IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 2 February 2024, accepted 28 February 2024, date of publication 4 March 2024, date of current version 5 April 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3373430

RESEARCH ARTICLE

Gender of Fetus Identification Using Modified Mel-Frequency Cepstral Coefficients Based on Fractional Discrete Cosine Transform

MOHAMED MOUSTAFA AZMY
Biomedical Engineering Department, Medical Research Institute, Alexandria University, Alexandria 21526, Egypt
e-mail: drmazmi@alexu.edu.eg

ABSTRACT Fetal heart sounds is measured to follow the developing status of fetus. The used database of fetal heart sounds is obtained from Physionet challenge. In this paper, novel models are created to extract features from fetal heart sounds; to identify the gender of fetus (male or female), using modified Mel-Frequency Cepstral Coefficients (MMFCC). The high pass filter which is used in the pre-emphasis step of MMFCC, used to remove noise from fetal heart sounds. These models are compared; to obtain high classification parameters (i.e. the accuracy rate and the area under curve). Classification method is based on deep learning, using recurrent neural networks (RNN) based on bidirectional long short-term memory (BiLSTM). The programming code of 13 models is created by the author, using Matlab software environment. Both accuracy rate and area under curve (AUC) are obtained using MMFCC based on fractional discrete cosine transform (FRDCT), overperform the same results of other models. So, this model is selected to extract features of fetal heart sounds. The contribution here is using novel model. FRDCT is new fractional transform, which is decomposed from fractional Fourier transform (FRFT) and discrete cosine transform (DCT). The secret behind selecting FRDCT based on MMFCC as a model; FRDCT decomposes signals perfectly in time-frequency domain. The best model can be programmed on a portable device; to know the gender of fetus by a caregiver in rural areas, without needing to use ultrasound.

INDEX TERMS DL, MFCC, MMFCC, DCT, FRFT, FDM, DCT.

I. INTRODUCTION

Fetal heart sounds used for many years to detect murmurs in fetal heart sounds [1]. Fetal heart sounds before birth used to detect any change about the status of the fetus. The abnormal fetal heart sounds may refer to lack of oxygen in fetus. Fetal heart sounds are important to monitor the developing of the fetus through the maternal gestation. Many researches are conducted about signal processing of fetal heart sounds and fetal electrocardiography (ECG). In 2010, Sameni and Clifford made review about signal processing of fetal ECG signal [2]. In 2015, Samieinasab and Sameni used empirical mode decomposition and non negative matrix factorization to extract features of fetal phonocardiogram (PCG) [3]. In 2017, Sutha and Jayanthi obtained feature extraction of fetal ECG (FECG) using two methods, least mean square and wavelet transform; to cancel noise for fetal QRs detection [4]. In 2021, Ponsiglione et al. made a review about techniques for analyzing signals of fetal heart rate [5]. In 2021 Khanmohammadi et al. obtained features of fetal PCG to get the gender of fetal using deep learning [6]. They obtained accuracy percent of 91%, which is less than the calculated accuracy rate in this paper. In 2021, Namazi and Krejcar made analysis of pregnancy development based on signals of fetal PCG [7]. In 2022, Farahi et al. used portable device and wavelet transform to made heart rate fetal analysis [8]. In 2022, Alkhodari et al. used deep learning to identify cardiac coupling between mother and fetus during gestation [9]. In 2023, Abburi and others used artificial intelligence algorithms for remote fetal heart rate monitoring and classification [10]. In 2023, Alkhodari and his colleagues used deep learning to extract features of fetal PCG [11].

The associate editor coordinating the review of this manuscript and approving it for publication was Sotirios Goudos.

II. MATERIAL AND METHODS

A. DATABASE

The database of fetal heart sounds is collected from Physionet challenge [12]. The gender of fetus in this database is male or female. Twin and non-identifying cases are excluded. The number of female and male cases is 47 and 53 cases respectively. The ratios 50%, 40%, 10% of the database are used for training percent, testing percent and validation percent respectively. All the database of training, validation and testing are separated from each other. There is no overlap between them. In this database, the mean of pregnancy weeks of mothers is 36.5 weeks. The maximum and minimum weeks of pregnancy are 41 and 26 weeks respectively.

B. MODELS OF FEATURE EXTRACTION

1) MEL FREQUENCY CEPSTRAL COEFFICIENTS (MFCC)

Mel Frequency Cepstral Coefficients (MFCC) is used widely in getting features of speech. It is used also to get the features of fetal heart sounds. It composed from eight steps. Pre-emphasis is applied in the first step using high pass filter (HPF). The coefficients a and b are used for HPF. a equals 1. The matrix of b is [1, -0.97]. Then framing is applied in the second step. Duration of frames is 0.025 seconds. The used sampling frequency is 16000 Hz. Number of samples per frame is 320. The time-step of each frame is 0.01 second. The hamming window is used during the third step. $X(k)$ is used fast Fourier transform (FFT), which is applied in the fourth step. Number of FFT points is 64. In the fifth step, Mel-scale is calculated from the frequency f in the linear scale, as shown in equation (1). 20 filter banks are created. In the sixth step, S(m) is the output log of multiplied power spectral energy and Mel filter banks, as shown in equation (2). Then discrete cosine transform (DCT) is obtained in the seventh step, as shown in equation (3). Finally, thirteen coefficients of MFCC are obtained including log frame energy as the first coefficient [13], [14].

$$Mel\left(f\right) = 2595 * log(1 + \frac{f}{700}) \tag{1}$$

$$S\left(m\right) = \log(\sum\nolimits_{k=1}^{N} \left|X\left(k\right)\right|^{2} H_{m}(k)) \tag{2}$$

$$C\left(n\right) = DCT\left(S\left(m\right)\right), n = 0, 1, \ldots M \tag{3}$$

DCT can be calculated, as shown in equations (19)-(20). Where $H_m(k)$ is the transfer function of filter-banks and m is the number of filter-banks. M is the total number of filter-banks, which equals 20.

2) METHOD OF STOCKWELL TRANSFORM

Stockwell Transform (ST) is mixed between wavelet transform and short-time Fourier transform (STFT). It gathers between the advantages of wavelet transform and STFT. Fixed window in STFT is replaced with Gaussian function in the frequency domain as shown in equation (4). Where τ is time index. ST is suitable for feature extraction of fetal heart sounds because it is non-stationary signals. ST describes signals in time-frequency domain [15]. The inverse ST is shown in equation (5).

$$S\left(\tau,f\right) = \int_{-\infty}^{\infty} x\left(t\right) \frac{|f|}{\sqrt{2\pi}} e^{\frac{-f^{2}(t-\tau)^{2}}{2}} e^{-2\pi ift} dt \tag{4}$$

$$x\left(t\right) = \sqrt{2\pi} \int_{-\infty}^{\infty} \frac{S\left(\tau,f\right)}{|f|} e^{+2\pi ift} df \tag{5}$$

3) METHOD OF MODIFIED STOCKWELL TRANSFORM

Modified Stockwell Transform (MST) is a novel method for feature extracting of fatal heart sounds. MST is the same as ST, but the Gaussian function is replaced with $e^{\frac{-f^{p}(t-\tau)^{p}}{2}}$ as shown in equation (6). Index p is changed, which values are 0.5, 1, 1.5, 2.5 and 3 [15].

$$S\left(\tau,f\right) = \int_{-\infty}^{\infty} x\left(t\right) \frac{|f|}{\sqrt{2\pi}} e^{\frac{-f^{p}(t-\tau)^{p}}{2}} e^{-2\pi ift} dt \tag{6}$$

4) METHOD OF FRACTIONAL FOURIER TRANSFORM

Fractional Fourier Transform (FRFT) represents signal in frequency domain and fractional domain [16], [17]. It is more suitable for non-stationary signals as fetal heart sounds than Fourier transform (FT). FRFT overcomes some disadvantages of FT. FT represents signals in the frequency domain only. So, FT is not suitable for non-stationary signals. FRFT of x(t) is $F^{\alpha}(u)$ as shown in equation (7). Discrete FRFT of x(n) is shown in equation (8). The fractional domain is u. Where $0 < |\alpha| < 2$. csc(Φ)=1/sin(Φ). cot(Φ)=1/tan(Φ). α is the order of FRFT. K_{α}(t,u) is the kernel function of FRFT as shown in equation (9).

$$F^{\alpha}\left(u_{\emptyset}\right) = \int_{-\infty}^{\infty} K_{\propto}\left(t, u_{\emptyset}\right) x\left(t\right) dt \tag{7}$$

$$F_{x}^{\propto}\left(k, m\right) = \sum\nolimits_{n=-N}^{N} x\left(n\right) K_{\propto}\left(n, m\right) \tag{8}$$

$$K_{\propto}\left(t, u_{\emptyset}\right) = \begin{cases} \sqrt{\dfrac{1 - i.cot(\emptyset)}{2\pi}} e^{i.\frac{(u_{\emptyset}{}^{2}+t^{2})\cot(\emptyset)}{2} - i.u_{\emptyset}.t.csc(\emptyset)} \\ \delta\left(t - u_{\emptyset}\right) if \emptyset = 2N\pi \\ \delta\left(t + u_{\emptyset}\right) if \emptyset = (2N+1)\pi \end{cases}$$

$$if \emptyset \neq N\pi \tag{9}$$

$\emptyset = \alpha\pi/2$. Where $\emptyset$ is rotational angle of FRFT. If it equals $\pi/2$, FRFT will convert to FT. In this paper, the fractional power is changed between 0.5 and 1.

5) METHOD OF FRACTIONAL STOCKWELL TRANSFORM

Fractional Stockwell Transform (FRST) is decomposed from FRFT and ST. It has the advantages of both transforms. It avoids the disadvantages of ST (fixed Gaussian function) and FRFT (power spectrum is not available in time-domain) [18]. The equation which describes FRST is presented in equation (10). Discrete FRST is shown in equation (11).

$$FRST_{x}^{\alpha}\left(\tau, u_{\emptyset}\right) = \int_{-\infty}^{\infty} x\left(t\right) g(\tau - t, u_{\emptyset}) K_{\propto}\left(t, u_{\emptyset}\right) dt \tag{10}$$

where K_α (t,$u_\emptyset$) is defined in equation (9). The fractional power of FRST is α.

$$FRST_x^{\propto}(k,m) = \sum_{n=-N}^{N} x(n)\, g(k-n,m)\, K_{\propto}(n,m) \quad (11)$$

$$g(t,u_\emptyset) = \frac{|u_\emptyset csc\emptyset|^p}{q\sqrt{2\pi}} e^{\frac{-t^2(u_\emptyset csc\emptyset)^{2p}}{2q^2}} \quad (12)$$

$\emptyset = \alpha\pi/2$. $0 < |\alpha| < 2$. If $\emptyset$ is $\pi/2$, FRST will be converted to ST. The parameters of p and q are used for controlling window shape of FRST. The value of p or q in this paper is 1. *g(t,u)* is Gaussian function, which depends on time t and fractional domain u, as shown in equation (12).

6) METHOD OF EMPIRICAL FOURIER DECOMPOSITION (EFD)

In Empirical Fourier Decomposition (EFD), the signal is decomposed in the Fourier domain using zero-phase filter banks. It is based on Fourier theory [19]. In the first step of EFD, discrete Fourier transform is calculated for signal *f(m),* as shown in equation (13). Then zero phase filter bank is calculated, as shown in equation (14). Where k is the boundary of improved segmentation technique, which is built on Fourier transform. k_{n-1} and k_n, are cutoff frequencies of bandpass filter for the zero-phase filter bank. The signal $\check{X}_n(k)$ is Fourier transform of $f(m)$ between γ_n and γ_{n+1}, as shown in equation (15). In zero-phase filter bank $[\gamma_1 \ldots \gamma_n]$ are the descending orders in the sorting process, for the series of the frequency spectrum magnitudes. $\gamma_0 = 0, \gamma_{n+1} = \pi$. Then filtered Fourier Transform is obtained, as shown in equation (16). Inverse Fourier Transform is obtained for filtered Fourier Transform; to obtain decomposed components $f_n(m)$, as shown in equation (17). Finally, the reconstructed signal is computed, as shown in equation (18).

$$F(k) = \sum_{k=1}^{N} f(m)\, e^{-jkm} \quad (13)$$

$$\mu_n(k) = \begin{Bmatrix} 1 \; k_{n-1} \le |k| \le k_n \\ 0 \;\; otherwise \end{Bmatrix} \quad (14)$$

$$k_n = \begin{cases} argmin\left(\check{X}_n(k)\right) & 0 \le n \le N \; and \; \gamma_n \ne \gamma_{n+1} \\ \gamma_n & 0 \le n \le N \; and \; \gamma_n = \gamma_{n+1} \end{cases} \quad (15)$$

$$\hat{F}_n(k) = F(k)\,\mu_n(k) = \begin{Bmatrix} F(k)\, k_{n-1} \le |k| \le k_n \\ 0 \;\; otherwise \end{Bmatrix} \quad (16)$$

$$f_n(m) = \sum_{k=1}^{N} \hat{F}_n(k)\, e^{jkm} \quad (17)$$

$$\tilde{f}(m) = \sum_{n=1}^{N} f_n(m) \quad (18)$$

7) METHOD OF EMPIRICAL DISCRETE COSINE DECOMPOSITION

In empirical discrete cosine decomposition (EDCD), the signal is decomposed in the discrete cosine domain using zero-phase filter banks. It is a new method of empirical decomposition methods. It is based on DCT theory [20], [21]. In the first step of EDCD, DCT is calculated for signal *f(m),* as shown in equation (19). Where $w(k)$ is calculated, as shown in equation (20). Then zero-phase filter bank is calculated as shown in equations (14). The signal $\check{X}_n(k)$ is presented in equation (15). It is DCT of $f(m)$ between γ_n and γ_{n+1}. Then filtered DCT is obtained, as shown in equation (21). Inverse DCT is obtained for filtered DCT; to calculate obtained decomposed components $f_n(m)$, as shown in equation (22). Finally, the reconstructed signal is computed, as shown in equation (23).

$$F(k) = w(k)\sum_{m=1}^{N} f(m)\cos\left(\frac{\pi(2m-1)(k-1)}{2N}\right) \quad k = 1,\ldots,N \quad (19)$$

$$w(k) = \begin{cases} \frac{1}{\sqrt{N}} & k = 1 \\ \sqrt{\frac{2}{N}} & 2 \le k \le N \end{cases} \quad (20)$$

$$\hat{F}_n(k) = F(k)\,\mu_n(k) = \begin{Bmatrix} F(k)\, k_{n-1} \le |k| \le k_n \\ 0 \; otherwise \end{Bmatrix} \quad (21)$$

$$f_n(m) = \sum_{k=1}^{N} w(k)\,\hat{F}_n(k)\cos\left(\frac{\pi(2m-1)(k-1)}{2N}\right) \quad m = 1,\ldots,N \quad (22)$$

$$\tilde{f}(m) = \sum_{n=1}^{N} f_n(m) \quad (23)$$

8) METHOD OF FACTIONAL DISCRETE COSINE TRANSFORM

Fractional discrete cosine transform (FRDCT) is a new fractional transform. It is decomposed from FRFT and DCT. It has the advantages of both transforms. It avoids the disadvantages of both DCT (it represents the high frequency components of signals poorly) and FRFT (the power spectrum of FRFT is not available in time-domain) [16]. FRDCT is presented in equation (24). Discrete FRDCT is shown in equation (25).

$$\mathrm{FDCT}_x^{\alpha}(u_\emptyset) = \int_{-\infty}^{\infty} x(t)\, L_{\propto}(t,u_\emptyset)\, dt \quad (24)$$

$$\mathrm{FDCT}_x^{\propto}(k,m) = \sum_{n=-N}^{N} x(n)\, L_{\propto}(n,m) \quad (25)$$

$$L_{\propto}(s,u_\emptyset) = \begin{cases} \sqrt{\frac{1-i.cot(\emptyset)}{2\pi}}\cos\left(\frac{(u_\emptyset{}^2+s^2)\cot(\emptyset)}{2} - u_\emptyset.s.csc(\emptyset)\right) \\ \delta(s-u_\emptyset) \;\; if\,\emptyset = 2N\pi \\ \delta(s+u_\emptyset) \;\; if\,\emptyset = (2N+1)\pi \end{cases}$$

$$if\,\emptyset \neq N\pi \tag{26}$$

where $L_\alpha(t,u_\emptyset)$ is defined in equation (26). The fractional power of FRDCT is α. Where s=ktπ/N. If $\emptyset$ is $\pi/2$, FRDCT is converted to DCT. The value of p or q in this paper is 1.

9) MODIFIED MEL FREQUENCY CEPSTRAL COEFFICIENTS (MMFCC)

Modified Mel Frequency Cepstral Coefficients (MMFCC) is used to get the features of fetal heart sounds using different methods. These several methods replace FFT in the fourth step, as described later [13], [14].

C. CLASSIFICATION METHODS

Deep learning is used as a classification method. It is a subclass of machine learning. It overcomes the disadvantages of conventional neural networks; which needs a lot of data for training. The large number of artificial neurons is applied; so the word 'deep' is used. The most famous used network of deep learning is recurrent neural network (RNN). It suitable for analysis sequence data as heart sounds. It contains a memory unit, which has three gates (i.e. input, output and forget gates). Here, the input gate receives the information of extracted features of fetal heart sounds. The number of input gates is 13. In this paper, the output layer describes the gender of fetus (i.e. male or female). So, the number of output gates is 2. Bi-directional long short-term memory (BiLSTM) is used. It is a sequence model, which takes two long short-term memories, one in the forward direction and the other in the backward direction. Number of nodes in each hidden layer is 100. Fully connected, softmax, dropout and classification layers are used. In fully connected layer, each neuron of the input is connected to each neuron in the output layer. In the softmax layer, the real numbers are converted into probability distribution; to accelerate the training process. Dropout layer is used after softmax layer; to avoid over-fitting. There are nine options, used through learning process. The first option is Adam (adaptive moment estimation) optimizer. It adjusts the first and the second moment of weights of RNN. The second option is applying maximum number of epochs to be 100; to minimize errors. In the third option, the minimum batch size is 150. Where the batch size is the number of samples propagated through the network. In the fourth option, the initial learning rate is 0.01; to control the updating errors of weights. In the fifth option, the gradient threshold is 1. The gradient is the changing in weights regarding to changing in errors. In the sixth option, the plotting of training process is applied. Validation process is used in the seventh option; to avoid over-fitting. The frequency of validation process is 10, which is chosen in the eighth option. Validation process isn't a part of testing process. Verbose is false in the ninth option; to disappear the training progress for each epoch [22].

III. RESULTS

Thirteen methods are compared to obtain the features of fetal heart sounds. The used first of these methods is MFCC; to calculate the features of fetal heart sounds. The applied second method is MMFCC based on ST instead of FFT. The used third method is MMFCC, using MST instead of FFT. The value of the factor p is 0.5, as shown in equation (6). The applied fourth method is MMFCC using MST instead of FFT. Where p equals 1 in equation (6). The used fifth method is MMFCC using MST instead of FFT. The value of p, which is used in equation (6), is 1.5. The applied sixth method is

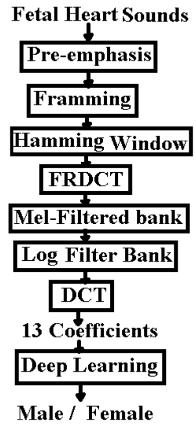

FIGURE 1. Steps of feature extraction using the model 13.

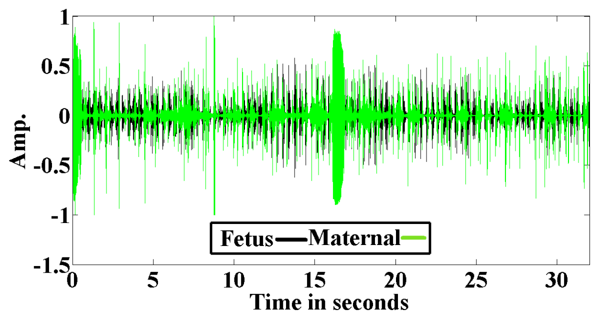

FIGURE 2. Heart sound signals of fetus and maternal heart sounds.

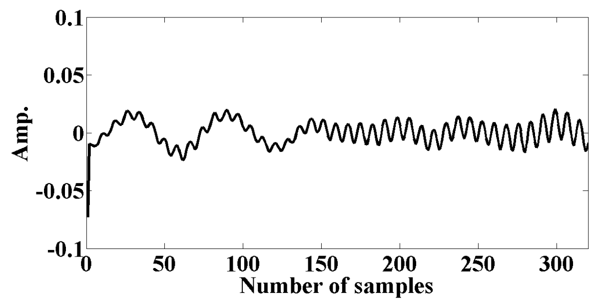

FIGURE 3. The output of high pass filter in Pre-emphasizing step.

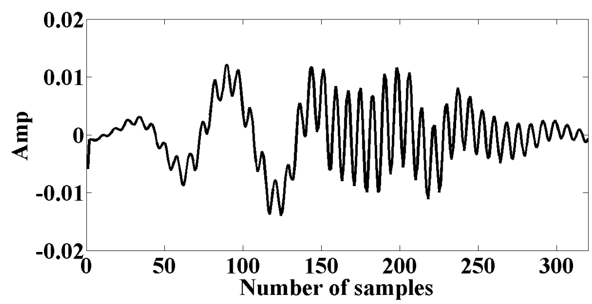

FIGURE 4. The output of hamming window step.

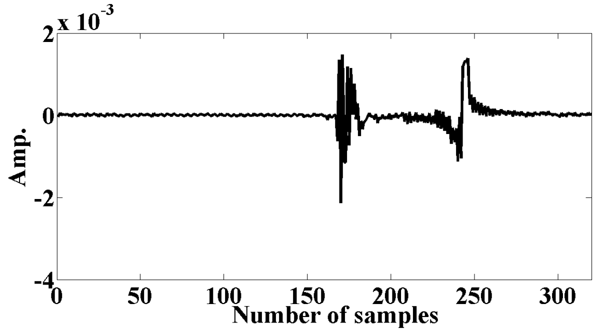

FIGURE 5. The output of FRDCT step.

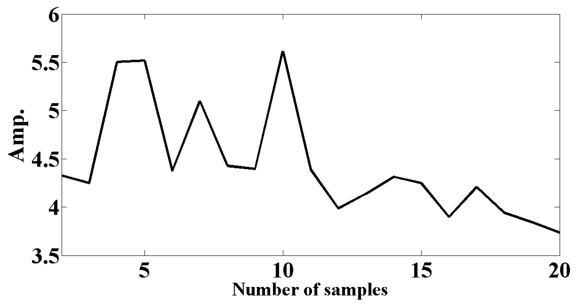

FIGURE 6. The output of Log Mel-scale filter step.

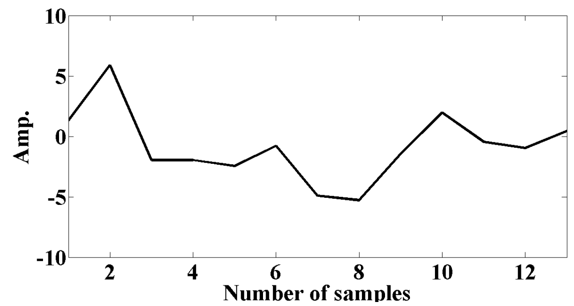

FIGURE 7. Coefficients of MMFCC based on FRDCT.

MMFCC using MST instead of FFT. Where p equals 2.5 in equation (6). The used seventh method is MMFCC using MST instead of FFT. Where p equals 3 in equation (6). The applied eighth method is MMFCC using EFD instead of FFT. The used ninth method is MMFCC using FRFT instead of FFT. The fractional power is 0.5. The applied tenth method is MMFCC using FRST instead of FFT. The fractional power of FRST is 1. The used eleventh method is MMFCC using FRFT instead of FFT. The fractional power of FRFT is 1. The applied twelfth method is MMFCC using EDCD instead of FFT. The used thirteenth method is MMFCC using FRDCT instead of FFT. The fractional power of FRDCT is 1.

The feature extraction steps are represented at Fig.1. FRDCT is used the fourth step; because it achieves the highest AUC and accuracy rate, as described later. Female fetus and her maternal heart sounds are represented against time, as shown in Fig.2. The output of high pass filter in the pre-emphasizing step is shown in Fig.3. The output of hamming window step is presented in Fig.4. The output of FRDCT step is shown in Fig.5. The output of log Mel-scale filter is presented in Fig.6. As shown in Fig.7, coefficients of MFCC based on FRDCT are presented.

TABLE 1. Tp, Fp, Fn and Tn of different models.

	Tp	Fp	Fn	Tn
Model 1	20	1	2	17
Model 2	20	1	0	19
Model 3	13	8	0	19
Model 4	21	0	5	14
Model 5	21	0	5	14
Model 6	20	1	0	19
Model 7	13	8	0	19
Model 8	12	9	1	18
Model 9	13	8	1	18
Model 10	13	8	0	19
Model 11	21	0	6	13
Model 12	12	9	0	19
Model 13	21	0	1	18

The number of true positive (Tp) cases predicts the positive classification correctly. The number of true negative (Tn) cases predicts the negative classification correctly. However, the number of false positive (Fp) cases predicts the positive classification incorrectly, and the number of false negative (Fn) cases predicts the negative classification incorrectly. These values are used for classifying the gender of the fetus.

As shown in Table 1, values of Tn, Fn, Fp and Tp, are presented. Parameters of classification can be calculated, as shown in equations (27)-(33). Where t_{pr} is true positive rate (sensitivity) and f_{pr} is false positive rate (1-specificity). These parameters are used to classify the gender of fetus. As shown in Table 2, classification parameters rates of different models are represented. They used to evaluate the presented models. Acc, sens, spec, pres, Fm, Gm, and Au are accuracy rate, sensitivity rate, specificity rate, precision rate, F-measure rate, G-mean rate and AUC.

Although the accuracy rates in models 2, 6 and 13 are the most accuracy rates, but the AUC of model 13 is the biggest; so model 13 is selected.

After using deep learning as classification method, the Receiver operating curve (ROC) is applied. The ROC is used to measure the performance of the classification method in different models. Since the ROC approaches from 1, as shown in Fig. 8, the performance of this model is stable. It is preferred for classifying fetal heart sounds.

$$\text{Accuracy rate} = 100 \cdot \frac{TP + TN}{FP + FN + TP + TN} \tag{27}$$

TABLE 2. Classification parameters rates of different models.

	Acc	Sens	Spec	Pres	Fm	Gm	Au
Model 1	92.5	95.2	89.5	90.9	93	92.3	0.927
Model 2	97.5	95.2	100	100	97.7	97.6	0.975
Model 3	80	61.9	100	100	76.5	78.7	0.852
Model 4	87.5	100	73.7	80.8	89.4	85.8	0.904
Model 5	87.5	100	73.7	80.8	89.4	85.8	0.904
Model 6	97.5	95.2	100	100	97.6	97.6	0.975
Model 7	80	61.9	100	100	76.5	78.7	0.852
Model 8	75	57.1	94.7	92.3	70.6	73.6	0.795
Model 9	77.5	61.9	94.7	92.9	74.3	76.6	0.81
Model 10	80	61.9	100	100	76.5	78.7	0.852
Model 11	85	100	68.4	77.8	87.5	82.7	0.889
Model 12	77.5	57.1	100	100	72.7	75.6	0.839
Model 13	97.5	100	94.7	95.5	97.7	97.3	0.977

$$\text{Sensitivity rate} = 100 \cdot \frac{\text{TP}}{\text{FN} + \text{TP}} \tag{28}$$

$$\text{Specificity rate} = 100 \cdot \frac{\text{TN}}{\text{FP} + \text{TN}} \tag{29}$$

$$\text{Precision rate} = 100 \cdot \frac{\text{TP}}{\text{FP} + \text{TP}} \tag{30}$$

$$\text{F\_measure} = 2 \cdot \frac{\text{precision} * \text{sensitivity}}{\text{precision} + \text{sensitivity}} \tag{31}$$

$$G - mean\ rate = 100.\sqrt{\frac{Tp}{Tp+Fn} . \frac{Tn}{Tn+Fp}} \tag{32}$$

$$AUC = \int_0^1 t_{pr}\left(f_{pr}\right) df_{pr} \tag{33}$$

IV. DISCUSSION

Many researches are conducted using Stockwell transform of fetal ECG or fetal heart sounds. Ulusar et al. used Stockwell transform to analyze signals of brain activity [23]. Lui et al. removed noise from fetal heart sounds using empirical mode decomposition [24]. Krishna extracted features of fetal ECG using Stockwell transform or Short Time Fourier Transform (STFT). Results proved that using Stockwell transform for feature extraction of fetal ECG overperform STFT [25]. Koutsiana et al. detected fetal heart sounds using fractal dimention and wavelet transform [26]. Abduh and others classified heart sounds using FRFT based on Mel-frequency spectral coefficients [27]. Bajaj and others denoised ECG signals using FRST [28]. Gupta and Mittal detected arrhythmia in ECG signal using fractional wavelet transform with

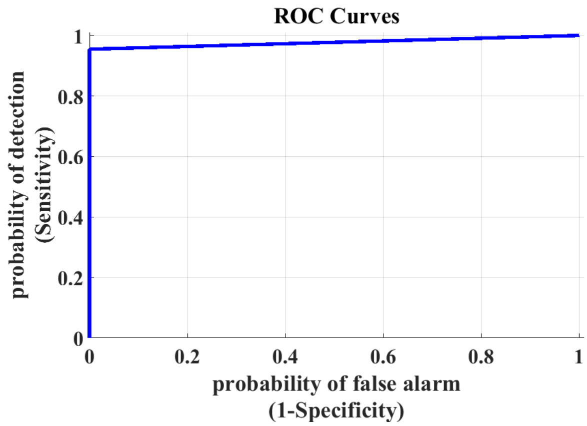

FIGURE 8. ROC of MMFCC based on FRDCT using deep learning.

principal component analysis [29]. Jallouli and his colleagues used Clifford wavelet entropy to extract features of fetal ECG signals [30]. Krupa and his colleagues estimated fetal heart rate using FRFT and wavelet transform [31]. Krupa et al. used joint time-frequency analysis to extract fetal ECG [32]. Krupa and Dhanalakshmi made researches about automatic detection of fetal QRS complex using ST based on deep learning [33]. Krupa et al. detected fetal QRS using IoT-based on deep learning [34]. Gupta et al. used Stockwell Transform and empirical mode decomposition for detecting arrhythmic of fetus [35].

In this paper, 13 methods are discussed. The successful method, which achieves high accuracy and high AUC, is selected using MFCC based on FRDCT. The model of FRDCT decomposes fetal heart sound signals perfectly in time and frequency domains; so this model achieves the highest accuracy rate and highest AUC.

The drawback of this research is using small database of fetal heart sounds. This database represents only weeks' gestation between 26^{th} and 36^{th} week.

In future, novel fractional transforms will be used to extract features of fetal heart sounds; they may achieve high accuracy rates. The used database of fetal heart sounds will be increased. The measured fetal heart sounds during maternal gestation will be obtained in earlier weeks.

V. CONCLUSION

Fetal heart sounds used to follow up development of fetus in the uterus. Thirteen methods are compared to obtain features of fetal heart sounds. Deep learning is used as a classification technique. Model 13 is selected; to obtain features of fetal heart sounds. The obtained accuracy rate is 97.5%. The obtained AUC is 0.977. In future, novel fractional transforms will be used to extract features of fetal heart sounds; they may achieve high accuracy rates. The used database of fetal heart sounds will be increased. The measured fetal heart sounds during maternal gestation will be obtained in earlier weeks.

REFERENCES

[1] J. J. Sampson, R. L. McCalla, and W. J. Kerr, "Phonocardiography of the human fetus," *Amer. Heart J.*, vol. 1, no. 6, pp. 717–734, Aug. 1926, doi: 10.1016/s0002-8703.

[2] R. Sameni, "A review of fetal ECG signal processing issues and promising directions," *Open Pacing, Electrophys. Therapy J.*, Jan. 2010, doi: 10.2174/1876536x01003010004.

[3] M. Samieinasab and R. Sameni, "Fetal phonocardiogram extraction using single channel blind source separation," in *Proc. 23rd Iranian Conf. Electr. Eng.*, May 2015, pp. 78–83, doi: 10.1109/IranianCEE.2015.7146186.

[4] P. Sutha and V. Jayanthi, "Fetal electrocardiogram extraction and analysis using adaptive noise cancellation and wavelet transformation techniques," *J. Med. Syst.*, vol. 42, no. 1, p. 21, Dec. 2017, doi: 10.1007/s10916-017-0868-3.

[5] A. M. Ponsiglione, C. Cosentino, G. Cesarelli, F. Amato, and M. Romano, "A comprehensive review of techniques for processing and analyzing fetal heart rate signals," *Sensors*, vol. 21, no. 18, p. 6136, Sep. 2021, doi: 10.3390/s21186136.

[6] R. Khanmohammadi, M. S. Mirshafiee, M. M. Ghassemi, and T. Alhanai, "Fetal gender identification using machine and deep learning algorithms on phonocardiogram signals," 2021, *arXiv:2110.06131*.

[7] H. Namazi and O. Krejcar, "Analysis of pregnancy development by complexity and information-based analysis of fetal phonocardiogram (PCG) signals," *Fluctuation Noise Lett.*, vol. 20, no. 4, Aug. 2021, Art. no. 2150028, doi: 10.1142/s0219477521500280.

[8] M. Farahi, A. Casals, O. Sarrafzadeh, Y. Zamani, H. Ahmadi, N. Behbood, and H. Habibian, "Beat-to-beat fetal heart rate analysis using portable medical device and wavelet transformation technique," *Heliyon*, vol. 8, no. 12, Dec. 2022, Art. no. e12655, doi: 10.1016/j.heliyon.2022.e12655.

[9] M. Alkhodari, N. Widatalla, M. Wahbah, R. Al Sakaji, K. Funamoto, A. Krishnan, Y. Kimura, and A. H. Khandoker, "Deep learning identifies cardiac coupling between mother and fetus during gestation," *Frontiers Cardiovascular Med.*, vol. 9, Jul. 2022, doi: 10.3389/fcvm.2022.926965.

[10] R. Abburi, I. Hatai, S. Samanta, R. Jaros, R. Martinek, T. A. Babu, and S. A. Babu, "Adopting artificial intelligence algorithms for remote fetal heart rate monitoring and classification using wearable fetal phonocardiography," *Authorea Preprints*, Oct. 2023.

[11] M. Alkhodari, M. Almadani, S. K. Ghosh, and A. H. Khandoker, "Fhsu-net: Deep learning-based model for the extraction of fetal heart sounds in abdominal phonocardiography," in *Proc. IEEE 33rd Int. Workshop Mach. Learn. Signal Process. (MLSP)*, Sep. 2023, pp. 1–6, doi: 10.1109/mlsp55844.2023.10285907.

[12] R. Sameni, M. Samieinasab. (2021). *Shiraz University Fetal Heart Sounds Database*. PhysioNet. [Online]. Available: https://doi.org/10.13026/42eg-8e59

[13] P. Lubaib and K. V. A. Muneer, "The heart defect analysis based on PCG signals using pattern recognition techniques," in *Proc. Int. Conf. Emerg. Trends Eng., Sci. Technol.*, 2015, pp. 1024–1031.

[14] O. F. Omogbenigun, *Speech Core Toolbox*. London, U.K.: London Metrop. Univ., 2007.

[15] M. Schimmel, J. Gallart, and C. Simon, "An ilternative inverse S-transform for filters with time-frequency localization," in *Proc. 4th Int. Symp. Image Signal Process. Anal.*, 2005, pp. 424–429.

[16] D. G. Bhalke, C. B. R. Rao, and D. S. Bormane, "Fractional Fourier transform based features for musical instrument recognition using machine learning techniques," in *Proc. Int. Conf. Frontiers Intell. Comput., Theory Appl. (FICTA)*, vol. 247, 2014, pp. 155–163.

[17] H. M. Ozaktas, M. A. Kutay, and G. Bozdagi, "Digital computation of the fractional Fourier transform," in *IEEE Trans. Sig. Proc.*, vol. 44, 1996, pp. 2141–2150.

[18] D. Wei and Y. Zhang, "Fractional stockwell transform: Theory and applications," *Digit. Signal Process.*, vol. 115, Aug. 2021, Art. no. 103090, doi: 10.1016/j.dsp.2021.103090.

[19] W. Zhou, Z. Feng, Y. F. Xu, X. Wang, and H. Lv, "Empirical Fourier decomposition: An accurate signal decomposition method for nonlinear and non-stationary time series analysis," *Mech. Syst. Signal Process.*, vol. 163, Jan. 2022, Art. no. 108155, doi: 10.1016/j.ymssp.2021.108155.

[20] A. K. Jain, *Fundamental of Digital Image Processing,*. Upper Saddle River, NJ, USA: Prentice-Hall, 1989.

[21] W. B. Pennebaker, and J. L. Mitchell, *JPEG Image Data Compression Standard, Van Nostrand Reinhold*. New York, NY, USA: Van Nostrand Reinhold, 1993, ch. 4.

[22] O. Faust, Y. Hagiwara, T. J. Hong, O. S. Lih, and U. R. Acharya, "Deep learning for healthcare applications based on physiological signals: A review," *Comput. Methods Programs Biomed.*, vol. 161, pp. 1–13, Jul. 2018.

[23] U. D. Ulusar, J. D. Wilson, P. Murphy, R. B. Govindan, H. Preissl, C. L. Lowery, and H. Eswaran, "Bio-magnetic signatures of fetal breathing movement," *Physiol. Meas.*, vol. 32, no. 2, pp. 263–273, Feb. 2011.

[24] Q. Liu, Z. Tan, Y. Zhang, H. Wang, "Denoising of fetal heart sound based on empirical mode decomposition method," *J. Biomed. Eng.*, vol. 32, no. 4, pp. 5–740, Aug. 2015.

[25] B. T. Krishna, "Fetal ECG extraction using time-frequency analysis techniques," in *Proc. Int. Conf. Robot. Autom. Sci. (ICRAS)*, Aug. 2017, pp. 167–171, doi: 10.1109/ICRAS.2017.8071938.

[26] E. Koutsiana, L. J. Hadjileontiadis, I. Chouvarda, and A. H. Khandoker, "Fetal heart sounds detection using wavelet transform and fractal dimension," *Frontiers Bioeng. Biotechnol.*, vol. 5, Sep. 2017, doi: 10.3389/fbioe.2017.00049.

[27] Z. Abduh, E. A. Nehary, M. Abdel Wahed, and Y. M. Kadah, "Classification of heart sounds using fractional Fourier transform based mel-frequency spectral coefficients and traditional classifiers," *Biomed. Signal Process. Control*, vol. 57, Mar. 2020, Art. no. 101788, doi: 10.1016/j.bspc.2019.101788.

[28] A. Bajaj and S. Kumar, "A robust approach to denoise ECG signals based on fractional stockwell transform," *Biomed. Signal Process. Control*, vol. 62, Sep. 2020, Art. no. 102090, doi: 10.1016/j.bspc.2020.102090.

[29] V. Gupta and M. Mittal, "Arrhythmia detection in ECG signal using fractional wavelet transform with principal component analysis," *J. Inst. Engineers (India), Ser. B*, vol. 101, no. 5, pp. 451–461, Oct. 2020, doi: 10.1007/s40031-020-00488-z.

[30] M. Jallouli, S. Arfaoui, A. Ben Mabrouk, and C. Cattani, "Clifford wavelet entropy for fetal ECG extraction," *Entropy*, vol. 23, no. 7, p. 844, Jun. 2021, doi: 10.3390/e23070844.

[31] A. Jaba Deva Krupa, S. Dhanalakshmi, N. L. Sanjana, N. Manivannan, R. Kumar, and S. Tripathy, "Fetal heart rate estimation using fractional Fourier transform and wavelet analysis," *Biocybernetics Biomed. Eng.*, vol. 41, no. 4, pp. 1533–1547, Oct. 2021, doi: 10.1016/j.bbe.2021.09.006.

[32] A. Jaba Deva Krupa, S. Dhanalakshmi, and R. Kumar, "Joint time-frequency analysis and non-linear estimation for fetal ECG extraction," *Biomed. Signal Process. Control*, vol. 75, May 2022, Art. no. 103569, doi: 10.1016/j.bspc.2022.103569.

[33] A. J. Deva Krupa and S. Dhanalakshmi, "Automatic detection of fetal QRS complex using time-frequency image based features and deep learning architecture," in *Proc. 3rd Int. Conf. Electron. Sustain. Commun. Syst. (ICESC)*, Aug. 2022, pp. 778–782, doi: 10.1109/ICESC54411.2022.9885294.

[34] A. Jaba Deva Krupa, S. Dhanalakshmi, K. W. Lai, Y. Tan, and X. Wu, "An IoMT enabled deep learning framework for automatic detection of fetal QRS: A solution to remote prenatal care," *J. King Saud Univ. Comput. Inf. Sci.*, vol. 34, no. 9, pp. 7200–7211, Oct. 2022, doi: 10.1016/j.jksuci.2022.07.002.

[35] K. Gupta, V. Bajaj, and I. A. Ansari, "Integrated S-transform-based learning system for detection of arrhythmic fetus," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–8, 2023, doi: 10.1109/TIM.2023.3271739.

MOHAMED MOUSTAFA AZMY received the B.Sc., master's, and Ph.D. degrees from the Faculty of Engineering, Alexandria University, in 1999, 2004, and 2008, respectively. His master's thesis was about wideband microstrip antenna and the Ph.D. thesis was about robust automatic Arabic speech recognition using syllables. Since 2009, he has made many researches about biomedical signal processing using machine learning.