Deep Learning-Based Non-invasive Fetal Cardiac Arrhythmia Detection

Kamakshi Sharma and Sarfaraz Masood

Abstract Non-invasive fetal electrocardiography (NI-FECG) has the possibility to offer some added clinical information to assist in detecting fetal distress, and thus it offers novel diagnostic possibilities for prenatal treatment to arrhythmic fetus. The core aim of this work is to explore whether reliable classification of arrhythmic (ARR) fetus and normal rhythm (NR) fetus can be achieved from multi-channel NI-FECG signals without canceling maternal ECG (MECG) signals. A state-of-the-art deep learning method has been proposed for this task. The open-access NI-FECG dataset that has been taken from the PhysioNet.org for the present work. Each recording in the NI-FECG dataset used for the study has one maternal ECG signal and 4–5 abdominal channels. The raw NI-FECG signals are preprocessed to remove any disruptive noise from the NI-FECG recordings without considerably altering either the fetal or maternal ECG components. Secondly, in the proposed method, the time–frequency images, such as spectrogram, are computed to train the model instead of raw NI-FECG signals, which are standardized before they are fed to a CNN classifier to perform fetal arrhythmia classification. Various performance evaluation metrics including precision, recall, F-measure, accuracy, and ROC curve have been used to assess the model performance. The proposed CNN-based deep learning model achieves a high precision (96.17%), recall (96.21%), $F1$-score (96.18%), and accuracy (96.31%). In addition, the influence of varying batch size on model performance was also evaluated, whose results show that batch size of 32 outperforms the batch size of 64 and 128 on this particular task.

Keywords Non-invasive fetal ECG · Arrhythmia detection · Convolutional neural network · Deep learning

---

K. Sharma
Department of Applied Sciences and Humanities, Jamia Millia Islamia, New Delhi 110025, India

S. Masood (✉)
Department of Computer Engineering, Jamia Millia Islamia, New Delhi 110025, India
e-mail: smasood@jmi.ac.in

A. Choudhary et al. (eds.), *Applications of Artificial Intelligence and Machine Learning*, Lecture Notes in Electrical Engineering 778,
https://doi.org/10.1007/978-981-16-3067-5_38

1 Introduction

Fetal cardiac arrhythmia basically refers to the abnormality in fetal heart rate (FHR) and/or cardiac rhythm, i.e., the condition in which fetal heart rhythm is either too fast or too slow. A healthy fetus has a heartbeat of 120–160 beats/min, beating at a regular rhythm, which is significantly higher than the heart rate of an adult (50–70 beats/min). Fetal arrhythmias are diagnosed in 1–3% pregnancies [1], out of which about 10% are considered as probable sources of morbidity. Benign fetal arrhythmias, such as PACs with less than 11 bpm and sinus tachycardia, usually do not require a treatment before or after birth. The hemodynamic fluctuations-related postnatal fetal arrhythmias involve interventions, as in few cases these may lead to preterm deliveries [2]. Sustained fetal arrhythmias prompt to the possibility of hydrops fetalis (a serious condition in which fluid builds up in two or more areas of the baby's body, causing severe swelling), cardiac dysfunction, or even fetal demise [3]. Thus, if cardiac arrhythmia is not diagnosed or left untreated, it can pose a risk to mother as well as fetus, including congestive heart failure. Earlier mild cases of arrhythmias were thought as benign but now prenatal cardiologists assert that any kind of irregular heartbeat should be identified and monitored closely to prevent any fatal fetal distress.

NI-FECG is an encouraging non-invasive alternative to fetal diagnosis and monitoring, which is performed by placing surface electrodes on a pregnant woman's abdomen to obtain a FECG signal. NI-FECG promises to assist fetal arrhythmia diagnosis by the means of uninterrupted analysis of the fetal heart rate (FHR) for the beat-to-beat variations. It can also assist in morphological analysis of the PQRST complex [4, 5] and thus present a number of advantages over the existing invasive modalities: reduced cost, analysis at local level (pregnant women not going over long distances for analysis), motion estimation, information on ventricular and atrial activity, and opportunity of long-term continuous remote monitoring. The capacity of the NI-FECG to deliver a precise estimation of the fetal heart rate has been shown by several recent studies. However, till today, the clinical usability of NI-FECG has rarely been explored. The NI-FECG extraction is also a challenge due to the temporal and frequency overlap between the fetal and the maternal electrocardiograms as they need modern signal processing methods [6, 7].

Although there are various methods of automatic classification methods of adult ECGs that have been proposed, little or no work has been done to analyze the NI-FECG signals, as they usually exhibit as a combination of substantial noise, fetal activity and a greater amplitude of maternal activity. This makes the precise extraction and further analysis of the FECG waveform a perplexing task to perform.

Deep learning (DL) is the mainstream of machine learning, which provides a structure where tasks like extraction of features and classification are executed together. With the development of artificial intelligence (AI), deep learning methods, such as feedforward artificial neural network (ANNs) and the recurrent neural networks (RNNs), long short-term memory (LSTM) and gate recurrent unit (GRU) are widely applied to the medical data [8]. The current study aims to automate the classification process of NI-FECG signals into arrhythmic (ARR) and normal rhythm (NR) using the state-of-the-art deep learning-based convolutional neural network.

2 Literature Review

This section takes two different sets of literature into account: (i) the non-invasive FECG analysis and (ii) to understand how machine learning and deep learning have been employed in detecting cardiac arrhythmia.

2.1 Literature Overview of the Non-invasive FECG Analysis

Clifford et al. [5] and Behar et al. [4] analyzed the clinical attributes extraction from the NI-FECG morphology. The fetuses included in these studies did not consist of any reported cardiac condition. This limited the research conclusions of whether the estimation of these physiological attributes was precise enough to offer actionable medical information.

The study [9] demonstrated the viability of the non-invasive FECG as a supplementary technique to identify the fetal atrioventricular block and hence could support clinical decisions.

In the research work [10], a systematic review is carried out to highlight normal fetal CTIs using NI-FECG and all the outcomes including fetal CTIs (P wave duration, PR interval, QRS duration and QT interval) were assembled as early pre-term ($\leq$ 32 weeks), moderate to late preterm (32–37 weeks) and term (37–41 weeks), concluding that NI-FECG establishes efficacy to quantify CTIs in the fetus, mainly at advanced gestations.

The study has established that NI-FECG assists in the identification of fetal arrhythmias. The diagnosis based on the extracted NI-FECG recordings was compared with the reference fetal echocardiography diagnosis to establish that NI-FECG and fetal echocardiography established the existence of an arrhythmia or not. This research work shows, for the first time, that NI-FECG allows to recognize fetal arrhythmias and also offers added evidence on the rhythm disturbance than echocardiography.

2.2 Literature Overview Cardiac Arrhythmia Detection (Adult and Fetus)

Bengio Y. in his work [11] advocated the popularity of deep learning methods, stating that deep learning architectures learn features at multiple levels of abstraction (i.e., layers) which allows in mapping the input to the output without being provided with hand-engineered features.

Karpagachevi et al. [12] classified the ECG signals taken from the PhysioNet arrhythmia database into five types of abnormal waveforms and normal beats, using extreme learning machine (ELM) and support vector machine (SVM). Experiment results show that ELM offers enhanced accuracy for all the classes, thus strongly recommended the use of the ELM-based method for classifying ECG. Though, the study gives good results on the classification but is only applied to adult ECG.

In the study, [13], a deep learning-based neural network with six hidden layers was proposed to recognize premature ventricular contraction (PVC) beats from the ECG recordings. The network was trained with six features which were extracted from ECGs for the purpose of classification. Although the researchers used a deep learning technique, they still used the hand-engineered features from the ECG data.

In another study by Pourbabaee et al. [14], a deep CNN was trained to extract features from the raw signals and to classify the paroxysmal atrial fibrillation (PAF) and the normal beats.

Andreotti et al. [15] compared the state-of-the-art feature-based classifier with a deep learning-based CNN. The short segments of the ECG were classified into four classes (AF, normal, other rhythms, or noise) and thus establishing that deep learning algorithms are proficient of categorizing short ECG recordings. It is also established in the study that deep learning models are aided from the augmented dataset while feature-based classifiers did not benefit from dataset augmentation.

In the study [16], Alin Isin and Selin Ozdalili proposed a deep learning-based structure previously trained on a general dataset which was used to carry out automatic ECG arrhythmia diagnostics on the MIT-BIH arrhythmia database. In the experiment, AlexNet, a transferred deep CNN, was used as the feature extractor. The study concluded that ECG arrhythmia detection approaches based on non-deep learning methods were outperformed by the transferred deep learning feature extraction approach.

The research work of Gao et al. [17] proposed a long short-term memory (LSTM) recurrent neural network with focal loss (FL) for detecting arrhythmia on a heavily imbalanced dataset using the MIT-BIH arrhythmia dataset. The supremacy of using LSTM with FL was recognized by analyzing with the cross-entropy loss function-based LSTM.

All the abovementioned works demonstrated the application of many widely used machine learning and deep learning network structures in detecting arrhythmia from adult ECG, but deep learning techniques for NI-FECG signals are yet to be explored. The recent study conducted by Zhong et al. [18] proposed a deep learning approach to detect fetal QRS. This study is the first of its kind where a deep learning technique is applied on non-invasive FECG signals. In the proposed work, fetal QRS complex was identified from single-channel raw NI-FECG signals without canceling the maternal ECG signals using a deep learning-based CNN model. A precision of 75.33, recall of 80.54%, and F-1 score of 77.85% were achieved by the proposed CNN.

Another study conducted by Lee et al. [19] proposed a much deep CNN architecture to detect fetal QRS using the NI-FECG signals, without the channel selection presented the positive predictive value of 92.77% with a mean sensitivity of 89.06%.

3 Experimental Design

The CNN classifier developed in the work ran on the deep learning toolbox and signal processing toolbox of MATLAB2018. The computer system had a 64 bit Microsoft Windows 10 operating system, configured with an 8 GB RAM and Intel®Core ™ i5-9300H processor. The epochs were set to 30 with five iterations per epoch. Model approximately took 11.53 s per epoch for training, though the respective epoch setting is not guaranteed to be the best configuration for the CNN network.

3.1 Data Source and Description

The training data used in this study is non-invasive fetal ECG arrythmia database (NIFEA DB) (February 19, 2019) taken from physionet.org. Dataset has been provided with an open access to the users. The dataset contains 26 samples including 12 arrthymic and 14 normal rhythmic fetal samples, obtained from 24 pregnant women, out of which two had normal rhythmic twins. This dataset contains 500 NI-FECG recordings, which were recorded constantly for varying periods ranging from a minimum of 7 min to a maximum of 32 min. Each of these records has a sample frequency either of 500 or 1000 Hz and is indicated in the header of each file. NI-FECG records contain one chest lead and four to five abdominal leads (recorded by placing five–six abdominal electrodes on maternal abdomen and two chest electrodes). A sample of normal rhythm (NR) and arrhythmic (ARR) signal is shown (see Fig. 1).

3.2 Signal Preprocessing

The raw NI-FECG signals are preprocessed to remove any disruptive noise from the NI-FECG recordings without considerably distorting the fetal or the maternal ECG components. For this, the mean removal technique is applied in which the mean of the signal is subtracted from every sample point, resulting in the removal of the unwanted DC component (noise) in the NI-FECG signals. Then, a ten-point moving average filter is applied to remove high-frequency noise. The low-frequency noise components are removed with the help of a high-pass filter after the removal of high-frequency noise. Each of these steps is applied to all the records collected from the chest, and abdominal channels and filtered signals are acquired for the next step.

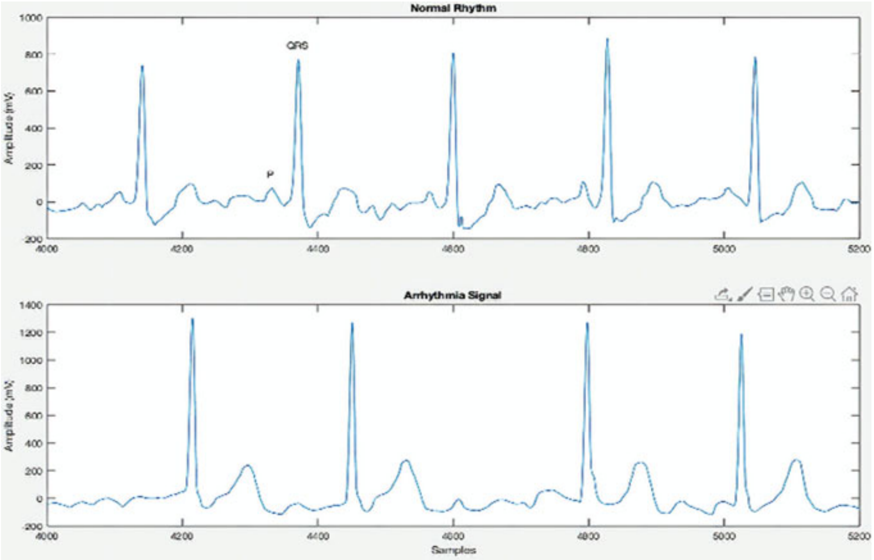

Fig. 1 Sample of normal rhythm (NR) signal and arrhythmia (ARR) signal

3.3 Feature Extraction

Extracting features from the data can assist in improving the train as well as test accuracies of the classifiers. Here, these time–frequency images, such as spectrogram, are computed, and then they are used to train the models. Time–frequency moments extract information from the spectrogram (see Fig. 2).

For the CNN network, each 1D feature is converted into a 2D feature to be as an input. Two time–frequency moments in time domains are spectral entropy (see Fig. 3) and instantaneous frequency (see Fig. 4).

Spectral entropy and instantaneous frequency vary by nearly one order or magnitude. Instantaneous frequency mean for the CNN can be a bit on the higher side for an effective learning. Since large inputs may slow down the network convergence rate, thus the mean and the standard deviation, the train set was used to standardize the train and test data. Standardization is a powerful method to improve the network performance during training. The 500 NI-FECG recordings were distributed into train and test sets with a ratio of 85:15. Since CNNs need a fixed window size, so the frame size is set to 100 ms.

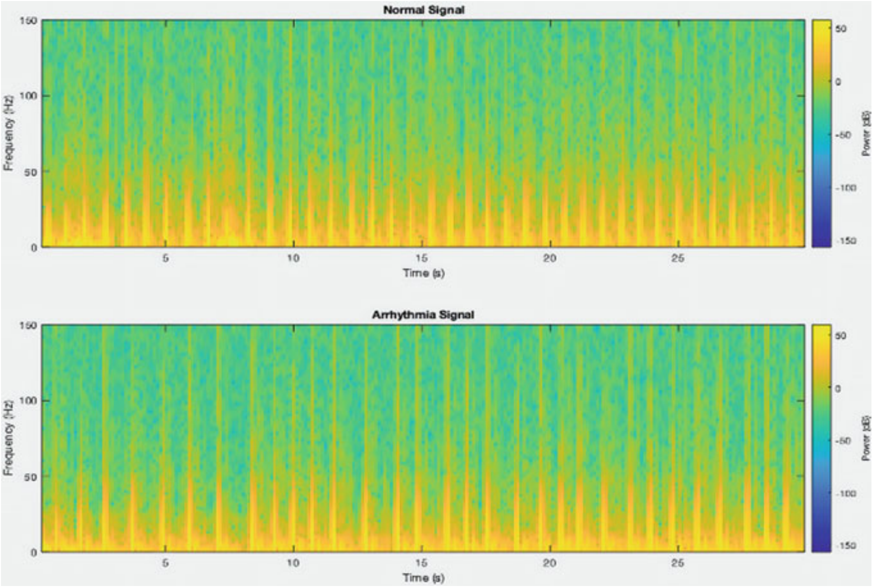

Fig. 2 Spectrogram of normal and arrhythmia signals

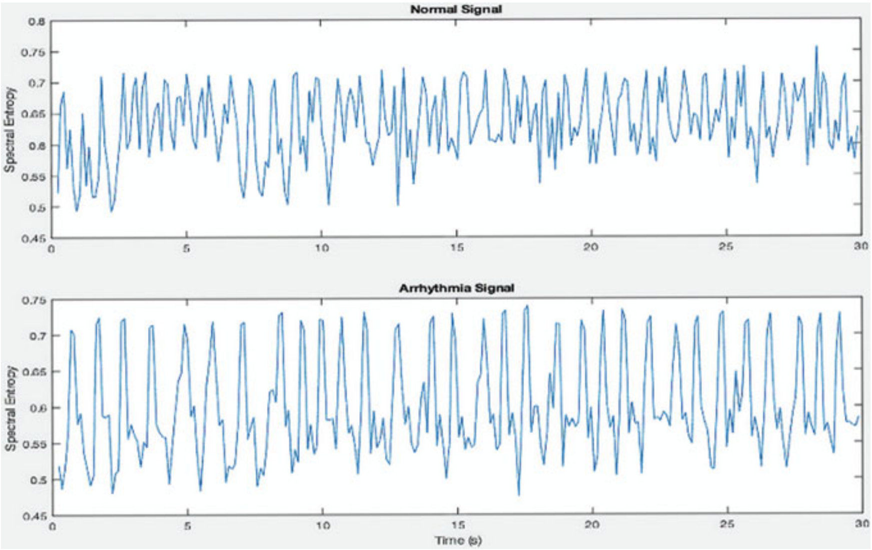

Fig. 3 Spectral entropy of normal signal and arrhythmia signal

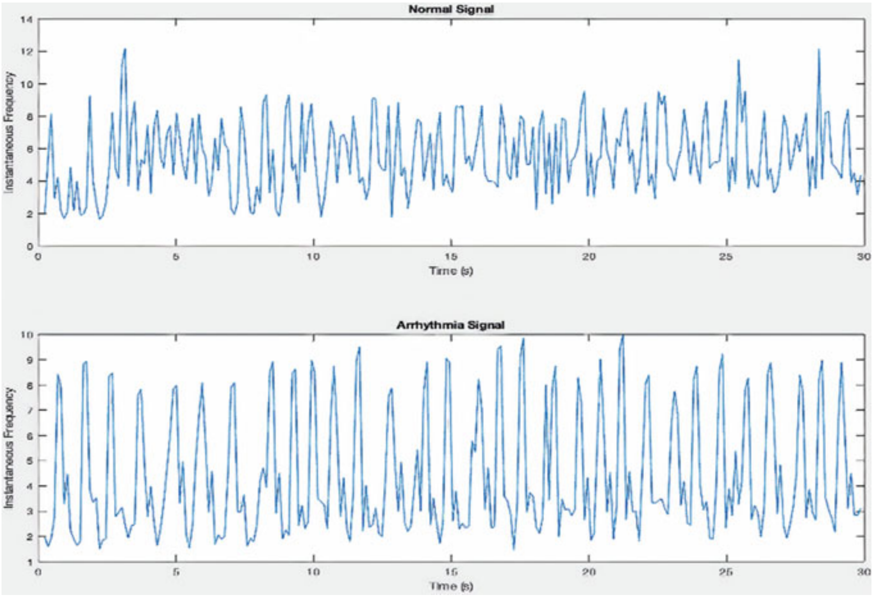

Fig. 4 Instantaneous frequency of normal signal and arrhythmia signal

3.4 CNN Architecture

An NI-FECG signal is fed as an input the CNN model, which is employed for the learning task. A sequence of labels (ARR and NR) are the outputs of the model. The schematic of the CNN architecture being used is shown (see Fig. 5). It contains seven convolutional layers, which are followed by two fully connected layers, a softmax layer, and a final classification layer.

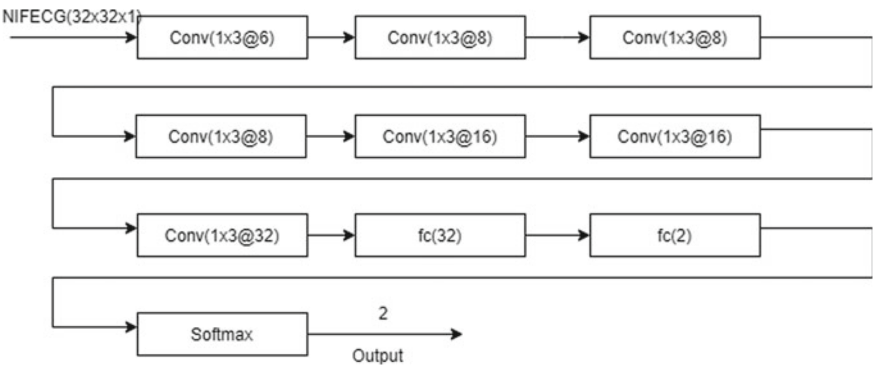

Fig. 5 Proposed CNN architecture

The first convolutional block contains a convolutional 2D layer (having six filters with a kernel size of 3), a batch normalization layer, and an activation function layer (ReLu) and a max pooling 2D layer. The second, third, and fourth convolutional blocks contain a convolutional 2D layer (having six filters with a kernel size of 3), a batch normalization layer and an activation function layer (ReLu), and a max pooling 2D layer. The fifth and sixth convolutional blocks contain a convolutional 2D layer (having 16 filters with a kernel size of 3), a batch normalization layer, and an activation function layer (ReLu). The seventh convolutional block contains a convolutional 2D layer (having 32 filters with a kernel size of 3), a batch normalization layer, and an activation function layer (ReLu). All the max pooling layers had a pool size of 2. The distribution over the two selected classes, i.e., ARR and NR, was produced by the concluding classification layer and softmax.

Batch normalization layer is used between the convolutional layers and the ReLu layers to speed up the training of convolutional neural networks. ReLu layer performs a threshold operation to each element of the input, but it does not change the size of its input. Max pooling layer divides the input into rectangular pooling regions and then computes the maximum of each region. Fully connected layer combines all the features learned by the previous layers and thus classifies the input. It acts independently on each time step in the case of sequence data, as is the case in this study. A softmax layer, following the fully connected layers, applies a softmax function to the input. Following the softmax layer is the classification layer, which computes the cross-entropy loss for the classification problems with mutually exclusive classes.

4 Result and Discussion

In this study, we tried to investigate whether reliable fetal arrhythmia classification could be attained using the deep learning approach without canceling the MECG signals from the NI-FECG signals. Deep supervised learning technique and convolutional neural networks (CNN) structure are used to achieve the goal of classification of fetal arrythmia, using the NI-FECG signals.

4.1 Performance Metrics

Performance measures, including accuracy, precision, recall, and $F1$-score, which were evaluated using a confusion matrix, were used for each class (NR and ARR). They were defined as follows:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

Accuracy gives the ratio of no. of accurate predictions to the total no. of input samples, which reflects the consistency among the real and the test results.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

Precision determines how precise or accurate the model is in predicting that from the total predicted positives, how many actually belonged to the positive class.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

Recall calculates how many of the actual positives our model has captured labeling it as true positive.

$$F1 = \frac{2 \times (\text{Precision} \times \text{Recall})}{\text{Precision} + \text{Recall}}$$

$F1$-score is the harmonic mean of recall and the precision and is helpful in seeking a balance between precision and recall.

4.2 Result

The performance indices of Zhong et al. [18] and Lee et al. [19] and the proposed model at 85:15 train–test ratios have been shown in Table 1.

The model performance has been evaluated considering three different batch sizes, as shown in Table 1. It shows the classification accuracy on the testing set, when the train–test ratio is set to 85:15 and learning rate 0.01 and Adam optimizer is used. The precision, recall, and $F1$-score are shown in Table 2.

Table 1 Performance metrics of all compared models

	Precision	Recall	$F1$-score	Accuracy
Zhong et al. [18]	89.03	91.57	90.28	91.33
Lee et al. [19]	92.89	90.27	91.56	93.27
Proposed model	*96.17*	*96.21*	*96.18*	*96.31*

Italics signifies that it belongs to the proposed model

Table 2 Classification accuracy of different batch sizes

Parameter	Value 1	Value 2	Value 3
Batch size	32	64	128
Accuracy on test set (%)	**96.31**	93.23	91.47

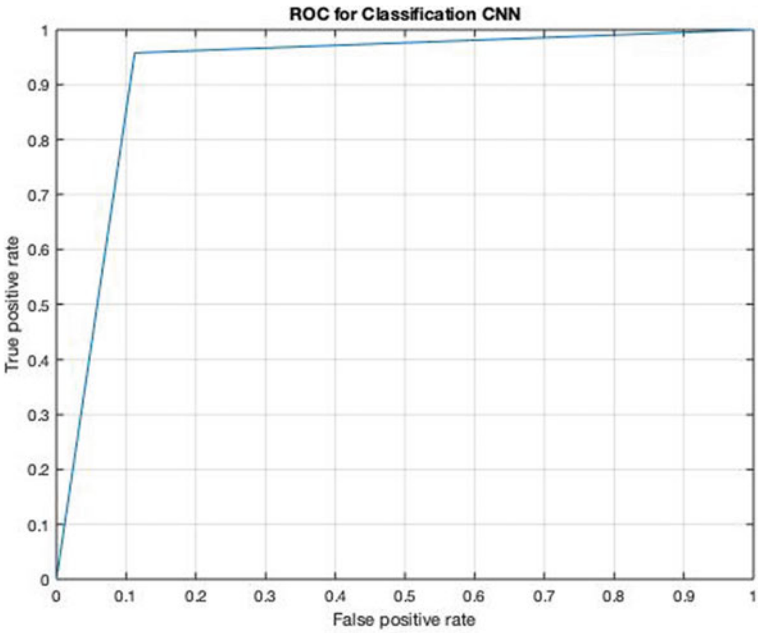

Fig. 6 ROC curve for the arrhythmic class

For the dataset and network structure used in the current research work, the optimal batch size is 32. To more intuitively compare the effectiveness of the proposed method, we analyzed the results using the ROC curve (see Fig. 6).

4.3 Discussion

The model was also analyzed by varying dropouts for the proposed network with an increasing dropout proportion and on comparing the results. It was found that the performance of the network did not improve. To confirm the effectiveness of the proposed model, the comparison is carried out with the algorithm proposed by Zhong et al. and that by Lee et al., respectively, and the results are graphically presented (see Fig. 7).

5 Conclusion

This work proposes a deep learning-based CNN model for classifying the fetal arrhythmia based only on NI-FECG signals without canceling the MECG signals.

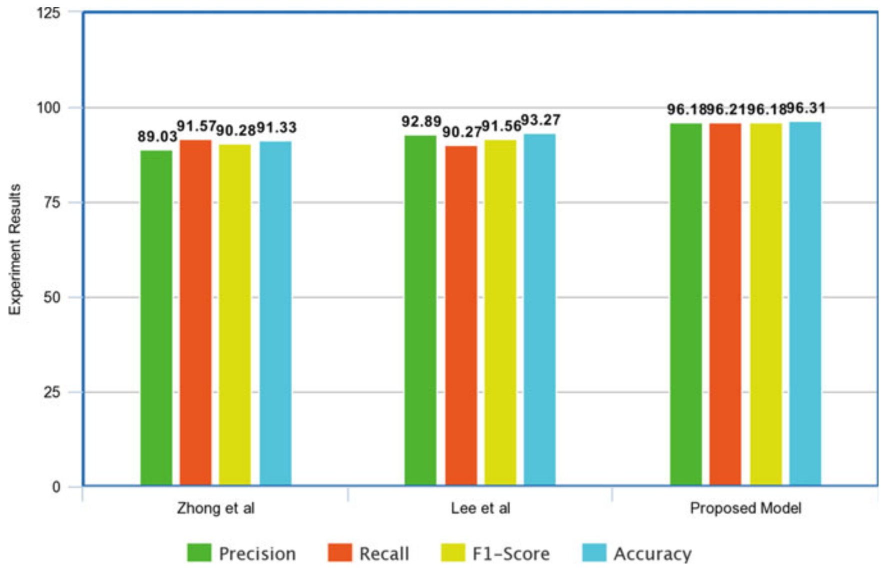

Fig. 7 Comparison of proposed model with Zhong et al. [18] and Lee et al. [19] models

The fetal arrhythmia can be classified with 96.31% accuracy, which is a good performance achieved by the proposed model. Our result shows that deep learning algorithms are capable of classifying the NI-FECG recordings. The DL method has an advantage that they do not require hand-engineered features, over other traditional methods. Moreover, if such pre-trained models are available, then it facilitates the imitation of those approaches, which can be fine-tuned to work for other databases or scenarios. The proposed method can efficiently assist the prenatal cardiologists to diagnose, analyze, and classify the NI-FECG signals in a more accurate way.

The present research work was conducted only on two NI-FECG signal types. So, in order to generalize the results, various NI-FECG beat types should be incorporated in the future works. Larger database with more patients and longer recordings is needed in the near future to provide the researchers with the opportunity to identify if their proposed algorithms are efficient in extracting features without any clinical distortion and classify the fetal cardiac arrhythmias.

References

1. Aggarwal S, Czaplicki S, Chintala K (2009) Hemodynamic effect of fetal supraventricular tachycardia on the unaffected twin. Prenat Diagn 29:292–293
2. Capuruço CA, Mota CC, Rezende GD, Santos R (2016) P06.03: Fetal tachyarrhythmia: diagnosis, treatment and outcome. Ultrasound Obstetrics Gynecol 48:182–182
3. Yaksh A, van der Does LJ, Lanters EA, de Groot NM (2016) Pharmacological therapy of tachyarrhythmias during pregnancy. Arrhythmia Electrophysiol Rev 5(1):41–44

4. Behar J, Zhu T, Oster J, Niksch A, Mah DY, Chun T, Greenberg J, Tanner C, Harrop J, Sameni R, Ward J (2016) Evaluation of the fetal QT interval using non-invasive fetal ECG technology. Physiol Meas 37(9):1392
5. Clifford G, Sameni R, Ward J, Robinson J, Wolfberg AJ (2011) Clinically accurate fetal ECG parameters acquired from maternal abdominal sensors. Am J Obstet Gynecol 205(1):47
6. Behar J, Andreotti F, Zaunseder S, Oster J, Clifford GD (2016) A practical guide to non-invasive foetal electrocardiogram extraction and analysis. Physiol Meas 37(5):R1
7. Clifford GD, Silva I, Behar J, Moody GB (2014) Non-invasive fetal ECG analysis. Physiol Measur 35(8):1521
8. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S (2017) Correction: corrigendum: dermatologist-level classification of skin cancer with deep neural networks. Nature 546(7660):686–686
9. Lakhno I, Behar JA, Oster J, Shulgin V, Ostras O, Andreotti F (2017) The use of non-invasive fetal electrocardiography in diagnosing second-degree fetal atrioventricular block. Maternal Health Neonatol Perinatol 3(1):14
10. Smith V, Arunthavanathan S, Nair A, Ansermet D, da Silva Costa F, Wallace EM (2018) A systematic review of cardiac time intervals utilising non-invasive fetal electrocardiogram in normal fetuses. BMC Pregnancy Childbirth 18(1):370
11. Bengio Y Learning deep architectures for AI. Now Publishers Inc. Foundations Trends Mach Learn 2(1):1–127
12. Karpagachelvi S, Arthanari M, Sivakumar M (2012) Classification of electrocardiogram signals with support vector machines and extreme learning machine. Neural Comput Appl 21:1331–1339
13. Jun TJ, Park HJ, Minh NH, Kim D, Kim YH (2016) Premature ventricular contraction beat detection with deep neural networks. In: 15th IEEE international conference on machine learning and applications (ICMLA). IEEE, pp 859–864
14. Pourbabaee B, Roshtkhari MJ, Khorasani K (2016) Feature learning with deep convolutional neural networks for screening patients with Paraoxysmal atrial fibrillation. In: 2016 IEEE international joint conference on neural networks. IEEE, pp5057–5064
15. Andreotti F, Carr O, Pimentel MA, Mahdi A, De Vos M (2017) Comparing feature-based classifiers and convolutional neural networks to detect arrhythmia from short segments of ECG. In: 2017 computing in cardiology (CinC). IEEE, pp 1–4
16. Isin A, Ozdalili S (2017) Cardiac arrhythmia detection using deep learning. Proc Comput Sci 120:268–275
17. Gao J, Zhang H, Lu P, Wang Z (2019) An effective LSTM recurrent network to detect arrhythmia on imbalanced ECG dataset. J Healthcare Eng 2019
18. Zhong W, Liao L, Guo X, Wang G (2018) A deep learning approach for fetal QRS complex detection. Physiol Measur 39(4):045004
19. Lee JS, Seo M, Kim SW, Choi M (2018) Fetal QRS detection based on convolutional neural networks in noninvasive fetal electrocardiogram. In: 4th international conference on frontiers of signal processing (ICFSP). IEEE, pp 75–78