

Interlead Conversion of Single-Lead Blindly-Segmented Electrocardiogram Signals

Sofia C. Beco^{1,2}, João Ribeiro Pinto^{*1,2}, Jaime S. Cardoso^{1,2}

¹ Faculdade de Engenharia, Universidade do Porto, Porto, Portugal.

² Centre for Telecommunications and Multimedia, INESC TEC, Porto, Portugal.

*corresponding author (joao.t.pinto@inesctec.pt)

Keywords: autoencoder, conversion, electrocardiogram (ECG), leads, U-net.

Abstract. The standard configuration's set of twelve electrocardiogram (ECG) leads is optimal for medical diagnosis of diverse cardiac conditions. However, it requires ten electrodes on the patient's limbs and chest, which is uncomfortable and cumbersome. Interlead conversion methods can reconstruct missing leads and enable more comfortable acquisitions that still allow for adequate diagnoses. This work contributes toward this goal by studying interlead conversion using single-lead non-aligned input ECG segments, exploring three architectures based on the encoder-decoder structure. Despite the significantly more challenging scenario, the proposed methodology was able to achieve state-of-the-art results both on the PTB database and in cross-database tests.

1 Scientific Background

The electrocardiogram (ECG) is the measurement of electrical potentials that make the heart contract and relax as intended. The morphology of the ECG signal depends on the location of the electrodes used for acquisition: different electrode placement results in different perspectives over the heart [1]. For medical purposes, the standard configuration acquires the ECG over twelve leads for more information, but it requires ten electrodes placed on the patient's arms, legs, and chest. Using fewer electrodes allows for more comfortable and inexpensive acquisitions, at the expense of certain leads that could be ideal for a more accurate diagnosis of certain conditions.

To get the best of both worlds, researchers have proposed methods for the automatic interlead conversion of ECG signals [2–6]. These transform short ECG segments to mimic other perspectives, using acquired leads to reconstruct any leads that were not recorded. However, these methods still present limited applicability, since they typically require multiple leads as input. Even the most advanced methods [4,5], that only use one input lead, still require the inputs to be single heartbeat segments aligned in time, which makes them dependent on separate processes and, overall, less flexible and robust.

This paper presents a study on the feasibility of ECG interlead conversion using short segments from just one lead without any kind of time alignment (blindly-segmented). With such input, the proposed methodology based on an encoder-decoder structure is trained to reconstruct other leads as faithfully as possible. This opens up new possibilities for more comfortable ECG acquisition in clinical scenarios without giving up the benefits of multi-lead recordings for medical diagnosis.

2 Methodology

2.1 General overview

The proposed methodology for interlead ECG conversion follows the encoder-decoder structure typically used for deep image segmentation. The encoder receives an input signal and processes it to create a compressed representation that retains relevant information for the task at hand. The decoder receives this representation and processes it so that the output matches the ground truth as closely as possible. Here,

the input to the encoder is a short ECG segment of one lead (X) and the ground-truth is the corresponding segment in a different lead (Y). Thus, the encoder is in charge of selecting the information from X that is needed for Y, and the decoder will use that information to reconstruct the corresponding lead Y signal.

2.2 Model architectures

The general encoder-decoder structure allows for diverse specific model architectures. In this work, three architectures were implemented and compared, based on convolutional autoencoders (AE), U-Nets, and Label Refinement Networks (LRN):

- **Convolutional Autoencoder** (see Fig. 1(a)): This architecture receives an input segment of lead X, which initially goes through a chain of three sequential blocks, each with half the signal resolution of the previous block. Each block includes two convolutional layers (each followed by batch normalisation and ReLU activation) and ends with a max-pooling layer. Between the encoder and the decoder, two convolutional layers compose the latent space or bottleneck block, which corresponds to the maximum point of information compression. The decoder mirrors the encoder, with three similar blocks composed of an upsampling layer and two transposed convolutional layers. The last transposed convolutional layer outputs a single-channel signal whose size corresponds to the input segment. Its activation function is the hyperbolic tangent for an output signal with amplitudes in $[-1, 1]$;
- **U-Net** (see Fig. 1(b)): The U-Net was proposed by Ronneberger *et al.* [7] as a tool for biomedical image segmentation. In this case, it was used for interlead conversion. The architecture is very similar to the autoencoder, however, the feature maps from the encoder blocks are directly routed to the corresponding decoder blocks through skip connections. This allows the model to propagate context information from higher resolution between the encoder and the decoder;
- **Label Refinement Network** (see Fig. 1(c)): The Label Refinement Network (LRN) was originally proposed by Islam *et al.* [8] for semantic image segmentation. As implemented for interlead conversion in this work, its architecture is identical to the aforementioned U-Net. The singularity of the LRN lies in the supervision strategy: while the U-Net only uses the output of the last decoder block in the reconstruction loss, the LRN computes the loss at the outputs of every decoder block. This results in supervision at several resolution levels, leading the decoder to offer a coarse reconstruction right after the first block, which should be gradually refined by the subsequent blocks for improved results at higher resolutions.

3 Experimental Setup

3.1 Data

The experiments conducted in this work used mainly the data provided in the PTB database [9], available on Physionet [10]. The PTB database includes data from 16 channels, including all 12 standard leads, sampled at 1 kHz. It contains a total of 549 records from 290 individuals, with one to five records per subject. Recordings were cropped into segments of 5 s (5000 samples). A second-order Butterworth bandpass filter with cutoff frequencies $f_c = [1, 40]$ Hz was applied to each segment to remove noise while retaining the most useful ECG information. The amplitudes of the n values of each signal x were then min-max normalised to the interval $[-1, 1]$:

$$x_n = 2 \times \frac{x_n - x_{min}}{x_{max} - x_{min}} - 1 \quad (1)$$

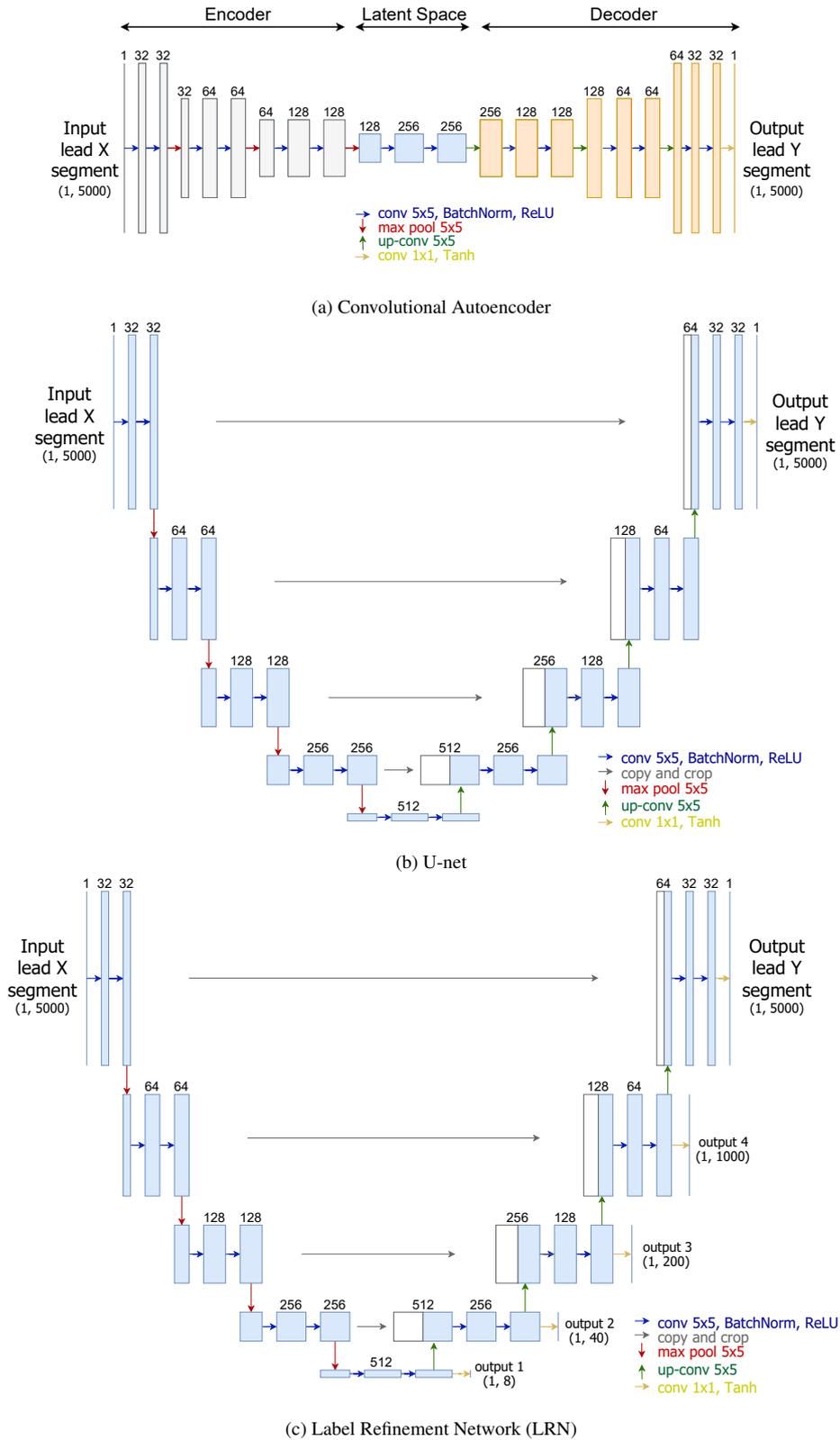


Figure 1: Structure of the implemented encoder-decoder architectures.

The data from PTB was divided into train and test sets, with approximately 63%, 7% and 30% of the segments, respectively, for a total of 7086, 787, and 3509 ECG segments for each set. For a more thorough and challenging evaluation, subjects are divided between the train/validation and test sets: the latter had recordings from subjects 1 to 50 while the former had recordings from subjects 51 to 290.

The INCART database, also available on Physionet, was used to test the performance of trained models on cross-database scenarios. This database contains 75 Holter recordings from 32 subjects undergoing tests for coronary artery diseases. Each record is 30

minutes long and contains twelve standard leads sampled at 257 Hz. Recordings from this database were resampled to 1 kHz and processed as described above for PTB.

3.2 Model training and evaluation

The models were trained using the $l1$ -loss between the model outputs and the corresponding ground-truth signals as the objective function. The Adam optimiser was used with an initial learning rate of 1×10^{-3} , over a maximum of 500 epochs with batch size 64 and early stopping patience of 50 epochs.

To compare conversions with the corresponding ground-truth signals, this work used the following metrics: the average and median Pearson correlation coefficient (r , used in the majority of the related literature), the average RMSE, the average Structural Similarity Index Measure (SSIM), and the average Dynamic Time Warping (DTW).

4 Results and Discussion

4.1 One-to-one lead conversion

To compare the selected architectures, the first experiment entailed the conversion from lead II to lead I, two of the most used ECG leads for medical purposes (see Table 1). According to most metrics, the U-net offers the best results. Its skip-connections give it the capability to send more information (and at more resolution levels) from the encoder to the decoders. The multi-resolution supervision of the LRN, expected to improve overall performance, appears to excessively draw the model’s attention away from the details, which ultimately harmed performance.

Table 1: Test results of one-to-one lead conversion.

Model	r (avg.)	r (median)	DTW	RMSE	SSIM
Autoencoder	0.67	0.78	542.78	8.33	0.58
U-Net	0.69	0.78	579.25	8.89	0.65
LRN	0.65	0.75	562.43	8.80	0.63

4.2 One-to-many leads conversion

Not all leads can be converted equally: the correlation between leads depends on their perspectives of the heart. While some leads (see Table 2), such as aVF or aVR are highly (positively or negatively) correlated with lead II, aVL is almost orthogonal due to their electrode placement. Hence, aVL would be much harder to convert from lead II than any other lead.

Table 2: Average correlation between lead II signals and other leads.

	I	III	aVR	aVL	aVF
r (avg.)	0.45	0.36	-0.71	0.01	0.77

This is verified in the results of multi-lead conversion using the U-net model (see Table 3). Conversion from lead II to aVF consistently offers the best results, while the conversions to lead III or aVL are overall the least successful. That is also visible in the examples of Fig. 2 where the model is unable to capture the finer details of the signals in lead III and lead aVL. Results also show that using an individual encoder for each decoder instead of a common shared encoder enables achieving considerably better results, aligned with those obtained in one-to-one conversion. This is because the former allows the model to learn to encode the specific information needed for each output lead instead of common representations.

4.3 Cross-database evaluation

Cross-database performance was evaluated using the previously trained U-net for lead II to lead I conversion on the INCART database. The results (see Table 4) are

Table 3: Test results of the U-net used for multi-lead conversion, with a shared encoder or with one encoder for each decoder.

Lead	Shared Encoder					Individual Encoders				
	r (avg.)	r (median)	DTW	RMSE	SSIM	r (avg.)	r (median)	DTW	RMSE	SSIM
I	0.37	0.47	652.74	10.34	0.48	0.69	0.78	579.25	8.89	0.65
III	0.42	0.58	646.96	10.55	0.54	0.57	0.68	742.87	11.57	0.60
aVR	-0.62	-0.75	942.31	16.76	0.15	0.91	0.95	341.35	5.38	0.78
aVL	-0.62	-0.75	777.45	12.75	0.29	0.49	0.65	769.51	12.28	0.66
aVF	0.79	0.87	434.61	7.09	0.73	0.84	0.90	446.08	6.99	0.70

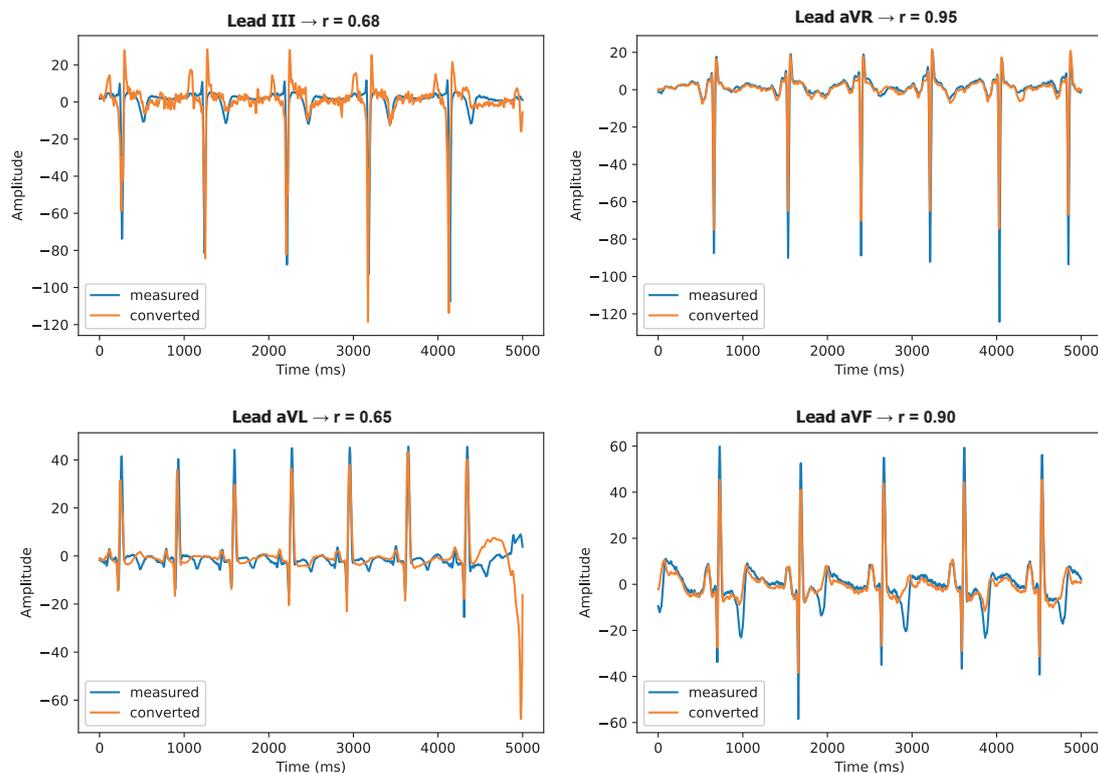


Figure 2: Examples of multi-lead conversions from lead II corresponding to the median r of the model (ground-truth in blue, conversion in orange).

slightly worse than those with the PTB database, as expected since the models were trained on PTB data and the INCART database is arguably more challenging regarding noise and variability. Such results would likely improve with further signal preprocessing. Nevertheless, in Figure 3, it is noticeable that the overall morphology of the signals was mostly successfully converted even for the lowest r quartile.

Table 4: Cross-database test results for INCART lead II to lead I conversion.

r (avg.)	r (median)	DTW	RMSE	SSIM
0.56	0.61	732.55	11.15	0.56

5 Conclusion

This work implemented and compared the performance of three deep learning architectures for interlead conversion of ECG signals. Unlike the literature, this work focused on the more challenging scenario of single-lead blindly-segmented ECG inputs. Despite this, the proposed methodology based on a U-net achieved state-of-the-art results in one-to-one and one-to-many experiments and promising results in cross-database scenarios. Nevertheless, further efforts should be devoted to expanding this study to all twelve standard leads, to use other leads as input, to perform experiments in larger databases, and to investigate the effect of medical conditions in the performance of ECG conversion.

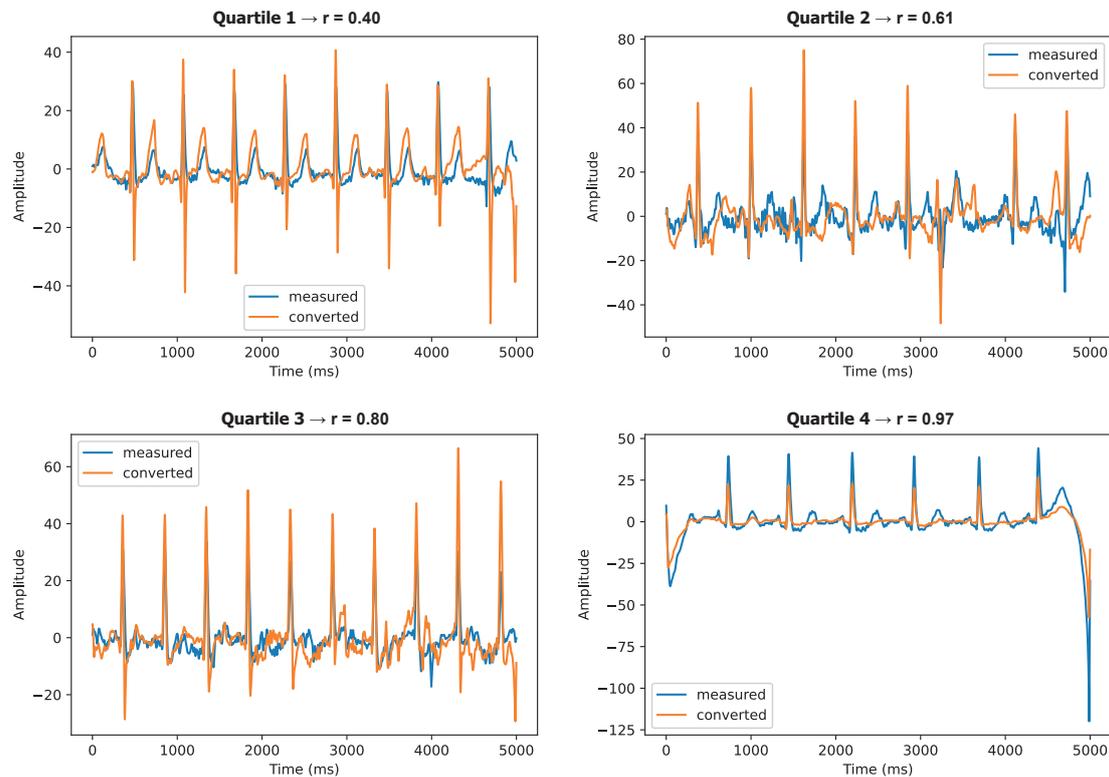


Figure 3: Cross-database results for example r quartiles (ground-truth in blue, conversion in orange).

Acknowledgments

This work is financed by National Funds through the Portuguese funding agency, FCT - Fundação para a Ciência e a Tecnologia within project UIDB/50014/2020 and the PhD grant “SFRH/BD/137720/2018”.

References

- [1] J. R. Pinto, J. S. Cardoso, and A. Lourenço, “Evolution, Current Challenges, and Future Possibilities in ECG Biometrics,” *IEEE Access*, vol. 6, pp. 34746–34776, 2018.
- [2] I. dos Santos Silva, J. R. Barbosa, R. D. de Sousa, I. F. B. de Souza, R. de Aguiar Hortegal, and C. D. M. Regis, “Comparison of spatial temporal representations of the vectorcardiogram using digital image processing,” *Journal of Electrocardiology*, vol. 59, pp. 164–170, Mar. 2020.
- [3] J. Sohn, S. Yang, J. Lee, Y. Ku, and H. C. Kim, “Reconstruction of 12-Lead Electrocardiogram from a Three-Lead Patch-Type Device Using a LSTM Network,” *Sensors*, vol. 20, p. 3278, June 2020.
- [4] J.-E. Lee, K.-T. Oh, B. Kim, and S. K. Yoo, “Synthesis of Electrocardiogram V-Lead Signals From Limb-Lead Measurement Using R-Peak Aligned Generative Adversarial Network,” *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 5, pp. 1265–1275, 2020.
- [5] M. Matyschik, H. Mauranen, P. Bonizzi, and J. Karel, “Feasibility of ECG reconstruction from minimal lead sets using convolutional neural networks,” in *Computing in Cardiology*, 2020.
- [6] G. H. Smith, D. J. Van den Heever, and W. Swart, “The reconstruction of a 12-lead electrocardiogram from a reduced lead set using a focus time-delay neural network,” *Acta Cardiologica Sinica*, vol. 37, no. 1, p. 47, 2021.
- [7] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI 2015)*, p. 234–241, 2015.
- [8] M. A. Islam, S. Naha, M. Rochan, N. Bruce, and Y. Wang, “Label Refinement Network for Coarse-to-Fine Semantic Segmentation,” *arXiv e-prints*, Mar. 2017. arXiv:1703.00551.
- [9] R. Boussejot, D. Kreiseler, and A. Schnabel, “Nutzung der EKG-Signaldatenbank CARDIODAT der PTB über das Internet,” *Biomedizinische Technik*, vol. 40, no. s1, pp. 317–318, 1995.
- [10] A. L. Goldberger, L. Amaral, L. Glass, J. M. Hausdorff, P. 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, Jun. 2000.