Comparative Analysis and Deep Learning Approaches for ECG-Derived Respiration Signal Estimation using the Fantasia Dataset

Lana Dominkovic

lana.caldarevic@student.fis.unm.si

Abstract—In this seminar work, we established baselines for ECG-derived respiration signal analysis using the Fantasia dataset and compared these with a deep learning autoencoder inspired by the U-Net model. Our model, comprising convolutional layers, ReLU activations, Batch Normalization, MaxPooling, and up-sampling, achieved an average validation CC of 0.51 and MSE of 0.046, outperforming 4 out of 6 traditional signal processing algorithms. Future work will explore different deep learning architectures and datasets to enhance generalizability and robustness in respiratory signal estimation.

1 INTRODUCTION

In the realm of data science and machine learning, the importance of establishing robust baselines cannot be overstated. For this seminar work, we aim to set strong baselines using well-established signal processing methods for ECG-derived respiration signal problem. These baselines will serve as reference points to effectively compare the performance of our deep learning solutions, ensuring that any improvements are not just apparent but quantifiable.

Signal processing methods have long been the backbone of data analysis across various domains, providing tools to extract, interpret, and transform data into meaningful insights. Methods such as Fourier Transform, Wavelet Transform, and Principal Component Analysis (PCA) offer a time-tested approach to understanding and manipulating signal data. By applying these techniques to our datasets, we can establish performance metrics and benchmarks that reflect a deep understanding of the inherent patterns and structures within the data.

These methods are described in RRest toolbox [1], which is a comprehensive, publicly available toolkit designed for the estimation of respiratory signal and rate (RR) from electrocardiogram (ECG) and photoplethysmogram (PPG) signals. It includes over 300 algorithms that combine various signal processing techniques, enabling reproducible and systematic comparisons of RR estimation methods. The toolbox facilitates the advancement of research by providing a standardized framework and data for evaluating the performance of different algorithms under consistent conditions.

By systematically applying these methods, we establish a comprehensive baseline that captures the effectiveness of traditional signal processing techniques. These baselines are crucial for benchmarking as they provide a clear standard against which our deep learning models can be evaluated. The comparison will highlight areas where deep learning excels, particularly in handling complex patterns and large-scale data, and areas where traditional methods still hold their ground.

For this work, we identified the Fantasia dataset, a publicly available resource from Physionet [2], which we will use for deriving ECG respiration signals. Baselines will be established using the mentioned toolbox for future reference.

The novel contributions of this work are as follows: (i) establishing baselines derived from novel respiratory signal estimation algorithms, which can serve as benchmarks for future research in this area, (ii) presenting the first results on the Fantasia dataset, and (iii) achieving superior results on the Fantasia dataset using deep learning algorithms compared to current novel signal processing methods.

2 DATASET DESCRIPTION

In our seminar work, we utilized Fantasia dataset to establish baselines for respiratory signal analysis using traditional signal processing methods and to compare with our deep learning solutions.

The Fantasia dataset contains long-term ECG recordings from healthy young and elderly subjects, designed to study heart rate variability and the effects of aging on cardiac function. This dataset provides high-quality ECG signals recorded under controlled conditions, making it ideal for research into age-related physiological changes and heart rate dynamics.

- **Number of Subjects**: 40 (20 young, 20 elderly)
- **Ages**: Young subjects (21-34 years), Elderly subjects (68-85 years)
- **Conditions**: Healthy
- **Sampling Frequency**: 250 Hz
- **Duration**: 120 minutes

From ECG and respiration signals, 32-seconds length windows were taken with the 16-seconds overlap, normalized to [0, 1] range and downsampled to 1024 samples (32 Hz).

3 PERFORMANCE OF SIGNAL PROCESSING METHODS

The performance of the algorithms in the RRest Toolbox was assessed using two metrics: Mean Squared Error (MSE) and Mean Cross-Correlation (CC). The signals were segmented into 32-second windows with a 50% overlap, and MSE and CC were then computed for each segment.

MSE measures the average of the squares of the errors, which is the average squared difference between the estimated values and the actual value. Lower MSE values indicate better performance.

CC measures the similarity between two signals. Higher CC values indicate that the estimated signal is more similar to the reference signal, signifying better performance.

Method	Mean CC	Mean MSE
ELF_RSlinB_FMeam_FPt_RDtGC_EHF	0.59	0.073
ELF_RSlinB_FMebw_FPt_RDtGC_EHF	0.50	0.069
ELF_RSlinB_FMefm_FPt_RDtGC_EHF	0.56	0.070
flt_BFi	0.37	0.083
flt_Wam	0.38	0.093
flt_Wfm	0.44	0.081

4 DEEP LEARNING METHODS

Research on extracting respiratory signals from ECG and PPG signals using deep learning is still in its infancy. Most existing studies primarily focus on estimating respiratory rate rather than extracting the complete respiratory waveform. This limited scope overlooks the potential wealth of information that can be derived from full respiratory signals, which could offer deeper insights into various physiological conditions.

Deep learning methods, like the RespNet model [3], which employs a U-Net architecture, have demonstrated high accuracy in predicting respiration from PPG signals. Additionally, in [4], the RespNet architecture was adapted for a custom dataset to extract respiration from ECG signals, yielding promising results.

The current emphasis on respiratory rate estimation, while useful, does not fully exploit the capabilities of deep learning models in biosignal analysis. Future research should aim to extend beyond respiratory rate to extract full respiratory signals, which could enhance diagnostic and monitoring capabilities in healthcare.

To this end, several promising deep learning approaches can be explored. Recent advancements in Transformer architectures and Generative Adversarial Networks (GANs) open up new avenues for innovative solutions. For example, Cycle GANs have been effectively used to derive respiratory rates from PPG signals [5], and similar techniques could be adapted for extracting respiratory signals from ECG data. Additionally, the Reservoir Computing framework [6], known for its efficiency in handling dynamic systems with faster training times and reduced data requirements, offers a promising approach for this problem. Exploring these advanced architectures and techniques could lead to breakthroughs in extracting detailed respiratory signals from ECG data, ultimately advancing the state of healthcare monitoring and diagnostics.

5 APPLICATION OF DEEP LEARNING

5.1 Model architecture

The network implemented is a fully convolutional autoencoder network designed to estimate respiratory signals from ECG data and is shown in Figure 1. This network's architecture is inspired by the U-Net model, commonly used for image segmentation tasks. Our implementation is a simplified version of U-Net, retaining the essential concept of shortcut connections but with fewer layers and parameters.

The network comprises an encoder and a decoder, both fully convolutional. The encoder captures feature representations of the ECG signal by applying various filters. Specifically, the network has three levels, each with an increasing number of filters: 4 in the first level, 8 in the second, and 16 in the third. Each filter is a 1D convolution filter with a length of 27. In the encoder, each convolutional layer is followed by ReLU activation and Batch Normalization, ensuring stable and effective training. MaxPooling layers are used to down-sample the signal, reducing its dimensionality and capturing important features.

The decoder mirrors the encoder's structure but performs up-sampling to reconstruct the signal. The decoder layers also include convolutional layers with ReLU activation and Batch Normalization. Additionally, the network uses dropout layers, with a dropout rate of 0.6, to prevent overfitting. Overall, the network consists of several convolutional layers, pooling layers, up-sampling layers, and dropout layers, resulting in a robust architecture for respiratory signal estimation from ECG data.

The training procedure involves several steps to optimize the model. First, hyperparameters such as learning rate, L2 regularization factor, dropout rate, number of filters, kernel size, and batch size are set. These configurations are essential for model performance. The model is trained for up to 200 epochs with a batch size of 256, using the Adam optimizer for its efficiency and effectiveness. The learning rate for Adam is 3e-4. The model is uses MSE as the loss function, and additional metrics, like correlation are used to evaluate performance during training.

Several callbacks are used during training, including logging callbacks to record training and validation metrics at each epoch, early stopping to monitor validation loss and stop training if it doesn't improve for a specified number of

epochs, preventing overfitting, and model checkpointing to save the best model weights during training. This process ensures that the model is trained effectively to achieve optimal performance in estimating respiratory signals from ECG data. The network was implemented using TensorFlow and trained on appropriate hardware to handle the computational requirements. The model contains 23,409 trainable parameters.

5.2 Evaluation

We employed a sliding window technique for data preprocessing. The entire recording was divided into 32-second long segments, with each segment overlapping by 16 seconds with the previous one. Each of these segments was then down-sampled to 1024 samples, which served as the input to our network. The network's output was a 32-second long estimation of the respiratory signal.

For model evaluation, we used cross-validation, a standard technique in machine learning that ensures each record in the dataset is used as a test sample exactly once. This approach helps in verifying the model's ability to generalize to new, unseen data. Specifically, we implemented 5-fold cross-validation, meaning that in each iteration, the model was trained on 80% of the dataset and tested on the remaining 20%.

5.3 Results

To quantitatively evaluate the performance of our methods, we used the correlation coefficient (CC) and Mean Squared Error (MSE) metrics. The average CC and MSE were measured across each test fold to provide a comprehensive assessment of model performance. The performance of our methods, as measured across the test folds, resulted in an average validation correlation coefficient (CC) of 0.51 and an average validation Mean Squared Error (MSE) of 0.046. Using CC as the primary metric, our method outperformed 4 out of 6 signal processing algorithms.

In Figure 2 and 3 we show examples of measured and ECG-derived respiratory signal with high correlation and lower correlation.

Method	Mean CC	Mean MSE
ELF_RSlinB_FMeam_FPt_RDtGC_EHF	0.59	0.073
ELF_RSlinB_FMebw_FPt_RDtGC_EHF	0.50	0.069
ELF_RSlinB_FMefm_FPt_RDtGC_EHF	0.56	0.070
flt_BFi	0.37	0.083
flt_Wam	0.38	0.093
flt_Wfm	0.44	0.081
DL	0.51	0.046

6 CONCLUSION

In this work, we implemented a fully convolutional autoencoder network designed to estimate respiratory signals from ECG data. The network architecture, inspired by the U-Net model, was simplified to retain essential shortcut connections while reducing the number of layers and parameters. The encoder and decoder components
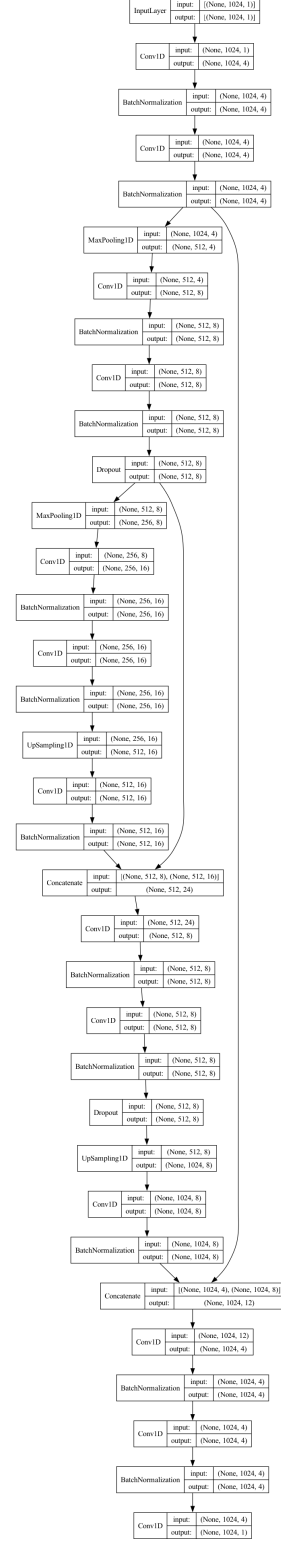

Figure 1. Model architecture

of the network utilized convolutional layers, ReLU activations, Batch Normalization, MaxPooling, and up-sampling techniques to effectively capture and reconstruct respiratory signals.

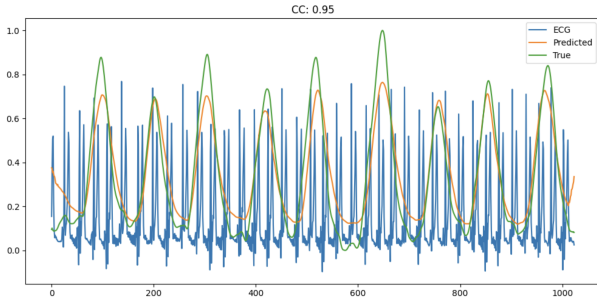We employed a sliding window technique for data preprocessing, segmenting the recordings into 32-second

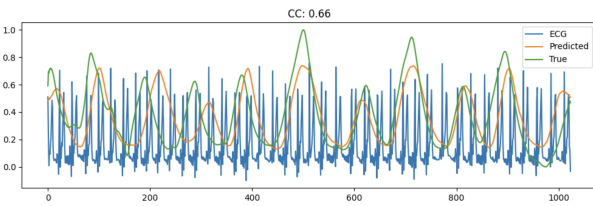Figure 2. Examples of measured and ECG-derived respiration with high correlation CC = 0.95

Figure 3. Examples of measured and ECG-derived respiration with lower correlation CC = 0.66

long segments with 16-second overlaps and down-sampling each segment to 1024 samples. The model was trained using a comprehensive procedure involving hyperparameter optimization, the Adam optimizer, Mean Squared Error (MSE) as the loss function, and additional metrics like correlation to evaluate performance. Several callbacks, including logging, early stopping, and model checkpointing, were used to ensure effective training.

For model evaluation, we utilized 5-fold cross-validation to verify the model's ability to generalize to new data. Our methods were quantitatively assessed using the correlation coefficient (CC) and MSE metrics. The performance results showed an average validation CC of 0.51 and an average validation MSE of 0.046. Using CC as the primary metric, our deep learning method outperformed 4 out of 6 signal processing algorithms.

By leveraging these datasets and establishing robust baselines with traditional signal processing methods, we provided a comprehensive comparison with our deep learning approaches. This demonstrated the effectiveness of advanced algorithms in respiratory signal estimation from ECG data. Future work will include exploring different deep learning architectures to further improve results and trying different datasets for ECG-derived respiration to enhance the generalizability and robustness of the models.

REFERENCES

[1] Peter H Charlton, Timothy Bonnici, Lionel Tarassenko, David A Clifton, Richard Beale, and Peter J Watkinson. An assessment of algorithms to estimate respiratory rate from the electrocardiogram and photoplethysmogram. *Physiological measurement*, 37(4):610, 2016.

[2] A. Goldberger, L. Amaral, L. Glass, J. Hausdorff, P. C. Ivanov, R. 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 [Online]*, 101(23):e215–e220, 2000.

[3] V Ravichandran et al. Respnet: A deep learning model for extraction of respiration from photoplethysmogram. *arXiv preprint arXiv:1902.04236*, 2019.

[4] E. Merdjanovska and A. Rashkovska. Respiration extraction from single-channel ecg using signal-processing methods and deep learning. In *2020 43rd International Convention on Information, Communication and Electronic Technology (MIPRO)*, pages 848–852. IEEE, 2020.

[5] S. A. H. Aqajari, R. Cao, A. H. A. Zargari, and A. M. Rahmani. An end-to-end and accurate ppg-based respiratory rate estimation approach using cycle generative adversarial networks. In *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pages 2029–2032. IEEE, 2021.

[6] D. J. Gauthier, E. M. Bollt, A. Griffith, and W. a. S. Barbosa. Next generation reservoir computing. *Nature Communications*, 12(1):1–8, 2021.