Transfer Learning for ECG-Based Age Estimation from Adult to Pediatric Populations

Sara Battiston[1], Niccolò Gonzato[1], Md Moklesur Rahman[1], Massimo Walter Rivolta[1], Antonio Sanzo[2], Irene Raso[3], Sara Santacesaria[3], Gianvincenzo Zuccotti[4,5], Savina Mannarino[3], Roberto Sassi[1]

[1] Computer Science Department, University of Milan, Milan, Italy
[2] Fondazione IRCCS Policlinico San Matteo, Pavia, Italy
[3] Pediatric Cardiology Unit, Buzzi Children's Hospital, Milan, Italy
[4] Pediatric Department, Buzzi Children's Hospital, Milan, Italy
[5] Department of Biomedical and Clinical Science, University of Milan, Milan, Italy

Abstract

Electrocardiogram (ECG)-based age prediction has emerged as a promising tool in medical AI, providing insights into physiological aging and potential health risks. While existing deep learning models have shown strong performance on adult populations using 10-second ECG recordings, their applicability to pediatric subjects remains largely unexplored. In this study, we tackled restrains set by the limited availability of pediatric ECG data by adopting a transfer learning approach: we first trained a convolutional neural network on single heartbeats from adult ECGs taken from the PTB-XL database. Then, we fine-tuned it on pediatric ECGs collected at the Buzzi Children Hospital, Milan, Italy. Our model achieved a RMSE of 10.32 years and a MAE of 8.03 years on adult data, which were found comparable to prior works trained on longer segments of ECG signals. In the pediatric dataset, the model achieved a RMSE of 2.67 years and a MAE of 1.88 years. These results suggest that meaningful age-related features can be extracted even from single heartbeats and that transfer learning enables effective adaptation across age groups, offering a practical solution for pediatric age estimation or in other contexts where available data might be typically more scarce.

1. Introduction

Amplitude and duration of waveforms, as measured on the electrocardiogram (ECG), are known to be affected by age. This phenomenon was leveraged to estimate the functional age of the heart, using modeling techniques ranging from statistics [1] to Deep Learning (DL) [2, 3]. Large differences between the functional and chronological age ($>$7 years) of the patients, estimated through neural networks, were found associated with higher risk of mortality [3] or cardiovascular comorbodities [2]. A review on the topic can be found in [4]. While existing studies have demonstrated the clinical relevance of ECG-based age prediction in adults, their applicability to pediatric populations remains largely unexplored. Adapting these models to younger age groups could enhance diagnostic and prognostic capabilities, providing developmental monitoring and early detection of heart conditions. However, pediatric ECGs present unique challenges due to rapid age-related changes in morphology, hormonal shifts, and the maturation of the cardiac conduction system. These factors complicate the development of reliable age predictors for younger patients. So far, only a few AI models have been specifically developed for pediatric ECGs. A recent example is the work of Dutenhefner *et al.* [5], who proposed a ResNet-based model for pediatric age regression, demonstrating that error in age estimation $>$ 2.5 years were linked to underlying pathologies. A further significant challenge is the usually limited availability of pediatric ECG data. In this study, we trained a DL age regressor on adult ECG data, and then applied transfer learning (TL) to adapt the model so to be able to deal with pediatric ECGs. In particular, instead of using the entire 10 s diagnostic ECG, we used single 12-lead beats as input of the age regressor. The work could be of interest for other situations where data availability is not homogeneous.

2. Methods

2.1. Dataset and preprocessing

Two different datasets were employed in the study. ECGs of adult subjects ("AD") were obtained from the

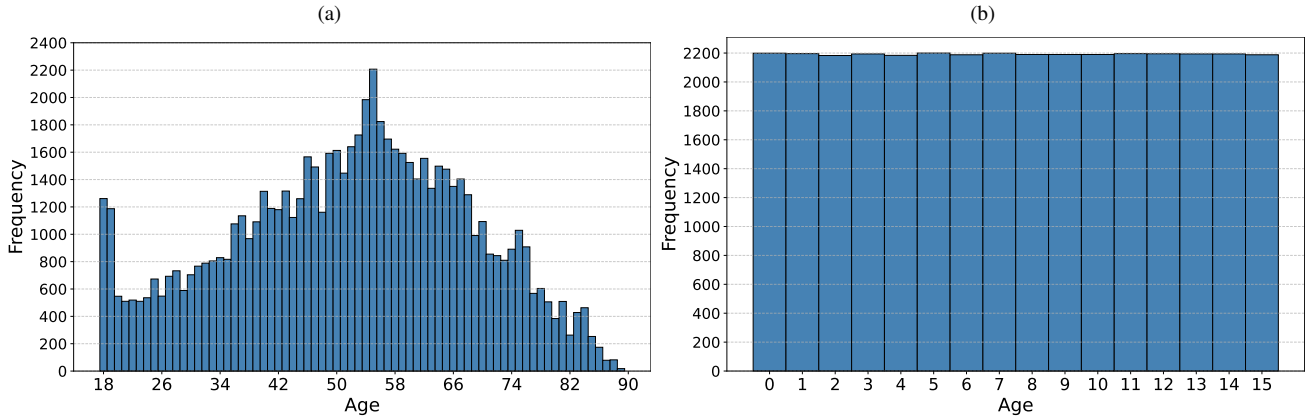

Figure 1: Distribution of the QRTS windows (heartbeats) used across different ages, as obtained for the study from the PTB-XL dataset (AD, panel a) and the Buzzi Children's Hospital pediatric ECG dataset (PE, panel b).

PTB-XL ECG dataset [6, 7], which includes standard 12-lead 10-second ECGs, sampled at 100 Hz, from $18,885$ different patients. We selected a subset of $6,776$ ECGs of healthy patients, reported as in normal sinus rhythm with at least 80% confidence, and spanning an age range $18-89$ years. A second pediatric dataset ("PE") was instead collected at the Buzzi Children's Hospital in Milan, Italy. A total of 54,399 standard 12-lead 10 seconds ECGs were digitally collected between 2011 and 2022 (sampling rate 500 Hz) from healthy children between 0 to 15 years.

Given that the PE data were sampled at a higher rate, we subsampled each pediatric ECG to 100 Hz. For both PE and AD data, we applied a zero-phase, order-3 Butterworth band-pass filter with low and high cutoff frequencies of 0.5 Hz and 40 Hz, respectively. Then, we identified the heartbeats using the `gqrs` detector included in the `WFDB` library [7, 8]. Lastly, we divided each of the PE and AD signal into QRST windows of duration 0.43 s, ranging from -60 ms to 370 ms relative to each detected R peak. The data in the PE dataset were collected in a pediatric hospital and therefore, due to clinical practice, around half the patients are under 1 year and the population is in general skewed towards younger ages. Given the fact that in this work the focus is on verifying how effectively transfer learning is capable of adapting an age regressor model from a large adult population to a smaller pediatric population, we downsampled each pediatric age group to about 2,200 QRST windows, roughly matching the maximum class size in the AD dataset. Overall, the total number of QRST windows extracted were 72,420 for AD, and 35,080 for PE. Figure 1 shows their number across age. For training the DL models, the set of QRST windows for each of AD and PE was split roughly into 90% for training and 10% for testing, with stratification, ensuring that beats belonging to the same patients were contained entirely within one of the splits.

2.2. Regression models & transfer learning

The problem of age estimation was framed as a regression task, where the target variables were integer-valued ages (for coherence with the the PTB-XL ECG where only integer-valued ages are available). The DL model consisted into two convolutional blocks, each made by a 2-d convolutional layer (Conv2d), a batch normalization layer (BatchNorm2d), a leaky relu (LeakyReLu) activation function, a max pooling layer (MaxPool2d), and a dropout layer (Dropout2d). The output was then set in input to another convolutional block, and then to a series of fully connected layers (Linear layer - ReLU activation - Dropout layer - Linear layer - ReLU activation - Linear layer) which had as output a single value for the predicted age. The input of the network was a 43×12 matrix, that is a single ECG QRST window (12 leads of 0.43 s). A scheme of the model is shown in Figure 2. As loss function, we used the mean squared error between the predicted age and the known age. The model was trained on the PTB-XL train dataset (65,166 QRTS windows or heartbeats). The training went on for 15 epochs, with a batch size of 32 and a learning rate of 10^{-4}. Adam optimizer was configured with hyperparameters $\beta = (0.9, 0.999)$, $\epsilon = 10^{-8}$, $\text{weight\_decay} = 10^{-3}$.

To establish whether the model architecture was sufficiently effective, we also retrained for comparison on the same AD training set the models proposed in [2] and [3], and compared their performance with our model. The models described in [2] and [3] were here adapted to accept in input a single heartbeat instead of 10-second ECG signals, as originally proposed.

After training the age regressor on the adult population, we transferred the learned features to pediatric ECGs using transfer learning.

Specifically, we fine-tuned the entire network without

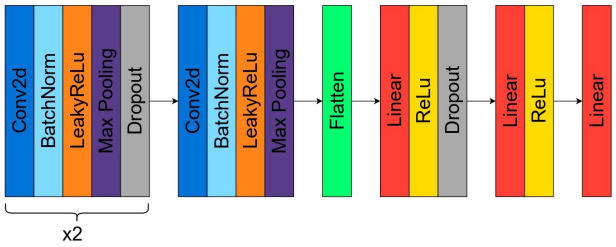

Figure 2: Architecture of the DL age regressor model including convolutional blocks and linear layers.

freezing any parameters. The training process involved 31,629 QRST windows from the PE training dataset and it was conducted for 3 epochs, with a batch size of 16 and a learning rate of $2 \cdot 10^{-5}$, to ensure that the network's previously acquired knowledge was effectively retained and adapted to the new data.

3. Results

In order to evaluate the goodness of the predictions of our age regression model, we selected the mean absolute error (MAE), the root mean square error (RMSE) and the Pearson correlation coefficient (ρ) between the predicted ages and the actual age. Specifically, on the PTB-XL test set (7,254 QRST windows), we quantified a MAE of 8.03 years, a RMSE of 10.32 years and a ρ of 0.77. A scatterplot showing the real and predicted ages on the adults test set against the main diagonal line (perfect prediction, in red) is shown in Fig. 3(left). Table 1 reports the comparison between ρ and RMSE values of our age regressor and those of Attia *et al.*'s and Lima *et al.*'s models trained and tested on the same AD data.

After the transfer learning, the DL regressor proposed in this study achieved a MAE of 1.88 years, a RMSE of 2.67 years and a ρ of 0.84 on the pediatric test set (3,451 signals). The scatterplot of the corresponding actual vs. predicted ages is reported in Fig. 3 (right).

4. Discussion

The DL model proposed in this study, when trained on single 12-leads heartbeats from the PTB-XL dataset achieved a good predictive performance over a 71-year age range, which we considered acceptable. Prediction errors were more pronounced above 75 years (underestimation), likely due to data scarcity in the dataset for this age range (Fig. 1a).

Despite previous studies suggested that heart rate variability might be a relevant features for age prediction [1], our model was able to capture meaningful age-related patterns using only ECG data. In addition, the model achieved results in line with previous studies using DL for the same task on adults. In fact, despite using only ECG data coming from a single heartbeat, our model reached a MAE of 8.03 years, while Attia *et al.* [2] and Lima *et al.* [3] reported MAEs of 6.9 and 8.38 years, respectively, using all the 10 s of diagnostic ECGs. Moreover, when we retrained these architectures on the same dataset, the performances of the three models became undistinguishable, as shown in Table 1. This suggests that the information learned in these DL models are mainly associated with the characteristics of the ECG waveforms, and not their variability in time.

After fine-tuning the proposed model on the pediatric ECGs using transfer learning, we achieved a RMSE of 2.67 years and a correlation coefficient ρ similar to the adult model, indicating effective adaptation to the younger age group. This lower RMSE, compared to the adult model, was consistent with the narrower pediatric age range (0–15 years), where smaller errors were expected. Our model also reached a MAE of 1.88 years, which was comparable to the 2.65 years reported by Dutenhefner *et al.* [5] using entire 10 s ECGs. This seems to conform that, even in children, age-related information is mainly contained in the shapes of the waveform, not their evolution in time. As shown in the scatterplot in Fig. 3, predictions closely distributes along the main diagonal (in red, dashed), though a slight underestimation persisted in the 12–15 age range, possibly inherited from the adult-trained representation of information in the model (not due to the data imbalance, as the age group were balanced in this dataset).

Overall, the transfer learning approach proved mostly effective in preserving key features learned from adult ECGs and adapting them to pediatric data. Retraining on the PE dataset required only 3 epochs (versus 16 for adults) and used a dataset roughly half of AD. This efficacy highlights the method's potential in situations with limited data availability, or in which it is necessary to adapt the model to new datasets.

Nonetheless, our approach had its limitations. While the primary aim was to develop a pediatric-specific age regressor, a more comprehensive solution would have needed ECG data spanning the full age range (0–17 years). However, our PE dataset included too few samples from adolescents aged 16–17, limiting the model's ability to learn a proper representation in this age group, which was excluded by design.

5. Conclusion

The results we obtained highlight the ability of transfer learning to adapt a well-functioning model of age regression from single-beat 12-lead ECG signals, trained on an adult population, to pediatric age groups, despite the limited data and the large variability of the pediatric ECG signals. Possible future works will focus on extending the model for processing 10 s ECGs (*e.g.*, to also include heart

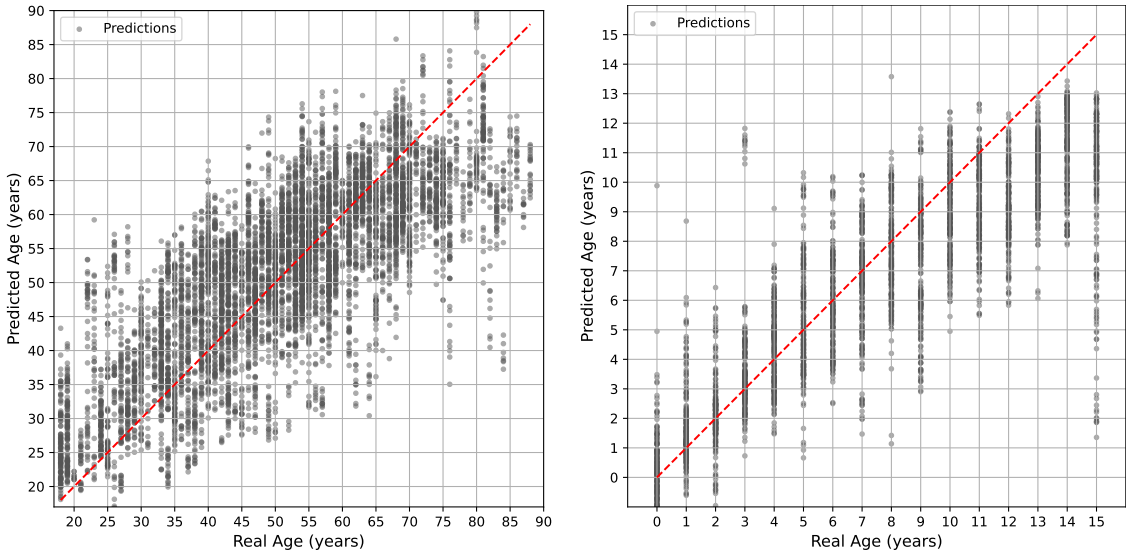

Figure 3: Scatterplots of the actual subject's age vs. the predicted age in the adult (left) and pediatric (right) populations.

Table 1: Performance comparison between the age predictor proposed in Attia *et al.*, Lima *et al.* and in this study, when all models are trained and tested on the AD dataset.

Model	Testing Loss (y)^2	ρ	RMSE (y)
This study	106.57	0.77	10.32
Attia *et al.* [2]	116.40	0.75	10.79
Lima *et al.* [3]	108.00	0.77	10.39

rate variability in the model), and evaluating the functional age for cardiac risk prediction in the pediatric population.

Acknowledgments

This work was partially supported by the project FAIR (Future Artificial Intelligence Research) project, funded by the NextGenerationEU program within the PNRR-PE-AI scheme (M4C2, Investment 1.3, Line on Artificial Intelligence).

References

[1] Starc V, Leban MA, Šinigoj P, Vrhovec M, Potočnik N, Fernlund E, Liuba P, Schlegel TT. Can functional cardiac age be predicted from the ECG in a normal healthy population? Comput Cardiol 2012;39:101–104.

[2] Attia ZI, Friedman PA, Noseworthy PA, Lopez-Jimenez F, Ladewig DJ, Satam G, Pellikka PA, Munger TM, Asirvatham SJ, Scott CG, et al. Age and sex estimation using artificial intelligence from standard 12-lead ECGs. Circ Arrhythm Electrophysiol 2019;12(9):e007284.

[3] Lima EM, Ribeiro AH, Paixão GM, Ribeiro MH, Pinto-Filho MM, Gomes PR, Oliveira DM, Sabino EC, Duncan BB, Giatti L, et al. Deep neural network-estimated electrocardiographic age as a mortality predictor. Nat Commun 2021; 12(1):5117.

[4] Ansari MY, Qaraqe M, Charafeddine F, Serpedin E, Righetti R, Qaraqe K. Estimating age and gender from electrocardiogram signals: A comprehensive review of the past decade. Artif Intell Med 2023;146:102690. ISSN 0933-3657.

[5] Dutenhefner PR, Lemos G, Rezende T, Fernandes JG, Tuler D, Pappa GL, Paixao GM, Ribeiro ALP, Meira Jr W. ECG-ResNeXt: Age prediction in pediatric electrocardiograms and its correlations with comorbidities. In Encontro Nacional de Inteligência Artificial e Computacional (ENIAC). SBC, 2024; 49–60.

[6] Wagner P, Strodthoff N, Bousseljot R, Kreiseler D, Lunze F, Samek W, Schaeffter T. PTB-XL, a large publicly available electrocardiography dataset (version 1.0.3), 2022. PhysioNet.

[7] Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. Circulation 2000;101(23):e215–e220.

[8] Xie C, McCullum L, Johnson A, Pollard T, Gow B, Moody B. Waveform Database Software Package (WFDB) for Python, 2022. PhysioNet.

Address for correspondence:

Sara Battiston
Department of Computer Science
University of Milan
Via Celoria 18, 20133, Milan, Italy
sara.battiston@unimi.it