Deep learning applied to non-invasive ECG pregnancy monitoring: fetal ECG recovery and QRS complex detection

Julia C. Remus[1], **Thiago L. T. da Silveira**[1,2]

[1]Institute of Informatics – Universidade Federal do Rio Grande do Sul
Porto Alegre – RS – Brazil
[2]Department of Applied Computing – Universidade Federal de Santa Maria
Santa Maria – RS – Brazil

juliacremus@gmail.com, thiago@inf.ufsm.br

Abstract. *Non-invasive electrocardiography (niECG) is an option to mitigate the limitations of the current pregnancy monitoring methods. Still, one of the main challenges in its use is the extraction of the fetal signal (fECG) from the measured abdominal signal (aECG) due to the low signal-to-noise ratio and the overlap of the maternal and fetal QRS complexes in time and frequency domains. We present two encoder-decoder deep learning models that use regions of interest (RoI) to perform simultaneous fetal QRS complex (fQRS) detection and fECG recovery. We show that such a RoI-based end-to-end approach leads to robust state-of-the-art results in inter-subject and cross-dataset evaluations.*

1. Introduction

Non-invasive electrocardiography (niECG) is a promising method for pregnancy monitoring, providing insights into both fetal and maternal cardiac health [Kahankova, R. *et al.* 2020]. This technique involves placing electrodes on the pregnant abdomen to capture the abdominal ECG (aECG), which consists of maternal ECG (mECG), fetal ECG (fECG), uterine contractions (UC), and noise components [Kahankova, R. *et al.* 2020]. Since niECG measures fECG indirectly, the main challenge lies in isolating the fECG from the aECG. This task is particularly difficult due to the overlap (in frequency and time domains) of maternal and fetal QRS complexes – key component of ECG signals – as well as the inherently low signal-to-noise ratio (SNR) of the fECG within the aECG [Kahankova, R. *et al.* 2020].

The fECG signal is used to assess fetal heart rate (fHR) and enables morphological analysis, which can help detect congenital heart diseases [Mendis, L. *et al.* 2023]. Additionally, when combined with UC data, fHR provides critical information about the fetus's ability to tolerate oxygen-deprived conditions [Vullings 2010]. Despite its potential for fetal health monitoring, the only highly accurate method for measuring fECG remains direct ECG acquisition using a scalp electrode during labor (sECG), though the resulting signal is affected by movement artifacts [Behar, J. *et al.* 2016]. This last method is not routinely used due to its invasiveness and risk of infection [Kahankova, R. *et al.* 2020].

Various deep learning (DL) models have been developed for fECG recovery [Zhong, W. *et al.* 2019, Mohebbian, M. R. *et al.* 2022, Ghonchi and Abolghasemi 2022, Rahman, A. *et al.* 2023, Wang, X. *et al.* 2023], but they all rely on external QRS complex detectors. Additionally, previous approaches preprocess both aECG and sECG signals, providing a smoothed and standardized input during training. As a result, these

models do not focus on fECG recovery under noisier conditions. Therefore, this work aims to develop DL models that retrieve morphologically precise fECG from aECG with simultaneous detection of fQRS – without relying on external algorithms. Besides aiming for real-time applications, we explored approaches for cost reduction in DL models.

2. Proposed Methodology

The proposed models are convolutional neural network (CNN)-based encoder-decoders (EDs), hereafter called RoINet and RoINet Lite – a lightweight version. Our methodology involves defining regions of interest (RoI) that improve the model's comprehension during training and define fQRS complexes locations. The following subsections detail our proposed methods. The project is open-source and available at GitHub[1].

2.1. Datasets and Data Preparation

The datasets used here are publicly available and have at least three aECG electrodes. The training process was performed using the extended Abdominal and Direct Fetal ECG dataset (XADFECG) [Matonia, A. *et al.* 2020], which has the invasive fetal signal (sECG), used as the ground truth for the fECG recovery task, the fR-peak annotations, and the aECG signals. Plus, the NI-FECG dataset [Sober and Marco 2007], containing data from one subject throughout the pregnancy, is used to evaluate the models.

The input data is structured as (b, t, c), where b represents the batch size, t is the segment length, and c denotes the number of stacked electrode signals. The input data and ground truth are divided into segments of $0.5s$ (t = 250 samples) to ensure the presence of at least one fQRS per segment. After segmentation, regular segment-level normalization is applied. Three abdominal electrodes ($c = 3$) are used as input signals. The batch size was set to four ($b = 4$). The ground truth consists of the normalized sECG stacked with the RoI. The RoI is defined as a Gaussian function $\mathcal{N}(\mu, \sigma^2)$, where μ corresponds to the fR-peak position, and σ^2 represents the standard deviation, equivalent to 25 ms based on the fQRS interval analysis from [Chivers, S. C. *et al.* 2022]. Due to the limitations of the XADFECG dataset, the sampling frequency used was 500Hz for both aECG and sECG.

2.2. ED Architectures

Based on LinkNet [Chaurasia and Culurciello 2017], we proposed two ED architectures to specialize the two target tasks: fECG recovery and fQRS detection. We consider shared encoder but split the decoder into two branches to deal with multi-tasking. In RoINet, the decoder was fully divided so the training process could better fit the parameters. To reduce the model size and inference time, we set up RoINet Lite, where we allocated only one convolutional layer to each task at the decoder's end. Batch normalization, used in LinkNet, was removed after the convolution layers due to the non-stochastic nature of the ECG signals. Dropout regularization was added after each encoder and decoder block to overcome overfitting. Skip connections were kept as in LinkNet. As a result, RoINet has a total of 21 million parameters while RoINet Lite has 4,6 million.

2.3. Training Details

The proposed loss function comprises three terms involving the reconstruction of the fECG signal, the RoI mask, and the fECG within the RoI. Each of these terms has one hyperparameter – α, β, or γ – ranging from 0 to 1, keeping their sum equal to one. The loss

[1]https://github.com/tlts-lab-ufrgs/fecg-research-roi-encoder-decoder

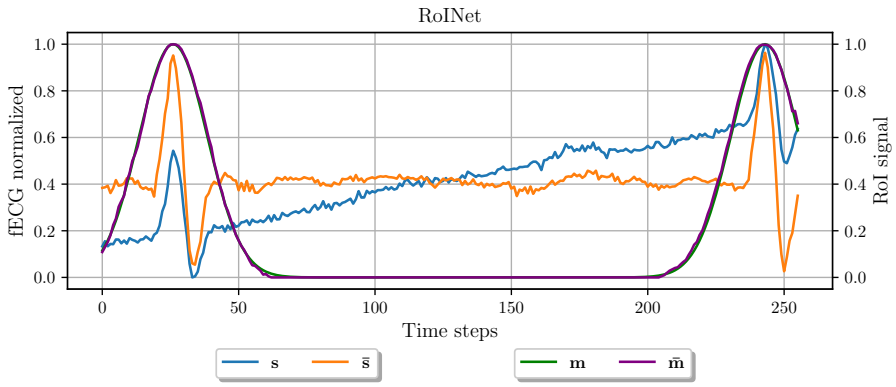

Figure 1. Recovered fECG signals and fQRS-related RoI masks on signals from the XADFECG dataset by RoINet model. The sECG signal ($\mathbf{s}$) is shown in blue, the predicted fECG signal ($\bar{\mathbf{s}}$) in orange, and almost entirely overlapped the original and predicted RoI masks ($\mathbf{m}$ and $\bar{\mathbf{m}}$) in green and purple. Two fQRS are present in this segment.

function is defined as $L(\bar{\mathbf{s}}, \bar{\mathbf{m}}, \mathbf{s}, \mathbf{m}) = \alpha \cdot \frac{1}{t}\sum^{t} \log\cosh(\bar{\mathbf{s}} - \mathbf{s}) + \beta \cdot \frac{1}{t}\sum^{t} \log\cosh(\bar{\mathbf{m}} - \mathbf{m}) + \gamma \cdot \frac{1}{t}\sum^{t} \left[\log\cosh(\bar{\mathbf{m}} \odot \mathbf{s} - \mathbf{m} \odot \mathbf{s}) + \log\cosh(\mathbf{m} \odot \bar{\mathbf{s}} - \mathbf{m} \odot \mathbf{s})\right]$, where $\odot$ is the Hadamard product. The terms of the loss function enforce the recovered signal ($\bar{\mathbf{s}}$) and RoI mask ($\bar{\mathbf{m}}$) to match their ground-truth versions ($\mathbf{s}$ and $\mathbf{m}$, respectively). Aiming to reconstruct the fECG signal under the RoI mask while ignoring environmental noises present in the sECG, the grid search applied in the loss function hyperparameters results in optimal values of $\alpha = 0.1$, $\beta = 0.3$, and $\gamma = 0.6$.

To reduce training overfitting and improve data variability in datasets, we applied data augmentation (DA) online during the training. Three techniques already used in ECG analysis [Rahman, M. M. *et al.* 2023] that encompass real niECG situations, such as baseline shift by body movements, electric and sensor noise, and electrode disconnection, were used. Namely, we applied low- and high-frequency signal addition and segment cutoff. Each technique has a probability of 50% for being applied to the segment.

3. Results and Discussion

To evaluate the models in the fQRS detection task, precision, recall, and F1-score were used, while for the fECG signal regression task, the $L1$ error was calculated. The methods were assessed using the leave-one-subject-out cross-validation. The confidence interval (CI) set at 95% is presented for all results. Cross-dataset tests are also performed.

3.1. fECG Recovery Results

Figure 1 presents one fECG recovering scenario, where in blue and in green are the ground truth data – the sECG signal ($\mathbf{s}$) and the generated RoI ($\mathbf{m}$) – compared with the predicted fECG signal ($\bar{\mathbf{s}}$) (in orange) and RoI ($\bar{\mathbf{m}}$) (in purple). Incorporating the mask into the loss function enabled the model to ignore signal variations unrelated to the cardiac cycle, acting as a filter. In this specific segment, there are two fQRS complexes reconstructed while ignoring the linear slope presented by the raw sECG signal.

To further investigate potential filtering effects, a preprocessing step from [Mohebbian, M. R. *et al.* 2022] was applied into sECG, including band-pass filtering and

power-line interference removal. The results show improved performance for both models, confirming that the predicted fECG signal aligns more closely with a filtered sECG than the raw (noisy) signal. Notably, the reconstruction error lowered from 0.142 to 0.106 for the whole signal using the RoINet model—the results for RoINet Lite are similar.

3.2. fQRS Complex Detection Results

Our fQRS complex detector reduces to finding the Gaussian peaks on the RoI masks generated by the models. Furthermore, the Gaussian distribution acts as an acceptance region for the R-peak localization, with its center aligned to the predicted fR-peak position. For validation, we compare our fQRS complex detection methods with the well-known P&T algorithm [Pan and Tompkins 1985], used in the recent works previously mentioned. The algorithm is applied in the predicted fECG signal ($\bar{\mathbf{s}}$). We note that the P&T has limited adaptability to signal variations. Our results for both datasets are better than those from P&T (discussed next); hence, we provide a more robust fQRS complex detection option.

In the XADFECG dataset, the RoINet model achieves 96.2% of precision and 95.85% of recall while the RoINet Lite model reaches 96.15% and 94.33% of precision and recall. The average results for the NI-FECG dataset show that both models have good responses in recall metric, with more than 99% each; however, the precision indicates non-existent fR-peaks, with 85.53% and 87.82% for the RoINet and RoINet Lite respectively. When analyzing the results per subject on the XADFECG dataset, we note that the models' performances are directly affected when trained on data with lower variability, leading to lower metrics, as seen for the subject `r03`. The models' responses to different XADFECG subjects show that RoINet recall is less affected by the interferences than the RoINet Lite – retrieving the fR-peaks in noisier environments. The number of false positives followed the same pattern in both models. While RoINet better detects fQRS complexes in noisy signals, it might be more susceptible to random signal variations.

When evaluating our methods on the NI-FECG dataset, we note that the methods' precision improves over pregnancy time. One possible reason is that not all t-segment intervals contain fQRS and, after normalization, the noise is amplified. Combined with a model susceptible to variations (RoINet), it leads to more false positives compared with the RoINet Lite. Meanwhile, its recall is stable throughout the interval (with values larger than 99%) possibly due to the model's robustness when training with noisy scenarios.

3.3. Comparison with the State-of-the-Art

Compared to state-of-the-art methodologies [Mohebbian, M. R. *et al.* 2022, Ghonchi and Abolghasemi 2022, Rahman, A. *et al.* 2023, Wang, X. *et al.* 2023], we propose a model that recovers the fECG signal while simultaneously detecting fQRS complexes in noisy environments, without relying on preprocessing steps. Notably, the previous methodologies use different preprocessing steps, and there are no standardized evaluation protocols or metrics.

The inference time per fECG minute recorded by our methods is lower than [Mohebbian, M. R. *et al.* 2022], [Ghonchi and Abolghasemi 2022], and [Wang, X. *et al.* 2023]. which report 1.9 $\pm$ 0.4s, 1.32s, and 0.97s, respectively. RoINet and RoINet Lite take only 0.066s and 0.034s. Also, RoINet Lite is the second lightest option, being surpassed only by [Mohebbian, M. R. *et al.* 2022] in model size.

For fQRS complex detection, previous methodologies have surpassed our results using leave-one-subject-out cross-validation (XADFECG), where we achieved F1-score of 96.0% $\pm$ 2.37% and 95.12% $\pm$ 2.62% for RoINet and RoINet Lite respectively. However, our metrics remain comparable to other approaches within the CI range. When evaluated on different datasets (NI-FECG), our model outperforms [Mohebbian, M. R. *et al.* 2022] and [Ghonchi and Abolghasemi 2022] in recall with more than 99% on both models; but lags in precision, with in 85.53% $\pm$ 2.16% and 87.82% $\pm$ 2.28% for the RoINet and RoINet Lite model. The lack of precision is attributed to the definition of t, which needed to be higher for earlier gestational weeks.

Finally, our numerical results for the fECG recovery task are comparable with [Rahman, A. *et al.* 2023] and exceed those of [Wang, X. *et al.* 2023]. However, our comparison is based on sECG and predicted fECG, whereas the referenced authors compare its predicted fECG with the preprocessed sECG. Moreover, not all studies assess the morphological aspects of fECG extraction as we do.

4. Conclusion

We presented two ED architectures for recovering the fECG signal and detecting fQRS complexes from the maternal aECG signals. The mean F1-score values obtained were 96% and 95.12% for RoINet and RoINet Lite, respectively, when trained and evaluated on the XADFECG database. In the NI-FECG database, RoINet and RoINet Lite could detect fQRS complexes with high recall ($> 99\%$) but presented lower precision values – 85.53% and 87.82%, respectively, due to the segment size (t) previously defined. Furthermore, compared with the P&T algorithm, our methods have higher robustness to noise and better detect fQRS complexes in both databases.

The use of a RoI introduced an ability to the models to filter high-frequency noise and changes in the signal baselines, resulting in fECG signals highlighing the fQRS complexes. Considering the preprocessed sECG signal (after a 1-100Hz bandpass filter and removal of 50Hz frequencies), the RoINet model returned an $L1$ reconstruction error of 0.1049, while the RoINet Lite model obtained an error of 0.1032.

The inference time of the models is compatible with real-world applications and shorter than the time required for the measurement. Even though the RoINet network achieves higher recalls (with fewer false negatives), the 5-fold reduction in the number of parameters proposed in the RoINet Lite version does not significantly change the results, making the method viable, especially considering embedded hardware applications. Future work should focus on validating the methodology with datasets exhibiting greater variability as well as evaluating its performance on hardware implementations.

The Dissertation associated with this paper has achieved the maximum grade and received distinction honors. Results from the RoINetwere published in [Remus and da Silveira 2024], while findings from RoINet Lite are described in the paper "Complexity-Reduced End-to-End Fetal ECG Signal Recovery and QRS Complex Detection" accepted in the 25th *Simpósio Brasileiro de Computação Aplicada à Saúde*.

Acknowledgments

The authors thank the *Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul* for partially funding this research.

References

Behar, J. *et al.* (2016). A practical guide to non-invasive foetal electrocardiogram extraction and analysis. *Physiological Measurement*, 37(5):R1–R35.

Chaurasia, A. and Culurciello, E. (2017). Linknet: Exploiting encoder representations for efficient semantic segmentation. In *IEEE Visual Communications and Image Processing*, pages 1–4.

Chivers, S. C. *et al.* (2022). Measurement of the cardiac time intervals of the fetal ECG utilising a computerised algorithm: A retrospective observational study. *JRSM Cardiovasc. Dis.*, 11:20480040221096209.

Ghonchi, H. and Abolghasemi, V. (2022). A dual attention-based autoencoder model for fetal ECG extraction from abdominal signals. *IEEE Sensors Journal*, 22(23):22908–22918.

Kahankova, R. *et al.* (2020). A review of signal processing techniques for non-invasive fetal electrocardiography. *IEEE Reviews in Biomedical Engineering*, 13:51–73.

Matonia, A. *et al.* (2020). Fetal electrocardiograms, direct and abdominal with reference heartbeat annotations. *Scientific Data*, 7(1).

Mendis, L. *et al.* (2023). Computerised cardiotocography analysis for the automated detection of fetal compromise during labour: A review. *Bioengineering*, 10(9):1007.

Mohebbian, M. R. *et al.* (2022). Fetal ECG extraction from maternal ECG using attention-based CycleGAN. *IEEE Journal of Biomedical and Health Informatics*, 26(2):515–526.

Pan, J. and Tompkins, W. J. (1985). A real-time QRS detection algorithm. *IEEE Transactions on Biomedical Engineering*, BME-32(3):230–236.

Rahman, A. *et al.* (2023). Fetal ECG extraction from maternal ECG using deeply supervised LinkNet++ model. *Engineering Applications of Artificial Intelligence*, 123:106414.

Rahman, M. M. *et al.* (2023). A systematic survey of data augmentation of ECG signals for AI applications. *Sensors*, 23(11):5237.

Remus, J. C. and da Silveira, T. L. T. (2024). An end-to-end RoI-based encoder-decoder for fetal ECG recovery and QRS complex detection. *IEEE International Symposium on Medical Measurements and Applications*, pages 1–6.

Sober, M. M. and Marco, J. G. (2007). Non-invasive fetal electrocardiogram database. Available at `https://physionet.org/content/nifecgdb/`.

Vullings, R. (2010). *Non-invasive fetal electrocardiogram:analysis and interpretation*. PhD thesis, Technical University of Eindhoven.

Wang, X. *et al.* (2023). Correlation-aware attention CycleGAN for accurate fetal ECG extraction. *IEEE Transactions on Instrumentation and Measurement*, 72:1–13.

Zhong, W. *et al.* (2019). Fetal electrocardiography extraction with residual convolutional encoder–decoder networks. *Australasian Physical; Engineering Sciences in Medicine*, 42(4):1081–1089.