Morphology Extraction of Fetal Electrocardiogram by Slow-Fast LSTM Network

This paper was downloaded from TechRxiv (<https://www.techrxiv.org>).

LICENSE

CC BY 4.0

SUBMISSION DATE / POSTED DATE

13-07-2020 / 14-07-2020

CITATION

Zhou, Ziqun; Huang, Kejie; Shen, Haibin; Qiu, Yue (2020): Morphology Extraction of Fetal Electrocardiogram by Slow-Fast LSTM Network. TechRxiv. Preprint. <https://doi.org/10.36227/techrxiv.12645878.v1>

DOI

[10.36227/techrxiv.12645878.v1](https://doi.org/10.36227/techrxiv.12645878.v1)

Morphology Extraction of Fetal Electrocardiogram by Slow-Fast LSTM Network

Ziqun Zhou, *Student Member, IEEE*, Kejie Huang, *Senior Member, IEEE*, Yue Qiu and Haibin Shen

Abstract—The morphology of Fetal Electrocardiogram (FECG) plays an important role in the early diagnosis of fetal health condition. However, it is intractable to extract the clean morphology of FECG signals, which are usually contaminated by Maternal ECG (MECG) and various noises. To extract the clean morphology of FECG signals from noninvasive abdominal ECG records, a high-performance and high-efficient two-stage Slow-Fast Long Short Term Memory (SFLSTM) based architecture is proposed. The MECG elimination and the FECG enhancement are realized by the elaborately designed slow LSTM and fast LSTM to filter out the MECG and the residual noise components, respectively. Qualitative and quantitative experiments are conducted on the records from two public databases. The experimental results show that our proposed MECG elimination and FECG enhancement schemes improve the Signal-to-Noise Ratio (SNR) by 3.09 dB and 1.81 dB, respectively. The proposed fast LSTM reduces the amount of computation by approximately 50%, without any degradation in performance. Our proposed method may leverage the noninvasive FECG monitoring for the early detection of fetal heart diseases.

Index Terms—adaptive filter, fetal electrocardiogram, FECG extraction, FECG enhancement, long short-term memory, recurrent neural network

I. INTRODUCTION

FETAL Electrocardiogram (FECG) presents the electrical activity of the fetal heart, which is one of the most important biomedical signals to help diagnose fetal heart health condition. FECG signals are extracted from Abdominal Electrocardiogram (AECG) signals, which are usually sensed by the noninvasive sensors attached to the mother's abdomen. Noninvasive FECG provides a stable long-term continuous heartbeat monitoring and morphological analysis [1].

Unfortunately, the weak FECG signals are susceptible to the strong Maternal ECG (MECG) signals from the mother's heartbeat. Moreover, AECG signals are usually interfered and deformed by various noises, such as the maternal Electromyogram (EMG), Power Line Interference (PLI), baseline wandering, and random electronic noises. Moreover, because of the overlap between the MECG and FECG components in time and frequency and the short of ground truth annotation, the FECG morphology extraction will be full of challenges and difficulties. As a result, it is an intractable job to extract the clean morphology of FECG signals from AECG signals. In the majority of the existing works, the R-peak detection algorithm has been widely used to extract the Fetal Heartbeat

Rate (FHR) by enhancing the fetal R-peak or eliminating the noise components [2], [3]. Most of the existing works, such as the most well-known wavelet transform methods, aim to extract accurate FHR, which may distort the original waveform of FECG signals [4]. The deformation is usually caused by the R-peak enhancement to obtain accurate FHR and the overlap between the MECG and FECG components in time and frequency.

In this way, there is still some space to improve the performance of conventional methods in the task of the morphology extraction of FECG. Some methods such as Template Subtraction (TS) [5], [6] may achieve a poor performance when dealing with the appreciable overlap between FECG and MECG and highly depend on the mQRS detection performance. Moreover, many parameters need to be set carefully in the parts of existing methods such as the Wavelet Transform (WT). Some methods, such as Blind Source Separation (BSS) techniques [7], need a large number of abdominal channels which may cause the discomfort to the mothers and the challenge of clinical utilization. Besides, obtaining the real data for training is difficult for the learning methods such as Adaptive Neuro-Fuzzy Inference Systems (ANFIS) and Least Square Support Vector Machine (LSSVM).

To address the issues mentioned above, a two-stage architecture based on Slow-Fast Long Short Term Memory (SFLSTM) is proposed, where one stage is to eliminate the MECG components adaptively through an elaborately designed Long Short Term Memory (LSTM) unit, and the other stage is to enhance FECG through a fast LSTM unit (also named as Compact LSTM). Compared with LSTM, the proposed fast LSTM reduces the amount of computation by approximately 50%, with no performance degeneration in the stage of enhancement.

Qualitative and quantitative experiments are conducted on the Database for the Identification of Systems (DaISy) and the Fetal ECG Synthetic Database (FECGSYNDB). Our proposed system can adapt to different noise conditions, and thus the performance is much improved. The experimental results show that our method extracts much cleaner FECG components. Compared with the conventional FECG extraction methods, the proposed method improves the Signal-to-Noise Ratio (SNR) by 3.09 dB in MECG elimination and another 1.81 dB in FECG enhancement and get better performance in lower Root Mean Square Error (RMSE) and latency time. Approximately 4.9 dB gain in total will significantly increase the accuracy of the fetal heart condition diagnosis.

The rest of this paper is organized as follows. Section II briefly introduces the ground works related to our work.

Z. Zhou, K. Huang, Y. Qiu and H. Shen are with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou, 310027 People's Republic of China. Email: {zhouziquan, huangkejie, yueq235, shen_hb}@zju.edu.cn.

TABLE I: Related Works and Description

Adaptive Methods		Non-adaptive Methods	
EKF	EKF, which is usually used for long term monitoring [24], can adaptively estimate the changing of the input signal and filter out FECG components.	TS	FECG is extrapolated by the templates from the input mixed signals, which depends on the quality of mQRS detection [4]. The parametric formulation got from a regular quasi-periodic heart rate pattern would hamper the detection of events such as extra systoles.
LMS	Only the current stage is used to minimize the Mean Square Error (MSE) between the output and the target.		
RLS	Historical signals are used to reduce the total squared error.	WT	It is limited by cross-over in the spectral domain [25]. Parameters must be selected carefully, including wavelet basis, thresholding rules and level of decomposition.
ESN	It works as a nonlinear medium for the reference signals to propagate through. MECG signals are estimated and subtracted by neural network. [4]		
ANFIS	A Sugeno fuzzy inference system based neuro-adaptive learning algorithm is utilized to determine the relationship between MECG and FECG.	BSS	The most popular spatial method [4] to detect and extract ECG patterns [7]. It requires a large number of abdominal channels, which cause discomfort to the mother and the clinical utilization challenging.

Section III describes our proposed two-stage LSTM network in detail. Section IV provides the experimental results. Finally, the conclusion is drawn in Section V.

II. RELATED WORKS

The research works with respect to the FECG signal extraction are under intensive development. Numerous approaches have been utilized in FECG extraction in recent years. They can be divided into adaptive methods and non-adaptive methods. Adaptive methods adjust the coefficients of the FIR filter based on the error signals, such as Adaptive Noise Cancelling (ANC) [8], Extended Kalman Filtering (EKF) [9], [10], Least Mean Squares (LMS) [11], [12], Recursive Least Squares (RLS) [12], Echo State neural Network (ESN) [12], and ANFIS [13]. The non-adaptive methods include TS [5], [6], WT [14], neural network [15], polynomial networks [16], Event-Synchronous interference Canceller (ESC) [17], v-support vector regression (v-SVR) [18], LSSVM [19], and BSS [7], such as Principle Component Analysis (PCA) [20], Independent Component Analysis (ICA) [21], Periodic Component Analysis (π CA) [22] and FastICA [23]. These methods are described briefly in Table I. Most of these non-adaptive methods can be combined with adaptive methods to form a variety of variant structures to improve performance.

Most of the existing works aim at the accurate extraction of FHR. The preprocessing steps for better FHR extraction always significantly distort the original waveform of FECG signals. Therefore, the morphology extraction of FECG signals is still an unsolved problem. The two obstacles to the realization of the task are the residual MECG components and the inevitable residual noises. We found the existing methods are still suffering from extra systoles, maternal and fetal beats overlapping, low SNR, or low quality of mQRS detection [4]. The resultant FECG components usually contain plenty of noises at the same order of magnitude and frequency as the desired FECG components, which will severely disrupt the morphology analysis. Therefore, the enhancement of FECG to eliminate resided noises is vital. In summary, two stages of processing are indispensable for the morphology extraction of FECG signals. The first stage is the MECG elimination to eliminate the MECG components from the AECG records without disturbing the FECG components. The second stage is the FECG enhancement, to extract the clean morphology of

FECG components from the filtered results contaminated by residual noises.

The ANC technique is a widely used method to estimate signals corrupted by additive noise or interference [8], i.e., eliminating MECG from the AECG records. It requires neither the prior knowledge of the MECG components nor the accurate locations of R peaks in MECG for signal slicing and alignment. The vanilla ANC only supports linear transformation, whereas the correlation between the abdominal MECG and the thoracic MECG is usually nonlinear. Hence, the defect of ANC will cause that the filtered FECG contains non-negligible MECG components. A method used to ameliorate the problem is to involve a “primary” input with a corrupted signal and a “reference” input containing pure noise signals correlated in some unknown way with the noise components in the primary input. The method transforms the reference input adaptively and estimates the signal by subtracting the transformed reference input from the primary input [26].

The principle ANC plays a role like a discriminator to eliminate the components from the primary input, which are linearly correlated to the reference input. It adjusts an adaptive filter to remove the MECG components from the AECG records by sending the thoracic MECG signal to the reference input [8]. However, since the characteristic of linear correlation, the vanilla ANC demands the reference input morphologically similar to the abdominal MECG waveform in the AECG records to work properly. Previous works have confirmed this phenomenon in the literature [8]. Unfortunately, the signal mapping from the maternal heart to the mother’s abdomen is nonlinear in reality, and the morphology of the sensed ECG waveforms are highly related to the locations of electrodes and the intrinsic characteristics of the heart. Hence, the required similarity cannot be guaranteed. As a result, ANC with linear filters has limited performance in MECG elimination.

Numerous works have been done to address the defects of ANC [27], [28]. Since the approximation of nonlinear transformation is required in the process of MECG elimination, these methods tried to extend the linear transformation provided by the linear adaptive filter to a nonlinear transformation, with the help of differential nonlinear mapping tools. These nonlinear ANC filter out the signal components in the primary input nonlinearly correlated to the reference input and retain the “independent” components. Nevertheless, through the experi-

ments, we found that the filtered FECG signals by the existing methods still have non-negligible residual MECG components, which will severely interfere with the morphology analysis of FECG.

The process of FECG enhancement eliminates the residual noises that reside in the filtered FECG, which is an indispensable step for the morphology extraction. Despite the numerous works done for the MECG elimination, very few works have been done to enhance FECG signals. [29] proposed a time-sequenced ANC method for the FECG enhancement. However, this method requires the accurate location of R peaks in FECG, which will significantly affect the performance. Moreover, it needs to adjust hundreds of linear filters, which would challenge the memory space and computation complexity.

In summary, the conventional methods for MECG elimination and FECG enhancement are still unqualified, more efficient methods are desired.

III. METHODS

A. Adaptive Nonlinear Noise Cancellation

The concept of ANC is shown in Fig. 1. Signal s is transmitted over a channel to an electrode sensor, which will also inevitably receive an unrelated noise n_0 . $s + n_0$ forms the primary input of the canceller. A second sensor receives a noise n_1 , which is uncorrelated with the signal s but correlated in some unknown way with the noise n_0 . This sensor provides the reference input to the canceller. The reference input is transformed through an adaptive filter to produce an output y . The adaptive filter is expected to make y as close as to n_0 . Then the output y is deducted from the primary input $s + n_0$ to produce the system output $z = s + n_0 - y$, which is also the estimation of signal s . More details of the ANC concept can be found in [8], including the theoretical proof.

Fig. 1: The concept of Adaptive Noise Cancellation.

However, the strict independence is only theoretically true if the processed signal has finite length. In this case, it is more suitable to describe the relationship “independent” as the non-linear correlation with high complexity, which is very difficult to be approximated. Thus, the relationship between the two records is either linearly correlated or nonlinearly correlated.

For the MECG elimination and FECG enhancement, the correlation is approximated nonlinearly in reality. Thus, the purpose of elimination and enhancement is redefined as to distinguish the complexity of the correlation between the signal components and the reference input. Hence, a “screener” is expected, whose screening criteria is the complexity of non-linear correlation.

To assess and demonstrate the complexity of correlation, we test the correlation or independent relationship between different signal components such as Abdominal MECG (AMECG), Thoracic MECG (TMECG), Abdominal FECG (AFECG), Thoracic FECG (TFECG) and noise through two indexes: Pearson product-moment correlation coefficient (PPMCC) and Hilbert-Schmidt Independent Criterion (HSIC) [30]. HSIC is used to measure the statistical dependence between two random vectors. [31] We randomly select signals from different channels in FECGSYNDB of all individual models for many times to guarantee the completeness. The signal correlation coefficients and the results of HSIC between different components are calculated in two cases. The first case is that indexes between pure maternal and fetal signals (AMECG, TMECG, AFECG, TFECG) got from abdomen and thorax are calculated without noise. In the other case, extra noise is added in correlation analysis. A strong correlation between MECG signals from abdominal and thorax or between FECG signals from these two parts can be observed in Table II. It is also obvious that MECG signals and FECG signals have very weak correlation. It means that the correlation between MECGs (AMECG and TMECG) or between FECGs (AFECG and TFECG) is much greater than the correlation between MECG and FECG.

TABLE II: The complexity of correlation

signal components 1&2	PPMCC		HSIC	
	case1	case2	case1	case2
AMECG,TMECG	8.11E-01	7.54E-01	3.68E-02	9.86E-03
AFECG,TFECG	6.14E-01	6.14E-01	4.68E-02	4.68E-02
AMECG,AFECG	8.60E-03	8.50E-03	1.14E-03	5.45E-04
TMECG,AFECG	8.23E-03	8.14E-03	1.22E-03	5.58E-04
AMECG,TFECG	9.11E-03	8.94E-03	1.33E-03	6.36E-04
TMECG,TFECG	8.52E-03	8.42E-03	1.14E-03	5.45E-04

The original concept of ANC is based on linear correlation. With the linear adaptive filters, the vanilla ANC distinguishes the linear correlation from the non-linear correlation between the signal components and the reference. Hence, an adaptive non-linear noise canceller is required to address the non-linear problem. Since the screening criteria of ANC is dependent on the adaptive filter, it is natural to replace the linear filter with non-linear mapping tools. An adaptive non-linear noise canceller is derived to distinguish non-linear correlations with different complexity, and the vanilla ANC is transformed into the expected screener. The rationality of this conceptual expansion has been verified in some degree by the success of former works [3], [32], which have tried to combine ANC with different non-linear mapping tools.

Non-linear mapping tools with sufficient capacity and flexibility are desired to get an ideal screener. Furthermore, the intrinsic sequential characteristic of signal data also contains useful information, which has not been fully utilized by the previous works. In this paper, we apply LSTM as the non-linear mapping tool, which is adept in solving the problems involving sequential data and is capable of approximating any non-linear relationships with sufficient model capacity [33]. A fast LSTM cell is also proposed to improve significantly

the computational efficiency, which will be discussed in Section III-C.

B. Long Short-Term Memory

Recurrent Neural Network (RNN) has been widely used in sequential nonlinear problems in recent years. The recurrent connections between neurons provide the network with inherent recurrent and causal architecture, which enables the model to transform the acquired historical information into desired results. However, the vanilla RNN usually has poor performance, which is caused by the vanishing gradient problem [34]. Fortunately, LSTM was proposed as a unique architecture of RNN cells to solve this problem, which maintains a more constant error to allow the model to learn over many steps, thereby opening a channel to link causes and effects remotely [33]. The architecture of the LSTM unit is shown in Fig. 2a. The formulas of LSTM (including the principle of RNN) are described in the following composite functions:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (3)$$

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \quad (4)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (5)$$

$$h_t = o_t \times \tanh(C_t) \quad (6)$$

where x_t is the input vectors, σ is the logistic sigmoid function, f_t , i_t , o_t and C_t , which have the same size as the hidden vector h_t , are the forget gate, input gate, output gate and cell activation vectors, respectively. W_f , b_f , W_i , b_i , W_o , b_o , W_C and b_C are the weight and bias matrices from the cell to forget gate, input gate, output gate and cell activation vectors, respectively. It is worth noting that in the architecture of LSTM, the requirement of the input dimension is flexible. The RNN model with LSTM is applicable for computations of multi-dimension vectors, which is an instrumental feature for our method.

We have proposed a novel method based on the technique of LSTM to eliminate the PLI from the ECG records [35], which is a critical pre-processing step for the morphology extraction of FECG signals. The performance we obtained reflects the potential of these techniques in the field of signal processing. The problem of eliminating PLI from ECG signals can be attributed to the category that extracting the signal components with known characteristics from non-stationary signals. However, in an AECG record, both FECG and MECG components are non-stationary. Thus in this paper, the problem is totally different, and a specific system scheme is required to solve this problem.

C. Fast LSTM

The architecture of LSTM was originally designed for long term dependencies in data sequences, such as speech recognition and machine translation, which frequently utilize the dependencies between words with long word span [36]. However, the internal mapping system of the FECG signals

Fig. 2: Architecture comparisons of cell units. (a) The original architecture of LSTM; (b) The proposed architecture of the fast LSTM.

is causal and locally consistent on the timeline. Thus, the attenuation coefficient of the hidden state should be smoothly variant or constant, which is different from the mechanism controlled by the original forget gate.

TABLE III: An Example of Variable Values in LSTM during the process of FECG enhancement

Variables	Mean $\pm$ Std.	Std. / Mean
i_t	0.4756 ± 0.0486	10.22%
f_t	$0.6939 \pm \mathbf{0.0303}$	4.37%
o_t	0.4959 ± 0.0558	11.25%
x_t	0.5590 ± 0.2289	40.95%

To verify our hypothesis, we employed the LSTM cell stated above for the experiment of the FECG enhancement. Concretely, we recorded all the values of the three gates during the process. The value of the input and output gates after the convergence of the LSTM model is tabulated in Table III. Table III illustrates the mean and standard deviation of the values at the input and the output of the three gates. Despite the dramatic change at input vector x_t , it can be observed that the output of the forget gate has a much smaller variance than the other two. It means that the value of f_t fluctuates within a very small range. This result is automatically learned by the LSTM model, which could be regarded as strong evidence for our hypothesis.

To improve the computing efficiency, a fast LSTM is proposed. More concretely, the forget gate is switched to a relatively constant value, which obtains the gate value without taking x_t and h_{t-1} as input. We also abandoned the output gate, since we found that the output gate has little effect on the function of LSTM in the FECG enhancement stage, which has also been proved in other applications [37]. The definition of the input gate is kept intact to maintain the function and flexibility of LSTM. The architecture of the fast LSTM is shown in Fig. 2b and the formulas of the modification we

TABLE IV: Overall of cell unit

Cell Unit	Parameters	Forward Ops	Backward Ops
slow LSTM	$4 \times [n \times (n + m) + n]$	$4 \times n \times (n + m) + 3 \times n$	$4 \times [n \times (n + m) + 2 \times m] + 3 \times (n + n)$
fast LSTM	$2 \times [n \times (n + m) + n] + n$	$2 \times n \times (n + m) + 2 \times n$	$2 \times [n \times (n + m) + 2 \times m] + n + n$

made are provided in (7) and (8).

$$f_t = \sigma(b_f) \quad (7)$$

$$o_t = 1.0 \quad (8)$$

The overall complexity comparisons of these two cell architectures are presented in Table IV. The term n represents the dimension of the input hidden vector h_{t-1} (the number of the cell units), and m denotes the dimension of the input vector x_t . It can be observed that fast LSTM reduces the number of recurrently updating parameters by approximately 50%. The computation complexity of each architecture is presented by the number of multiplication for one propagation in the cell unit. These indexes show that fast LSTM only has approximately half of the computation complexity compared with slow LSTM in both forward and backpropagation. Our proposed fast LSTM not only greatly reduces the computations but also outperforms or keeps the same performance as the slow LSTM in the enhancement stage. One disadvantage of our proposed fast LSTM is that it is inferior to the slow LSTM in the extraction stage, where the strong nonlinear mapping ability is required. Therefore, the slow LSTM structure is used in the extraction stage for high performance.

Moreover, the information contained in each data point is influenced by the sampling rate in the task of signal processing, which is very different from the data processed by the LSTM in conventional applications. The information in each signal data point will be attenuated if the sampling rate is too high. Despite the function provided by LSTM, the model could still suffer from the vanishing gradient problem if the input sequence is too long. The computational complexity will also be challenged. Thus, for this ECG processing problem, we downsampled the original records to 250 Hz if necessary, which is still higher than the Nyquist rate of FECG signals.

D. MECG Elimination

Generally, noninvasive AECG signals $A(t)$ sensed from the maternal abdominal wall consists of several signal components, which can be modeled by the following equation:

$$A(t) = M(t) + F(t) + w(t), \quad (9)$$

where t is the time index, $M(t)$ represents the maternal ECG signal components, $F(t)$ represents the pure FECG, and $w(t)$ represents all the noise signals, such as power line interference, baseline wandering, motion artifact, and electromyogram.

From the perspective of ANC, the component $M(t) + w(t)$ in $A(t)$ is the noise resided in the primary input. However, in the process of MECG elimination, usually only the reference input correlated to $M(t)$ is available. Thus, we leave the cancellation of $w(t)$ to the FECG enhancement stage.

To eliminate $M(t)$ from $A(t)$, we use a set of signals $\vec{r}(t)$ as the reference input to approximate the MECG components

$M(t)$ in $A(t)$ through the nonlinear function $R[\cdot]$ supplied by the proposed slow LSTM based model. The block diagram of our method for MECG elimination is shown in Fig. 4. After the approximation, the desired FECG $F(t)$ could be filtered by subtracting the estimated MECG components $\hat{M}(t)$ from $A(t)$. Similar to ANC methods, the reference input is set to be the maternal ECG signal synchronously obtained from the mother's thoracic wall, and the FECG components resided in it should be ignored. In some cases, more than one thoracic channels are available. In our method, we make the definition that the reference input $\vec{r}(t)$ contains all the available thoracic ECG signals inspired by [38]. This definition is made possible by the flexible input dimension requirement of our model. All the signals in $\vec{r}(t)$ are simultaneously and sequentially sent into the model. Thus, $x(t)$ and $\vec{r}(t)$ are derived as follows:

$$x(t) = A(t) \quad (10)$$

$$\vec{r}(t) = \begin{bmatrix} T_1(t) \\ T_2(t) \\ \vdots \\ T_m(t) \end{bmatrix} \quad (11)$$

where $T(t)$ is the available maternal thoracic ECG signals and m is the dimension of $\vec{r}(t)$ (corresponding to x_t in the second stage). Since the noninvasive ECG signals sensed from the body surface are the projections of vectorcardiographic loop, sending more than one channel could provide more information of the cardiac electric activities, which could make the transformation be approximated by the model easier.

The flow of the model computation is shown in Fig. 3. Similar to the traditional filtering method, we need a segment of historical signal points to estimate the desired output. Analogy to the definition of the "order" of a digital filter, we supply a segment with the length of t_o to function $R[\cdot]$ to estimate $\hat{x}(t)$. The formula is shown in (12). The error $\epsilon(t)$ from the subtraction between the estimated $\hat{x}(t)$ and $x(t)$ leads to the adaptation process of the model, as shown in (13). For the MECG elimination, the estimated $\hat{x}(t)$ and the MECG components $M(t)$ in $A(t)$ after the convergence of the model should be equal. Thus, $\epsilon(t)$ also gives the desired filtered FECG output.

$$\hat{x}(t) = R[\vec{r}(t - t_o), \vec{r}(t - t_o + 1), \dots, \vec{r}(t)], \quad (12)$$

$$\epsilon(t) = x(t) - \hat{x}(t). \quad (13)$$

The parameter adaptation of our model needs an objective function for guidance. Considering the concept of ANC, the objective function for the adaptation is defined as the MSE between $\hat{x}(t)$ and $x(t)$. Moreover, referring to the RLS adjustment method broadly applied by ANC [8], we defined

the objective function as follows:

$$J(\Theta, t) = \frac{1}{t_e} \sum_{\tau=t-(t_e-1)/2}^{t+(t_e-1)/2} [\epsilon(\tau)]^2, \quad (14)$$

The step length of the signal segment for evaluation is denoted

Fig. 3: Flow of the model computation.

Fig. 4: System architecture of our method for MECG elimination.

as t_e . It can be seen in (14) and Fig. 3 that we have defined the evaluation scopes both before and after the processed point. The reason is based on the fact that the desired mapping is non-stationary. If we only evaluate the objective function on the data points before the processed one like in the conventional RLS, the resultant mapping approximation will be biased. Furthermore, through the experiments, we found that the input $x(t)$ contains unpredictable noises and distortions in reality. By adjusting the model based on the evaluation results after the processed point, the model is given the ability to adapt to the unwanted interferences in advance. As a side effect, a fixed delay will be introduced to the output.

Fig. 5: Block diagram of our proposed model architecture based on RNN.

The architecture of our model is shown in Fig. 5, which is unrolled to demonstrate the recurrent relationship in the time domain. We need to adapt the model to approximate the desired nonlinear function $R[\cdot]$. The model contains three layers in total, including an input layer, a hidden layer, and an output layer. The input layer holds the sequential input signals. The hidden vector $\tilde{h}(t)$ is the representation of the historical information of the reference input. These hidden states are

combined linearly in the output layer to produce $\hat{x}(t)$. W_{hx} and b_x are the weight and bias matrices from the hidden vector to the output, respectively. The formula of the output layer is provided in (15).

$$\hat{x}(t) = W_{hx}\tilde{h}(t) + b_x \quad (15)$$

E. Adaptive Updating Process of Model Parameters

Parameters of our model are only adaptively updated for the current record under processing, without the usual separation of training and testing phases, which is a challenging machine learning task. The reason is that most of the desired approximations in different records are disparate, and the converged models are non-universal. This mechanism is also consistent with the concept of ANC. The RNN based model can be abstracted as the function f with parameters Θ :

$$\hat{x}(t) = f(\tilde{r}(t); \Theta), \quad (16)$$

where Θ represents all the adjustable parameters in the model including b_f , W_i , b_i , W_C , b_C , W_{hx} and b_x which are defined as:

$$\Theta = [b_f, W_i, b_i, W_C, b_C, W_{hx}, b_x], \quad (17)$$

Θ is continuously optimized to minimize the objective function $J(\Theta, t)$ and to track the variance of the non-stationary mapping. By applying the training algorithm Back-Propagation Through Time (BPTT) [39], the following numerical computation is performed to update the parameters to lower the objective function:

$$\tilde{\Theta} = \Theta - \gamma \frac{\partial J(\Theta)}{\partial \Theta}, \quad (18)$$

where $\tilde{\Theta}$ denotes the updated Θ and γ denotes the learning rate.

We use Adam [40] to adjust γ automatically. In the process of adaptive updating, the learning rate is an important variable that would impact the performance of tracking the non-stationary system. The system with faster variance requires a larger learning rate. The optimization algorithm Adam also has a hyperparameter Γ to configure the basis of automatic adjustment, which could be used to set the adaptation speed of our model.

Considering the principle of mini-batch Stochastic Gradient Descent (SGD) method and enormous former works [41], the convergence of the model can be guaranteed by the adaptive updating process stated above. The construction and the adjustment process of the model are both implemented on the open-source framework Tensorflow [42].

F. FECG Enhancement

In the FECG enhancement, we expect to remove the noise components $w(t)$ resided in the filtered FECG signals, which would always severely corrupt the morphology of FECG and affect the diagnosis results.

We propose to use the subtly changed method introduced in Section III-C and Section III-D, to enhance the FECG signals. In this case, $\tilde{r}(t)$ and $x(t)$ are the filtered FECG signals from different channels, which contain nonlinearly correlated FECG

Fig. 6: System architecture of our method for FECG enhancement.

components $F(t)$, and “independent” noise components $w(t)$. Concretely, the filtered FECG signal to be enhanced is sent to the primary input $x(t)$ and the available ones from the other channels are sent to the reference input $\tilde{r}(t)$. The formulas are shown as follows:

$$x(t) = F_1(t) + w_1(t), \quad (19)$$

$$\tilde{r}(t) = \begin{bmatrix} F_2(t) + w_2(t) \\ F_3(t) + w_3(t) \\ \vdots \\ F_m(t) + w_m(t) \end{bmatrix}, \quad (20)$$

where signals $F_1(t), F_2(t), \dots, F_m(t)$ are nonlinearly correlated FECG components from different channels and $w_1(t), w_2(t), \dots, w_m(t)$ are the corresponding noises. The system architecture for FECG enhancement is shown in Fig. 6. In the process of FECG enhancement, the model is adaptively updated to produce the output as close as possible to $F_1(t)$. The main difference in the FECG enhancement is that $\hat{x}(t)$ is the desired output $\hat{F}_1(t)$, instead of the error signal $\epsilon(t)$. Again, the proposed adaptive nonlinear noise canceller is more like a screener, which sorts out the signal components from the primary input $x(t)$ nonlinearly correlated to the reference input $\tilde{r}(t)$ with lower complexity.

IV. EXPERIMENT RESULTS

A. Database

We use the data from the realtime Database for the Identification of Systems (DaISy) [43], and synthesized database Fetal ECG Synthetic Database (FECGSYNDB) [44], [45].

As shown in Fig. 7, there are five AECG channels and three Thoracic ECG (TECG) channels in DaISy. All eight channels are sampled at 250 Hz synchronously. The duration of each signal is 10 seconds. The AECG signals are denoted as “ $A_i(t)$ ”, and the TECG signals are denoted as “ $T_i(t)$ ”, where the subscript i is used to represent the index of the channel.

FECGSYNDB is a large database comprised of simulated adult and noninvasive FECG signals, generated by the FECGSYN simulator [44]. Ten different pregnancies were created by simulating maternal and fetal heartbeats as punctual dipoles with different magnitudes and spatial positions. For each simulated pregnant subject, seven different physiological events (one of these is the base signal without noise) were simulated, with five different levels of additive noise (SNR = 00, 03, 06, 09, and 12 dB) and five repetitions for each

Fig. 7: The eight channels of signals from DaISy are normalized to the range of 0 to 1.

combination of settings. As a result, this database contains 1750 synthetic signals in total.

B. Experiment Settings

There are four hyper-parameters in our proposed method: t_o , t_e , the number of the slow LSTM and fast LSTM cells in the LSTM layer n , and the hyper learning rate Γ . We summarize the settings in Table V. Through the experiments, we found that there is a trade-off between the speed and the extent of convergence, which is depended on the principle of the adaptive updating process, as mentioned above. Larger Γ helps to speed up the convergence and enables the model to track fast variances of the target non-stationary mapping. However, it will introduce a worse extent of convergence and result in ringing output. All the hyperparameter settings are optimized by experiments.

TABLE V: Summary of Hyper Parameter Settings

Parameters	MECG Elimination	FECG Enhancement
t_o	120 ms	120 ms
t_e	240 ms	240 ms
n	32	32
Γ	0.001	0.001

To compare the performance, we implemented Template Subtraction CERUTTI (TS-CERUTTI), Template Subtraction Extended Kalman Filter (TS-EKF), Template Subtraction Principle Component Analysis (TS-PCA), LMS adaptive filter, RLS adaptive filter, and ESN. These methods were implemented with the FECGSYN toolbox [46] to extract the FECG components from the same MECG records as baselines. FastICA [23] is used to filter out the MECG components for FECG extraction where the output is named as FastICA.. To show the benefit of our proposed fast LSTM, FastICA was also used to replace our proposed FECG enhancement method in

the second stage (named as FastICA-Fil). ANFIS and LSSVM are added only in FECGSYNDB where the source signals are available for learning-based methods. The bell-shaped membership function was used in ANFIS to achieve the best extraction, as described in [47]. The nonlinear transformation of MECG was realized by LSSVM, and then the FECG was obtained by subtracting MECG components from the original signals [48]. When train both methods, the training data and labels were from the first ten seconds of MECG and FECG signals. The testing data and labels were from the same signals, with an interval of more than ten seconds between the training and testing data. Both methods were trained separately for each simulated pregnant subject with 35000 times to achieve convergence. Moreover, to show the performance of MECG elimination and FECG enhancement, we present the outputs of these two stages in all experimental results.

Fig. 8: Time-Sequence Diagram

(a) Other Methods; (b) Our Proposed;

Here t_s is the sampling interval time of $4ms$; t_w is the waiting time for sampling data of enough length to start system; t_d is the time for displaying the waveform of output sequences, which is equal to the length of the output data multiplied by t_s ; t_c is the time for calculating the output; $t_{c/point}$ is the computing time for each point, which can be calculated by dividing t_c by the total length of one output; and t_l is the system latency, which is equivalent to the delay time between output data points and corresponding input points. The blue block represents the input of sampled data points. The purple block represents the corresponding output.

All experiments are conducted on Intel Xeon CPU E5-2678 v3 @ 2.50GHz with Python v3.5 and Matlab 2019b. The time-sequence diagram in Fig. 8 shows the feasibility of our system in real-time practice. As illustrated in Fig. 8a, the conventional methods have to accumulate sufficient data points (a few seconds) for each processing, resulting in long latency from the AECG signal capturing to the FECG morphology extraction. In contrast, our proposed SFLSTM shown in Fig. 8b can extract FECG signals at every data point, which significantly reduces the latency caused by processing length of the input

TABLE VI: The System Time for Methods.

Here $t_l(last)$ is the system's delay for the last point, which is equivalent to the delay time between the last output data point and corresponding input point. The same $t_l(first)$ is for the first output point.

Methods	$t_l(last)$	$t_l(first)$	$t_{c/point}$	Platform
FastICA	1.14s	11.14s	0.46ms	MATLAB
TS-CERUTTI	0.39s	10.39s	0.15ms	MATLAB
TS-EKF	0.32s	10.32s	0.13ms	MATLAB
TS-PCA	0.61s	10.61s	0.24ms	MATLAB
LMS	0.72s	10.72s	0.29ms	MATLAB
RLS	0.70s	10.70s	0.28ms	MATLAB
ESN	0.85s	10.85s	0.34ms	MATLAB
ANFIS	5.02s	15.02s	2.01ms	MATLAB
LSSVM	0.83s	10.83s	0.33ms	MATLAB
SFLSTM	82.67ms	82.67ms	2.67ms	PYTHON

data. Therefore, our proposed method can efficiently connect to the component or template analysis or making inductive inference, which is more suitable for real-time FECG extraction. As shown in Table VI, though the processing speed of our proposed method is slower than the conventional counterparts, it is still much faster than the data capturing, because $t_{c/point}$ of our proposed method is smaller than the sampling interval. Therefore, the processing throughput is sufficient for real-time processing.

C. Evaluation Indexes

In our experiments, since the pure FECG components are unavailable in DaISy, we can't deliver the accurate quantitative measure results on the filtered signals, e.g., the SNR. Instead, we employ the Signal Quality Index (SQI) to quantify the quality of the extracted FECG components. Concretely, we adopt the $kSQI$, who is the fourth moment (kurtosis) of the signal [49], which is defined as follows:

$$kSQI = \frac{\mathbf{E}\{(X - \mu_X)^4\}}{\sigma_X^4}, \quad (21)$$

where X is the signal vector considered as the random variable, μ_X and σ_X are the mean and the standard deviation of X , $\mathbf{E}\{(X - \mu_X)^4\}$ is the expected value of the quantity $(X - \mu_X)^4$. Good quality ECG is considered as highly non-Gaussian since it is not a random signal. Whereas random noises such as muscle artifacts and baseline wander are prone to have normal distributions, and their kurtosis is less than five [49].

Besides, to quantitatively evaluate the performance of our method, the SNR, RMSE, Peak Signal-to-Noise Ratio (PSNR) and SNR based on cross correlation SNR_CC of the output signals are adopted.

The SNR is derived as:

$$SNR = 10 \log \left[\frac{P_{signal}}{P_{noise}} \right], \quad (22)$$

$$P_X = \sum_{t=0}^{L-1} X(t)^2,$$

where P_{signal} and P_{noise} represent the power of the original noise-free FECG components and the residual noise

components in the output, respectively. The residual noise components are obtained by the subtraction between the output and the corresponding original FECG signal from the FECGSYNDB. The definition of the power of signals (noted as X) is employed as the second line, where L is the data length of the signal $X(t)$.

The RMSE is derived as:

$$\text{RMSE} = \sqrt{\frac{1}{L} \sum_{t=0}^{L-1} [X(t) - Y(t)]^2}, \quad (23)$$

where L , X and Y are the data length and the matrices of output and label, respectively.

The PSNR is derived as:

$$\begin{aligned} \text{PSNR} &= 10 \log \left[\frac{\text{MAX}^2}{\text{MSE}} \right], \\ \text{MSE} &= \frac{1}{L} \sum_{t=0}^{L-1} [X(t) - Y(t)]^2, \end{aligned} \quad (24)$$

where L , MAX , X and Y are the length of data, the max value of the label matrix and the matrices of output and label, respectively.

The cross correlation $R_{(S1,S2)}^{CC}$ between the matrices $S1$ and $S2$ is derived as:

$$\begin{aligned} R_{(S1,S2)}^{CC}(k) &= \sum_{i=0}^{L-1} S1(i) * S2(k+i), k \in [0, L-1], \\ S2(L+j) &= S2(j), \end{aligned} \quad (25)$$

where $S1$, $S2$ and L represent the signal sequences and the data length.

The SNR_{CC} is derived as:

$$\begin{aligned} \text{SNR}_{CC} &= 10 \log \left[\frac{P_{CC_{signal}}}{P_{CC_{noise}}} \right], \\ P_{CC_*} &= \sum_{i=0}^{L-1} CC_*(i)^2, \\ CC_{signal}(i) &= R_{(label,label)}^{CC}(i), \\ CC_{noise}(i) &= CC_{signal}(i) - R_{(output,label)}^{CC}(i), \end{aligned} \quad (26)$$

where $P_{CC_{signal}}$ and $P_{CC_{noise}}$ represent the power of the cross correlation sequences. The definition of the power (noted as P_{CC_*}) is employed as the second line. The $CC_{signal}(i)$ represents the signal of cross correlation sequence, which is equal to the autocorrelation coefficient of the original noise-free FECG component label, $CC_{noise}(i)$ represents the noise of cross correlation sequence, which is equal to the subtraction of $CC_{signal}(i)$ and the cross correlation coefficient $R_{(output,label)}^{CC}(i)$ between the output and label. The $R_{(S1,S2)}^{CC}(i)$ can be calculated by (25).

Besides, we use SNR or $k\text{SQI}$ as the monitoring condition in the iterative process. When the hyper-parameters were set as described above, we found that SNR and $k\text{SQI}$ would not increase significantly after ten iterations, which can be seen in Fig. 9. Therefore, the convergence condition is defined as

when the mean of the SNR or $k\text{SQI}$ have a sufficiently small change η :

$$\begin{aligned} |\overline{\text{SNR}_{i \sim (i+3)}} - \overline{\text{SNR}_{(i-3) \sim i}}| &\leq \eta_{\text{SNR}}, i \geq 5 \\ |\overline{k\text{SQI}_{i \sim (i+3)}} - \overline{k\text{SQI}_{(i-3) \sim i}}| &\leq \eta_{k\text{SQI}}, i \geq 5 \end{aligned} \quad (27)$$

$\overline{\text{SNR}_{i \sim (i+3)}}$ or $\overline{k\text{SQI}_{i \sim (i+3)}}$ means the average of SNR or $k\text{SQI}$ of the iterations from i to $(i+3)$. In the figure on the lower left corner, it needs to be explained that the high $k\text{SQI}$ index at the beginning is due to the initialization of the model, which results in small initial coefficients and makes the output signal contain high MECG residual. As can be seen in Fig. 9, the system usually reaches convergence in the first ten to fifteen iterations.

Fig. 9: The Convergence Curves

D. Qualitative Experiment

For qualitative evaluation, we implemented the processes of MECG elimination and FECG enhancement on the ECG records from the DaISy and FECGSYNDB. In the first database of DaISy, the three channels of TEGG signals were employed as one reference signal expressed as $\vec{r}(t)$ as well as two TEGG channels in the database of FECGSYNDB.

The waveform results from the two databases are shown in Fig. 10. The performance of the five AECG channels from DaISy and FECGSYNDB under different case or noise condition can be observed in Fig. 10a and Fig. 10b, respectively. $\hat{F}(t)$ and $Y(t)$ are the FECG signals after MECG elimination and FECG enhancement, respectively. It can be observed from Fig. 10, our proposed method achieves a much better performance with lower residual noise and MECG and less deformation of FECG.

The effect of each method to do the FECG extraction can be represented intuitively in Fig. 10. From the comparison of source signals (MECG or pure FECG) and outputs extracted by methods, it can be seen that our proposed method achieves much lower residual noise and MECG and much less deformation of FECG at the situations of different channels, noise or cases.

In addition, to quantify the quality of the extracted FECG components, the $k\text{SQI}$ of output records from DaISy calculated by (21) is shown in Table VII. Though some methods have the effect in filtering out random noise, but the poor performance in MECG elimination leads to large residual MECG components, as shown in Fig. 10a. Therefore, the methods like

Fig. 10: Experimental results of the FECG morphology extraction.

(a) DaISy: $A(t)$ are sampled from all five AECG channels. All outputs are normalized by the same certain value to ensure $\hat{F}(t)$ is in the range of 0 to 1; (b) FECGSYNDB: $A(t)$ are sampled from 32 AECG channels. All outputs are normalized by the same certain value to ensure the source of FECG signals is in the range of 0 to 1. The waveform of the pure FECG signals is repainted under all outputs extracted by different methods with red color for comparison.

FastICA, TS-CERUTTI, TS-EKF, TS-PCA, and LMS may have a higher value of $kSQI$ than RLS and ESN because of the influence of residual MECG components. Though RLS and ESN can remove MECG components more completely than the other methods mentioned above, they still contain some

random noise, resulting in low $kSQI$, as shown in Table VII. The results $\hat{F}(t)$ and $Y(t)$ filtered and enhanced by our method have much better performance because we take both residual MECG signal and FECG signal quality into consideration, especially in those MECG signal or noise dominated cases.

TABLE VII: The average k SQI of the extracted FECG components from all five channels of DaSy.

Methods	Average of Channel 1-5
FastICA	6.92
TS-CERUTTI	7.56
TS-EKF	10.38
TS-PCA	6.42
LMS	8.36
RLS	4.02
ESN	4.83
Proposed (MECG Elimination)	6.23
FastICA-Fil	7.99
Proposed (FECG Enhancement)	11.37

Both waveform and k SQI show that the enhanced FECG signal extracted by our SFLSTM network contains much cleaner morphological characteristics for medical diagnosis than the conventional methods.

E. Quantitative Experiment

Besides the qualitative evaluation, we also provided quantitative analysis. Only the synthetic ECG recordings in FECGSYNDB were conducted in this experiment, because there is no source signal provided for SNR, RMSE, PSNR, and SNR_CC calculation in DaSy [44], [45].

The mean value of SNR is calculated by (22), which is the average of SNR from the output records of all subjects mentioned in Table VIII. The results in different cases and

levels of additive noise are tabulated in Table VIII. It can be observed that there are a few irregularities in the results generated by the conventional methods because the FECG components are overwhelmed by the noise in some randomly selected channels. In contrast, our method has noticeably higher SNR than the counterparts in both elimination and enhancement processes. The mean and standard deviation of RMSE, PSNR and SNR_CC calculated by (23),(24),(26) are tabulated in Table IX. It can be seen from Table IX, the output of our proposed method achieves much smaller mean of RMSE and higher mean of PSNR SNR_CC, which is an evidence to show the better performance of our method. The Coefficient of Variation (C.V) can be calculated by the ratio of the standard deviation to the mean. In this way, the smaller standard deviation and the smallest C.V obtained through our method show the greater adaptability under different situations.

To minimize the computation cost, we need only to carry out the enhancement when the k SQI after MECG elimination is smaller than 10. Compared with ESN, ANFIS, and LSSVM, our proposed SFLSTM network obtains more than 3.09 dB and 1.81 dB SNR improvement in MECG elimination and FECG enhancement, respectively. The approximate 4.9 dB gain in total will significantly enhance the clinical usefulness of noninvasive FECG monitoring and help the early detection of fetal heart diseases. Our SFLSTM achieves higher means of SNR, lower means of RMSE, and smaller standard deviations, indicating the highly adaptive ability.

TABLE VIII: SNR_{out} of Methods on Records from FECGSYNDB

The records are sampled from the subjects of number 01, 04, 07, 10, the SNR of 00, 03, 06, 09, 12dB, the case of c0 to c5 with noise, and the repetition of number three for our experiment. Each synthetic record consists of 32 channels of AECG and two channels of thoracic MECG. We employ two thoracic MECG channels and randomly pick two channels of AECG from the 32 channels for each record, which can guarantee the completeness of the database.

Methods	SNR	C0	C1	C2	C3	C4	C5	Mean $\pm$ Std.	Methods	C0	C1	C2	C3	C4	C5	Mean $\pm$ Std.
Fast ICA	00	-2.99	-1.90	-2.57	-7.32	0.72	-2.11	-2.70 $\pm$ 2.38	TS-CERUTTI	-1.43	-0.92	-1.66	-6.11	1.98	-1.55	-1.62 $\pm$ 2.37
	03	-1.30	0.30	-3.94	-5.54	-0.77	-2.07	-2.22 $\pm$ 1.97		0.32	2.10	-2.78	-3.98	0.55	-0.27	-0.68 $\pm$ 2.07
	06	0.09	0.61	-0.11	-5.11	1.74	-0.51	-0.55 $\pm$ 2.16		1.04	1.19	2.90	-3.42	1.17	1.51	0.73 $\pm$ 1.96
	09	0.66	1.63	2.18	-2.43	0.72	1.42	0.70 $\pm$ 1.49		2.12	1.97	3.61	-1.70	0.70	1.82	1.42 $\pm$ 1.63
	12	-0.25	0.92	-0.17	-4.20	-0.67	-0.15	-0.76 $\pm$ 1.62		4.05	5.83	1.68	-1.13	-1.57	4.20	2.18 $\pm$ 2.77
TS-PCA	00	-1.36	-0.71	-1.35	-5.60	0.88	-1.41	-1.59 $\pm$ 1.96	TS-EKF	-1.11	-0.65	-1.40	-4.34	-1.73	-1.71	-1.82 $\pm$ 1.19
	03	0.16	1.82	-2.24	-3.31	0.04	0.05	-0.58 $\pm$ 1.70		0.29	1.07	-2.09	-2.61	0.49	-0.01	-0.48 $\pm$ 1.37
	06	0.74	0.80	2.41	-2.20	0.70	1.32	0.63 $\pm$ 1.40		1.26	0.17	1.47	-4.71	0.17	0.64	-0.17 $\pm$ 2.09
	09	1.87	1.77	2.51	-2.31	0.12	1.55	0.92 $\pm$ 1.61		0.67	1.01	1.25	-0.98	-0.09	1.06	0.49 $\pm$ 0.79
	12	2.56	3.56	0.79	-1.57	-1.75	2.55	1.02 $\pm$ 2.07		1.54	1.89	0.79	-0.32	0.74	1.41	1.01 $\pm$ 0.72
LMS	00	-3.19	-1.55	-2.19	-7.35	0.55	-2.88	-2.77 $\pm$ 2.38	RLS	-1.72	-0.25	-0.91	-6.16	0.95	-1.51	-1.60 $\pm$ 2.22
	03	-2.29	-1.12	-4.55	-5.76	-1.16	-2.08	-2.82 $\pm$ 1.74		-0.61	0.52	-2.64	-4.31	-0.65	-0.24	-1.32 $\pm$ 1.64
	06	-1.04	-1.04	-2.77	-7.24	1.02	-2.83	-2.32 $\pm$ 2.55		0.55	0.74	-0.39	-5.13	1.17	-0.19	-0.54 $\pm$ 2.12
	09	1.05	-1.09	-0.69	-5.55	0.06	-1.13	-1.22 $\pm$ 2.07		3.03	1.41	1.71	-3.50	0.41	1.17	0.71 $\pm$ 2.04
	12	-2.10	-1.20	-2.09	-6.00	-2.97	-2.20	-2.76 $\pm$ 1.54		0.44	1.49	0.62	-3.58	-2.55	0.82	-0.46 $\pm$ 1.89
ESN	00	-0.61	0.73	0.18	-4.78	1.67	-0.42	-0.54 $\pm$ 2.04	ANFIS	-1.90	-0.32	-0.88	-5.28	-0.67	-0.88	-1.65 $\pm$ 1.69
	03	0.79	1.54	-1.45	-3.29	0.34	1.01	-0.18 $\pm$ 1.68		-0.44	1.97	-1.62	-4.13	-0.04	-1.52	-0.96 $\pm$ 1.85
	06	1.46	2.36	0.82	-3.20	1.53	1.29	0.71 $\pm$ 1.81		3.55	2.24	3.35	-1.55	0.49	2.72	1.80 $\pm$ 1.80
	09	3.94	2.81	3.00	-1.77	1.30	2.79	2.01 $\pm$ 1.86		2.35	2.70	4.75	-0.01	0.63	3.96	2.40 $\pm$ 1.68
	12	2.44	2.95	1.37	-1.04	-0.98	2.50	1.21 $\pm$ 1.64		6.30	7.51	7.21	2.25	0.74	6.17	5.03 $\pm$ 2.58
LSSVM	00	-0.12	-0.39	-0.38	-3.39	0.46	0.46	-0.74 $\pm$ 1.23	$\hat{F}(t)$	0.03	1.89	0.48	-3.19	5.20	1.01	0.90 $\pm$ 2.49
	03	2.15	1.02	-2.57	-3.35	0.80	-0.27	-0.37 $\pm$ 1.97		3.74	3.81	1.41	-0.27	5.92	3.76	3.06 $\pm$ 1.98
	06	5.57	0.96	1.51	-3.20	1.42	1.26	1.26 $\pm$ 2.54		5.96	7.24	6.13	2.06	8.69	5.49	5.93 $\pm$ 2.03
	09	8.90	4.32	4.41	-0.08	1.60	3.97	3.85 $\pm$ 2.79		8.02	7.10	8.05	3.65	9.79	6.66	7.21 $\pm$ 1.87
	12	8.15	7.43	6.92	3.52	5.08	6.75	6.31 $\pm$ 1.55		8.78	9.37	9.68	5.24	9.68	9.19	8.66 $\pm$ 1.56
Fast ICA-Fil	00	0.63	2.65	1.98	-0.01	5.56	2.38	2.20 $\pm$ 1.78	$Y(t)$	2.84	6.07	5.33	3.61	6.01	5.02	4.81 $\pm$ 1.20
	03	5.18	4.45	2.27	1.68	5.17	3.95	3.79 $\pm$ 1.36		6.67	5.76	5.99	5.46	6.19	6.03	6.02 $\pm$ 0.38
	06	6.32	7.14	4.47	1.48	8.19	4.90	5.42 $\pm$ 2.17		5.93	7.94	6.11	4.94	8.49	5.80	6.54 $\pm$ 1.25
	09	7.66	7.11	7.83	5.20	8.04	6.08	6.99 $\pm$ 1.02		7.78	8.95	9.09	6.67	9.13	8.01	8.27 $\pm$ 0.89
	12	8.65	8.30	8.19	4.94	8.04	8.56	7.78 $\pm$ 1.29		8.48	9.45	10.34	8.34	8.73	9.75	9.18 $\pm$ 0.73

TABLE IX: RMSE PSNR SNR_CC of Methods on Records from FECGSYNDB

Methods	RMSE		PSNR		SNR_CC	
	Mean $\pm$ Std.	C.V	Mean $\pm$ Std.	C.V	Mean $\pm$ Std.	C.V
FastICA	1.99E-01 $\pm$ 2.03E-01	1.02	16.77 $\pm$ 6.27	0.37	-1.71 $\pm$ 8.33	-4.87
TS-CERUTTI	1.74E-01 $\pm$ 1.72E-01	0.99	17.59 $\pm$ 5.84	0.33	-3.22 $\pm$ 7.63	-2.37
TS-PCA	1.71E-01 $\pm$ 1.51E-01	0.89	17.28 $\pm$ 5.19	0.30	-2.96 $\pm$ 6.80	-2.29
TS-EKF	1.56E-01 $\pm$ 1.09E-01	0.70	17.44 $\pm$ 4.39	0.25	-2.08 $\pm$ 4.27	-2.05
LMS	2.10E-01 $\pm$ 1.81E-01	0.86	15.45 $\pm$ 5.12	0.33	-6.95 $\pm$ 6.97	-1.00
RLS	1.74E-01 $\pm$ 1.53E-01	0.88	17.18 $\pm$ 5.20	0.30	-3.01 $\pm$ 7.45	-2.48
ESN	1.40E-01 $\pm$ 1.03E-01	0.74	18.47 $\pm$ 4.35	0.24	-0.27 $\pm$ 6.49	-24.41
ANFIS	1.03E-01 $\pm$ 7.20E-02	0.70	21.57 $\pm$ 5.62	0.26	2.99 $\pm$ 9.19	3.07
LSSVM	1.10E-01 $\pm$ 9.91E-02	0.90	21.45 $\pm$ 5.82	0.27	5.00 $\pm$ 5.90	1.18
$\hat{F}(t)$	1.04E-01 $\pm$ 1.17E-01	1.13	22.72 $\pm$ 6.56	0.29	7.59 $\pm$ 8.74	1.15
FastICA-Fil	9.05E-02 $\pm$ 7.77E-02	0.86	22.94 $\pm$ 5.50	0.24	7.99 $\pm$ 6.77	0.85
$Y(t)$	6.55E-02 $\pm$ 3.15E-02	0.48	24.58 $\pm$ 3.88	0.16	10.25 $\pm$ 6.17	0.60

V. CONCLUSION

In this paper, a two-stage SFLSTM network has been proposed to extract the morphology of FECG. The MECG signals and FECG enhancement are done by the slow LSTM and fast LSTM, respectively. The fast LSTM is designed to reduce the computation cost while keeping the performance of FECG enhancement. Through the experiment results, we can conclude that the proposed method has greatly improved the performance and shown the highly adaptive ability under different noise or case conditions. The extracted clean FECG components could efficiently support the FECG morphology-based medical diagnosis.

REFERENCES

- [1] R. G. Sæderup, H. Zimmermann, D. H. Eiríksdóttir, J. Hansen, J. J. Struijk, and S. Schmidt, "Comparison of cardiocography and fetal heart rate estimators based on non-invasive fetal ecg," in *2019 Computing in Cardiology (CinC)*. IEEE, 2019, pp. Page–1.
- [2] C. Lin, C. Yeh, C. Wang, W. Shi, B. M. F. Serafico, C. Wang, C. Juan, H. Vincent Young, Y. Lin, H. Yeh, and M. Lo, "Robust fetal heart beat detection via r-peak intervals distribution," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 12, pp. 3310–3319, 2019.
- [3] F. Andreotti, F. Gräßer, H. Malberg, and S. Zaunseder, "Non-invasive fetal ecg signal quality assessment for multichannel heart rate estimation," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 12, pp. 2793–2802, 2017.
- [4] R. Kahankova, R. Martinek, R. Jaros, K. Behbehani, A. Matonia, M. Jezewski, and J. A. Behar, "A review of signal processing techniques for non-invasive fetal electrocardiography," *IEEE reviews in biomedical engineering*, 2019.
- [5] S. M. Martens, C. Rabotti, M. Mischi, and R. J. Sluijter, "A robust fetal ECG edtection method for abdominal recordings," *Physiological measurement*, vol. 28, no. 4, p. 373, 2007.
- [6] S. Cerutti, G. Baselli, S. Cividari, E. Ferrazzi, A. M. Marconi, M. Pagani, and G. Pardi, "Variability analysis of fetal heart rate signals as obtained from abdominal electrocardiographic recordings," *Journal of Perinatal Medicine-Official Journal of the WAPM*, vol. 14, no. 6, pp. 445–452, 1986.
- [7] L. De Lathauwer, B. De Moor, and J. Vandewalle, "Fetal electrocardiogram extraction by blind source subspace separation," *IEEE transactions on biomedical engineering*, vol. 47, no. 5, pp. 567–572, 2000.
- [8] B. Widrow, J. R. Glover, J. M. McCool, J. Kaunitz, C. S. Williams, R. H. Hearn, J. R. Zeidler, J. E. Dong, and R. C. Goodlin, "Adaptive noise cancelling: Principles and applications," *Proceedings of the IEEE*, vol. 63, no. 12, pp. 1692–1716, 1975.
- [9] F. Andreotti, M. Riedl, T. Himmelsbach, D. Wedekind, N. Wessel, H. Stepan, C. Schmieder, A. Jank, H. Malberg, and S. Zaunseder, "Robust fetal ecg extraction and detection from abdominal leads," *Physiological measurement*, vol. 35, no. 8, p. 1551, 2014.
- [10] R. Sameni, "Extraction of fetal cardiac signals from an array of maternal abdominal recordings," Ph.D. dissertation, Institut National Polytechnique de Grenoble-INPG; Sharif University of Technology (SUT), 2008.
- [11] S. Murawwat, A. Batool, A. Ahmed, A. Ansar, and A. Iqbal, "Fecg extraction using adaptive lms," in *2018 International Conference on Computing, Electronic and Electrical Engineering (ICE Cube)*. IEEE, 2018, pp. 1–6.
- [12] J. Behar, A. Johnson, G. D. Clifford, and J. Oster, "A comparison of single channel fetal ecg extraction methods," *Annals of biomedical engineering*, vol. 42, no. 6, pp. 1340–1353, 2014.
- [13] K. Assaleh, "Extraction of fetal electrocardiogram using adaptive neuro-fuzzy inference systems," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 1, pp. 59–68, 2006.
- [14] F. Mochimaru, Y. Fujimoto, and Y. Ishikawa, "Detecting the fetal electrocardiogram by wavelet theory-based methods," *Progress in Biomedical Research*, vol. 7, pp. 185–193, 2002.
- [15] G. Camps-Valls, M. Martinez-Sober, E. Soria-Olivas, R. Magdalena-Benedito, J. Calpe-Maravilla, and J. Guerrero-Martinez, "Foetal ECG recovery using dynamic neural networks," *Artificial Intelligence in Medicine*, vol. 31, no. 3, pp. 197–209, 2004.
- [16] K. Assaleh and H. Al-Nashash, "A novel technique for the extraction of fetal ecg using polynomial networks," *IEEE Transactions on Biomedical Engineering*, vol. 52, no. 6, pp. 1148–1152, 2005.
- [17] G. M. Ungureanu, J. W. Bergmans, S. G. Oei, A. Ungureanu, and W. Wolf, "The event synchronous canceller algorithm removes maternal ecg from abdominal signals without affecting the fetal ecg," *Computers in biology and medicine*, vol. 39, no. 6, pp. 562–567, 2009.
- [18] L. Han, X.-j. Pu, and X.-j. Chen, "Method of fetal electrocardiogram extraction based on ν -support vector regression," *IET Signal Processing*, vol. 9, no. 5, pp. 430–439, 2015.
- [19] X.-J. Pu, L. Han, Q. Liu, and A.-Y. Jiang, "Fecg extraction based on least square support vector machine combined with fastica," *J. Inf. Sci. Eng.*, vol. 33, no. 6, pp. 1595–1609, 2017.
- [20] J. Behar, J. Oster, and G. D. Clifford, "Combining and benchmarking methods of foetal ecg extraction without maternal or scalp electrode data," *Physiological measurement*, vol. 35, no. 8, p. 1569, 2014.
- [21] R. Martinek, R. Kahankova, J. Jezewski, R. Jaros, J. Mohylova, M. Fajkus, J. Nedoma, P. Janku, and H. Nazeran, "Comparative effectiveness of ica and pca in extraction of fetal ecg from abdominal signals: Toward non-invasive fetal monitoring," *Frontiers in physiology*, vol. 9, p. 648, 2018.
- [22] R. Sameni, C. Jutten, and M. B. Shamsollahi, "Multichannel electrocardiogram decomposition using periodic component analysis," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 8, pp. 1935–1940, 2008.
- [23] S. Nikam and S. Deosarkar, "Fast ica based technique for non-invasive fetal ecg extraction," in *2016 Conference on Advances in Signal Processing (CASP)*. IEEE, 2016, pp. 60–65.
- [24] R. Vullings, B. De Vries, and J. W. Bergmans, "An adaptive kalman filter for ecg signal enhancement," *IEEE transactions on biomedical engineering*, vol. 58, no. 4, pp. 1094–1103, 2010.
- [25] G. D. Clifford, I. Silva, J. Behar, and G. B. Moody, "Non-invasive fetal ecg analysis," *Physiological measurement*, vol. 35, no. 8, p. 1521, 2014.
- [26] M. Ungureanu and W. Wolf, "Basic aspects concerning the event-synchronous interference canceller," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 11, pp. 2240–2247, 2006.
- [27] K. Assaleh, "Extraction of fetal electrocardiogram using adaptive neuro-fuzzy inference systems," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 1, pp. 59–68, 2007.

- [28] J. Behar, A. Johnson, J. Oster, and G. Clifford, "An echo state neural network for foetal ecg extraction optimised by random search," in *Machine Learning for Clinical Data Analysis and Healthcare NIPS Workshop*, 2013.
- [29] E. R. Ferrara and B. Widrow, "Fetal electrocardiogram enhancement by time-sequenced adaptive filtering," *IEEE Transactions on Biomedical Engineering*, no. 6, pp. 458–460, 1982.
- [30] A. Gretton, O. Bousquet, A. Smola, and B. Schölkopf, "Measuring statistical dependence with hilbert-schmidt norms," in *International conference on algorithmic learning theory*. Springer, 2005, pp. 63–77.
- [31] Z. Ma, Z. Zhan, X. Ouyang, and X. Su, "Nonlinear dimensionality reduction based on hsc maximization," *IEEE Access*, vol. 6, pp. 55 537–55 555, 2018.
- [32] Y. Ma, Y. Xiao, G. Wei, and J. Sun, "Foetal ecg extraction using non-linear adaptive noise canceller with multiple primary channels," *IET Signal Processing*, vol. 12, no. 2, pp. 219–227, 2017.
- [33] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [34] Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," *IEEE transactions on neural networks*, vol. 5, no. 2, pp. 157–166, 1994.
- [35] Y. Qiu, K. Huang, F. Xiao, and H. Shen, "A segment-wise reconstruction method based on bidirectional long short term memory for power line interference suppression," *Biocybernetics and Biomedical Engineering*, vol. 38, no. 2, pp. 217–227, 2018.
- [36] A. Graves, A.-r. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," in *Acoustics, speech and signal processing (icassp), 2013 IEEE International Conference on*. IEEE, 2013, pp. 6645–6649.
- [37] R. Jozefowicz, W. Zaremba, and I. Sutskever, "An empirical exploration of recurrent network architectures," in *Proceedings of the 32nd International Conference on Machine Learning*, vol. 37. Lille, France: PMLR, 07–09 Jul 2015, pp. 2342–2350.
- [38] L. Y. Taha and E. Abdel-Raheem, "Extraction of fetal electrocardiogram signals using blind source extraction based parallel linear predictor filter," in *2018 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*. IEEE, 2018, pp. 569–572.
- [39] M. C. Mozer, "A focused backpropagation algorithm for temporal pattern recognition," *Complex Systems*, vol. 3, no. 4, pp. 349–381, 1989.
- [40] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *the 3rd International Conference for Learning Representations*, San Diego, 2015.
- [41] L. Bottou, "Large-scale machine learning with stochastic gradient descent," in *Proceedings of COMPSTAT'2010*. Springer, 2010, pp. 177–186.
- [42] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, "TensorFlow: Large-scale machine learning on heterogeneous systems," 2015. [Online]. Available: <http://tensorflow.org/>
- [43] B. D. Moor, P. D. Gersem, B. D. Schutter, and W. Favoreel, "Daisy: A database for identification of systems," *Journal A*, vol. 38, pp. 4–5, 1997.
- [44] J. Behar, F. Andreotti, S. Zaunseder, Q. Li, J. Oster, and G. D. Clifford, "An ecg simulator for generating maternal-foetal activity mixtures on abdominal ecg recordings," *Physiological measurement*, vol. 35, no. 8, p. 1537, 2014.
- [45] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.
- [46] F. Andreotti, J. Behar, S. Zaunseder, J. Oster, and G. D. Clifford, "An open-source framework for stress-testing non-invasive foetal ecg extraction algorithms," *Physiological measurement*, vol. 37, no. 5, p. 627, 2016.
- [47] R. Martinek, H. Skutova, R. Kahankova, P. Koudelka, P. Bilik, and J. Koziorek, "Fetal ecg extraction based on adaptive neuro-fuzzy interference system," in *2016 10th International Symposium on Communication Systems, Networks and Digital Signal Processing (CSNDSP)*. IEEE, 2016, pp. 1–6.
- [48] X. Pu, X. Zeng, L. Han, and J. Chen, "Extraction of fetal electrocardiogram signal using least squares support vector machines," *J Electron Inf Technol*, vol. 31, no. 12, pp. 2941–2947, 2009.
- [49] G. D. Clifford, F. Azuaje, and P. McSharry, "Advanced methods and tools for ecg data analysis," 2006.

Ziqun Zhou (S'19) received the B.S. degree in Information and Communication Engineering from Harbin Engineering University, Harbin, China, in 2018. Now, he is pursuing the Ph.D. degree in Information Science & Electronic Engineering at Zhejiang University. His research interests include deep learning and medical biological information processing.

Kejie Huang (M'13-SM'18) received his Ph.D degree from the Department of Electrical Engineering, National University of Singapore (NUS), Singapore, in 2014. He has been a principal investigator at College of Information Science Electronic Engineering, Zhejiang University (ZJU) since 2016. Prior to joining ZJU, he spent five years in the IC design industry including Samsung and Xilinx, two years in the Data Storage Institute, Agency for Science Technology and Research (A*STAR), and another three years in Singapore University of Technology and Design (SUTD), Singapore. He has authored or coauthored more than 30 scientific papers in international peer-reviewed journals and conference proceedings. He holds four granted international patents, and another eight pending ones. His research interests include low power circuits and systems design using emerging non-volatile memories, architecture and circuit optimization for reconfigurable computing systems and neuromorphic systems, machine learning, and deep learning chip design. He currently serves as the Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS-PART II: EXPRESS BRIEFS.

Yue Qiu received the B.S. and Ph.D degree in Information Science & Electronic Engineering from Zhejiang University, Zhejiang, China, in 2013 and 2018, respectively. His main field of interest lies in deep learning and medical biological information processing.

Haibin Shen is currently a Professor with Zhejiang University, a member of the second level of 151 talents project of Zhejiang Province, and a member of the Key Team of Zhejiang Science and Technology Innovation. His research interests include learning algorithm, processor architecture, and modeling. His research achievement has been used by many authority organizations. He has published more than 100 papers on academic journals, and he has been granted more than 30 patents of invention. He was a recipient of the First Prize of Electronic Information Science and Technology Award from the Chinese Institute of Electronics, and has won a second prize at the provincial level.