# Fetal State Classification Using Cardiotocography Analysis

**Abstract**—Cardiotocography (CTG) is a foetal monitoring technique that is widely used to assess the fetus's well-being during pregnancy and childbirth.The method of surveillance is based on the examination of specific foetal heart rate patterns and uterine contractions. Computerized analysis can reduce examination time and the requirement for extra testing for foetal health by reducing intra-observer and inter-observer variability in visual CTG recording explanation.Several studies also shown that signal processing techniques could aid in the detection of foetal heart rate (FHR) trends, although they were all limited to physicians. In April 2008, we created a LabVIEW-based FHR and uterine contraction (UC) pattern analysis software based on the criteria and consensuses of the National Institute of Child Health and Human Development. This software has a lot of potential in the field of home healthcare. The median filter and the peak/valley detection method were used in this work for signal processing. Data from nineteen pregnant women was used to verify the analysis' performance. FHR baseline, baseline variability, early deceleration, UC frequency, and NST all have a 90 percent accuracy rate.

**Index Terms**—Cardiotocography, Fetal Heart Rate

## 1 Introduction

CARDIOTOCOGRAPHY (CTG) is a foetal monitoring technique that is widely used to assess the fetus's well-being during pregnancy and childbirth. The technique relies on the analysis of foetal heart rate (FHR) patterns and uterine contractions to keep track of the baby. The Non-stress test (NST) is a common and successful screening test used in CTG to determine whether foetuses are at risk of hypoxia. The foetal autonomic nervous system's sympathetic and parasympathetic tone interact to produce FHR reactivity, which is measured by NST [1].If two or more accelerations exceeding 15 beats per minute (bpm) amplitude and lasting 15 seconds or more in a 20-minute window after 32 gestational weeks, and two or more accelerations exceeding 10 beats per minute amplitude and lasting 10 seconds or more before 32 gestational weeks, the NST is considered reactive.

Early detection reduces complications and foetal death episodes while also reducing the need for intrusive procedures [2]. Monitoring the foetal heart rate is also crucial for reducing the risk of foetal morbidity and mortality and determining the best delivery time. Heart rate variability (HRV) analyses can be used to diagnose foetal pathological disorders like intrauterine growth restriction (IUGR), diabetes, infections, and placental abruption.IUGR is one of the most serious causes of neonatal morbidity and mortality: it is defined as a pathological suppression of foetal growth, with the foetus failing to reach its full growth potential as a result. It's all about prenatal hypoxia and asphyxia. About 5 percent of all pregnancies have IUGR [3].As a result, prenatal monitoring is critical for detecting IUGR or other potentially dangerous abnormalities in the foetus and assisting clinicians in their decision-making.

Visual interpretation of CTG recordings has a lot of intra-observer and inter-observer variability. Computerized analysis can eliminate these limitations, reducing examination time and reducing the need for extra foetal health testing. Several studies also revealed that signal processing techniques could aid in the identification of FHR patterns, although they were all limited to physicians [4].Several technical approaches can be used to assess FHR. Cardiotocography (CTG) is the most widely used technique, which combines FHR measurement with uterine contraction detection via a Doppler ultrasound probe. By measuring foetal cardiac movements using a Doppler ultrasonography, heartbeat events can be detected. During the screening procedure, the gadget provides a trace that reports heart rate fluctuations.

Although less accurate than internal FECG in detecting FHR, [5]the CTG method is widely used in antepartum surveillance since it is noninvasive and detects FHR better than abdominal ECG.

In the realm of prenatal well-being assessment, new ideas have developed in recent years with the goal of lowering invasiveness and enabling continuous and distant foetal monitoring. These efforts have primarily been made in the technological realm, with novel monitors and gadgets based on various[6] techniques (Echo4D, ST segment analysis (STAN) for intrapartum electronic foetal monitoring) being designed and developed.

The American College of Obstetricians and Gynecologists (ACOG) attempted to set the criteria for foetal monitoring terminology and interpretation in 2009 . The ACOG guidelines also consolidate all data on the effectiveness of foetal monitoring and highlight the major issues in this sector. In fact, most diagnostic findings drawn from CTG recordings are based on qualitative visual evaluation of CTG traces and, as a result, on the clinician's experience. In the interpretation of CTG traces, this results in widespread intraobserver and interobserver variability[7].

The study's ultimate goal is to include the new proposed index to a list of factors that can help with early prenatal diagnosis. These indicators are critical for reducing inter- and intra-observer variability while also identifying and diagnosing dangerous abnormalities in the foetus and assisting doctors in their decision-making.

## 2 LITERATURE REVIEW

The foetal heart rate (FHR) is one of the most crucial pieces of information about the foetus during the pregnancy cycle. Cardiotocography (CTG) is a technique used by obstetricians to obtain information about the fetus's heart rate (FHR) and uterine contractions (UC). Not only was the CTG used to obtain FHR, but it was also used to monitor the mother's contractions and other foetal monitoring procedures.CTG is a medical test that records UC and FHR and is used all throughout pregnancy. External or internal approaches can be used to carry out this examination. A catheter is inserted into the uterus when a specific amount of enlargement has occurred in an internal test. The mother's stomach is fitted with a pair of sensor nodes for external testing. In most cases, CTG data is split into two lines. The FHR in beats per minute is recorded on the top line.

The information obtained from a CTG is used to detect a pathological state early on, which can assist the obstetrician in anticipating future difficulties and preventing lasting foetal impairment before it occurs. Hypoxia can cause temporary disability or death during the delivery of a baby who is showing signs of it. More than half of these deaths occur as a result of incorrect FHR pattern recording diagnosis and unsuitable foetal therapy.

## 3 METHODOLOGY

### 3.1 Dataset

At a tertiary referral facility, we conducted this study to evaluate the visual interpretation of intrapartum foetal heart rate tracings with our novel analysis software. After admission to the labour and delivery facility, 21 tracing records were obtained from features. The following were the criteria for inclusion: (1) baseline value, (2) accelerations, and (3) percentage-of-time-with-abnormal-long-term-variability.

The pregnant woman's abdomen is strapped with a continuous-wave Doppler ultrasonography transducer. It's used to aim an ultrasonic beam at the foetal heart and detect Doppler-shifted echoes from moving cardiac components. A uterine fundus[8] is used to measure the relative pressure within the uterus. The FHR and UC signals were captured in the device and can be transferred to a computer.

| | baseline value | accelerations | fetal_movement |
|---|---|---|---|
| **count** | 2126.000000 | 2126.000000 | 2126.000000 |
| **mean** | 133.303857 | 0.003178 | 0.009481 |
| **std** | 9.840844 | 0.003866 | 0.046666 |
| **min** | 106.000000 | 0.000000 | 0.000000 |
| **25%** | 126.000000 | 0.000000 | 0.000000 |
| **50%** | 133.000000 | 0.002000 | 0.000000 |

Fig1. Dataset

### 3.2 Category indexes

In 2008, the National Institute of Child Health and Human Development (NICHD) updated the most recent indices and standards for classifying foetal heart rate and uterine contraction. FHR baseline, baseline variability, acceleration, deceleration, and NST are only a few of the category metrics available. The categories are divided into three levels. In Table I, we organised and listed each index's definitions as well as the category classifications.

The FHR baseline and variability are measures that show the fetus's main active condition. FHR baseline should be between 110 and 160 beats per minute. Tachycardia is defined as an FHR baseline greater than 160 bpm for more than 10 minutes, and bradycardia is defined as an FHR baseline less than 110 bpm for more than 10 minutes [9]. Fetal distress is indicated by tachycardia and bradycardia. The foetus' activity is represented by baseline variability. The variance of FHR in one minute is used to determine baseline variability. Minimum baseline variability is defined as undetectable to 5 bpm, moderate variability is defined as 6 bpm to 25 bpm, and noticeable variability is defined as greater than 26 bpm. Minimal and severe symptoms indicate that the foetus is in a low activity or distress state, which is harmful. Acceleration in FHR is described as a hilly increasing region. Accelerations must meet the following criteria: the onset to peak variance must be greater than 15 bpm and persist longer than 15 seconds.The acceleration frequency has an impact on NST. If two accelerations occur in twenty minutes, the NST is regarded as reactive; otherwise, it is defined as non-reactive. The UC peak is frequently associated with a slowdown, which appears in FHR signals as a valley. Decelerations can be classified into three types: early, late, and varied. An early deceleration is defined as a period of more than 30 seconds from start to nadir, with a minimum of 18 seconds from nadir to UC peak. A late deceleration is defined in a similar way as a late acceleration[10].

Table I
The identification of category index and classification [15]

| Index | Tier | Definition | Code |
|---|---|---|---|
| Baseline rate | bradycardia | <110 bpm | 1 |
| | tachycardia | >160 bpm | 2 |
| | normal | 110-160 bpm | 3 |
| Baseline variability | Absent | >Undetectable & <5 bpm | 4 |
| | Minimal | <=5 bpm | 5 |
| | Moderate | 6-25 bpm | 6 |
| | Marked | >25 bpm | 7 |
| Acceleration | | 15bpm & >15sec & >2 times/20min | 8 |
| Deceleration | Late | Onset to nadir >=30sec & lag time>= 18sec | 9 |
| | Early | Onset to nadir >=30sec & no lag time | 10 |
| | Variable | Onset to nadir <30sec | 11 |
| NST | | Reactive | 12 |
| UC | | appearance | 13 |
| Category I | Including all of 3, 6, 12; absent of 10, 11 | | |
| Category II | Including any of 1-(1∩4), 2, 4, 5, 7, 11∩(5or6),9∩6 | | |
| Category III | 4∩(9or11or1) | | |

### 3.3 Bagging Ensemble Classifier

Bootstrap aggregating (Bagging) is a simple ensemble learner that combines the base models for building and aggregation in which the base models are constructed using bootstrap samples from the training set and prediction is done by voting or averaging. There may be no improvement or even a little loss in accuracy for even stable algorithms, the basis models not necessarily diversified.If this makes it more unbalanced, it might be simplified and used for single model formation. Overfitting safeguards, which are utilised in some algorithms, may have to be abandoned. Models that have been overfitted to their specific bootstrap instances have a higher chance of varying. Creating many models based on unique data samples without combining them into an ensemble and then picking the one that appears to be the most appealing is not a satisfactory method. Because model assessment should be the foundation for model selection. Reiterating training and assessment rounds numerous times is required to produce low variance performance estimations that can be used to pick a model[11]. For models that simply differ in their training samples, this is challenging.

Input: Data set D=(x1, y1),( x2, y2),. . . .( xm, ym);

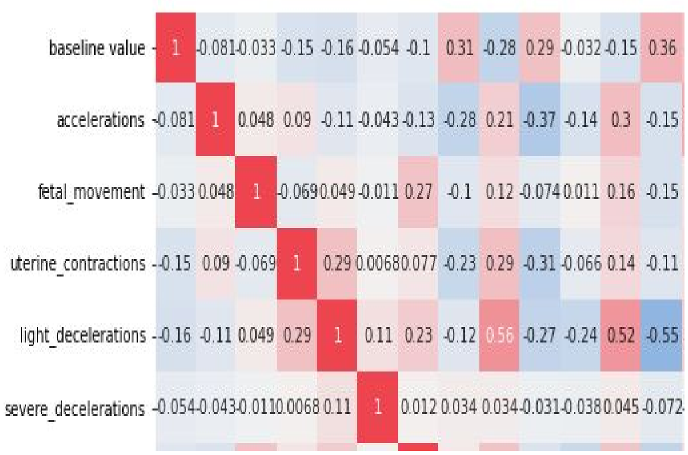


Fig4.

## 4 Results

K-fold cross validation is used to evaluate the performance of each learner. The number k is set to 10 in this research. As a result, the CTG dataset has been broken down into ten sections. The residual component is utilised for testing, while nine subsets are used for training. The method is performed ten times in order to ensure that each group of data is used as testing data. All of the classifiers in this study were trained on data without the use of features. During the training phase, the training parameters of the classifiers are selected using a k-fold cross-validation approach. The approach also calculates the area under the ROC curve and the F-measure for each learner.

Compares the testing performance of individual and ensemble learners.[11] As shown in table 2, the Random Forest outperformed all other single classifiers because it uses a greedy strategy to find the optimum answer when selecting classification structures that can be used for tests at each tree node.

Pruning the tree is one solution. All classifiers were improved by bagging. Bagging with Random Forest ensemble classifier has the highest classification accuracy (99.02 percent). F-measure, which is comparable to classification accuracy, is another performance measurement approach. The total accuracy and F-measure findings appear to be similar, as shown in Table 1. This holds true for all types of education. The ROC area can be used to assess the performance of classifiers. Bagging ensemble models outperform single classifiers in terms of ROC area[12].

Table 2. Performance of Different Machine Learning Techniques for CTG Classification

| | Accuracy | | F-Measure | | ROC Area | |
|---|---|---|---|---|---|---|
| | Single | Bagging | Single | Bagging | Single | Bagging |
| SVM | 98.63% | 98.69% | 0.986 | 0.987 | 0.957 | 0.973 |
| k-NN | 97.60% | 97.65% | 0.975 | 0.975 | 0.972 | 0.979 |
| ANN | 98.36% | 98.47% | 0.983 | 0.984 | 0.992 | 0.997 |
| Random Forest | 98.96% | 99.02% | 0.99 | 0.99 | 0.999 | 0.999 |
| CART | 98.80% | 98.85% | 0.988 | 0.988 | 0.944 | 0.976 |
| C4.5 | 98.58% | 98.91% | 0.986 | 0.989 | 0.933 | 0.985 |
| REP Tree | 98.20% | 98.09% | 0.982 | 0.981 | 0.941 | 0.993 |
| Random Tree | 98.36% | 98.96% | 0.984 | 0.99 | 0.955 | 0.987 |

F-measure, which is comparable to classification accuracy, is another performance measurement approach. The total accuracy and F-measure findings appear to be similar, as shown in Table 1. This holds true for all types of education. The ROC area can be used to assess the performance of classifiers. Bagging ensemble models outperform single classifiers in terms of ROC area.

```
Accuracy  : 0.8849765258215962
Recall    : 0.8849765258215962
Precision : 0.8809569858814997
Confusion Matrix:
[[317  13   3]
 [ 21  37   6]
 [  1   5  23]]
```

Fig5.

## 5 Discussion

The CTG data has been categorised in several research. Huang used three machine learning techniques to create classification models that predicted foetal distress from CTG data. Bagging ensemble learners are proposed in this research to improve the performance of classifiers for CTG data classification.[13] In addition, we used a publically available dataset to develop an ensemble classifier model for diagnosing foetal anomalies in babies in an intelligent and automated manner. The study's key contribution is the use of a Bagging ensemble learner to improve CTG data classification performance.

In terms of classification accuracy, F-measure, and ROC area, the classifiers' performance is increased by using the Bagging ensemble learner. The experimental results show that utilising a Bagging ensemble model, the presented model can classify CTG data with a 99.02 percent accuracy. With a suitable combination of Bagging ensemble classifiers, our model can detect foetal anomalies in a baby. The CTG data recognition performance is improved by using our proposed Bagging with Random Forest model[14].

## 6 Conclusions

Obstetricians can use CTG data to detect foetal anomalies and decide whether or not to intervene medically before the baby suffers irreversible damage. However, the obstetrician's visual analysis of the CTG data was not objective or accurate. In the medical field, using decision support systems to identify or forecast aberrant circumstances is becoming more common. The purpose of this research was to use CTG data to diagnose prenatal hazards. Using the CTG dataset as a decision support system, machine learning algorithms may be used to identify prenatal anomalies.A 10-fold cross validation procedure is used to assess the trainees' classification performance. With CTG data, the Bagging with Random Forest method had a 99.02 percent accuracy. The findings of this study's experiments show that Bagging ensemble with Random Forest can be used to classify normal and pathological CTG cases.

## References


[1] H. Sahin and A. Subasi, "Classification of the cardiotocogram data for anticipation of fetal risks using machine learning techniques," Appl. Soft Comput., vol. 33, pp. 231–238, 2015.

[2] M. Cesarelli, M. Romano, and P. Bifulco, "Comparison of short term variability indexes in cardiotocographic foetal monitoring," Comput. Biol. Med., vol. 39, no. 2, pp. 106–118, 2009.

[3] R. M. Grivell, Z. Alfirevic, G. M. Gyte, and D. Devane, "Antenatal cardiotocography for fetal assessment," Cochrane Database Syst. Rev., no. 9, 2015.

[4] F. H. Boehm, and M. F. Haire, "Xerox telecopier transmission of fetal monitor tracings: a 4-year experience," Obstet. Gynecol.vol 53, pp. 520-523, Apr. 1979.

[5] R. W. Naef, J. C. Morrison, J. F. Washburne, B. N. McLaughlin, K. G. Jr Perry, and W. E. Roberts, "Assessment of fetal well-being using the nonstress test in the home setting," Obstet. Gynecol., vol. 84, pp. 424-426, Dec. 1994.