

Leveraging Large-Scale ECG Representations for Interpretable Detection of Cardiac Amyloidosis in a Clinical Cohort

Gandin Ilaria⁽¹⁾, Indennitate Carla^(1,2), Treu Riccardo⁽¹⁾, Porcari Aldostefano⁽²⁾, Merlo Marco^(1,2)

(1) Department of Medical Sciences, University of Trieste, Trieste, Italy

(2) Cardiovascular Department, Azienda Sanitaria Universitaria Giuliano Isontina, Trieste, Italy

CORRESPONDING AUTHOR: Gandin Ilaria, igandin@units.it

INTRODUCTION

Deep neural networks (DNNs) have shown strong performance in diagnostic classification of raw electrocardiogram (ECG) signals. However, their clinical utility remains limited due to the scarcity of large annotated datasets and the inherent lack of interpretability. Variational autoencoders (VAEs), a class of unsupervised deep learning models based on encoder–decoder convolutional neural networks, can address these challenges by compressing input signal into a lower-dimensional latent space, obtaining latent representations that preserve key information and that can be explored using explainability techniques.

In this study, we investigate the application of this approach to the detection of transthyretin cardiac amyloidosis (CA)—a progressive and often underdiagnosed disease caused by the deposition of misfolded amyloid proteins in the myocardium, leading to substantial morbidity and mortality. Clinical suspicion typically arises from a combination of specialized examinations, often beginning with the ECG, but a definitive diagnosis requires bone tracer scintigraphy. Notably, only a subset of individuals with suspected CA are ultimately confirmed to have the disease.

AIMS

Among cardiology patients with suspected CA, we aimed to develop a diagnostic model using a two-step approach: first, by compressing ECG signals into a reduced set of independent and explanatory generative factors using a VAE; and second, by applying an interpretable classification algorithms to predict the presence of CA.

METHODS

In this retrospective cross-sectional study, we included patients referred to the Trieste Cardiovascular Department with clinical suspicion of CA who subsequently underwent bone tracer scintigraphy to confirm the diagnosis. For each patient, the ECG closest in time to the scintigraphy was selected for analysis. All ECGs were exported from the Mortara system in raw voltage format. From these, 1.2-second median beats were extracted and reformatted into the MUSE ECG format required by the variational autoencoder (VAE) model developed by van de Leur et al. that was trained on 1.1 million ECGs [1].

The VAE compressed each ECG waveform into 21 latent variables (ECG factors), which were initially assessed for univariate associations with CA status. These factors were then used as input features in a multivariable classification model. The dataset was randomly split into training and test sets. A LASSO (least absolute shrinkage and selection operator) logistic regression model was trained using various combinations of input variables. The optimal regularization parameter (λ) was selected via 10-fold cross-validation using the `glmnet` package in R. Model performance in terms of discrimination and calibration was evaluated in the holdout test set.

RESULTS

A total of 370 individuals were included in the study, of whom 119 (32%) were diagnosed with CA. The median age was 79 (IQR=[72, 83]), and 142 (28%) were females. In univariate analysis, 3 ECG factors showed significant association with CA status: F11 (OR=66 per 1 SD increase, 95% CI [0.51, 0.83]), previously linked to subtle QRS- and T-wave

changes; F25 (OR=1.63, [1.28, 2.09]), linked to longer QRS duration; and F30 (OR=0.71, [0.56, 0.89]), linked to longer QT-interval.

Compared to a simple model based on sex and age that show an ROC-AUC of 0.67 (95% CI [0.56, 0.77]), the model including all ECG factors as predictor led to an ROC-AUC of 0.69 (95% CI [0.59, 0.79]). The full model combining sex, age and ECG factors reached an ROC-AUC of 0.70 (95% CI [0.60, 0.80]). Calibration metrics for models including ECG factors were acceptable, with calibration slopes between 0.77 and 1.07 and intercepts ranging from 0.08 to 0.31.

CONCLUSIONS

These findings highlight the potential of VAE models to leverage representations learned from large-scale datasets and apply them to smaller, specialized diagnostic tasks such as the detection of CA. Although the discriminatory performance observed in this study was modest, the model's interpretability offers valuable insights into ECG patterns associated with CA, which may guide and inform future research into its underlying electrophysiological characteristics.

REFERENCES

1. van de Leur et al. Improving explainability of deep neural network-based electrocardiogram interpretation using variational auto-encoders. *Eur Heart J Digit Health* . 2022 Jul 25;3(3):390-404