Journal of Population Therapeutics & Clinical Pharmacology

E-NN: FETAL HEART RATE ANALYSIS DURING PREGNANCY USING AN ENSEMBLE DENSENET-BC AND CONVOLUTIONAL NEURAL NETWORK-BASED FRAMEWORK

Yousuf Iqbal[1], Shabir Hussain[2*], Muhammad Ayoub[3], Muhammad Faial Idrees[4], Sana Munir[5], Umair Rafiq[6], Shumaila Majeed[7]

[1,5,6,7]Department of Computer Science, National College of Business Administration & Economics, Lahore, Punjab, Pakistan
[2*]Institute of Biopharmaceutical and Health Engineering, Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055, China
[3]Department of Computer Science and Engineering, Shanghai Jiao Tong University, China
[4]Department of Computer Science, Air University Multan Campus, Punjab, Pakistan

*Corresponding Author: Shabir Hussain
*Email: shabir.hussain@sz.tsinghua.edu.cn

ABSTRACT

Purpose: This study aims to develop and validate a deep learning framework based on a weighted voting mechanism to automatically analyze the fetal heart rate (FHR) in electronic fetal monitoring (EFM) during pregnancy. The aim is to accurately divide the FHR into normal or pathological states, thus reducing the dependence on doctors' experience and potentially preventing unnecessary interventions such as cesarean section.

Method: The study implemented CNN and DenseNet-BC models based on the weighted voting mechanism to analyze the FHR as normal or pathological. The multi-model training method based on a down-sampling algorithm deals with imbalanced data. The effectiveness of the proposed CNN combined with the multi-model training method is evaluated using an open database named CTU-UHB.

Results: The experiment results show that the proposed method performs well and is stable on the CTU-UHB dataset. The CNN model based on a weighted voting mechanism and multi-model training method achieved high accuracy of 97.67% in automatically analyzing the FHR as normal or pathological. This demonstrates the potential for using deep learning techniques to improve the accuracy of FHR monitoring and reduce the dependence on doctors' experience.

Conclusion: The implemented deep learning framework combined with the multi-model training method shows promising results in automatically analyzing the FHR as normal or pathological, indicating the potential for using deep learning techniques to improve the accuracy and objectivity of FHR monitoring during pregnancy. This study has implications for reducing unnecessary interventions and improving maternal and fetal outcomes.

Keywords: fetal heart rate monitoring, automatic analysis, convolutional neural network, DenseNet-BC, ensemble learning.

Introduction

Intrauterine hypoxia and acidosis can cause fetal discomfort before or during birth and affect the newborn's neurological system. Real-time monitoring of fetal development should be implemented to address these problems. When there are warning signals, doctors might immediately perform a cesarean section or other emergency procedures. Cardiotocography (CTG), a technological method of fetal development monitoring used in clinical settings, is carried out using an electronic fetal monitor (EFM) apparatus. Two components make up a typical CTG output: uterine contractions and fetal heart rate (FHR) (UC). These days, CTG monitoring is frequently used to determine the fetus's health [1]. To safeguard the fetus's health, clinicians can quickly provide appropriate medications when fetal distress is identified through CTG monitoring. Doctors' expertise is crucial when assessing CTG recordings with just their eyes, and different observers have different opinions [2]. Different doctors may have different opinions on the same CTG recording. Furthermore, long- term clinical practice is required to train a doctor with rich experience.
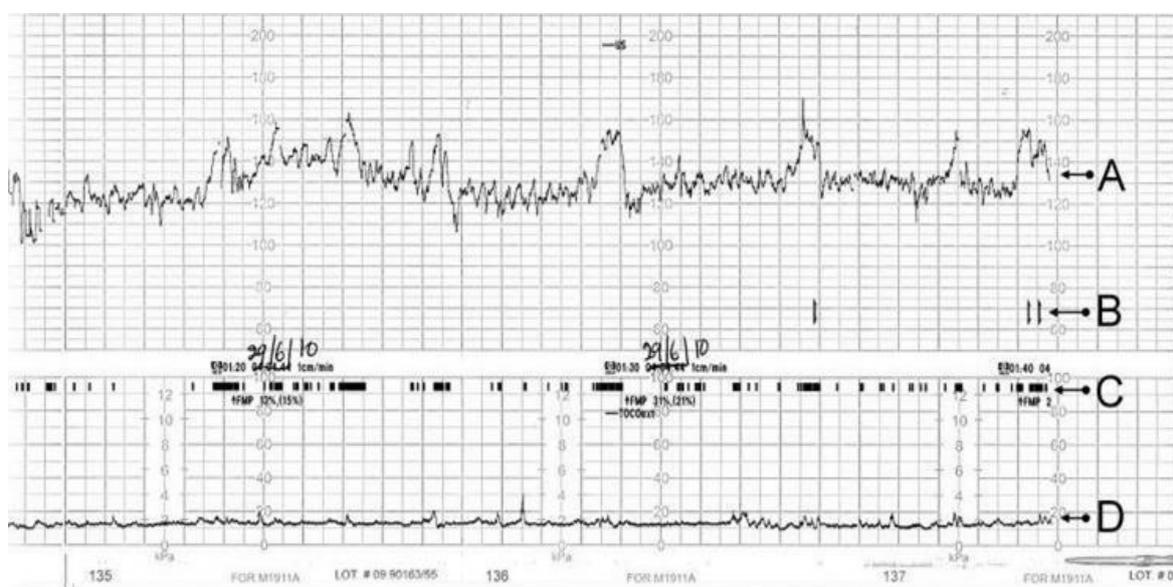
Figure 1: A typical CTG output. (A) Fetal heart rate. (B) Indicator showing movements felt by mother. (C) Fetal movement. (D) Uterine contractions.

Several medical organizations have worked for a long time to create standards for interpreting CTG that would aid physicians in reaching more accurate conclusions. The Intrapartum Fetal Monitoring Expert Consensus Panel of the International Federation of Gynecology and Obstetrics (FIGO) published the first draught of the recommendations in 1985. They unveiled their new CTG intrapartum categorization system in October 2015 [3]. This new method categorizes CTG into three groups: Normal CTG, Suspect CTG, and Pathological CTG. The final one indicates a substantial likelihood that the fetus may develop hypoxia or ketoacidosis and that reversible causes must be addressed immediately. In an emergency, delivery should take place right away. In this method, examining FHR is primarily used to collect CTG information. FHR baseline, variability, and deceleration make up the system. Every parameter is set to a value within the acceptable range. The CTG will be classified as suspicious or pathological when one or more signs go outside the normal range. These recommendations are merely meant to be helpful, and physicians will still make the ultimate decision.

With the advancement of computers, there is a pressing need for the automatic technical examination of CTG data. Visser et al. [4] utilized numerical approaches to investigate the connection between FHR and fetal condition. Many researchers created software to automatically assess CTG based on the FIGO guidelines after release, including Diogo et al. [5] and Bernardes et al. [6]. More and more pertinent characteristics are being retrieved from FHR and employed for

analysis because of an ongoing study on the technology. The extraction of features, including baseline, acceleration, and deceleration of FHR, was also done using artificial neural networks (ANN) [7, [8]). Gonçalves et al. [9] researched the association between linear and nonlinear properties and fetal state. The top 5 characteristics connected to the FHR and UA in categorization were discovered by Spilka et al. [10]. They also contrasted nonlinear features with conventional characteristics and found that adding them might increase classification accuracy [11]. According to the literature, there are the following research problems.

- FHR monitoring in clinical practice highly depends on doctors' experience, which can lead to incorrect judgments and unnecessary interventions such as cesarean section.
- The traditional approach to analyzing FHR is limited by subjectivity and requires significant expertise.
- There is a need for a computer-aided system to analyze fetal heart rate (FHR) and overcome the limitations of the traditional approach.

Regarding the selection of classifiers, several approaches were tested, including Support Vector Machine (SVM) [12]–[15] and ANN [16]. These techniques, however, included complicated feature extraction procedures, and the choice of features would impact the outcome. Convolutional Neural Network (CNN) technology has recently been effectively used in medical applications like Electrocardiograms (ECG) [17]. Petrozziello et al. [18] and Li J et al. [19] placed the FHR signals into a one-dimension CNN for categorization in the FHR classification zones. Comert [20], to transform the FHR data into a spectrogram, utilized the Short Time Fourier Transform (STFT). These spectrograms were fed into the pre-trained Alexnet [21] on ImageNet. Spectrograms were input into a 1-layer CNN model by Zhao et al. [22], who used the continuous wavelet transform (CWT) for spectrum transformation. Our contributions in this study are described below:

- The proposed convolutional neural network (CNN) model with a weighted voting mechanism can effectively classify FHR as a normal or pathological state.
- The multi-model training method based on a down-sampling algorithm helps to deal with imbalanced data.
- The proposed method was evaluated on an open database named CTU-UHB and performed well and be stable on this dataset.

We provide a CNN model for FHR analysis and classification in this study. A one-dimensional CNN receives the raw FHR signal and divides it into normal and abnormal states. Each FHR signal is split into a number of fragments, each of which has a varied weight according to the varying relevance of the information it contains at different points in time. Meanwhile, a multi-model training approach based on a down-sampling technique is also used to address the disparity between pathological and normal data.

Materials and Methods

The three components of the suggested methodology are signal preprocessing, classification with weights and a voting mechanism, and multi-modeling for skewed data. Figure 2 depicts the whole architecture, and the following sections will provide more in-depth explanations of each component.
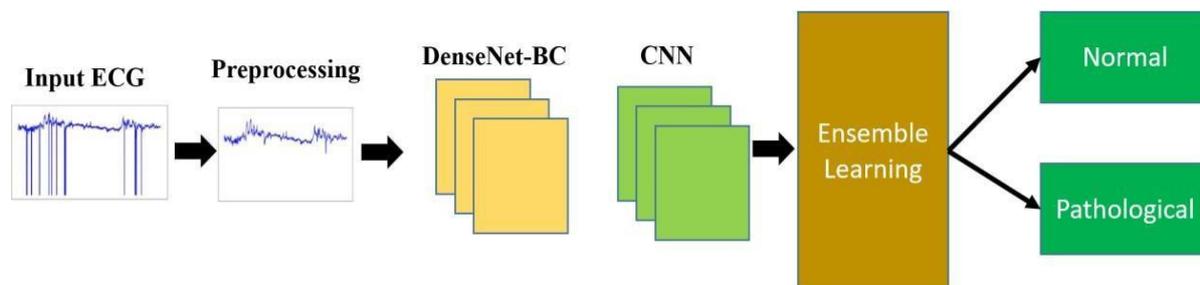
Figure 2: The proposed framework in this study is to classify FHR signals into two categories using ensemble learning.

Dataset Statistics

There are 552 unprocessed CTG recordings in the open database CTU-UHB, which was created by the Czech Technical University (CTU) and the University Hospital (UHB) [24]. Among 9164 intrapartum recordings, these recordings were carefully chosen. The following criteria, such as singleton pregnancy and gestational age more than 36 weeks, can only be met via recordings. Each CTG recording comprises FHR and UC time series and is sampled at a rate of 4 Hz. In this study, we concentrate on the final 30 minutes of the FHR time series prior to delivery.

The pH of the umbilical artery is used to categorize the 552 CTG recordings into two groups [25]. The pH threshold is set to 7.05. As a result, pathological samples are classified as samples with a pH value lower than 7.05, whereas the remaining samples are classed as normal samples. The selection indicates that 44 instances fall into the abnormal group while the remaining 508 cases fall into the normal category. The sample FHR data from the dataset is shown in Figure 3 (a).

Preprocessing

In this study, we process the raw FHR signals using a simple preprocessing strategy [26]. Initially, the extreme values (below 50 bpm and over 200 bpm) are identified and classified as missing data (the value is set to 0). Second, fragments with missing values that are spaced apart by more than 15 seconds are identified and removed from the FHR signal. Then, Hermite spline interpolation is used to fill in the missing values for the remaining fragments. Lastly, the FHR signal is smoothed using a conventional median filter.

After completing all these preprocessing, we take the final 30 minutes' worth of signals (7200 points) for additional study. The FHR signal will then be down-sampled to 1 Hz in the following step to make it shorter. In the end, the FHR signal is normalized by the baseline after the baseline of the token signal has already been determined. The sample FHR after preprocessing is shown in Figure 3(b).

(a)   **(b)**
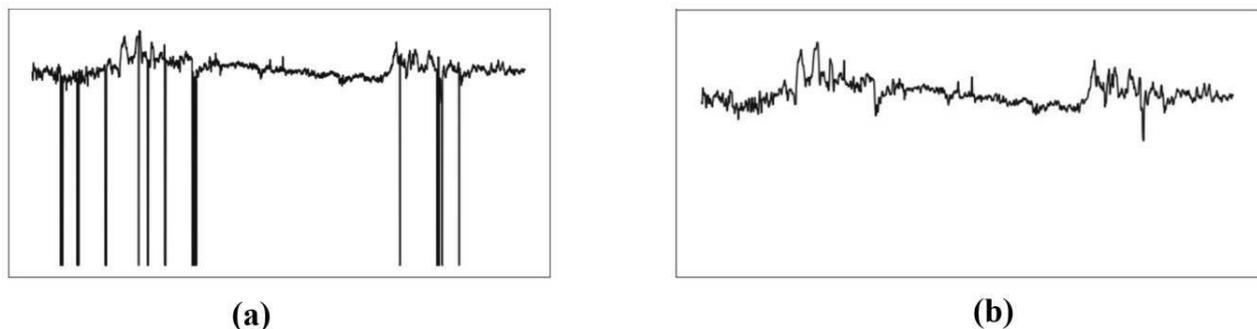
Figure 3: Sample FHR data (a) and the sample FHR after preprocessing (b).

Convolutional Neural Network

A convolutional neural network (CNN) is a type of deep learning neural network that is commonly used for image analysis tasks [33]. In the context of classifying fetal heart rate (FHR) into normal or pathological states, the CNN takes the FHR signal as input and performs a series of convolutions and pooling operations to extract important features from the data. These features are then passed through several fully connected layers, which finally classify the input into one of the two classes. The convolution operation is a mathematical operation that takes two functions, f and g, and produces a third function h, representing how much the shape of g "overlaps" with the shape of f at each point. Mathematically, the convolution of f and g is defined in Equation 1.

$$h(x) = (f * g)(x) = \int f(t)g(x - t)\, dt \quad (1)$$

In Equation 1, $*$ denotes the convolution operator, x is the input variable and t is a dummy variable of integration. In the context of CNNs, f is typically the input image or signal, and g is a small matrix

of weights called a kernel or filter. The pooling layers perform a downsampling operation by taking the maximum or average value over a local region of the input. The pooling layer operation by using max pooling is shown in Equation 2.

$$f_i(z) = \frac{e^{z_i}}{\sum_j z_j} \qquad (2)$$

The fully connected layers perform a matrix multiplication between the previous layer's output and a weight matrix and then apply a nonlinear activation function such as the rectified linear unit (ReLU) or sigmoid function to the result. Rectified Linear Unit (ReLU) activation function is shown in Equation 3.

$$f(x) = max(0, x) \qquad (3)$$

The complete structure of CNN used in this study is described in Figure 4.

Figure 4: The 2-layer Convolutional Neural Network.

DenseNet-BC

DenseNet-BC (DenseNet with bottleneck and compression) is an extension of the original DenseNet architecture that uses bottleneck layers and compression to reduce the number of parameters and improve computational efficiency [34]. The structure of the models is shown in Figure 5.
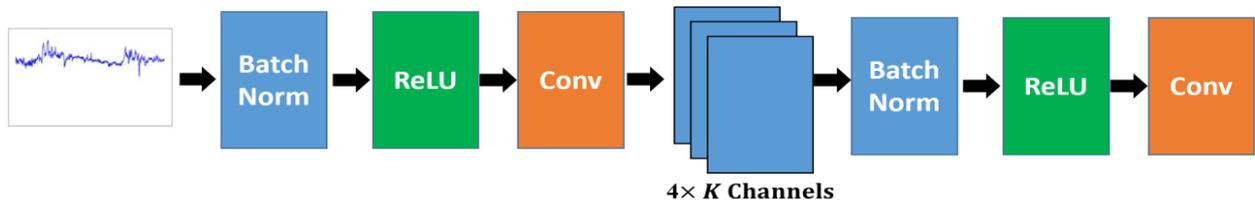
Figure 5: The structure of the model.

First, the output of each dense block is computed as follows:

$$\begin{aligned} x_0 &= X \\ x_k &= H_k([x_k - 1, f_k(x_k - 1)]), k = 1,2, \ldots, L \end{aligned} \qquad (4)$$

where X is the input image, $H_k(.)$ is the composite function of k convolutional layers with batch normalization and ReLU activation, and $f_k(.)$ is the bottleneck function that reduces the number of feature maps. Further, the transition layers are applied to the output of each dense block to reduce the spatial resolution and the number of feature maps as shown in Equation 5.

$$\begin{aligned} t_l &= BN(c_l(x_l)) \\ y_l &= Pool(t_l) \end{aligned} \qquad (5)$$

Where c_l is the composite function of a convolutional layer with batch normalization and ReLU activation, $BN(.)$ is the batch normalization function, $Pool(.)$ is the pooling function, and t_l and y_l are the output of the transition layer and the input to the next dense block, respectively. Finally, the output of the last dense block is fed into a global average pooling layer followed by a fully connected layer and a softmax activation function to obtain the class probabilities as described in Equations 6 and 7.

$$z = GAP(x_l) \qquad (6)$$

$$y = FC(z) \qquad (7)$$

$$\acute{y} = softmax(y) \qquad (8)$$

where GAP is the global average pooling function, FC is the fully connected layer, and $\acute{y}$ is the predicted probability distribution over the classes.

Ensemble Learning

Ensemble learning is a machine learning technique that combines multiple models to improve the system's overall performance. In ensemble learning, various models are trained on the same dataset, combining their predictions to generate a final prediction. Ensemble learning is based on the idea that a group of weak models can come together to form a strong model that is more accurate and robust than any of the individual models. Ensemble learning can be done in different ways, including bagging, boosting, etc. In this study, we used the weighted voting ensemble technique. Weighted voting is a technique used in ensemble learning where multiple models are combined to make a prediction. Each model is assigned a weight, and the weighted sum of their predictions is used as the final prediction. Mathematically, for instance, we have N models and their corresponding predictions are denoted as $y_1, y_2, ..., y_N$. Let $w_1, w_2, ..., w_N$ be the weights assigned to each model, such that $\sum w_i = 1$ and $w_i \geq 0$ for $i=1$ to N.

Then, the weighted voting formula can be represented as shown in Equation 9.

$$yfinal = w_1y_1 + w_2y_2 + \cdots + w_Ny_N \qquad (9)$$

In this Equation $yfinal$ is the final prediction made by the ensemble of models. The weights assigned to each model can be determined based on their performance on a validation dataset, such that better-performing models are assigned higher weights. The weights can also be tuned through grid or random search techniques. Weighted voting can improve the performance of a model by combining the strengths of multiple models and reducing the impact of individual model weaknesses. The complete voting mechanism using weighted voting is shown in Figure 6.
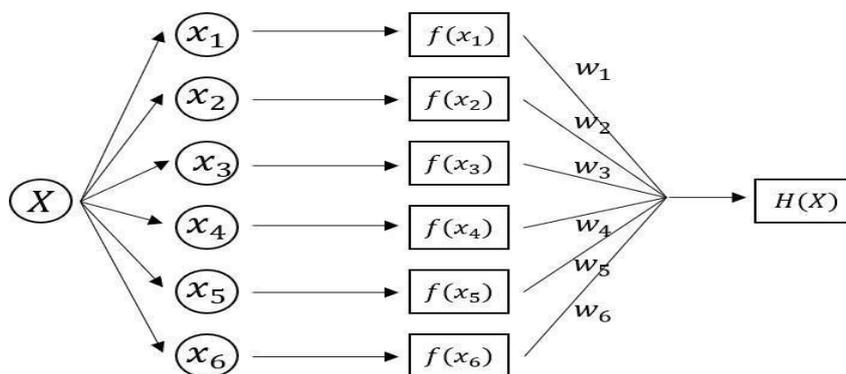
Figure 6: The voting mechanism

Each FHR signal was divided into six pieces since the original data was too lengthy, and each piece was input independently into the suggested CNN model. Meanwhile, from a medical perspective, a fragment more accurately represents the condition of the fetus the closer it is to delivery. As a result, we give each fragment a weight corresponding to its significance in the original data. Each fragment is fed into the proposed CNN model, and the output indicates the probability of the selected fragment for each category. The next step is to add all six output results together with

weights, and the final output result represents the probability of the original data for each category. Once the training set and testing set have been split, the negative samples are significantly less numerous than its positive ones. To create a new training set, we collect all negative samples and 1.5 times the positive samples. The standard under-sampling approach has the disadvantage that we are discarding potentially useful information. Under-fitting and poor generalization of the testing set may result from this. So, we create a new training set by using the same negative samples and the remaining positive samples. We end up with seven balanced training sets. The suggested CNN model receives data from each newly created training set, and as a result, we can finally have 7 trained models.

Results and Discussion

The 7 trained models are assessed against the testing set. We obtain seven projected values for each testing sample, after which we compute the percentage in the positive category. Lastly, using a threshold, we identify the category to which the sample belongs. Meanwhile 7-fold cross-validation is utilized to verify the stability of our proposed method. Meanwhile the accuracy and receiver operating characteristic (ROC) are also taken into account as shown in Figure 7.
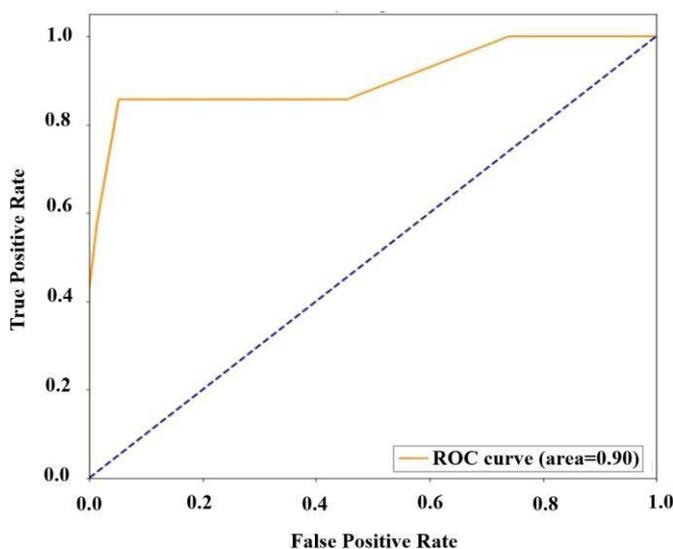
Figure 7: ROC curve for the proposed framework to classify

The most popular method for visually representing experimentally acquired statistical data is confusion matrices, which are also used to solve classification problems in deep learning and machine learning methods. Figure 8 (a) shows how well the CNN model performed, while Figure 8 (b) shows how well the DenseNet-BC performed, and similarly, Figure 8 (c) represents the performance of implemented ensemble approach using the 2×2 confusion matrix.
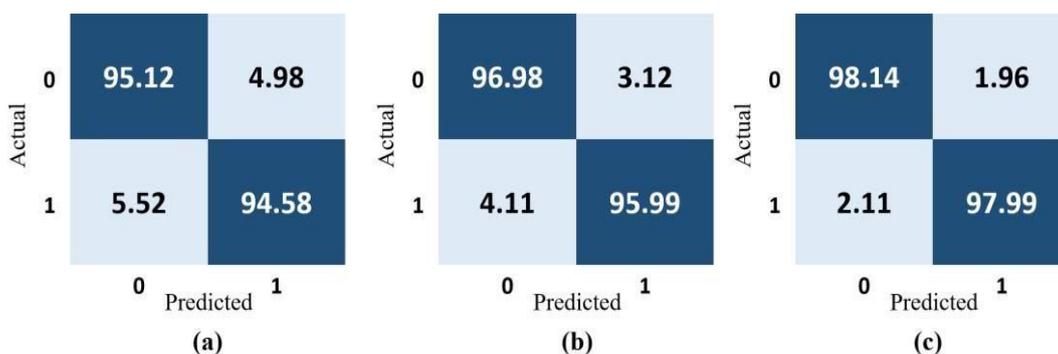
Figure 8: Confusion matrix representation of the performance of the (a) CNN model, (b) DenseNet-BC model, and (c) ensemble approach.

FSDop and FSMHR results are illustrated in Figure 9(a) during the beginning of the second delivery stage, even though there are no MHR FSs. We can observe that the MHR contains accelerations synced with contractions, even though it was taken discontinuously. The FHR, on the other hand, slows down and is challenging to study since it crosses the MHR curve. The genuine FHR can be seen in the scalp ECG signal at the beginning of the recording; the model missed this signal [23]. The model's predicted likelihood of an FS is displayed on a color scale. Because it perfectly fits the MHR, it can be shown that the algorithm correctly detected the period of FSs at minute 250. The model successfully predicted that the signals were TSs when the Doppler signal was superimposed on the scalp ECG signal. At minute 256, it seems as though another brief FS period was accurately recognized. If one looks very closely, one can see that the model failed to identify the FSs at minutes 246 and 249; this is only a tiny inaccuracy. Without an MHR signal, it would be quite challenging to comprehend this scenario.
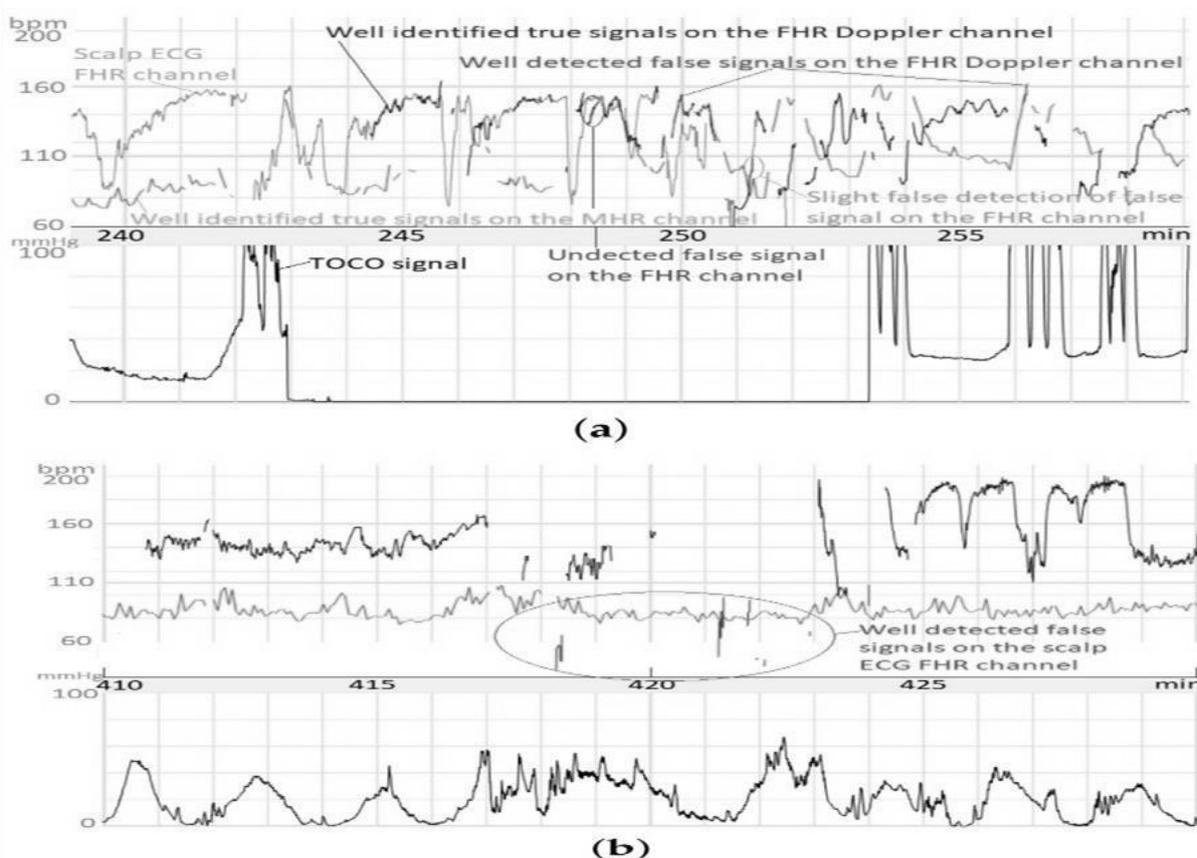
Figure 9: Results of the three models' outcomes (FSMHR, FSDop, and FSScalp). Results for FSScalp with a first-stage delivery recording (a), while (b) Results for FSScalp with a first-stage delivery recording.

Figure 9(b) shows an illustrative result for FSScalp over the first delivery stage. The few probable FSs appear to have been detected correctly. Even though there are no MHR FSs, Figure 10(c) depicts another exemplary result for FSDop (and FSMHR) throughout a period corresponding to the first delivery stage. The scalp electrode supported the MHR's accelerations and the FHR's decelerations. Sometimes, the MHR and FHR readings were the same. Even the MHR channel had extended periods of MS, FSDop could record. It's conceivable that the model did not identify a potential short FS at minute 27. Figure 10 (d) displays a second example result for FSDop and FSMHR during an epidural administration. The MS time on the Doppler channel is rather long. The MHR channel experienced FSs, most of which were picked up. These MHR FSs do not interfere with accurately detecting FSs on the Doppler FHR channel.
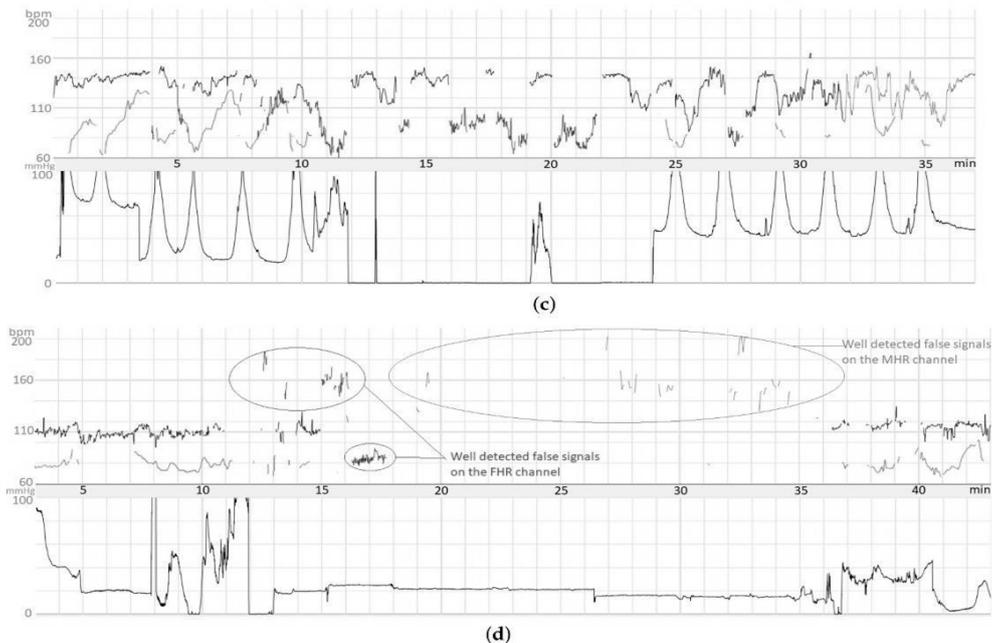
Figure 10: Results for the three models (FSMHR, FSDop, and FSScalp). (c) Results for FSDop and FSMHR with a recording from the first delivery stage. (d) Results for FSDop and FSMHR with a recording from the first delivery stage, during the administration of an epidural. The likelihood of an FS estimated by each model is represented as a color gradient.

The final interface within the FHRMA toolbox interface is shown in general in Figure 11. Additionally, it displays the outcomes of the weighted median filter baseline (WMFB) method-based morphological analysis (baseline, accelerations, decelerations, and UCs) [14]. The second stage of delivery, or the second half of this tape, is nearly entirely made up of FSs. Therefore, by identifying FSs, the approach prevented accelerations or decelerations during these times from being detected. The method also revealed the 110 Hz FHR signal, representing a protracted deceleration or bradycardia on edge.
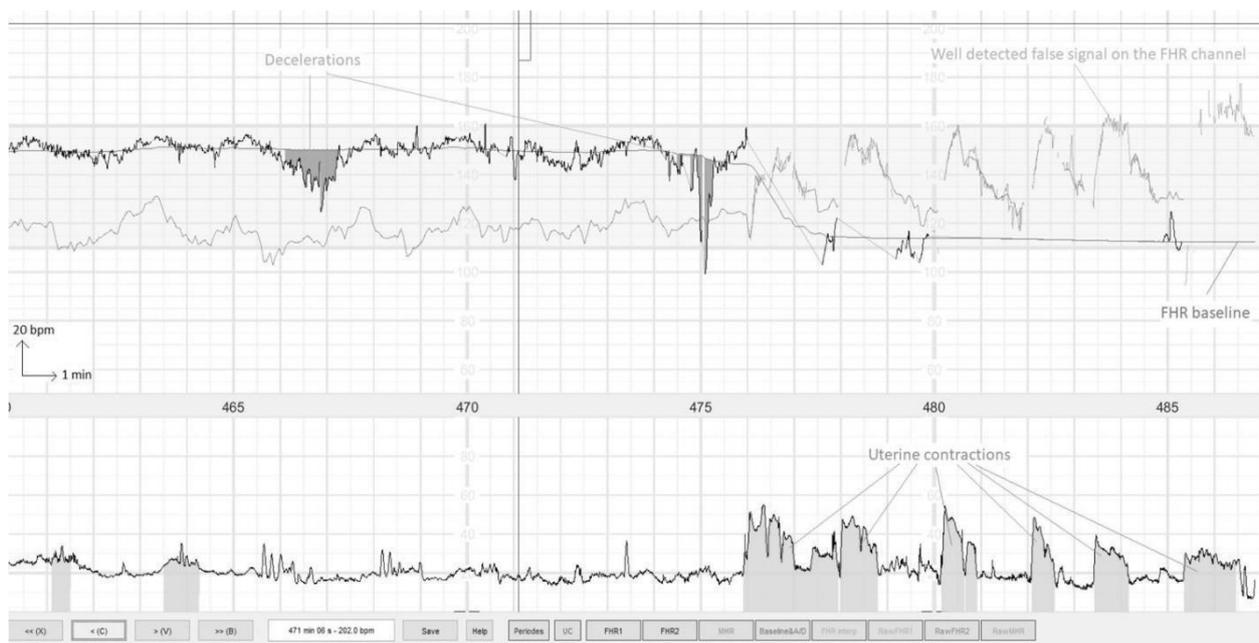
Figure 11: The FHRMA toolbox interface automatically displays both FSs and the results of morphological analysis (baseline, accelerations, decelerations, and UCs). The FHR signal comes from the Doppler sensor. The second stage of delivery starts at 475 min.

Table 1 provides a summary of all the findings for the training and test datasets. The datasets and models are represented in the table's left and right halves. The measure that makes the most sense is accuracy, often higher than 99%. Most signal samples are TSs; therefore, a simple categorization of all samples as TSs would also result in high accuracy. The accuracy of this simple model would be equal to the "Percentage of "true" among annotated" column. A model could still be helpful because FHR patterns are frequently redundant; we might also think deleting FS is more significant than leaving TS in place.

Table 1: Statistical results for the study's models and datasets.

Datasets	Training			Validation			
Model	FSDop	FSMHR	FSSCalp	FSDop	FSDop WO MHR	FSMHR	FSSCalp
Number of Recording	82	82	22	82	82	22	25
No. Analyzed Window	129	129	30	129	129	30	39
Avg Window Length	41	41	37	41	41	37	31
Missing Signal	16.3%	0.16%	0.39%	16.3%	0.16%	0.39%	0.19%
Annotated	68.8%	0.69%	0.39%	68.8%	0.69%	0.39%	0.98%
Accuracy	99.30%	0.9775	0.98%	99.30%	0.98%	0.99%	0.91%
Specificity	99.6%	0.991	0.99%	99.6%	0.99%	1.00%	1.00%
Positive Prediction	96.3%	0.92	0.96%	96.3%	0.92%	0.97%	0.99%
Negative Prediction	99.7%	0.984	0.98%	99.7%	0.98%	0.99%	0.99%
AUC	0.9992	0.9928	0.995	0.9992	0.9928	0.009971	0.01%
Cross Entropy	0.0213	0.0643	0.0395	0.0213	0.0643	0.000412	0.00%

This requirement should be satisfied for a usable model since the contingency table comprises the specificity (Sp), positive predictive value (PPV), and negative predictive value > 1 is similar to Acc > Percentage of TS. Some models, however, did not satisfy this requirement because of the highly skewed data, which made it clear that this is not a simple issue [24]. Regardless of the threshold for the output probability, the classifier's performance is shown by the AUC. The model outperforms the chance if the AUC exceeds 0.5 [25]. Performance on the validation dataset was, as to be expected, a little bit poorer than performance on the training data. Because the models' accuracy was substantially more significant than the proportion of TSs (89.5%, 89.3%, and 99.4%, respectively) and highly excellent (99.30% for FSDop, 98.68% for FSMHR, and 99.90% for FSScalp), the number of errors was multiplied by 15.0 for FSDop, 8.1 for FSMHR, and 6.0 for FSScalp. FSDop is probably used more frequently in practice because the FS rate was substantially lower in the FSDop dataset (10.5%) than in the FSScalp dataset (0.6%). The high (10.7%) FS rate for the MHR was mainly caused by selection bias during annotating. Finally, the comparison results are shown in Table 2.

Table 2: Comparative Performance of the proposed model with previous studies.

Study	Method	Accuracy (%)
det et al. [30]	Recurrent Unit	98.98
Fergus et al. [31]	Deep Neural Network	95.78
ld et al. [32]	Machin Learning Algorithms	92.23
Our Proposed Study	Ensemble Learning	99.90

Conclusion

In this study, we implemented CNN and DenseNet-BC models for automatically analyzing and categorizing FHR recordings based on ensemble learning. We use the under-sampling approach to handle skewed input in the ensemble learning technique based on weighted voting. The model is tested and trained using the public CTU-UHB database. Results of the experiments indicate that our approach is more reliable and stable than other ones currently in use. There are some drawbacks as well, such the scant amount of data. The processing techniques for imbalanced data sets should also be considered because most of these data sets are unbalanced in real life. Our technique must be tested and validated on additional and bigger data sets to be used with this technology in practice.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Conflict of Interests

The authors declare that the research was conducted without any commercial or financial relationships construed as a potential conflict of interest.

Data Availability

Data will be available on request from the authors due to privacy/ethical restrictions.

Ethics statement

The institutional review committee of our hospital approved ethics.

References
1. R. M. Grivell, Z. Alfirevic, G. M. Gyte, and D. Devane, "Antenatal car- diotocography for fetal assessment," Cochrane Database of Systematic Reviews, no. 9, 2015.
2. L. Hruban, J. Spilka, V. Chuda'cˇek, P. Janku˚, M. Huptych, M. Bursˇa, A. Hudec, M. Kacerovsky', M. Koucky', M. Procha'zka et al., "Agreement on intrapartum cardiotocogram recordings between expert obstetricians," Journal of evaluation in clinical practice, vol. 21, no. 4, pp. 694–702, 2015.
3. D. Ayres-de Campos, C. Spong, and E. Chandraharan, "for the figo intrapartum fetal monitoring expert consensus panel (2015b) figo con- sensus guidelines on intrapartum fetal monitoring: cardiotocography," Int J Gynecol Obstet, vol. 131, pp. 13–24.
4. G. Visser, G. Dawes, and C. Redman, "Numerical analysis of the normal human antenatal fetal heart rate," BJOG: An International Journal of Obstetrics & Gynaecology, vol. 88, no. 8, pp. 792–802, 1981.
5. D. Ayres-de Campos, J. Bernardes, A. Garrido, J. Marques-de Sa, and L. Pereira-Leite, "Sisporto 2.0: a program for automated analysis of cardiotocograms," Journal of Maternal-Fetal Medicine, vol. 9, no. 5, pp. 311–318, 2000.

6. J. Bernardes, D. Ayres-de Campos, A. Costa-Pereira, L. Pereira-Leite, and A. Garrido, "Objective computerized fetal heart rate analysis," International Journal of Gynecology & Obstetrics, vol. 62, no. 2, pp. 141–147, 1998.

7. P. Warrick, E. Hamilton, and M. Macieszczak, "Neural network based detection of fetal heart rate patterns," in Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005., vol. 4. IEEE, 2005, pp. 2400–2405.

8. C. Ulbricht, G. Dorffner, and A. Lee, "Neural networks for recognizing patterns in cardiotocograms," Artificial intelligence in Medicine, vol. 12, no. 3, pp. 271–284, 1998.

9. H. Gonc¸alves, A. P. Rocha, D. Ayres-de Campos, and J. Bernardes, "Linear and nonlinear fetal heart rate analysis of normal and acidemic fetuses in the minutes preceding delivery," Medical and Biological Engineering and Computing, vol. 44, no. 10, pp. 847–855, 2006.

10. V. Chuda'cˇek, J. Spilka, P. Janku˚, M. Koucky', L. Lhotska', and M. Hup- tych, "Automatic evaluation of intrapartum fetal heart rate recordings: a comprehensive analysis of useful features," Physiological Measurement, vol. 32, no. 8, p. 1347, 2011.

11. J. Spilka, V. Chuda´cˇek, M. Koucky´, L. Lhotska´, M. Huptych, P. Janku˚,G. Georgoulas, and C. Stylios, "Using nonlinear features for fetal heart rate classification," Biomedical signal processing and control, vol. 7, no. 4, pp. 350–357, 2012.

12. Hussain, S., Ayoub, M., Jilani, G., Yu, Y., Khan, A., Wahid, J. A., ... & Weiyan, H. (2022). Aspect2Labels: A novelistic decision support system for higher educational institutions by using multi-layer topic modelling approach. *Expert Systems with Applications*, *209*, 118119..

13. N. Krupa, M. Ali, E. Zahedi, S. Ahmed, and F. M. Hassan, "Antepartum fetal heart rate feature extraction and classification using empirical mode decomposition and support vector machine," Biomedical engineering online, vol. 10, no. 1, pp. 1–15, 2011.

14. P. A. Warrick, E. F. Hamilton, D. Precup, and R. E. Kearney, "Clas- sification of normal and hypoxic fetuses from systems modeling of intrapartum cardiotocography," IEEE Transactions on Biomedical En- gineering, vol. 57, no. 4, pp. 771–779, 2010.

15. Hussain, S., Yu, Y., Ayoub, M., Khan, A., Rehman, R., Wahid, J. A., & Hou, W. (2021). IoT and deep learning based approach for rapid screening and face mask detection for infection spread control of COVID-19. Applied Sciences, 11(8), 3495.

16. Bukhari, N., Hussain, S., Ayoub, M., Yu, Y., & Khan, A. (2022). Deep learning based framework for emotion recognition using facial expression. *Pakistan Journal of Engineering and Technology*, *5*(3), 51-57.

17. P. Rajpurkar, A. Y. Hannun, M. Haghpanahi, C. Bourn, and A. Y. Ng, "Cardiologist-level arrhythmia detection with convolutional neural networks," arXiv preprint arXiv:1707.01836, 2017.

18. A. Petrozziello, I. Jordanov, T. A. Papageorghiou, W. C. Redman, and A. Georgieva, "Deep learning for continuous electronic fetal monitoring in labor," in 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 18,pp. 5866–5869.

19. J. Li, Z.-Z. Chen, L. Huang, M. Fang, B. Li, X. Fu, H. Wang, and Q. Zhao, "Automatic classification of fetal heart rate based on convolutional neural network," IEEE Internet of Things Journal, vol. 6, no. 2, pp. 1394–1401, 2018.

20. Z. Co¨mert and A. F. Kocamaz, "Fetal hypoxia detection based on deep convolutional neural network with transfer learning approach," in Computer Science On-line Conference. Springer, 2018, pp. 239–248.

21. Hussain, S., Wahid, J. A., Ayoub, M., Tong, H., & Rehman, R. (2023). Automated Segmentation of Coronary Arteries using Attention-Gated UNet for Precise Diagnosis. Pakistan Journal of Scientific Research, 3(1), 124-129.

22. Z. Zhao, Y. Deng, Y. Zhang, Y. Zhang, X. Zhang, and L. Shao, "Deepfhr: intelligent prediction of fetal acidemia using fetal heart rate signals based on convolutional neural network," BMC medical informatics and decision making, vol. 19, no. 1, pp. 1–15, 2019.

23. W. Gao and Y. Lu, "Fetal heart baseline extraction and classification based on deep learning," in 2019 International Conference on Informa- tion Technology and Computer Application (ITCA). IEEE, 2019, pp. 211–216.

24. V. Chuda'cˇek, J. Spilka, M. Bursˇa, P. Jankuˆ, L. Hruban, M. Huptych, and L. Lhotska', "Open access intrapartum ctg database," BMC pregnancy and childbirth, vol. 14, no. 1, pp. 1– 12, 2014.

25. R. L. Goldenberg, J. F. Huddleston, and K. G. Nelson, "Apgar scores and umbilical arterial ph in preterm newborn infants," American journal of obstetrics and gynecology, vol. 149, no. 6,

26. pp. 651–654, 1984.

27. J. Spilka, G. Georgoulas, P. Karvelis, V. P. Oikonomou, V. Chuda'cˇek, C. Stylios, L. Lhotska', and P. Jankuˆ, "Automatic evaluation of fhr recordings from ctu-uhb ctg database," in International Conference on Information Technology in Bio-and Medical Informatics. Springer, 2013, pp. 47–61.

28. Hussain, S., Ayoub, M., Yu, Y., Wahid, J. A., Khan, A., Moller, D. P., & Weiyan, H. (2023). Ensemble Deep Learning Framework for Situational Aspects-Based Annotation and Classification of International Student's Tweets during COVID-19. *Computers, Materials & Continua*, *75*(3)..

29. 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.

30. F. Shaffer and J. Ginsberg, "An overview of heart rate variability metrics and norms," Frontiers in public health, vol. 5, p. 258, 2017.

31. Boudet, S., Houzé de l'Aulnoit, A., Peyrodie, L., Demailly, R., & Houzé de l'Aulnoit, D. (2022). Use of Deep Learning to Detect the Maternal Heart Rate and False Signals on Fetal Heart Rate Recordings. Biosensors, 12(9), 691.

32. Fergus, P., Hussain, A., Al-Jumeily, D., Huang, D. S., & Bouguila, N. (2017). Classification of caesarean section and normal vaginal deliveries using foetal heart rate signals and advanced machine learning algorithms. Biomedical engineering online, 16(1), 1-26.

33. Gold, N., Herry, C. L., Wang, X., & Frasch, M. G. (2021). Fetal cardiovascular decompensation during labor predicted from the individual heart rate tracing: a machine learning approach in near-term fetal sheep model. Frontiers in Pediatrics, 9, 593889.

34. Yu, Y., Carl, O., Hussain, S., Hou, W., & Weis, T. (2022). A privacy-protecting step-level walking direction detection algorithm based on floor vibration. IEEE Sensors Journal, 23(10), 10814-10824.

35. Iqbal, Y., Khan, A., Hussain, S., & Rafiq, U. (2024). Analysis of Remote Learning Challenges During COVID-19 Pandemic on Pakistan's Education Sector. *Pakistan Journal of Engineering and Technology*, *7*(2), 59-65.