Demographic-Conditioned State Space Models for ECG-Based Age Estimation

Benjamin Bracke[1,2⋆], Andreas Stang[2], Börge Schmidt[2], and Christoph M. Friedrich[1,2]

[1] Department of Computer Science, University of Applied Sciences and Arts Dortmund, 44227 Dortmund, Germany
[2] Institute for Medical Informatics, Biometry and Epidemiology, University Hospital Essen, 45130 Essen, Germany
benjamin.bracke@fh-dortmund.de

Abstract. This study introduces a novel demographic-conditioned architecture that leverages Mamba2 State Space Models (SSMs) and a dual Attention Aggregation mechanism to integrate 12-lead ECGs with patient sex metadata for biological age estimation (ECG-age). The architecture efficiently captures long-range physiological dependencies, while an attention-based aggregation mechanism renders the model invariant to lead availability and signal duration. Trained on the full CODE cohort ($N = 1.3$ million) and validated on the CODE-15 % test set ($N = 233k$), the model achieves a mean absolute error of 6.72 ± 5.81 years ($R^2 = 0.80$) in chronological age prediction. This performance surpasses a common used 1D-ResNet baseline (8.44 ± 7.12 years, $R^2 = 0.69$). This improved precision strengthens ECG-age as a prognostic biomarker: patient hearts predicted to be over 8 years older than their chronological age exhibited a higher Hazard Ratio (HR) for all-cause mortality (HR 2.06, 95 % CI $1.92 - 2.21$) compared to the 1D-ResNet baseline (HR 1.80, 95 % CI $1.68 - 1.92$) on the CODE-15 % test set. These results demonstrate that the SSM-based architecture yields an ECG-age marker that is prognostically superior to current baseline models. Code is available at: https://github.com/B-Bracke/demographic-ssm-ecg-age

Keywords: ECG · Biologically Age Estimation · State Space Models.

1 Introduction

Cardiovascular diseases (CVDs) remain the leading cause of death globally, necessitating accurate and accessible tools for early risk stratification [26]. While chronological age (the time elapsed since birth) is a fundamental component of clinical risk scores [7,9], it often fails to capture individual variations in physiological decline. Conversely, the "biological age" concept aims to quantify these individual aging trajectories, hypothesizing that biological aging is accelerated by adverse environmental conditions and specific pathologies [12]. While biomarkers

⋆ Corresponding Author

such as vascular stiffness or molecular profiles [6] effectively measure biological aging, they often require complex, invasive, or expensive procedures that limit their utility in widespread screening [12]. The 12-lead electrocardiogram (ECG), a ubiquitous and low-cost diagnostic tool, has emerged as a promising proxy for biological aging [1]. Aging induces subtle structural and functional adaptations in the myocardium, such as ventricular hypertrophy [3] and fibrosis [5], which manifest as distinct patterns in electrical activity. Recent advances in Deep Neural Networks (DNNs), particularly Convolutional Neural Networks (CNNs), have enabled "end-to-end" modeling of these patterns, predicting a patient's age directly from raw ECG traces [2,4]. The discrepancy between this predicted "ECG-age" and the patient's chronological age serves as a biomarker. Patients with an accelerated ECG-age (e.g., > 8 years above chronological age) exhibit higher risks of mortality and cardiovascular events [13,16].

However, CNN-based architectures face inherent limitations due to their restricted receptive fields, struggling to efficiently capture the long-range temporal dependencies and global waveform dynamics critical for modeling subtle, long-term physiological markers scattered across the cardiac cycle [1,13,16,18]. Furthermore, CNN-based models are restricted by fixed input dimensions, treating the 12-lead ECG as a static sequence rather than a dynamic series with varying durations or potentially noisy leads. To address these challenges, an architecture is introduced that leverages novel State Space Models (SSMs) [11]. Unlike CNNs, SSMs efficiently model long-range dependencies and can capture the global context of the ECG waveform, as demonstrated by ECGMamba [20], S2M2ECG [27], and MSECG [17]. Additionally, a dual attention aggregation mechanism dynamically fuses 12-lead ECG data, rendering the model invariant to signal duration and missing leads, which potentially enables deployment on real-world, lead-reduced, or noisy data. (i.e. from wearable devices). Simultaneously, the attention weights offer inherent explainability by isolating critical leads and time segments. Furthermore, the architecture features a multimodal-conditioned design, allowing it to deeply fuse patient metadata (e.g., demographic data such as sex) into the network to enhances feature extraction.

2 Methodology

The proposed approach formulates biological age estimation as a multimodal regression task using 12-lead ECG signals at 400 Hz (12×4096 samples, approx. 10 sec.) and sex. The objective is to learn a non-linear mapping from cardiac electrical activity to a continuous age value using the following novel architecture.

2.1 Model Architecture

The proposed architecture (Figure 1) integrates a convolutional front-end for local feature extraction with a deep, bidirectional State Space Model backbone for global context modeling. The pipeline processes 12-lead ECGs as spatio-temporal feature sequences, culminating in hierarchical attention aggregation

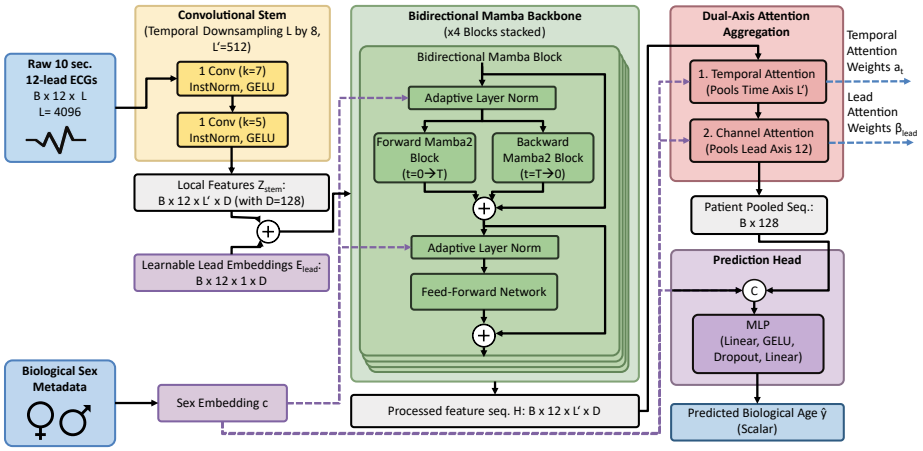

Fig. 1: Schematic of the proposed model architecture for ECG-age prediction and its core components: a convolutional stem to extract local features, a spatial context processed by a bidirectional Mamba2-based backbone with adaptive layer norm, and a dual-axis attention aggregation module to perform sequential temporal and channel-wise pooling for biological age prediction via a MLP head.

and a multi-layer perceptron for direct age regression.

Convolutional Stem The raw ECG input $X \in R^{B\times 12\times L}$ (where B is the batch size and $L = 4096$ is the sequence length) is tokenized by a stem of two 1D convolutional layers ($k = 7, k = 5$) to extract local features while reducing temporal redundancy. Each layer employs Instance Normalization and GELU activation. This stage projects inputs to a latent feature dimension $D = 128$ and downsamples the temporal resolution by a factor of 8, yielding the feature map $Z_{stem} \in R^{B\times 12\times L'\times D}$, where $L' = L/8 = 512$. This results in a temporal resolution of 20 ms per token (downsampled from the initial 2.5 ms at 400 Hz), which is sufficient to capture meaningful ECG morphology while reducing the computational complexity of the subsequent sequence modeling backbone.

Lead-Specific Embeddings To maintain parameter efficiency, the Mamba backbone shares weights across all 12 input leads. To preserve spatial context lost by sharing, learnable lead embeddings $E_{lead} \in R^{12\times D}$ are used. These vectors are broadcast and added to the latent features immediately after the stem via $Z_{input} = Z_{stem} + E_{lead}$. This explicitly encodes the anatomical viewing angle, enabling the shared backbone to learn lead-specific physiological markers.

Bidirectional Mamba Backbone Sequence modeling is performed by four stacked bidirectional Mamba-2 blocks [8]. Mamba leverages Selective State Space Models [10] (SSMs) to map sequences with linear complexity ($O(L)$), making it computationally superior to Transformers ($O(L^2)$) for high-resolution physio-

logical signals [25]. Since standard SSMs are inherently causal, they suffer from temporal smearing, where the latent state at time t is dominated by the accumulated history of $0 \to T$. This sequential compression blurs local morphological features, making it difficult for downstream attention modules to isolate specific events (such as the QRS complex) from the preceding context. To resolve this and restore local feature distinctiveness at every time step, a bidirectional design is employed. Two parallel SSM heads process the sequence in forward ($t = 0 \to T$) and backward ($t = T \to 0$) directions. The outputs of these heads are summed to the residual stream. This is followed by a Feed-Forward Network (FFN) for channel mixing. Adaptive Layer Normalization is applied immediately before both the SSM and FFN sub-layers, conditioning the signal flow at every stage.

Metadata Fusion Patient metadata (sex) is fused via Adaptive Layer Normalization [19] (AdaLN) within each bidirectional Mamba block. A learned projection of the sex embedding c generates scale (γ) and shift (β) parameters for the normalization layers. This conditions the feature distribution at every layer, enabling the model to learn distinct aging trajectories for male and female physiology within a unified model.

Dual-Axis Attention Aggregation The backbone-processed feature sequence H is mapped to a single patient-level vector via a hierarchical dual-axis attention pooling mechanism [15]. First, a **Temporal Attention** module learns to prioritize diagnostically relevant segments (e.g., ST-segments) over noise by computing weights α_t for each time step t. Second, a **Channel Attention** module aggregates the temporally pooled vectors by assigning importance weights β_{lead} to each lead. This dynamically down-weights noisy or missing leads, creating a robust, lead-invariant representation. Crucially, the sex embedding c serves as a context vector to modulate the attention scores, ensuring the aggregation strategy is biologically adaptive.

Prediction Head The final representation is concatenated with the sex embedding and passed through a multi-layer perceptron (two linear layers, GELU, dropout 0.2) to output the scalar predicted biological age $\hat{y}$.

2.2 Augmentations

To enhance generalization across diverse acquisition devices and out of distribution cohorts, a hierarchical augmentation pipeline is applied. **Signal-level artifacts** are simulated via random convolution ($k = 7$), additive Gaussian noise, and non-linear gamma scaling ($x' = \text{sign}(x)|x|^\gamma, \gamma \sim \mathcal{U}[0.9, 1.1]$) to mimic hardware variability and movement noise. **Structural robustness** is enforced through Time Masking (zeroing up to 20 % of the sequence) and Channel Masking (dropping up to 11 leads), compelling the attention mechanism to capture global context from partial views.

2.3 Loss Function

Since the divergence between the predicted "ECG-age" and chronological age often reflects a true biological age gap, a robust huber criterion [14] is applied for optimization. This minimizes prediction error while remaining resilient to legitimate physiological deviations, ensuring the model captures consistent aging patterns rather than overfitting to chronological labels. To stabilize optimization, target ages y were divided by 100 to normalize them to the range $[0, 1]$. The Huber δ parameter was set to 0.05 (equivalent to 5 years). To address class imbalances, which is concentrated around middle age ($50 - 60$ years) with fewer samples at the extremes ($< 30, > 80$ years), patient ages were discretized into 20 bins (0–5, . . . , 95–100 years), and sample weights based on inverse bin frequency. Finally, weights were normalized to a unit mean per batch, and the total loss was scaled by a factor $S = 200$ to prevent gradient underflow during mixed-precision (bfloat16) training.

2.4 Implementation Details

The input ECG recordings were resampled to 400 Hz. To ensure efficient training, ECG traces were fixed to a duration of 10 seconds (4096 samples). Recordings shorter than this duration were zero-padded, while longer recordings were cropped. Each lead was independently normalized to zero mean and unit variance. The model was optimized for 30 epochs on a single NVIDIA H100 GPU using the AdamW optimizer ($\beta_1 = 0.9$ and $\beta_2 = 0.95$) and mixed-precision (bfloat16) training. To ensure training stability with the Mamba-2 architecture, a large effective batch size of 1024 was combined with a cosine annealing learning rate schedule, featuring a linear warmup for the first 5 % of steps that peaked at $5 \cdot 10^{-3}$.

3 Experiments and Results

The proposed architecture was trained to integrate raw ECG signals with biologically sex metadata. To assess its reliance on this demographic data and evaluate scenarios where it is missing, two inference strategies were employed: a standard sex-conditioned prediction and a sex-agnostic approach. Rather than training a separate unimodal baseline with minor architectural changes, the sex-agnostic method marginalizes the sex variable by averaging the predicted ECG-age from male- and female-conditioned forward passes. These two configurations, hereafter termed "sex-conditioned" and "sex-agnostic" are benchmarked against the 1D-ResNet baseline proposed by [16]. To ensure a fair comparison, the 1D-ResNet baseline was re-evaluated on the exact same test datasets.

3.1 Datasets

The large-scale electrocardiographic cohort "Clinical Outcomes in Digital Electrocardiography" [22,21] (CODE) from the TeleHealth Network of Minas Gerais

Table 1: Baseline characteristics of the study cohorts (first ECG-exam only for each unique patient). Models were trained on CODE-dev and tested on CODE-15 % and SaMi-Trop.

Characteristics	**CODE-dev** ($n = 1,324,535$)	**CODE-15 %** ($n = 233,768$)	**SaMi-Trop**[a] ($n = 1,631$)
Sex, male, n (%)	532,229 (40)	94,807 (41)	550 (34)
Age (years), mean (s.d.)	52 (17)	51 (20)	60 (13)
Hypertension, n (%)	378,547 (31)	54,209 (28)	593 (36)
Diabetes, n (%)	78,031 (6)	10,930 (6)	161 (10)
Smoking, n (%)	82,709 (7)	11,743 (6)	498 (31)
Previous MI, n (%)	8,824 (0.7)	1,224 (0.6)	76 (5)
Follow-up (years), med (IQR)	3.5 (2.0–4.6)	3.6 (2.0–5.1)	2.1 (2.0–2.2)
Events, n (%)	39,102 (3.1)	6,569 (3.4)	104 (6.4)

[a]Characteristics for this dataset was derived from [16]. MI: Myocardial Infarction.

(TNMG) in Southeast Brazil was utilized for this study. Study details can be obtained from [22] while baseline characteristics are summarized in Table 1. Access to the training data was obtained through a formal data usage agreement with the CODE study investigators.

Development Dataset Model training was conducted using a CODE-dev split ($n = 1,324,535$ unique patients). This dataset comprises approx. 85 % of the full CODE cohort, excluding the official validation subset (CODE-15 %). The training population is characterized by a mean age of 52 ± 17 years, 40 % male representation, and a 3.1 % overall all-cause mortality rate over a median follow-up of 3.5 years. For internal validation of model architectural decisions, a stratified subset of $n = 96,223$ patients was used as validation dataset.

Test Datasets Model performance was evaluated on two distinct cohorts to assess both internal consistency and external validation. The officially designated hold-out subset CODE-15 % [23] of the CODE cohort ($n = 233,768$ unique patients) was utilized for internal validation. This split was pre-generated by the original study authors via stratified sampling to ensure a uniform age distribution (approx. 3,100 exams per year of age). Consequently, the mean age is slightly higher (51 ± 20 years) than the training set, while maintaining similar clinical characteristics. Generalization to severe pathology was assessed using the SaMi-Trop cohort [24] ($n = 1,631$), a prospective study of patients with chronic Chagas cardiomyopathy. As shown in Table 1, this population differs notably from the development data: patients are older (mean 60 ± 13 years), predominantly female (66 %), and present with a high-risk clinical profile, contributing to a near-double mortality rate of 6.4 % over a median follow-up of 2.1 years.

Table 2: Regression performance on the CODE-15 % and SaMi-Trop datasets. Metrics include Mean Absolute Error (MAE) ± standard deviation (STD), Pearson correlation (R), coefficient of determination (R^2), and mean predicted ages.

Metric	1D-ResNet Baseline	Proposed Model (sex-conditioned)	Proposed Model (sex-agnostic)
CODE-15 % Dataset[a]			
MAE ± STD	8.44 ± 7.12	**6.72±5.81**	6.87 ± 5.88
R; R^2	0.84; 0.69	**0.90; 0.80**	0.89; 0.79
Pred. Mean ± STD	52.1 ± 18.7	**51.6** ± 18.3	52.2 ± **18.1**
SaMi-Trop Dataset[b]			
MAE ± STD	9.96 ± 7.61	**7.83 ± 6.30**	8.06 ± 6.31
R; R^2	0.60; 0.04	**0.70; 0.38**	0.69; 0.35
Pred. Mean ± STD	62.6 ± 14.1	**62.0** ± 12.3	62.7 ± **12.0**

Real Mean ± STD for [a]CODE-15 %: 51 ± 20 and [b]SaMi-Trop: 60 ± 13.

3.2 Quantitative Results

Table 2 summarizes the biological age estimation results. The proposed architecture indicates performance improvements over the 1D-ResNet baseline, reducing the Mean Absolute Error (MAE) by approximately 1.7 to 2.1 years across both cohorts. Notably, it mitigates the baseline's severe correlation drop on the external SaMi-Trop dataset, improving R^2 from 0.04 to 0.38. Furthermore, the negligible performance difference between the sex-conditioned and sex-agnostic configurations suggests the network may not strictly require explicit metadata. This implies the raw ECG potentially captures sufficient sex-specific physiological variations, aligning with previous findings [2].

3.3 Prognostic Value

Divergence between predicted "ECG-age" and chronological age serves as a meaningful mortality predictor. To evaluate this, age- and sex-adjusted Cox regression (Table 3) stratified patients using an established 8-year threshold [16]: (1) **Overestimation** (ECG-age > age + 8 years), (2) **Underestimation** (ECG-age < age - 8 years), and (3) **Concordant** (absolute discrepancy ≤ 8 years). Internally (CODE-15 %), the proposed architecture yielded a higher hazard ratio for the overestimation group compared to the baseline, indicating a pronounced risk stratification capacity. The sex-agnostic model performed comparably, implying strong intrinsic prognostic value within the raw ECG. However, external validation (SaMi-Trop) revealed a domain gap. The sex-conditioned model retained a statistically significant association with mortality ($p = 0.02$) with tighter confidence intervals than the baseline, whereas the sex-agnostic configuration lost statistical significance ($p = 0.26$). This indicates that while explicit demographic metadata appears redundant internally, it provides critical statistical stabilization for mortality prediction when generalizing to unseen populations.

Table 3: Age- and sex-adjusted Cox regression for all-cause mortality. Hazard Ratios (HR) compare Overestimation (ECG-age > age +8 years) and Underestimation (ECG-age < age −8 years) groups against the Concordant baseline (ECG-age ≤ 8 years discrepancy).

	CODE-15 %		**SaMi-Trop**	
Model	**HR [95 % CI]**	**p-value**	**HR [95 % CI]**	**p-value**
1D-ResNet Baseline				
Underestimation	0.81 [0.76–0.86]	< 0.005	0.95 [0.55–1.64]	0.86
Overestimation	1.80 [1.68–1.92]	< 0.005	2.38 [1.51–3.74]	< 0.005
Proposed Model (sex-conditioned)				
Underestimation	0.78 [0.73–0.83]	< 0.005	0.86 [0.49–1.50]	0.59
Overestimation	2.06 [1.92–2.21]	< 0.005	1.76 [1.10–2.80]	0.02
Proposed Model (sex-agnostic)				
Underestimation	0.79 [0.74–0.84]	< 0.005	0.67 [0.36–1.23]	0.19
Overestimation	1.99 [1.85–2.13]	< 0.005	1.32 [0.81–2.13]	0.26

4 Conclusion

This study introduced a novel architecture for ECG-based biological age estimation that outperforms a 1D-ResNet baseline in estimation error on the internal CODE-15 % dataset while maintaining strong mortality risk stratification. Sex-agnostic inference yielded comparable results to the sex-conditioned approach, suggesting raw ECGs sufficiently encode sex-specific physiological variations, potentially rendering explicit sex metadata redundant. Although the proposed model showed incrementally higher hazard ratios over the baseline, its clinical impact requires further investigation. Performance drops on the external SaMi-Trop cohort highlighted a clear domain gap for both models. Future work should address this gap through broader cross-cohort validations, integrate non-redundant metadata (e.g., comorbidities), and conduct rigorous ablation studies against unimodal architectures. Finally, evaluating robustness against missing leads and utilizing explainability analyses will be crucial to link attention scores with known physiological markers of cardiac aging.

Acknowledgments. The work of Benjamin Bracke was partially funded by a PhD grant from the University of Applied Sciences and Arts Dortmund, Germany. Special thanks are extended to Antônio H. Ribeiro and the CODE study investigators for providing the data and starter scripts.

Ethics declarations. This secondary analysis of anonymized data complies with all relevant ethical regulations. The CODE study was approved by the Universidade Federal de Minas Gerais Ethics Committee (49368496317.7.0000.5149) and SaMi-Trop study by the Brazilian National Institutional Review Board (CONEP 179.685/2012).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Ansari, M.Y., Qaraqe, M., Charafeddine, F., Serpedin, E., Righetti, R., Qaraqe, K.: Estimating age and gender from electrocardiogram signals: A comprehensive review of the past decade. Artificial Intelligence in Medicine **146**, 102690 (Dec 2023). https://doi.org/10.1016/j.artmed.2023.102690
2. Attia, Z.I., Friedman, P.A., Noseworthy, P.A., Lopez-Jimenez, F., Ladewig, D.J., Satam, G., Pellikka, P.A., Munger, T.M., Asirvatham, S.J., Scott, C.G., Carter, R.E., Kapa, S.: Age and Sex Estimation Using Artificial Intelligence From Standard 12-Lead ECGs. Circulation. Arrhythmia and Electrophysiology **12**(9), e007284 (Sep 2019). https://doi.org/10.1161/CIRCEP.119.007284
3. Bacharova, L., Estes, E.H.: Left Ventricular Hypertrophy by the Surface ECG. Journal of Electrocardiology **50**(6), 906–908 (2017). https://doi.org/10.1016/j.jelectrocard.2017.06.006
4. Ball, R.L., Feiveson, A.H., Schlegel, T.T., Starc, V., Dabney, A.R.: Predicting "heart age" using electrocardiography. Journal of Personalized Medicine **4**(1), 65–78 (Mar 2014). https://doi.org/10.3390/jpm4010065
5. Biernacka, A., Frangogiannis, N.G.: Aging and Cardiac Fibrosis. Aging and Disease **2**(2), 158–173 (Apr 2011), https://pmc.ncbi.nlm.nih.gov/articles/PMC3153299/
6. Bulpitt, C.J., Rajkumar, C., Cameron, J.D.: Vascular compliance as a measure of biological age. Journal of the American Geriatrics Society **47**(6), 657–663 (Jun 1999). https://doi.org/10.1111/j.1532-5415.1999.tb01586.x
7. Cooney, M.T., Vartiainen, E., Laatikainen, T., De Bacquer, D., McGorrian, C., Dudina, A., Graham, I., SCORE and FINRISK investigators: Cardiovascular risk age: concepts and practicalities. Heart **98**(12), 941–946 (Jun 2012). https://doi.org/10.1136/heartjnl-2011-301478
8. Dao, T., Gu, A.: Transformers are SSMs: Generalized Models and Efficient Algorithms Through Structured State Space Duality. In: International Conference on Machine Learning. No. abs/2405.21060
9. Dhingra, R., Vasan, R.S.: Age as a risk factor. The Medical Clinics of North America **96**(1), 87–91 (Jan 2012). https://doi.org/10.1016/j.mcna.2011.11.003
10. Gu, A., Dao, T.: Mamba: Linear-Time Sequence Modeling with Selective State Spaces (May 2024), http://arxiv.org/abs/2312.00752, arXiv:2312.00752 [cs]
11. Gu, A., Goel, K., Re, C.: Efficiently Modeling Long Sequences with Structured State Spaces. No. abs/2111.00396 (2021)
12. Hamczyk, M.R., Nevado, R.M., Barettino, A., Fuster, V., Andrés, V.: Biological Versus Chronological Aging: JACC Focus Seminar. Journal of the American College of Cardiology **75**(8), 919–930 (Mar 2020). https://doi.org/10.1016/j.jacc.2019.11.062
13. Hempel, P., Ribeiro, A.H., Vollmer, M., Bender, T., Dörr, M., Krefting, D., Spicher, N.: Explainable AI associates ECG aging effects with increased cardiovascular risk in a longitudinal population study. npj Digital Medicine **8**(1), 25 (Jan 2025). https://doi.org/10.1038/s41746-024-01428-7
14. Huber, P.J.: Robust Estimation of a Location Parameter. The Annals of Mathematical Statistics **35**(1), 73–101 (Mar 1964). https://doi.org/10.1214/aoms/1177703732
15. Lee, J., Lee, Y., Kim, J., Kosiorek, A., Choi, S., Teh, Y.W.: Set Transformer: A Framework for Attention-based Permutation-Invariant Neural Networks. In: Proceedings of the 36th International Conference on Machine Learning. pp. 3744–3753 (May 2019), https://proceedings.mlr.press/v97/lee19d.html

16. Lima, E.M., Ribeiro, A.H., Paixão, G.M.M., Ribeiro, M.H., Pinto-Filho, M.M., Gomes, P.R., Oliveira, D.M., Sabino, E.C., Duncan, B.B., Giatti, L., Barreto, S.M., Meira Jr, W., Schön, T.B., Ribeiro, A.L.P.: Deep neural network-estimated electrocardiographic age as a mortality predictor. Nature Communications **12**(1), 5117 (Aug 2021). https://doi.org/10.1038/s41467-021-25351-7
17. Lin, J., Chiu, I., Wang, K.C., Liu, K.C., Wang, H.M., Yeh, P.C., Tsao, Y.: MSECG: Incorporating Mamba for Robust and Efficient ECG Super-Resolution. In: ICASSP 2025 - 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). pp. 1–5. IEEE, Hyderabad, India (Apr 2025). https://doi.org/10.1109/ICASSP49660.2025.10890523
18. Lindow, T., Maanja, M., Schelbert, E.B., Ribeiro, A.H., Ribeiro, A.L.P., Schlegel, T.T., Ugander, M.: Heart age gap estimated by explainable advanced electrocardiography is associated with cardiovascular risk factors and survival. European Heart Journal - Digital Health **4**(5), 384–392 (Oct 2023). https://doi.org/10.1093/ehjdh/ztad045
19. Peebles, W., Xie, S.: Scalable Diffusion Models with Transformers. In: 2023 IEEE/CVF International Conference on Computer Vision (ICCV). pp. 4172–4182. IEEE, Paris, France (Oct 2023). https://doi.org/10.1109/ICCV51070.2023.00387
20. Qiang, Y., Dong, X., Liu, X., Yang, Y., Hu, F., Wang, R.: ECGMamba: Towards ECG Classification with State Space Models. In: 2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). pp. 6498–6505 (Dec 2024). https://doi.org/10.1109/BIBM62325.2024.10822754
21. Ribeiro, A.H., Ribeiro, M.H., Paixão, G.M., Oliveira, D.M., Gomes, P.R., Canazart, J.A., Ferreira, M.P., Andersson, C.R., Macfarlane, P.W., Meira Jr., W., Schön, T.B., Ribeiro, A.L.P.: CODE dataset (2021). https://doi.org/10.17044/SCILIFELAB.15169716
22. Ribeiro, A.L.P., Paixão, G.M.M., Gomes, P.R., Ribeiro, M.H., Ribeiro, A.H., Canazart, J.A., Oliveira, D.M., Ferreira, M.P., Lima, E.M., Moraes, J.L.d., Castro, N., Ribeiro, L.B., Macfarlane, P.W.: Tele-electrocardiography and bigdata: The CODE (Clinical Outcomes in Digital Electrocardiography) study. Journal of Electrocardiology **57**, S75–S78 (Nov 2019). https://doi.org/10.1016/j.jelectrocard.2019.09.008
23. Ribeiro, A.H., Paixao, G.M., Lima, E.M., Horta Ribeiro, M., Pinto Filho, M.M., Gomes, P.R., Oliveira, D.M., Meira Jr, W., Schon, T.B.: CODE-15%: a large scale annotated dataset of 12-lead ECGs (Jun 2021). https://doi.org/10.5281/ZENODO.4916206
24. Ribeiro, A.H., Paixao, G.M., Lima, E.M., Horta Ribeiro, M., Pinto Filho, M.M., Gomes, P.R., Oliveira, D.M., Meira Jr, W., Schon, T.B., Sabino, E.C.: Sami-Trop: 12-lead ECG traces with age and mortality annotations (Jun 2021). https://doi.org/10.5281/ZENODO.4905618
25. Tiezzi, M., Casoni, M., Betti, A., Gori, M., Melacci, S.: State-space modeling in long sequence processing: a survey on recurrence in the transformer era. Neural Networks **193**, 108039 (Jan 2026). https://doi.org/10.1016/j.neunet.2025.108039
26. World Health Organization: World health statistics 2019: monitoring health for the SDGs, sustainable development goals (2019), https://apps.who.int/iris/handle/10665/324835
27. Zhang, H., Wang, R., Zhou, C., Shi, J., Ge, Y., Li, Z., Chang, S., Wang, H., He, J., Huang, Q.: S2M2ECG: Spatio-temporal bi-directional State Space Model Enabled Multi-branch Mamba for ECG (Sep 2025). https://doi.org/10.48550/arXiv.2509.03066