3D space shows promising results as well.

## 5. CONCLUSION

In this paper, we consider an unsupervised approach for improving the understanding of FHR recordings by exploiting uterine activity. The contraction-based FHR segments are transformed into a visual space for tracking the states of fetal responses to every valid uterine contraction. State-of-the-art machine learning techniques are applied on real-world cardiotocography recordings. The newly proposed metrics for evaluation of the trajectory of the recordings shows clear difference between the normal and abnormal classes.

## 6. REFERENCES

- [1] D. Ayres-de Campos, "Electronic fetal monitoring or cardiotocography, 50 years later: what's in a name?" *American Journal of Obstetrics & Gynecology*, vol. 218, no. 6, pp. 545–546, 2018.
- [2] A. Sweha, T. W. Hacker, and J. Nuovo, "Interpretation of the electronic fetal heart rate during labor," *American Family Physician*, vol. 59, no. 9, p. 2487, 1999.
- [3] G. A. Macones, G. D. Hankins, C. Y. Spong, J. Hauth, and T. Moore, "The 2008 National Institute of Child Health and Human Development workshop report on electronic fetal monitoring: update on definitions, interpretation, and research guidelines," *Journal of Obstetric, Gynecologic, & Neonatal Nursing*, vol. 37, no. 5, pp. 510–515, 2008.
- [4] D. Ayres-de Campos, C. Y. Spong, E. Chandrachud, and F. I. F. M. E. C. Panel, "FIGO consensus guidelines on intrapartum fetal monitoring: Cardiotocography," *International Journal of Gynecology & Obstetrics*, vol. 131, no. 1, pp. 13–24, 2015.
- [5] L. Hruban, J. Spilka, V. Chudáček, P. Janků, M. Hupčych, M. Burša, A. Hudec, M. Kacerovský, M. Koucký, M. Procházka *et al.*, "Agreement on intrapartum cardiotocogram recordings between expert obstetricians," *Journal of Evaluation in Clinical Practice*, vol. 21, no. 4, pp. 694–702, 2015.
- [6] M. J. Stout and A. G. Cahill, "Electronic fetal monitoring: past, present, and future," *Clinics in Perinatology*, vol. 38, no. 1, pp. 127–42, 2011.
- [7] Z. Zhao, Y. Zhang, and Y. Deng, "A comprehensive feature analysis of the fetal heart rate signal for the intelligent assessment of fetal state," *Journal of Clinical Medicine*, vol. 7, no. 8, p. 223, 2018.
- [8] K. Yu, J. G. Quirk, and P. M. Djurić, "Dynamic classification of fetal heart rates by hierarchical Dirichlet process mixture models," *PloS one*, vol. 12, no. 9, p. e0185417, 2017.
- [9] G. Feng, J. G. Quirk, and P. M. Djurić, "Supervised and unsupervised learning of fetal heart rate tracings with deep Gaussian processes," in *2018 14th Symposium on Neural Networks and Applications (NEUREL)*. IEEE, 2018, pp. 1–6.
- [10] A. Petrozziello, C. W. Redman, A. T. Papageorgiou, I. Jordanov, and A. Georgieva, "Multimodal convolutional neural networks to detect fetal compromise during labor and delivery," *IEEE Access*, vol. 7, pp. 112 026–112 036, 2019.
- [11] Z. Zhao, Y. Zhang, Z. Comert, and Y. Deng, "Computer-aided diagnosis system of fetal hypoxia incorporating recurrence plot with convolutional neural network," *Frontiers in Physiology*, vol. 10, p. 255, 2019.
- [12] G. Feng, J. G. Quirk, and P. M. Djurić, "Discovering causalities from cardiotocography signals using improved convergent cross mapping with Gaussian processes," in *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2020, pp. 1309–1313.
- [13] G. J. Warmerdam, R. Vullings, J. O. Van Laar, J. Bergmans, L. Schmitt, S. Oei *et al.*, "Using uterine activity to improve fetal heart rate variability analysis for detection of asphyxia during labor," *Physiological Measurement*, vol. 37, no. 3, p. 387, 2016.
- [14] G. Warmerdam, R. Vullings, J. Van Laar, J. Bergmans, L. Schmitt, S. Oei *et al.*, "Detection rate of fetal distress using contraction-dependent fetal heart rate variability analysis," *Physiological Measurement*, vol. 39, no. 2, p. 025008, 2018.
- [15] L. Yang, C. Heiselman, J. G. Quirk, and P. M. Djurić, "Identification of uterine contractions by an ensemble of Gaussian processes," in *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2021, pp. 1020–1024.
- [16] T. Przybyła, J. Jeżewski, and D. Roj, "Unsupervised clustering for fetal state assessment based on selected features of the cardiotocographic signals," *Journal of Medical Informatics & Technologies*, vol. 13, pp. 157–162, 2009.
- [17] L. Yang, M. Ajirak, C. Heiselman, J. G. Quirk, and P. M. Djurić, "Unsupervised detection of anomalies in fetal heart rate tracings using phase space reconstruction," in *2021 29th European Signal Processing Conference (EUSIPCO)*. IEEE, 2021.
- [18] N. Lawrence and A. Hyvärinen, "Probabilistic non-linear principal component analysis with Gaussian process latent variable models," *Journal of Machine Learning Research*, vol. 6, no. 11, 2005.
- [19] M. Titsias and N. D. Lawrence, "Bayesian Gaussian process latent variable model," in *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics. JMLR Workshop and Conference Proceedings*, 2010, pp. 844–851.
- [20] V. Chudáček, J. Spilka, M. Burša, P. Janků, L. Hruban, M. Hupčych, and L. Lhotská, "Open access intrapartum CTG database," *BMC Pregnancy and Childbirth*, vol. 14, no. 1, pp. 1–12, 2014.