2022 44th Annual International Conference of
the IEEE Engineering in Medicine & Biology Society (EMBC)
Scottish Event Campus, Glasgow, UK, July 11-15, 2022

2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) | 978-1-7281-2782-8/22/$31.00 ©2022 IEEE | DOI: 10.1109/EMBC48229.2022.9871826

Fetal Movement Cancellation in Abdominal Electrocardiogram Recordings Using Signal-to-Signal Translation*

Arash Shokouhmand, *Student Member, IEEE*, Negar Tavassolian, *Senior Member, IEEE*

Abstract— **This study addresses the cancellation of fetal movement in abdominal electrocardiogram (AECG) recordings through deep neural networks. For this purpose, a generative signal-to-signal translation model consisting of two coupled generators is employed to discover the relations between fetal movement-contaminated and clean AECG recordings. The model is trained on the fetal ECG synthetic database (FECGSYNDB) which provides AECG recordings from 10 pregnancies along with their ground-truth maternal and fetal ECG signals. The signals are initially segmented into 4-second segments and then fed into the network for denoising. It is demonstrated that the signal-to-signal translation method can reconstruct clean AECG signals with average mean-absolute-error (MAE), root-mean-square deviation (RMSD), and Pearson correlation coefficient (PCC) of 0.099, 0.124, and 99.12% respectively, between clean and denoised AECG signals. Furthermore, the robustness of the method to low signal-to-noise ratio (SNR) input values is shown by an RMSD range of (0.047, 0.352) for SNR values within the range of (-3, 3) dB.**

Clinical Relevance— **The proposed framework allows for the denoising of abdominal ECG signals for non-invasive fetal heart rate monitoring. The approach is accurate due to the use of advanced neural network techniques.**

I. Introduction

As reported by the World Health Organization (WHO), 2.6 million stillbirths occur globally every year which could be avoided by timely access to emergency obstetric care [1]. Proactive monitoring of fetal health conditions allows expectant mothers to be aware of the wellbeing of their babies. Fetal heart rate (FHR) represents an important vital sign to be monitored during intermittent clinical visits in the third trimester. Abnormal FHR patterns tend to precede adverse prenatal outcomes such as premature birth, hypoxia, or intrauterine growth retardation which may arise in between clinical visits [2]. Continuous monitoring of FHR allows for delivering timely medical interventions to reduce the rate of fetal mortality [2].

In recent years, a variety of technologies have been proposed for FHR monitoring. These include electrocardiogram (ECG) [3], phonocardiogram (PCG) [4], seismo-cardiogram (SCG) [5], and gyro-cardiogram (GCG) [6]. ECG and PCG measure FHR through electrical activity of cardiac muscle and heartbeat sounds respectively, whereas SCG and GCG respectively collect linear and rotational abdominal vibrations. Our observation in [6] has demonstrated that SCG and GCG modalities provide more accurate estimations of FHR when fused with ECG recordings. Fetal ECG (FECG) monitoring could be conducted both invasively and non-invasively [7]. The invasive method aims to acquire FECG data through an electrode attached to the scalp of the fetus during labor [7]. However, this procedure is limited to the period of labor and may induce infections and bruising of the baby's scalp [8]. As an alternative to fetal scalp monitoring, non-invasive fetal ECG (NIFECG) has been investigated recently, showing promise for FHR monitoring [9]. For this purpose, ECG electrodes are placed on the mother's abdomen for data acquisition [10]. However, there are still unexplored challenges complicating FHR extraction through NIFECG, an example of which is the impact of fetal motion on abdominal ECG recordings [11]. This undesired physiological event poses difficulties for estimating fetal ECG components which hold the same level of energy as motion components, as demonstrated in [11]. Hence, the detection and mitigation of fetal motion artifacts in abdominal recordings would allow for more robust non-invasive FHR estimation.

Conventional methods of non-invasive fetal ECG extraction primarily rely on three stages, namely signal pre-processing, maternal ECG estimation and subtraction, and fetal ECG extraction. The main focus of these methods is on the last two stages [9], whereas other physiological events such as fetal movement could account for inaccurate FHR estimation [11]. Furthermore, these methods are primarily based on deterministic approaches such as independent component analysis (ICA) and adaptive filters, which are less able to extract fetal movement patterns compared to deep neural networks [12].

In this work, a signal-to-signal translation framework for fetal movement cancellation is developed where the noisy abdominal ECG and its corresponding clean counterpart are employed as inputs to a one-to-one mapping neural network. This network is employed to discover cross-domain relations between fetal movement-contaminated and clean abdominal recordings, reinforcing the appearance of clean abdominal ECG signals. The proposed framework leverages *generative adversarial networks (GANs)* which are widely applicable in computer vision, and adapts them to 1-dimensional signal processing. *To the best of our knowledge, this is the first study addressing the cancellation of fetal movement in abdominal ECG recordings through deep neural networks.* The organization of the paper is as follows: In Section II, the pre-processing, the methodology for fetal movement cancellation, and the training procedure of the signal-to-signal translation network are explained. The evaluation metrics and

*This work was supported by the National Science Foundation (NSF) under award number 1855394.

A. Shokouhmand and N. Tavassolian are with the Department of Electrical and Computer Engineering at Stevens Institute of Technology, Hoboken, NJ 07030, USA. (e-mail: ashokouh@stevens.edu).

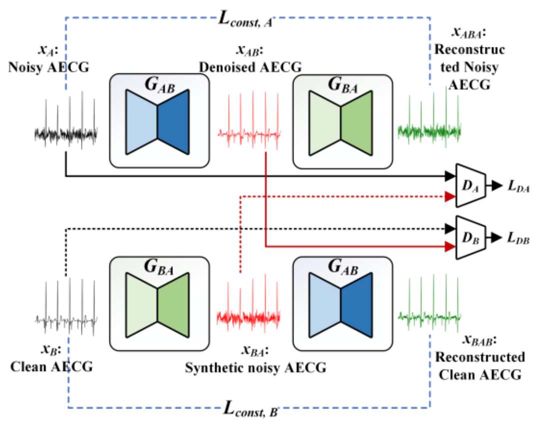

Fig. 1. The generative model for discovering cross-domain relations, including two generators (***G_{AB}*** and ***G_{BA}*** for noisy AECG-to-denoised AECG and clean AECG-to-noisy AECG translations respectively) and two discriminators (***D_A*** and ***D_B*** for scoring synthetic signals generated by ***G_{BA}*** and ***G_{AB}*** respectively).

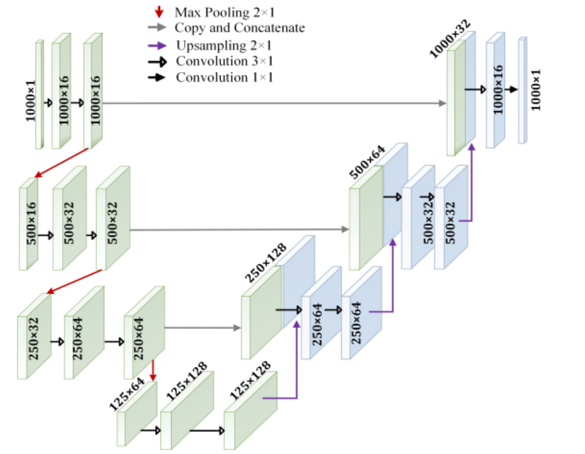

Fig. 2. The architecture of the generators adopted for one-to-one mapping in signal-to-signal translation network. This architecture is a 1-D U-Net with three levels of max-pooling and upsampling.

experimental results are discussed in Section III, and the paper is concluded in Section IV.

II. Methodology

In the following sub-sections, the dataset, signal-to-signal translation network, and the training procedure are described in detail.

A. Dataset

In this work, we use the fetal ECG synthetic database (FECGGSYNDB) which is publicly available on PhysioNet [13]. This dataset includes 32-channel artificial non-invasive abdominal ECG (AECG) recordings of 10 pregnancies, amounting to 145.8 hours of data. A distinctive characteristic of FECGSYNDB is the inclusion of the ground-truth signals corresponding to fetal ECG (FECG), maternal ECG (MECG), and non-stationary physiological events such as fetal movement and uterine contractions. Motion components provided by FECGSYNDB allow for simulating the non-stationary dynamics of fetal movement in realistic pregnancies. Each recording consists of 5 minutes of data sampled at 250 Hz. In this study, we used the data of 4 channels (27, 28, 29, and 30) which are located below the umbilicus according to [11]. In order to train our model, we need a clean AECG dataset and its corresponding fetal movement-contaminated version. To generate the clean AECG dataset, we combine only maternal ECG and fetal ECG signals. For the noisy version however, maternal ECG, fetal ECG, and fetal movement components are incorporated.

B. Data Pre-processing

As we employ a deep neural network for motion cancellation in an end-to-end manner, signal preparation merely consists of signal segmentation. As such, each channel of abdominal ECG recordings is segmented into 4-second overlapping windows with 90% overlap between consecutive segments. Hence, each segment represents 1,000 samples (4 seconds at 250 Hz) of data for fetal motion cancellation. These segments are then normalized into a standard distribution (zero-mean and unit variance) to ensure stable convergence when training the network. It is to be noted that signal preparation is applied to both fetal contaminated-AECG and the corresponding clean ground-truth components for supervised training.

C. Signal-to-Signal Translation Network

Fig. 1 depicts our developed 1-dimensional (1-D) generative model for fetal movement cancellation. The idea of this network was adopted from DiscoGAN proposed for image cross-domain translations [14], where a synthetic image is generated in one domain given attributes of an image in the other domain. We implemented a modified version of DiscoGAN for 1-D signals to translate noisy abdominal ECG (x_A) to its clean counterpart (x_{BA}) through generator ***G_{AB}***, and similarly a clean AECG (x_B) to its motion-contaminated counterpart (x_{AB}) using generator ***G_{BA}***. The architecture of the generator networks is illustrated in Fig. 2. This network is a three-level maxpooling/upsampling U-Net with rectified linear unit (ReLU) as the activation function, consisting of an encoder to transform the input signals to an embedding space and a decoder to reconstruct the desired output from the embedded features.

In order to have a one-to-one correspondence between the two domains, we constrain the relation by assuming ***G_{AB}*** as the inverse mapping of ***G_{BA}***. As such, we expect ***G_{BA}*** to be able to reconstruct x_A from x_{AB}, which is called x_{ABA}. The same assumption stands true for x_B. As such,

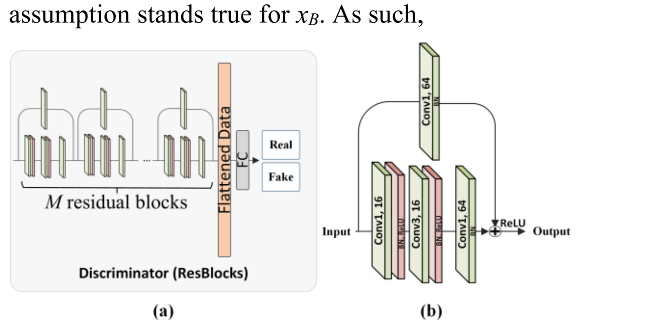

Fig. 3. (a) The residual discriminator network for scoring synthetic samples. (b) the residual block including three convolutional layers and a convolutional layer on the skip connection. ReLU and batch normalization (BN) are used to add non-linearity and stable convergence.

$$x_{ABA} = \boldsymbol{G_{BA}}(\boldsymbol{G_{AB}}(x_A)),$$
$$x_{BAB} = \boldsymbol{G_{AB}}(\boldsymbol{G_{BA}}(x_B)), \quad (1)$$

which implies the two generative networks, $\boldsymbol{G_{AB}}$ and $\boldsymbol{G_{BA}}$, are coupled to ensure our generative functions can map each domain to its counterpart. These two constraints can be related to the objective function using the L_1 distance between the ground-truth signal and its reconstructed version. Therefore, our generative objective function to be minimized during the training is:

$$L_{G,const} = L_{const,A} + L_{const,B} = \|x_{ABA} - x_A\| + \|x_{BAB} - x_B\|, \quad (2)$$

where $L_{G,const}$, $L_{const,A}$, and $L_{const,B}$ represent the total reconstruction loss of the generator, reconstruction loss of x_A, and reconstruction loss of x_B respectively. To guarantee inter-domain mapping, each reconstruction channel shown in Fig. 1 is scored with a discriminator, the structure of which is depicted in Fig. 3 (a). The discriminator consists of 5 residual layers (shown in Fig. 3 (b)) followed by a linear layer using the sigmoid function to generate the score within the range of (0,1). As such, the generated signals x_{AB} and x_{BA} are input to the discriminators of the clean AECG domain ($\boldsymbol{D_B}$) and the noisy ECG domain ($\boldsymbol{D_A}$) respectively, on an adversarial basis. This creates an adversarial loss for the generator networks as below:

$$L_{G,adv} = L_{adv,A} + L_{adv,B} = -\mathbb{E}_{x_A\sim P_A}\left[\log \boldsymbol{D_B}\left(\boldsymbol{G_{AB}}(x_A)\right)\right] -\mathbb{E}_{x_B\sim P_B}[\log \boldsymbol{D_A}\left(\boldsymbol{G_{BA}}(x_B)\right)], \quad (3)$$

where $L_{G,adv}$ denotes the adversarial loss which encourages the generators to learn the distribution of data. As a result, the total loss for the generator is the weighted sum of the reconstruction loss and adversarial loss:

$$L_G = \lambda L_{G,const} + (1-\lambda)L_{G,adv}. \quad (4)$$

where λ was set to 0.35 based on our observation.

The adversarial training requires discriminators to distinguish between real and fake distributions. As such, $\boldsymbol{D_A}$ and $\boldsymbol{D_B}$ attempt to learn the distribution of real data more accurately to avoid classifying synthetic data as real, or real data as fake. Hence, discrimination loss for $\boldsymbol{D_A}$ and $\boldsymbol{D_B}$ are defined as

$$L_{D_B} = -\mathbb{E}_{x_B\sim P_B}[\log \boldsymbol{D_B}\,(x_B)] -\mathbb{E}_{x_A\sim P_A}\left[\log(1-\boldsymbol{D_B}\left(\boldsymbol{G_{AB}}(x_A)\right))\right], \quad (5)$$

and

$$L_{D_A} = -\mathbb{E}_{x_A\sim P_A}[\log \boldsymbol{D_A}\,(x_A)] -\mathbb{E}_{x_B\sim P_B}\left[\log(1-\boldsymbol{D_A}\left(\boldsymbol{G_{BA}}(x_B)\right))\right], \quad (6)$$

respectively. The sum of the terms in (4) and (5) builds the total adversarial loss of discriminators (L_D) as follows:

$$L_D = L_{D_A} + L_{D_B}. \quad (7)$$

The training procedure continues until $\boldsymbol{G_{AB}}$ learns the mapping function from domain $\boldsymbol{A}$ to domain $\boldsymbol{B}$.

D. Training Procedure

The signal-to-signal translation network represents a supervised learning procedure. It thus requires access to raw abdominal ECG recordings and the corresponding clean signals. The training procedure is conducted based on subject-held-out cross-validation. As such, the data of 9 subjects are used for training and the performance is evaluated on the held-out subject. We train our model for 80 epochs on abdominal ECG segments. The Adam optimization algorithm [15] is used with an initial learning rate of 0.001, which is reduced by a factor of 0.98 for every five epochs without performance improvement. The model is implemented on an NVIDIA GeForce RTX 2070 with a batch size of 8.

III. Experimental Results and Discussion

In this section, the evaluation metrics are introduced, based on which the experimental results are presented.

A. Evaluation Metrics

In this part, we report the performance of the proposed signal-to-signal translation framework using root-mean-square-deviation (RMSD), mean-absolute-error (MAE), and Pearson correlation coefficient (PCC) between the denoised AECG and the input raw AECG signals. Furthermore, RMSD is used to evaluate the robustness to input signal-to-noise ratio (SNR) values. In this work, SNR is defined as the ratio of the power of fetal ECG components to that of fetal movement artifacts. For this, we re-scaled the constituent parts of the abdominal ECG signals to simulate AECG signals at SNR values within the range of (-3, 3) dB.

B. Performance Evaluation & Discussion

To evaluate the performance of the proposed method, signal segments of the held-out subject were initially normalized and then fed into the network. Fig. 4 shows an example of a denoised signal with the input SNR of 0 dB. As can be seen in Fig. 4 (b), fetal QRS complexes are visible after the cancellation of fetal movement in the noisy AECG shown in Fig. 4 (a). Comparing Fig. 4 (b) and (c), it is visually demonstrated that the network could mitigate the impact of noise on the AECG signal. Table I summarizes the performance of our signal-to-signal translation framework in terms of MAE, RMSD, and PCC ($\boldsymbol{\rho_{B,AB}}$(%)) at an input SNR of 0 dB. According to the table, MAE values vary within the

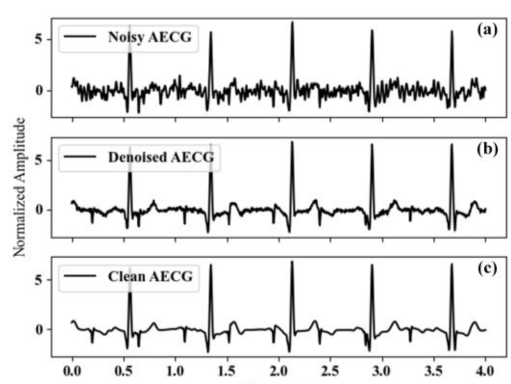

Fig. 4. (a) Input noisy abdominal ECG. (b) Denoised abdominal ECG. (c) Clean abdominal ECG.

TABLE I. PERFORMANCE OF THE DENOISING METHOD FOR SUBJECT-HELD-OUT CROSS-VALIDATION FOR SNR= 0 dB

Subject	MAE	RMSD	$\rho_{B,AB}$(%)
1	0.083	0.104	99.47
2	0.122	0.152	98.86
3	0.079	0.100	99.51
4	0.104	0.129	99.17
5	0.040	0.053	99.87
6	0.181	0.226	97.56
7	0.142	0.179	98.42
8	0.063	0.080	99.67
9	0.102	0.129	99.18
10	0.077	0.097	99.52
Mean±std.	**0.099±0.038**	**0.124±0.047**	**99.12±0.655**

range of 0.040-0.181 with an average (±standard deviation) value of 0.099 (±0.038). The MAE linearly scores the denoised abdominal ECG signal, whereas RMSD aims at considering larger weights for larger error values. As mentioned in Table I, RMSD values vary within the range of 0.053-0.226. The RMSD value for subject 6 suggests a relatively higher estimation error compared to other subjects. However, the average RMSD (±standard deviation) value was reported by 0.124 (±0.047), implying high morphological similarity between the ground-truth clean abdominal ECG signal (x_B) and its denoised counterpart (x_{AB}). As described in Table I, PCC values are reported in percentage. Subject 5 with 99.87% and subject 6 with 97.56% suggest the highest and lowest PCC values. The average PCC (±standard deviation) suggests 99.12% (0.655%) morphological consistency between ground-truth and estimated abdominal ECG segments.

To evaluate the robustness of the method against various input SNR values, the signal-to-signal translation network was trained for input SNR values within the range of (-3, 3) dB (-3 dB SNR implies the energy level of fetal movement components is double that of fetal ECG components). Fig. 5 illustrates average RMSD values over all ten subjects in terms of input SNR. The lowest RMSD was = 0.047 for SNR= 3 dB, whereas the largest error (0.352) was suggested by SNR= -3 dB. Small variations of RMSD in terms of SNR demonstrate the robustness of the signal-to-signal translation approach.

IV. CONCLUSION

This paper reports on the development of a signal-to-signal translation technique for the cancellation of fetal movement in abdominal ECG recordings. The proposed method performs based on a one-to-one mapping between two domains corresponding to noisy abdominal ECG signals and their clean counterparts. For this purpose, a generative adversarial network (GAN) is leveraged to ensure discovering cross-domain relations between the two domains. Two generators are trained, one for mapping raw abdominal ECG signals to their clean counterparts, and the other for translating clean abdominal ECG signals to their noisy versions. The model achieved average MAE, RMSD, and PCC values of 0.099, 0.124, and 99.12% respectively on the PhysioNet fetal ECG synthetic database. Furthermore, the

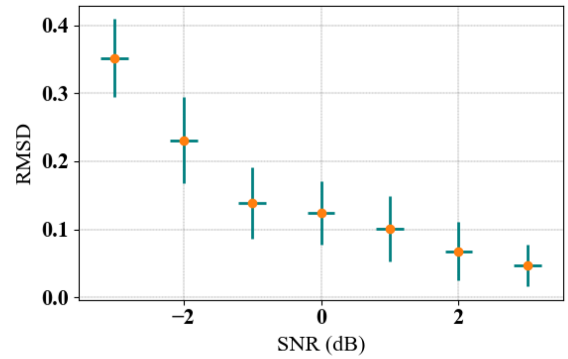

Fig. 5. Root-mean-square deviation (RMSD) of the estimated signal in terms of input signal-to-noise ratio (SNR).

robustness of the method is evaluated in terms of input SNR values, where the RMSD value varies within the range of (0.047, 0.352). This indicates high correlations between ground-truth clean signals and their denoised counterparts.

Future work will investigate cross-domain relations between two different modalities for fetal heart rate estimation. The information about inter-modality relations would allow the effective fusion of ECG with PCG and SCG data for fetal heart rate and movement monitoring respectively.

REFERENCES

[1] W. H. O. (WHO), "Stillbirth," 2018. [Online]. Available: https://www.who.int/health-topics/stillbirth.

[2] R. K. Freeman, T. J. Garite, M. P. Nageotte, and L. A. Miller, *Fetal heart rate monitoring*. Lippincott Williams & Wilkins, 2012.

[3] J. Behar, J. Oster, and G. D. Clifford, "Combining and benchmarking methods of foetal ECG extraction without maternal or scalp electrode data," *Physiol. Meas.*, vol. 35, no. 8, p. 1569, 2014.

[4] F. Kovács, C. Horváth, Á. T. Balogh, and G. Hosszú, "Fetal phonocardiography—past and future possibilities," *Comput. Methods Programs Biomed.*, vol. 104, no. 1, pp. 19–25, 2011.

[5] C. Yang, C. Antoine, B. K. Young, and N. Tavassolian, "A Pilot Study on Fetal Heart Rate Extraction from Wearable Abdominal Inertial Sensors," *IEEE Sens. J.*, vol. 19, no. 22, pp. 10773–10781, 2019.

[6] A. Shokouhmand, C. Antoine, B. K. Young, and N. Tavassolian, "Multi-modal Framework for Fetal Heart Rate Estimation: Fusion of Low-SNR ECG and Inertial Sensors," in *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 2021, pp. 1–4.

[7] R. Sameni and G. D. Clifford, "A review of fetal ECG signal processing; issues and promising directions," *Open Pacing. Electrophysiol. Ther. J.*, vol. 3, p. 4, 2010.

[8] T. Kawakita, U. M. Reddy, H. J. Landy, S. N. Iqbal, C. Huang, and K. L. Grantz, "Neonatal complications associated with use of fetal scalp electrode: a retrospective study," *BJOG An Int. J. Obstet. Gynaecol.*, vol. 123, no. 11, pp. 1797–1803, 2016.

[9] R. Kahankova *et al.*, "A review of signal processing techniques for non-invasive fetal electrocardiography," *IEEE Rev. Biomed. Eng.*, vol. 13, pp. 51–73, 2019.

[10] J. Behar, F. Andreotti, S. Zaunseder, J. Oster, and G. D. Clifford, "A practical guide to non-invasive foetal electrocardiogram extraction and analysis," *Physiol. Meas.*, vol. 37, no. 5, p. R1, 2016.

[11] 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," *Physiol. Meas.*, vol. 37, no. 5, p. 627, 2016.

[12] K. P. Murphy, *Machine learning: a probabilistic perspective*. MIT press, 2012.

[13] A. L. Goldberger *et al.*, "PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[14] T. Kim, M. Cha, H. Kim, J. K. Lee, and J. Kim, "Learning to discover cross-domain relations with generative adversarial networks," in *International Conference on Machine Learning*, 2017, pp. 1857–1865.

[15] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv Prepr. arXiv1412.6980*, 2014.