Long-Term Prediction of Myocardial Infarction Using Matrix-Based ECG Enabled by Convolutional Neural Networks

Eduardo Caballero[1], Antonio Miguel[1], Patricia B Munroe[2,3], Julia Ramírez[1,2,4], Juan Pablo Martínez[1,4]

[1]Aragon Institute of Engineering Research, University of Zaragoza, Zaragoza, Spain
[2]Clinical Pharmacology and Precision Medicine, William Harvey Research Institute, Queen Mary University of London, London, United Kingdom
[3]NIHR Barts Cardiovascular Biomedical Research Centre, Barts and The London Faculty of Medicine and Dentistry, Queen Mary University of London, London, United Kingdom
[4]Biomedical Research Networking Centre on Bioengineering, Biomaterials and Nanomedicine, Spain

Abstract

Risk stratification of myocardial infarction (MI) is crucial for early intervention and prevention. ECG-based models, enabled by artificial intelligence, have shown promise when predicting cardiovascular events. However, these methods are still limited to relatively short follow-up periods and may not be applicable to a general/ healthy population. We hypothesized that the performance of these AI models could improve if the input data was presented in a structured format, enabling the model to extract more relevant features from the same ECG signal. We designed and compared the performance of two Convolutional Neural Networks, designed for MI prediction across 8 follow-up times, from 1 to 12 years. The first model was designed to take 3 seconds from a 1-lead ECG as input (1D), while the second one takes a matrix representation (2D) of the same 1-lead ECG. We trained both models using 97,382 1-Lead ECGs taken from UK Biobank from subjects without previous history of MI. Our proposed 2D model showed a slightly better performance, sustained across all predictions, with an average AUC increase of 0.047. These results suggest that the proposed model better captures the difference between subjects with higher risk of MI. With this work we aimed to improve the risk prediction for MI to provide healthcare professionals more accurate and opportune information for decision making.

1. Introduction

Cardiovascular disease (CVD) is the leading cause of death worldwide, representing around 32% of global deaths annually. Ischemic heart disease (IHD), including Myocardial infarction (MI) being the main cause of these fatalities [1, 2]. Despite advances in diagnosis and treatment, predicting and preventing these events remains a significant challenge [3].

Electrocardiogram (ECG) models enabled by artificial intelligence (AI), have shown promise in predicting several cardiovascular diseases [4, 5]. Specifically for MI there have been several models proposed to predict these events, with different approaches [6-9]. For example, Gustafsson [9] proposed a model for MI prediction during hospitalization, using 8-Lead ECGs, sex and age, achieving good results. However, there is a lack of robust models for predicting CVDs at long term follow up periods. Furthermore, most models rely on 12-Lead ECG modalities, hindering a wider applicability when only 1 lead is available (i.e., wearable devices). One example of this being used is the model proposed by Sau [10] where they developed an AI model to predict mortality using the 8 independent leads of the ECG and then adapted it to use only 1 lead.

We aim to overcome these limitations by proposing a new AI-ECG model, based in convolutional neural networks (CNN), that inputs a matrix representation of the ECG signal. We hypothesize that with this format the neural network will be able to extract more easily relevant features, such as heart rate variability, to help improve risk prediction of CVD [11 – 13], thanks to the CNNs ability to recognize complex spatial patterns.

2. Materials and Methods

2.1. Materials

ECG recordings from the UK Biobank database were used (application 8256, the study was approved by an institutional review committee and all subjects gave informed consent).

Individuals were excluded if they had a prior diagnosis of MI (defined as primary or secondary cause of death or

admission to hospital with documented MI, using ICD9, ICD10 and OPCS4 codes), if the ECG peak to peak amplitude after pre-processing was lower than 0.25 mV or higher than 5 mV, if the total duration of the recording was less than 3 seconds, or if there was no follow-up for at least a year after the ECG recording.

We used a total of 97 382 ECG signals after exclusion, taken from two different cohorts, prioritizing the earliest available recording. 64 537 ECG signals came from a bicycle test that consisted of three stages, resting (15 seconds), activity (6 minutes) and recovery (1 minute); from this study a modified lead I was recorded, and we only selected an ECG segment during the first stage, i.e., when the subjects were at rest. The remaining 32 845 ECG signals came from the imaging visit, where they recorded a 12-lead resting ECG with variable duration; from this cohort we only take the Lead I from the recordings.

The average follow-up time after the ECG recording was 9,37 ± 3,84 years, while the median age at the time of the recording of the ECG was 60 years (IQR: 53-66). As we do not have information for all subjects for the full 12-year follow-up we only considered each recording until the follow-up time that we had, after this the subject is considered as censored and not used to calculate the loss in training.

2.2. ECG Preprocessing

The sampling frequency for all ECG signals was 500 Hz. All signals were filtered with a forward-backward band-pass Butterworth filter with cut-off frequencies of 0.5 and 40 Hz, and filter orders of 3 and 10, respectively, to remove baseline drift and high-frequency noise.

2.3. 1D ECG CNN Model

The first CNN model was designed to input a 3-second segment of the ECG signal and output an 8-element vector, corresponding to the prediction of MI at 8 timepoints (1, 2, 3, 4, 6, 8, 10 and 12 years) after the ECG recording. We used a three-block architecture; first a feature extractor, composed of six 1D convolutional layers, then a self-attention layer to remain with only one feature vector and finally a prediction block composed of 3 fully connected layers.

To deal with the data imbalance (Table 1) we used data augmentation to equally represent the three classes in all partitions (case, control and censored). To do this, for each fold, we generated 10,000 segments by randomly taking a 3 or 6 second segment of a random ECG, ensuring that one third of the resulting samples represented each class at the end of the 12-year follow-up.

For the training of the models, we used a 10 k-fold strategy, ensuring that each partition maintained the class-control balance from the full dataset, and reserving 20% of the total signals for the test set, also maintaining the balance. This strategy generated 10 models, one for each fold. The loss functions used during training was specially designed to include subjects up top their censoring date (date when we no longer have follow-up information), this was done to avoid that the lack of data induces a survival bias. The final predictions were obtained by averaging the predictions generated by the 10 distinct models. All model development and testing was done using the PyTorch package (Version 2.3.1+cu118) in Python (Version 3.8.10).

Table 1. Class balance for different follow-up years.

Years	Cases	Controls	Censored
1	204 (0,21%)	97 178 (99,79%)	-
2	439 (0,45%)	96 067 (98,65%)	876 (0,90%)
3	671(0,69%)	93 036 (95,54%)	3 675 (3,77%)
4	909 (0,93%)	82 023 (84,23%)	14 450 (14,84%)
6	1 312 (1,35%)	69 489 (71,36%)	26 581 (27,30%)
8	1 702 (1,75%)	62 447 (64,13%)	33 233 (34,13%)
10	2 135 (2,19%)	54 720 (56,19%)	40 527 (41,62%)
12	2 490 (5,56%)	51 684 (53,07%)	43 208 (44,37%)

2.4. 2D ECG CNN Model

The second model was designed to take a representative matrix of the 1-lead ECG recording as input, and, as the 1D ECG CNN model, produce an 8-element vector with the predictions of MI at the same timepoints. Similarly, we used the three-block architecture, but modifying the feature extractor block to handle a matrix using 2D convolutional layers, and changing the self-attention layer to a global average pooling layer to handle the extra dimension.

To create the representative matrix, all R-wave peaks were detected using a wavelet-based delineator, proposed in [4]. From each detected R-wave, we extracted a 1,6 seconds segment, encompassing 0,35 seconds before and 1,25 after the peak. These segments contained the full morphology of a typical heartbeat (including the P-wave, QRS complex, and T-wave) and, depending on the heart rate, a significant portion of the subsequent one, incorporating information about the heart rate variability. Each one of these segments became a row in our matrix, with all segments aligned to the R-wave peak. This created a representation where the vertical dimension captured the morphology of individual heartbeats, while the alignment allowed the network to learn subtle variations in the morphology of consecutive beats. In figure 1 we can see a representation of the transformation from ECG to matrix.

The same 10 k-fold strategy was used as the one used for the 1D ECG CNN Model, while also making sure we used the same subjects in each partition. To perform the data augmentation the generated matrices were randomly cropped to 4 or 8 consecutive rows, to represent approximately the same that a 3 or 6 second segment would

capture; and 10 000 matrices were generated, similar to the 1D ECG CNN. Both model architectures were chosen to be similar to ensure that performance differences stem from the signal pre-processing and signal transformation rather than from the network architecture itself.

We tested two distinct comparative evaluations of model input: the first paired a 3-second segment 1D ECG CNN with a 4-beat 2D ECG CNN; the second paired a 6-second segment 1D ECG CNN with an 8-beat 2D ECG CNN, to have fair comparisons.

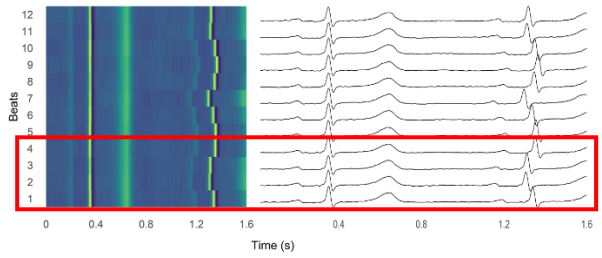


Figure 1. Visual representation of the matrix transformation. a) Image representation of the matrix, each row represents a heartbeat, intensity shown with colormap. b) Visual representation of each row containing one heartbeat and a fragment of the next.

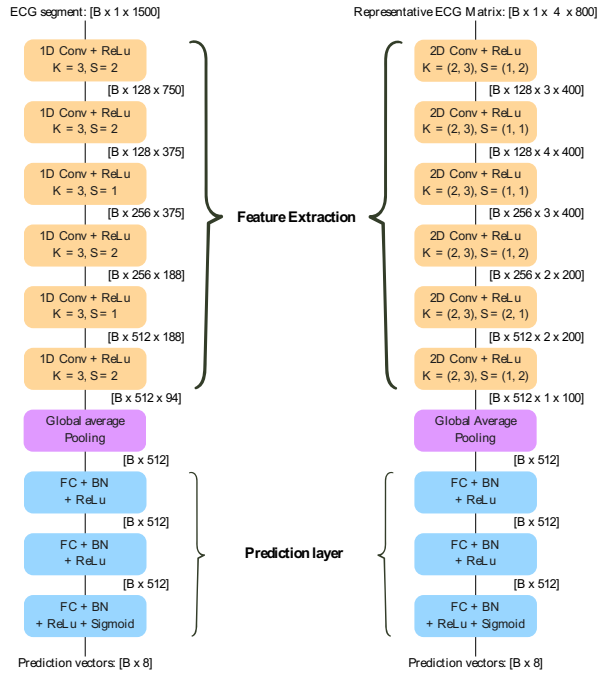


Figure 2. Comparison of architecture of CNN Models, 1D ECG CNN model (left) and 2D ECG CNN model (right).

3. Results

Model performance was evaluated by quantifying the area under the receiver operating characteristic curve (AUC–ROC) of the network output for all follow-up times. The evaluation was done using only the test partition, corresponding to 19 477 ECG recordings not used for either training nor validation.

For the first comparison, 3-second segment 1D ECG CNN vs 4-beat 2D ECG CNN, the 2D model showed a better performance across all follow up times, with an average increase in the AUC-ROC of 0,029. The difference between the two models was found statistically significant at all follow up times.

For the second comparison, 6-second segment 1D ECG CNN vs 8-beat 2D ECG CNN, the 2D model outperformed the 1D model at most follow-up times (2 to 12 years), with this difference being statistically significant across this range. However, the 1D model showed superior performance at the 1-year follow-up, though this difference was not statistically significant.

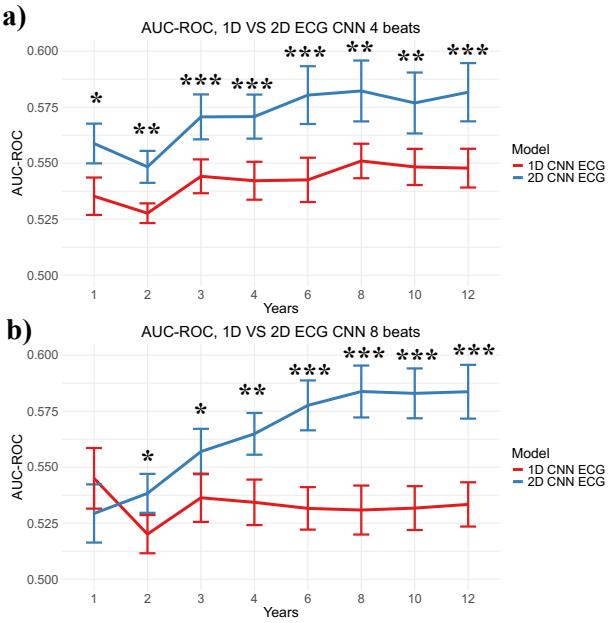


Figure 3. AUC comparison across the 8 timepoints. a) 3-second 1D ECG CNN vs 4-beat 2D ECG CNN. b) 6-second segment 1D ECG CNN vs 8-beat 2D ECG CNN. Note: Asterisks indicate statistical significance at the following levels: *: p < 0,05, **: p < 5x10-4, ***: p < 5x10-6.

4. Discussion

In this study, we developed, trained, and tested two distinct CNN models designed for the prediction of MI risk for eight timepoints along a 12-year follow-up. Our comparative analysis of the models performance, evaluated using AUC-ROC, revealed that the proposed model using matrix representations of the ECG signal consistently outperformed the 1D CNN across different follow-up time predictions. These results suggest that the

matrix representation, by capturing both the morphological characteristics of individual heartbeats and the temporal dynamics between them, is better able to discern subtle yet critical differences between subjects with a higher propensity for experiencing a future myocardial infarction.

With this approach we aim to improve the accuracy of MI risk prediction and, providing clinicians with more reliable and opportune information to guide their decision-making processes. Accurate long-term risk stratification is particularly challenging in the context of cardiovascular disease due to the inherent complexity of the underlying physiological processes and, often, the limited availability of comprehensively labeled longitudinal data spanning many years. Our investigation into these advanced deep learning techniques represents a step towards addressing this critical gap, potentially enabling earlier identification of at-risk individuals and facilitating more proactive and personalized preventative strategies.

4.1. Limitations and Future Work

The methodology's validation only on UK Biobank limits its generalizability to broader populations due to potential demographic and prevalence biases; additional testing is needed to validate this model. Additionally, the 2D-CNN architecture was adapted from a 1D-optimized network, thus it may not be optimal for processing a matrix, suggesting potential for further performance improvements with dedicated 2D network designs and optimization.

We also plan to adopt a more comprehensive evaluation. This will include time-dependent classification metrics like the F1 score and the Area Under the Precision-Recall Curve (AUPRC), which are more robust in the presence of imbalanced data. We will also incorporate the survival curves to provide a more intuitive and clinically-relevant characterization of the model's performance.

The proposed methodology presents an opportunity for predicting diverse cardiovascular endpoints beyond MI, this approach could be applied to identify patterns in ECG indicative of arrhythmias, heart failure, or other conditions.

Acknowledgments

This work has been supported by projects PID2022-140556OB-I00, PID2023-148975OB-I00, and CNS2023-143599, funded by the Spanish Ministry of Science and Innovation (MCIN) and FEDER, and by the group T39-23R2014-2020 on Biomedical Signal Interpretation and Computational Simulation (BSICoS). JR acknowledges the funding of the RYC2021-031413-I grant from MCIN. The calculations were performed using ICTS NANBIOSIS.

References

[1] N. Salari, et al. “The global prevalence of myocardial infarction: a systematic review and meta-analysis”, *BMC Cardiovasc. Disord.*, vol. 23, no. 206, Apr. 2023.

[2] A. Alexova, "ANNEX: Prevalence of STEMI", *ESC Press Office*, [Online]. Available: https://www.escardio.org/The-ESC/Press-Office/annex-prevalence-of-stemi. [Accessed: Aug. 9, 2025].

[3] E. Svennberg, et al., “State of the art of artificial intelligence in clinical electrophysiology in 2025: A scientific statement of the European Heart Rhythm Association (EHRA) of the ESC, the Heart Rhythm Society (HRS), and the ESC Working Group on E-Cardiology”, *EP Europace*, vol. 27, no. euaf071, May 2025.

[4] M. Z.H. Kolk, et al., “Prediction of sudden cardiac death using artificial intelligence: Current status and future directions”, *Heart Rhythm*, vol. 22, no. 3, pp. 756-766, Mar, 2025.

[5] CH. Lin, et al., “A multitask deep learning model utilizing electrocardiograms for major cardiovascular adverse events prediction”, *NPJ Digital Medicine*, vol. 8, no.1, Jan. 2025.

[6] O. J. Mechanic, M. Gavin, and S. A. Grossman, "Acute myocardial infarction," *StatPearls*, no. NBK459269, Oct. 2017. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/29083808/

[7] D. Chumachenko, et al., "Machine learning methods in predicting patients with suspected myocardial infarction based on short-time HRV data", *Sensors,* vol. 22, no. 7033, Sep. 2022.

[8] R. Liu, et al., "An artificial intelligence-based risk prediction model of myocardial infarction", *BMC Bioinform.*, vol. 23, no. 217, Jun. 2020.

[9] S. Gustafsson, et al., “Development and validation of deep learning ECG-based prediction of myocardial infarction in emergency department patients”, *Sci. Rep.*, vol. 12, no. 19615, Nov. 2022.

[10] A. Sau, et al., "Artificial intelligence-enabled electrocardiogram for mortality and cardiovascular risk estimation: a model development and validation study", *Lancet Digit. Health*, vol. 51, no e802, Nov. 2024.

[11] L. B. T. Yugar, et al., “The role of heart rate variability (HRV) in different hypertensive syndromes”, *Diagnostics*, vol. 13, no. 785, Feb. 2023.

[12] N. Chattipakorn, et al., "Heart rate variability in myocardial infarction and heart failure", *Int. J. Cardiol.*, vol. 120, no. 3, pp. 289-296, Sep. 2007.

[13] Y. Zhang, et al. "Value of heart rate variability on dynamic electrocardiogram in predicting ventricular fibrillation in elderly acute myocardial infarction patients", *Ann. Palliat. Med*., vol. 9, no. 5, Sep. 2020.

[14] J. P. Martinez, et al., "A wavelet-based ECG delineator: evaluation on standard databases", *IEEE Trans. Biomed. Eng.*, vol. 51, no. 4, pp. 570-581, Apr. 2004.

Address for correspondence:

Eduardo Caballero, Aragon Institute of Engineering Research, University of Zaragoza,
Mariano Esquillor s/n, Lab 5.1.02, 50018, Zaragoza, Spain
ecaballero@unizar.es