KalmanNet For ECG Analysis
Fetal Heart Rate Estimation

Mehdi Bakka   Guy Revach

Signal and Information Processing Laboratory (ISI)
ETHZ

November,2021

Electrocardiogram (ECG)

An Electrocardiogram (ECG) is a simple test that can be used to check your heart's rhythm and electrical activity.

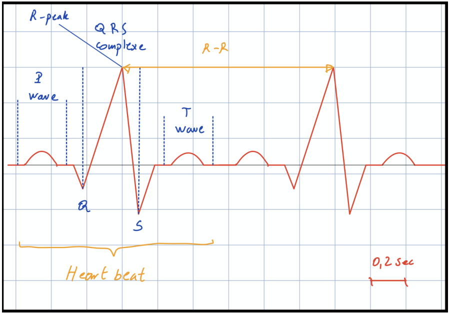


Figure 1: Illustration of an Electrocardiogram

Heart Rate (HR) :

$$HR = \frac{300}{\text{\# of R-R squares}} = \frac{300}{6} = 50bpm$$

Agenda

- Fetal Heart Rate (HR) Estimation - Problem Formulation
- List of Related Approaches
  - An Adaptive Kalman Filter for ECG Signal Enhancement 2011, by Rik Vullings, Bert de Vries, and Jan W. M. Bergmans.
  - Warmerdam et al (2018).
  - A dilated inception CNN-LSTM network for fetal heart rate estimation, by E.Fotiadou et al 2021.
  - Unsupervised Feature Extraction, Signal Labeling, and Blind Signal Separation in a State Space World, by Nour Zalmai et al 2017.

Motivation

Fetal Electrocardiogram (FECG) is essential to track the well being of a fetus. Thus the extensive research and high interest accounted to the FECG estimation. The invasive recording methods are quite precise, but they are only possible after the labor and may be harmful to the fetus. Therefore, the need to develop new accurate HR methods for a good estimation of the FECG signals given a noninvasive/external recording (ex: maternal abdominal recordings).

ECG Recordings

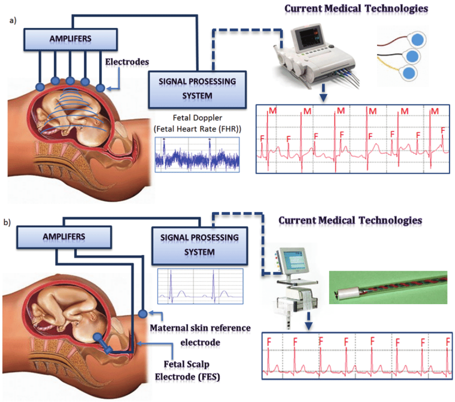


Figure 2: The fetal ECG monitoring a) external b) internal.[1]

Task Description

Extract the fetal HR given maternal abdomen recordings.

Questions

- Is it only the heart rate or also signal enhancement?
- Does it include removing the maternal ECG?
- Does it include peak detection?

Challenges

The noninvasive abdominal recordings are severely contaminated by electrical interferences such as :

- Maternal ECG
- Powerline interference
- Muscle noise
- Equipment noise
- Fetus position

Thus complicating the extraction of physiological informations.

ECG Recordings

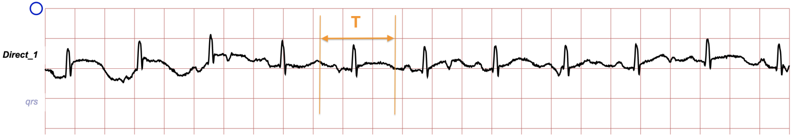


Figure 3: Direct electrocardiogram recorded simultaneously from fetal head (5 sec).

Figure 4: Abdominal fetal-ECG recording (5 sec).

Data

Abdominal recordings measured by 4 electrodes as shown in fig.4, aside with the ground truth considered to be the direct ECG recorded from the fetal's head as illustrated in fig.3.

Preprocessing

- The powerline inference as in Warmerdam et al (2016) and Varanini et al (2013).
- Maternal ECG as in Vullings et al (2009).
- For the adaptive kalman filter : the QRS complexes are detected, then the ECG complexes are subsequently defined as the signal within a predefined time window of length **T** around the QRS complex.

Mathematical Model

In a simplified form, both the relation between consecutive ECG complexes and the corruption of the recorded ECG can be described by means of a state-space model:

Mathematical Model (based on the adaptive KF)

$$\mathbf{x}_k = \mathbf{x}_{k-1} + \mathbf{v}_k, \qquad \mathbf{v}_k \sim \mathcal{N}(0, \Lambda_k), \qquad \Lambda_k = \lambda_k^2 \cdot \mathbf{I} \tag{1a}$$

$$\mathbf{y}_k = \mathbf{x}_k + \mathbf{w}_k, \qquad \mathbf{w}_k \sim \mathcal{N}(0, \Sigma_k), \qquad \Sigma_k = \sigma_k^2 \cdot \mathbf{I} \tag{1b}$$

Where :

- $\mathbf{x}_k$: [Tx1] ECG complex for heartbeat k.
- $\mathbf{y}_k$: [Tx1] recorded signal.
- $\mathbf{v}_k$: process noise, models the ECG complexes evolution between heartbeats
- $\mathbf{w}_k$: measurement noise
- T : length of the ECG complex

Evaluation and Metrics

- (normalized) Mean Square Error (MSE) between the filtered ECG signals $\hat{x}$ and the original ECG signals x.

$$MSE = \frac{\sum_k (\mathbf{x}_k - \hat{\mathbf{x}}_k)^\top (\mathbf{x}_k - \hat{\mathbf{x}}_k)}{\sum_k \hat{\mathbf{x}}_k^T \mathbf{x}_k}$$

- Mean Absolute Error (MAE):

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^{N} |\text{FHR}_{\text{target}} i - \text{FHR}_{\text{predict}} i|$$

Here, N is the length of the output sequence (N=240), $FHR_{predicted}$ is the estimated fetal HR and FHR_{target} is the fetal HR measured by the scalp electrode.

Reliability evaluation - (From CNN LSTM)

- The reliability of the method is evaluated in terms of **positive percent agreement (PPA)**, which is the percentage of fetal HR outputs that were within 10% of the actual HR value.
- Results of 5% tolerence is also presented (called PPA_5)
- **coverage** - a second reliability metric - corresponds to the time that the method outputted a nonzero fetal HR value.

Datasets I

1. A private dataset obtained in a collaboration between the Eindhoven University of Technology and the Maxima Medical Center, Veldhoven, TheNetherlands. Used in:
   - CNN-LSTM algorithm.
   - Warmerdam algorithm.
2. The set-A of the 2013 Physionet/Computing in Cardiology Challenge. `https://physionet.org/content/challenge-2013/1.0.0/`. Used in:
   - Adaptive Kalman algorithm.
   - CNN-LSTM algorithm.
   - Warmerdam algorithm.

Datasets II

3. B. De Moor, “DaISy: Database for the identification of systems,” http://homes.esat.kuleuven.be/ smc/daisy, eSAT/STADIUS, KU Leuven, Belgium, Used dataset: [96-012], Accessed: 2017-01-20. Used in:
   - Nour Zalmai et al algorithm.

Approach-1: Adaptive Kalman Filter.

Adaptive Kalman Filter

- As discussed earlier, the FECG are very noisy signals. One of the many ways to reduce the noise or equivalently enhance the SNR while preserving the clinically relevant morphological variations in the ECG signals is the **adaptive Kalman filter**.
- After preprocessing the given abdominal ECG recordings, the signal is divided into blocks of length 'T' (The value of T chosen here is 120% of the mean interval between consecutive heartbeats), each block represents a heartbeat.
- In the state-space description (1), the problem of enhancing the SNR of the ECG is reduced to the problem of sequentially estimating the model parameter vector x_k and the noise covariances Σ_k and Λ_k . Here, sequential estimation refers to the estimation of the relevant parameters based on the earlier estimate and all newly arriving data.

Note

The measurement and process noise are assumed to be are spatially uncorrelated. i.e:

$$\Sigma_k = \sigma_k^2 I \quad \text{and} \quad \Lambda_k = \lambda_k^2 I$$

Algorithm implementation

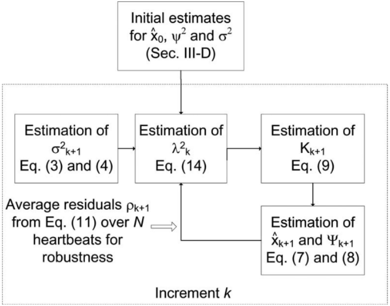


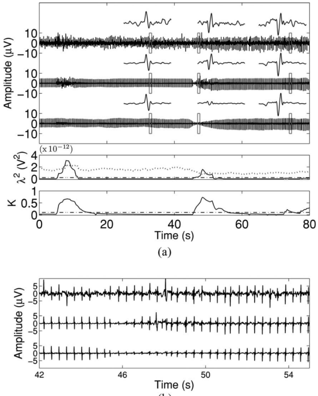
Figure 5: Illustration of the algorithmic implementation of the developed adapaptive Kalman filter.

Data & Loss function

Data : FECG signals recorded from the maternal abdomen.
Loss function: The performance is quantified by calculating ϵ, the normalized MSE between the filtered ECG signals $\hat{x}$ and the original ECG signals x:

$$\epsilon = \frac{\sum_k (\mathbf{x}_k - \hat{\mathbf{x}}_k)^T (\mathbf{x}_k - \hat{\mathbf{x}}_k)}{\sum_k \hat{\mathbf{x}}_k^T \mathbf{x}_k}$$

Results



(a)
(b)

(a) Top panel shows a FECG signal recorded from the maternal abdomen, before filtering, but after preprocessing (top graph), after filtering with the adaptive Kalman filter (center graph), and after filtering with the fixed Kalman filter (bottom graph). The ECG complexes indicated in the rectangles are shown zoomed in in the accompanying graphs. The second panel shows the estimated process noise covariance λ^2 for the adaptive Kalman filter (solid line), the process noise covariance for the fixed Kalman filter (dash-dot line), and the estimated measurement noise covariance σ^2 (dotted line). The bottom panel shows the Kalman gain K for both the adaptive (solid line) and fixed (dash-dot line) filters. For the estimation of λ^2 and K in the adaptive Kalman filter, N is chosen equal to 10. Since the process noise covariance is often estimated as 0 [see (15)], λ^2 cannot be expressed in decibel as in Fig.20. Hence, λ^2 is expressed here in an absolute sense (i.e., in V^2). In (b), a zoom of the top panel between 42 and 55 s is depicted.

Limitations

In the derivation of the adaptive Kalman filter, several assumptions are made for mathematical simplicity, but that might limit the applicability of the filter.

- Both the process and measurement noise are assumed to be Gaussian.
- Measurement and process noise are assumed to be uncorrelated.
- Fixed length T.

Approach-2: CNN-LSTM network.

CNN-LSTM network

The model, which is illustrated in this figure, is comprised of two main blocks. The DICNN network block (encoder) consists of six stacked dilated convolution inception modules. The DICNN network is used as a feature extractor and the extracted features are fed to the LSTM network block (decoder) that is responsible for estimating the fetal HR.

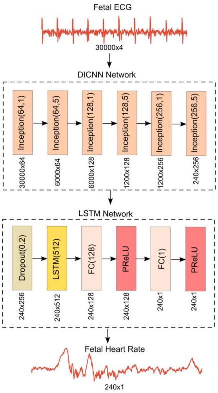


Figure 6: Overview of the proposed DICNN-LSTM network

Main Contributions

- A deep hybrid dilated inception CNN-LSTM(DICNN-LSTM) encoder-decoder network that extracts the fetal HR from noninvasive abdominal recordings. This paper is the first to employ deep learning to estimate the fetal heart HR without explicitly detecting the QRS complexes.
- The reliability of the method is reinforced by a classifier,which uses a CNN network to identify the periods of time that the extracted fetal HR is inaccurate, increasing the suitability of the method for clinical application.
- Experimental results demonstrate the advantage of this method over top-performing state-of-the-art algorithms.

Data - I

Two data set are used :

1. A **private** dataset obtained in a collaboration between the Eindhoven University of Technology and the Máxima Medical Center, Veldhoven, The Netherlands.
    - Information:
        - 28 Abdominal Recorings
        - Sampling rate: 500Hz
        - Total Duration: 91 Hrs
        - Each recording obtained Particulary from the first and second stage of labor.
        - Simultaneous scalp HP recordings are performed and stored at 4 Hz (used as ground truth).
        - Presence of clock drift between scalp fetal HR and fetal ECG.

Data - II

- Training & Testing
  - 16 out of 28 recordings are randomly selected and used for training.
  - 6 recordings are used as validation set to tune the parameters.
  - The remaining 6 recordings are kept as a test set to evaluate the performance of the network. The scalp fetal HR was used as the desired output of the network (labels).

Data - III

2. The set-A of the 2013 Physionet/Computing in Cardiology Challenge.
   `https://physionet.org/content/challenge-2013/1.0.0/`
   - Non invasive abdominal recordings.
   - Total duration: 751 min
   - Sampling rate: 1000 Hz
   - Data obtained from multiple sources using a variety of instruments.
   - Used as a test set.

Loss function

The used loss-function is the mean absolute error MAE:

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N} |\mathrm{FHR}_{\mathrm{target}} i - \mathrm{FHR}_{\mathrm{predict}} i|$$

where N is the length of the output sequence (N=240), $FHR_{predicted}$ is the estimated fetal HR and FHR_{target} is the fetal HR measured by the scalp electrode.

Evaluation Metrics & Results - I

- **Metrics** :
  The reliability of the method is evaluated in terms of positive percent agreement (PPA), which is the percentage of fetal HR outputs that were within 10% of the actual HR value. Results of 5% tolerence is also presented (called PPA_5). In addition, we used a second reliability metric, the coverage, that corresponds to the time that the method outputted a nonzero fetal HR value.
- **Reference method** :
  The performance of this method was compared to the performance of the algorithm of Warmerdam et al (2018), that suggests an adaptive multichannel R-peak detection method that combines HR information and ECG waveform.

Results:

Algorithm	MSE (bpm^2)	MAE (bpm)	PPA (%)	PPA_5 (%)	Coverage (%)
DICNN-LSTM	104.5/59.6[a]	3.3/2.4[a]	93.9/96[a]	91/93.6[a]	100
DICNN-LSTM + Fetal HR reliability classifier	49.4	2	97.3	95.4	87.9
Warmerdam *et al* (2018)	129.7	3.4	93.6	91.5	94.5

Figure 7: Fetal HR extraction performance on the private test dataset. Calculated only in the periods that Warmerdametal(2018) outputs a HR value.

Results

Algorithm	MSE (bpm^2)	MAE (bpm)	PPA (%)	PPA_5 (%)	Coverage (%)
DICNN-LSTM	14.2	1.6	98.6	95.6	100
DICNN-LSTM + Fetal HR reliability classifier	6.9	1.1	99.6	98.7	82
Warmerdam *et al* (2018)	30.8	1.5	97.9	96.5	100
Varanini *et al* (2013)	23.8	1	98.4	95.4	100
Behar *et al* (2014)	172	5.7	91	78.8	100

Figure 8: Fetal HR extraction performance on set-a of 2013 Physionet/Computing in Cardiology Challenge.

Evaluation Metrics & Results - II

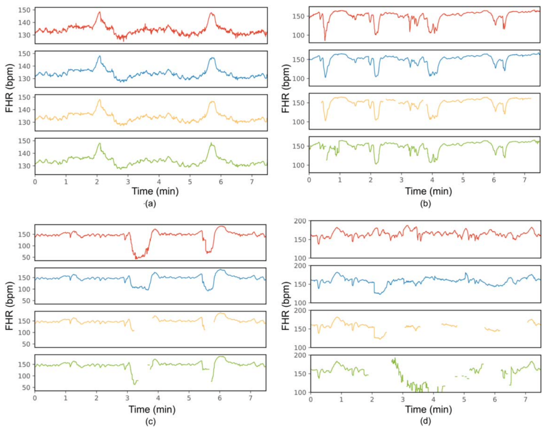


Figure 8. Four fetal heart rate (FHR) segments from our private test set of 7.5 min duration each. For each case a–d: the first subplot (red) depicts the scalp FHR, the second (blue) the estimated FHR from the DICNN-LSTM network, the third (yellow) the estimated FHR from the DICNN-LSTM network in combination with the reliability classifier (only reliable FHR values are displayed) and the fourth (green) the FHR obtained from the algorithm of Warmerdam *et al* (2018) (only when an FHR value is outputted).

Limitations

This approach has several limitations:

- The proposed fetal HR extraction network has many parameters such as the number of layers, dilation rates, number of nodes for each layer, type of layers, size of input and output signals etc. One better way to choose these parameters is is leave-on-out cross validation that is appropriate in cases of small datasets like ours, but dismissed because it is computationally very expensive.
- No sensitivity analysis was performed on the parameters due to their large number.