FMM-Head: Enhancing Autoencoder-based ECG anomaly detection with prior knowledge

Giacomo Verardo*[1][0000−0001−7367−9200], Magnus Boman[2,3][0000−0001−7949−1815], Samuel Bruchfeld[2][0000−0002−0284−5176], Marco Chiesa[1][0000−0002−9675−9729], Sabine Koch[2][0000−0001−7144−8740], Gerald Q. Maguire Jr.[1][0000−0002−6066−746X], and Dejan Kostic[1][0000−0002−1256−1070]

[1] KTH Royal Institute of Technology, Stockholm, Sweden
{verardo, mchiesa, maguire, dmk}@kth.se
[2] Karolinska Institutet, Stockholm, Sweden
{magnus.boman,samuel.bruchfeld,sabine.koch}@ki.se
[3] MedTechLabs, Stockholm, Sweden

Abstract. Detecting anomalies in electrocardiogram data is crucial to identify deviations from normal heartbeat patterns and provide timely intervention to at-risk patients. Various AutoEncoder models (AE) have been proposed to tackle the anomaly detection task with machine learning (ML). However, these models do not explicitly consider the specific patterns of ECG leads, thus compromising learning efficiency. In contrast, we replace the decoding part of the AE with a reconstruction head (namely, FMM-Head) based on prior knowledge of the ECG shape. Our model consistently achieves higher anomaly detection capabilities than state-of-the-art models, up to 0.31 increase in area under the ROC curve (AUROC), with as little as half the original model size and explainable extracted features. The processing time of our model is four orders of magnitude lower than solving an optimization problem to obtain the same parameters, thus making it suitable for real-time ECG parameters extraction and anomaly detection. The code is available at: https://github.com/giacomoverardo/FMM-Head

Keywords: Machine Learning · ECG anomaly detection · AutoEncoders.

1 Introduction

Cardiovascular conditions are the main causes of death worldwide [13]. Tools such as electrocardiogram (ECG) measurements are utilized to monitor and identify these conditions. An ECG records the heart activity by detecting electrical signals. Electrodes positioned on different parts of the body measure the signal propagation through different planes (*i.e., leads*), thus allowing the analysis of multiple heart sections. Collecting ECG data is standard procedure for both hospitalized patients and outpatients since it allows detection of various cardiovascular conditions, such as myocardial infarction and arrhythmia. In recent years, the amount of available ECG data has increased considerably due to wearables (*e.g.,* smart, low-powered, ECG-capable devices (*e.g.,* smartwatches

* Corresponding author

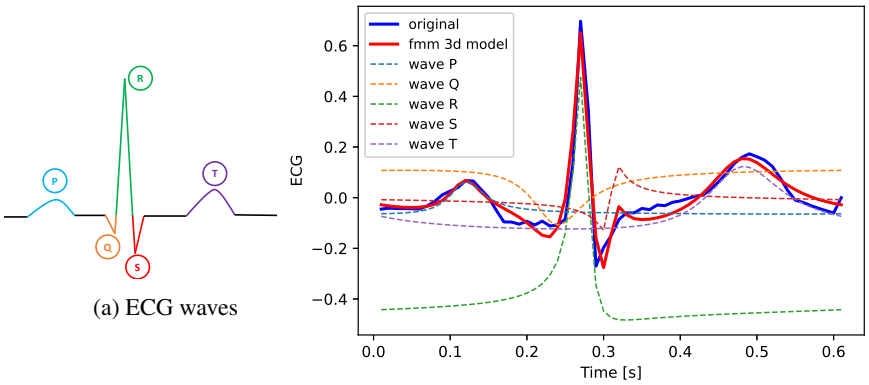
and wearable smart textiles [18])), institutional databases, and continuous ambulatory monitoring of high-risk patients. Continuous ambulatory monitoring produces a huge quantity of data, whose analysis can be difficult since it requires expert knowledge of cardiac conditions and their related effect on ECG measurements [25,26]. This increase in available data moves the bottleneck from *monitoring* to **processing** the collected data. Given the vast amount of available data, (*deep learning (DL)*) has been employed to tackle multiple ECG-related tasks. We propose including ECG prior knowledge in neural networks to improve anomaly detection and explainability in ECGs.

Anomaly detection through deep learning and (*ML*) models is a promising technique to improve care by spotting health records that deviate from the patterns of normal data *without* any knowledge of what the underlying conditions might be. Anomalies may be symptoms of major heart issues, like heart muscle failure [27]. In contrast, ML *classification* requires labeled data from different health conditions (*i.e.,* classes) that are used to train the model. AutoEncoders (AEs) are a family of ML models trained to reconstruct the original input signal. AEs are trained only on data which show no anomaly, so that during the testing and inference phases an anomaly alert is raised if the input sample is not normal. An anomaly can be detected when the reconstruction loss between original and predicted data is high [9]. Multiple rule-based ECG anomaly detection methods have been proposed [4]. Unlike ML models, these techniques rely on extracting well-known indicators for specific heart conditions. However, these methods lack generalization capabilities since they rely on strong *a priori* knowledge of what these parameters are; therefore, these assumptions hinder their usability for anomaly detection of *unknown diseases*, *i.e.,* there is no *a priori* knowledge of them.

Although the most prominent strength of AEs is the lack of assumptions regarding the classes and shapes of different inputs, the inclusion of *a priori* information about the structure of input data may be beneficial for the learning procedure. While ECG signals demonstrate different patterns depending on the underlying heart condition, their shape is composed of five waves (shown in Figure 1a), which correspond to different instants of the heart's electrical signal, as measured via the electrodes. For different heart conditions, the shape of these waves change, but the number of waves and their general structure are steady. This *weak a priori* knowledge is valid for almost all ECG classes, but this knowledge is currently not exploited by state-of-the-art anomaly detectors.

Recently, [23] proposed *Frequency Modulated Möbius (FMM)* waves to provide explainable parameters for ECG data [24]. They proposed an optimization algorithm to iteratively compute the amplitude, position, direction, and frequency parameters for the five waves composing the ECG signal through a cycle of polarization and depolarization. However, this optimization takes tens of seconds to be solved for a single heartbeat, thus making it unsuitable for real-time monitoring of critical patients and processing of voluminous quantities of ECG data. [31] have shown that a neural network (NN) can be used to approximate the FMM coefficients and correctly classify heartbeats, but did not apply it for anomaly detection.

Our contributions are threefold. Firstly, we develop FMM-Head, a first approach for incorporating *weak a priori* knowledge of the ECG leads' structure into an AE model. In particular, FMM-Head replaces the decoding sub-model of AEs, provides an explainable representation of the FMM parameters, and reconstructs the signal accordingly (see



(a) ECG waves

(b) FMM ECG Decomposition

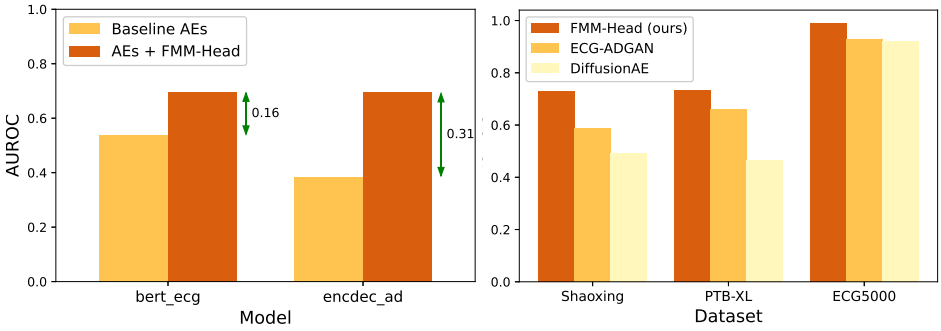
Fig. 1: (a) shows the ECG shape, (b) shows the FMM decomposition of an ECG wave

Figure 1b). We design a generic pooling layer to adapt FMM-Head to different hidden dimensions of the AE encoder. Secondly, we demonstrate FMM-Head's ease of use by incorporating it into five baseline AEs models, thus showing that our layer is flexible enough to handle the output of different kinds of encoders. FMM-Head significantly enhances the performance of these AEs. Moreover, as shown in Figure 2a, even low-performing models such as EncDecAd can be enhanced to be on par with other models. At the same time, FMM-Head outperforms baselines based on Generative Adversarial Networks (GANs) [8] and diffusion models [3], as shown in Figure 2b. Thirdly, we evaluate and compare our enhanced models to the baselines. Replacing the decoder with the FMM-Head almost halves the total number of trainable parameters of the AEs and leads to up to -77% reduction in inference time and -47% memory required to store the model. Using a fully connected AE with 6 layers, the execution time is 21 thousand times lower than the optimization solution to the FMM problem using the [23] code, which was not designed to perform anomaly detection. The 4 orders of magnitude lower time to process batches of heartbeats enables real-time anomaly detection and is suitable for analyzing huge amounts of ECG data. Our model also provides coefficients that are highly correlated with the original FMM coefficients, thus making its output more transparent than blackbox AEs, whose extracted features are not explainable. Although training for anomaly detection reduces the output similarity to the real FMM coefficients, it improves the *detection of anomalies* of five baseline models.

2 Background

2.1 Electrocardiograms

For over a hundred years, ECGs have been used to detect heart conditions, such as myocardial infarction and arrhythmia [2]. The main idea behind ECG monitoring is to repeatedly measure the electrical polarization and depolarization waves that propagate



(a) (b)

Fig. 2: (a) FMM-Head improves anomaly detection for transformer and LSTM-based models and (b) shows better performance than GANs and diffusion models

through the heart muscles and cause rhythmic contraction and relaxation. A standard 12-lead ECG machine uses 10 electrodes, which are combined in different pairwise combinations to measure the voltage through planes that intersect the heart with different orientations, thus giving insights into various parts of the heart (*e.g.,* inferior, superior anterior, posterior leads). In contrast, smartwatches produce one ECG lead, *i.e.,* the plane intersecting the heart through the arm, similarly to having 2 electrodes.

Figure 1a shows the shape of an ECG, composed of 5 waves corresponding to rhythmic polarization and depolarization phases. The P wave corresponds to the depolarization of the atria. The QRS complex depends on depolarization of ventricula before their contraction, whereas the T wave is determined by ventricular repolarization. Detecting a cardiac disturbance in conductance may involve different features of the ECG, such as the interval and slope between the peaks, their amplitude or underlying area. Depending on the task, a cardiologist uses this information to identify conditions and diseases. Instances of such conditions with relative ECG change are nodal tachycardia (hidden P-wave), sinus tachycardia (visible P wave with higher rate), hypertrophic cardiomyopathy (deep and narrow Q waves in specific leads), myocardial infarction (ST segment elevation or depression), and atrial fibrillation (wide QRS complex, absence of P wave).

Given the bulk of possible heart conditions and their corresponding relevant ECG features, anomaly detection and classification of ECG data requires *strong* a-priori knowledge [13]. The problem is exacerbated in real-time monitoring of patients [25,26]; hence, it requires automatic processing to scale with the quantity of available data.

2.2 FMM Waveforms

To provide a comprehensive way to parameterize the rhythmic behavior of the hearth, [24] proposed modeling each individual heartbeat as the sum of 5 FMM waves [23]. Each wave is modeled as per the following equation:

$$W(t_i) = \mu(t_i) + e(t_i) = M + A\cos(\varphi(t_i)) + e(t_i) \tag{1}$$

$$\varphi(t) = \beta + 2\arctan(\omega\tan(t-\alpha)) \tag{2}$$

$$\mathbf{e} = (e(t_1), \ldots, e(t_n))^T \sim \mathcal{N}(\boldsymbol{e}; 0, \sigma^2\boldsymbol{I}) \tag{3}$$

where, $A \in \mathbb{R}^+$ is the amplitude of the wave, $\alpha \in [0, 2\pi)$ represents the position of the peak, $\beta \in [0, 2\pi)$ is the peak direction, $\omega \in [0, 1]$ parametrizes the lobe width of the peak, $M \in \mathbb{R}$ is the constant offset of the wave, t_i is the timestep index. The tuple $\theta = (A, \alpha, \beta, \omega, M)$ represents the encoded parameters necessary to represent each wave. Their proposed 3D model includes the FMM waves formulation but makes assumptions of the parameters that are shared between leads, *i.e.,* the α and ω coefficients are shared among leads while A, β, and M are not. Therefore, the lead vector $\boldsymbol{X}$ for lead L is:

$$\boldsymbol{X}(t_i)^L = M^L + \sum_{j \in \{P,Q,R,S,T\}} W(t_i, A_j^L, \alpha_j, \beta_j^L, \omega_j) + e^L(t_i); \tag{4}$$

To estimate the optimal parameters, the following objective function is used:

$$\theta^* = \min_{\theta} \sum_{L} \sum_{i} (\boldsymbol{X}^L(t_i) - \hat{\boldsymbol{X}}^L(t_i, \theta))^2; \tag{5}$$

The estimated best tuple θ^* is obtained by repetitively iterating a fitting and wave assignation phase. During fitting, optimization is performed over a single FMM wave, and the single-lead parameters are extracted by solving a linear regression problem. More than 5 waves might be obtained. During wave assignation, a choice of the best ones is made: each peak (P, Q, R, S, T) is selected based on the α coefficient and thresholds on the main model's parameters. The proposed algorithm is inherently sequential, not parallelizable and it often requires minutes-order of magnitude of execution time.

2.3 AutoEncoders

AEs (Figure 3a) are a family of self-supervised NN models [9]. AEs have a dumbbell structure (*i.e.*, wide-narrow-wide as in Figure 3a), sequentially combining (i) an encoder, which transforms the inputs' features into a lower-dimensional latent representation, and (ii) a decoder, that reconstructs the input from the latent space. The loss function measures the error between original and output data. Hence, the aim of an AE is to exactly reconstruct the inputs. However, due to the dumbbell shape, irrelevant information is lost, thus enforcing a compact, semantically meaningful latent space.

AEs are state-of-the-art models for anomaly detection [5]. AEs are trained on normal data so that abnormal samples during test phases can straightforwardly be recognized. Samples unseen during training, so-called holdout data, will be wrongly encoded and decoded, thus causing a large loss. Different kinds of NN models have been proposed to tackle ECG anomaly detection, including Long-short term memory (LSTM)-based models [6,22,16,17,10], transformers [1] and variational AEs [12,14].

2.4 Circular Variables

A representation of circular data [15] is necessary whenever the direction of a measure is a crucial feature to understand the correspondent phenomenon. ECGs are intrinsically circular since the electrical signal passing through the heart is quasi-periodic and can be modeled as an oscillator. In the FMM formulation, α and β are circular variables since they respectively represent a position and a direction within the $[0, 2\pi]$ interval.

The circular mean of a circular random variable θ can be computed as: $\bar{\theta} = \arctan 2 \left(\sum_{i=1}^{n} \sin(\theta_i), \sum_{i=1}^{n} \cos(\theta_i) \right)$. The correlation between circular variables should be computed differently than linear correlation. For instance, the linear Pearson coefficient ρ between two random variables x and y, for whom we extracted n samples, can be computed as: $\rho = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}$. However, using the Pearson coefficient between two circular variables will return an incorrect estimate of the correlation. For example, a realization of $x, y = \epsilon, 2\pi - \epsilon$ should positively contribute to the correlation coefficient. To solve this problem, circular correlation [11] can be computed by using the circular mean instead of the linear mean: $\rho = \frac{\sum_{i=1}^{n} \sin(\theta_{1i} - \bar{\theta}_1) \cdot \sin(\theta_{2i} - \bar{\theta}_2)}{\sqrt{\sum_{i=1}^{n} \sin^2(\theta_{1i} - \bar{\theta}_1) \cdot \sum_{i=1}^{n} \sin^2(\theta_{2i} - \bar{\theta}_2)}}$.

3 Methodology

Our FMM-Head layer reconstructs the original ECG input and provides the corresponding explainable FMM coefficients. To do so and maintain FMM parameter explainability, we split the training procedure into a warm-up regression phase and an anomaly detection training phase. Combined with constraints imposed inside the NN, this two-phase approach provides meaningful FMM coefficients.

3.1 Preprocessing

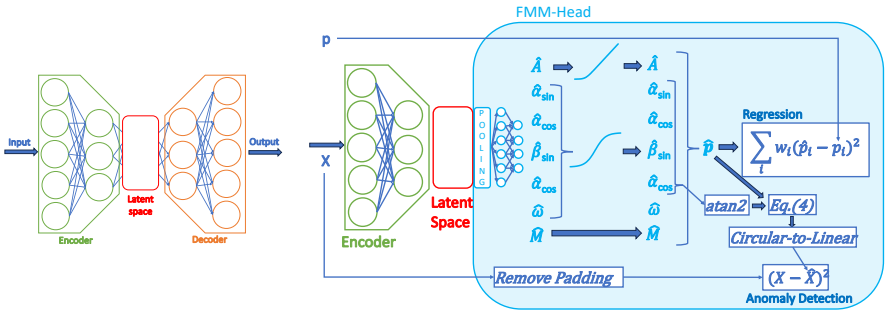
We preprocess the ECGs to extract heartbeats and provide correctly structured input data. We build on top of the preprocessing in [24], whose pipeline consists of (i) low pass filtering to remove baseline wandering, a common artifact in ECGs due to patient's breathing or movement, (ii) application of the Pan-Tompkins algorithm [19] to detect R-peaks, (iii) extraction of ECG heartbeats' interval around the R-peak with 40% of the distance from the previous peak and 60% from the next one. Additionally, we zero-pad each sequence to a constant length in order to be able to feed the input samples to the evaluated models. The original coefficients of the FMM-model are also preprocessed. In particular, circular variables such as α and β are split into their cosine and sine, so that they can be easily learnt by the NN. Also, the parameters are sorted according to the α parameter, so that the P wave corresponds to the first coefficients, etc.

3.2 FMM-Head

We design a novel layer, which we name FMM-Head since it resides on top of the NN and draws inspiration from the FMM formulation detailed in Section 2.2. As depicted in Figure 3b, the main idea behind FMM-Head is that any hidden representation encoded into the latent space can be mapped to meaningful FMM coefficients. This mapping is provided by a non-linear function that is implemented through one pooling layer and two fully connected layers, followed by suitable activation functions that

produce parameters in the correct range for the FMM formulation. The parameters obtained after the activation function are used to reconstruct an ECG time-series $\hat{X}$ for anomaly detection. We leverage this weak *a priori* knowledge drawn from the FMM formulation to better reconstruct the input ECG signal.

The main drawback of standard anomaly detection through FMM waves is that the reconstruction might not reflect the original meaningful pattern of the FMM optimization. Although it is straightforward to obtain FMM waves that approximate an ECG signal, obtaining peaks with meaningful shapes is challenging. Hence, we propose an initial warm-up regression phase using the original FMM coefficients. This procedure constrains the output of the NN to be in the range of the original FMM coefficients, thus steering the anomaly detection phase to correct and meaningful patterns.



(a) AE

(b) FMM-Head AE

Fig. 3: *(a)* Structure of standard AE and *(b)* AE with the FMM-Head decoder. The FMM-Head is usually smaller than the baseline decoders.

Pooling layer Depending on the employed AE, the encoder produces a latent space with different shapes. To handle different dimensions of hidden representations, a layer that maps them to a common output is needed. The pooling layers take as input the latent representation and generate a 2D output, which can be used as input to the following fully connected layers. The pooling layer is, in general, different for each encoder. Transformer and LSTM-based networks have output shapes consisting of a batch, time step, and feature dimension, which can be reduced to two by applying a linear transformation to each time step feature. For fully connected AEs and LSTM layers with state output (such as [17]) there is no need for pooling since the output is already 2D. For convolutional networks, we flatten the 3D output space into a 2D representation.

Fully connected layers Two fully connected layers are used to map the output of the pooling layer to the size of the FMM parameters. The first layer utilizes a non-linear function with $\tanh$ activation, while the second layer employs a linear activation. While the second layer has a fixed size (depending on the number of parameters N), the number of units of the first layer can be flexibly chosen. We tested multiple options (64, 128, and 256 units), and found only slight differences in performance. When the size is

too small, there is a drastic decrease in performance. Hence, we chose one hidden layer composed of 256 units and a second layer with N units for the fully connected network.

Activation functions for FMM coefficients The output of the fully connected network is unsuitable for direct generation of FMM waves since different parameters have different range requirements. We handle this by selecting an appropriate activation function for each parameter. In particular, a non-negative amplitude is obtained by applying *softplus*, while restricting the ω parameter to the range $(0, \omega_{max})$ with a sigmoid-like function. The sine and cosine components of the circular variables are also mapped to the $(0, 1)$ interval: $\hat{A}' = \ln(1 + e^{\hat{A}})$, $\omega' = \frac{\omega_{\text{max}}}{1 + e^{-\omega}}$, and $\alpha'_{\text{sin/cos}} = \frac{2}{1 + e^{-\alpha_{sin/cos}}} - 1$. The predicted α and β are obtained by computing the angle that corresponds to the predicted sines and cosines. The use of ω_{max} instead of the original unitary limit reduces the possibility of non-meaningful waves with good reconstruction. With high ω values, peaks can be too wide and hence negatively influence the pattern of other peaks.
The output of the layer includes parameters that may be shared or not between leads. We exploit the formulation from Section 2.2 to identify such parameters and provide single and multiple layer output variables for shared and not shared coefficients respectively. Hence, FMM-Head can inherently support multi-leads ECG inputs.

Regression and anomaly detection The predicted parameters can be used to directly compute the mean squared error loss. We directly inject the original FMM parameters, obtained by running FMM optimization, to compute the error and back-propagate it to train the AE. We employ the cosine and sine circular parameters to compute the loss and also apply a weight to each parameter to produce better alignment with the original time series. Specifically, we apply a $10\times$ higher weight for the parameters of the R-peak to obtain well-separated waves in the QRS complex. This specific multiplying factor was selected empirically, but it is not crucial to correctly perform warm-up, since lower values still provide distinct QRS waves. We perform a warm-up regression phase on the original coefficients by training the AE to produce results closer to the original optimized ones. Hence, although such a network cannot be employed for anomaly detection, the NN provides a prediction of the FMM coefficients.
After warm-up regression, the NN is trained as a standard AE. We employ Equation (4) to generate the 5 ECG waves, sum them, and obtain a reconstructed signal in the $[0, 2\pi)$ domain. We then map the signal to the original length through a linear transformation and compute the mean squared error between the input ECG, stripped of the zero-padding, and the predicted sum of waves. We point out that the original FMM coefficients are only needed to perform regression in the training phase, where time is not a constraint. In contrast, the time-critical testing phase does not require computing them.

4 Experimental Results

We evaluate the performance, inference time, and model size of our FMM-Head on three datasets: Shaoxing [32], PTB-XL [30], and ECG5000 [7]. We compare to seven baselines: five AEs (ECG-NET [22], CVAE [12], EncDecAD [17], a transformer model referred as BertECG [1], a fully connected AE), an ECG-specific GAN (ECG-ADGAN [21]) and one diffusion-based model (DiffusionAE [20]). The non-AE

baselines were motivated by reviews from ICLR2024. Although FMM-Head can handle more than one lead, we use lead 2 to train the NNs on normal data and then test the anomaly detection performance on a holdout set that includes both abnormal and normal classes, as done for the provided baselines. We run 5 experiments for multiple learning rates and report the one with the highest average area under the ROC curve (AUROC). If applicable, we ran the training session for 500 warm-up epochs followed by 500 epochs. For ECG-ADGAN we instead train for 20000 epochs, as in [21]. We set ω_{max} to 0.2 according to the distribution of the original ω_i. For AEs we perform early-stopping based on the validation loss. Our code is available at [29]. To enhance reproducibility, we provide the FMM coefficients for PTB-XL and Shaoxing at the same link, which were obtained by solving the FMM optimization problem [24].

With warm-up regression, FMM-Head consistently enhances the anomaly detection capabilities of the baseline models by up to 0.31 of AUROC. Table 1 shows the AUROC for the evaluated models and datasets and the gains compared to baselines for the different models. For the Shaoxing and PTB-XL datasets, our pipeline consists of warm-up regression followed by anomaly detection training. In these cases, the performance increase is consistent for all combinations of datasets and models. The AUROC enhancement is more evident for low-performant baselines, such as EncDecAd and BertECG, where the evaluated AUROC can increase by up to 0.31. The main reason some baselines perform poorly is the complexity of the task. Compared to simple datasets (such as ECG5000), PTB-XL and Shaoxing are among the largest publicly available ECG datasets, consisting of multiple patients and labeled conditions. We experimentally determined that one of the major sources of complexity is the variable length of the ECG time series. While most models do not consider this factor, our FMM-Head inherently maps the time series to the $[0, 2\pi)$ interval, thus equalizing the lengths of the samples in the last layer. Therefore, even low-performant baselines can achieve considerable AUROC for variable-length training sets.

The benefits of FMM-Head are noticeable for high-performing baselines, such as ECG-NET and CVAE. Replacing the decoder with a head based on prior knowledge better exploits the features extracted from the encoder. One exception is FMM-CAE, whose baseline counterpart CVAE performs slightly better for the PTB-XL dataset. We argue that convolutional AEs are mostly focused on recognizing patterns between adjacent time steps, therefore being less suitable for the extraction of features for FMM coefficients. Our belief is confirmed by the 0.02 decrease in AUROC for FMM-CAE in the Shaoxing dataset, which is the largest decrease in performance of all of the model-dataset combinations. However, for the ECG5000 dataset the AUROC is already very high and the addition of the FMM-Head makes little (at most 1.7%) difference.

Baselines with integrated FMM-Head can extract coefficients that are highly correlated with those extracted by solving an optimization problem in less than $\frac{1}{20000}$ of the time. Although our model is built for anomaly detection and *not* to exactly reproduce the FMM coefficients, we can nevertheless produce coefficients correlated with those obtained with the original FMM optimization. Thus, our method provides explainability by means of the interpretability benefits discussed in [24].

Figure 4a shows the linear or circular correlation between the coefficients extracted by [24] and those obtained after warm-up. The extracted parameters are usually highly

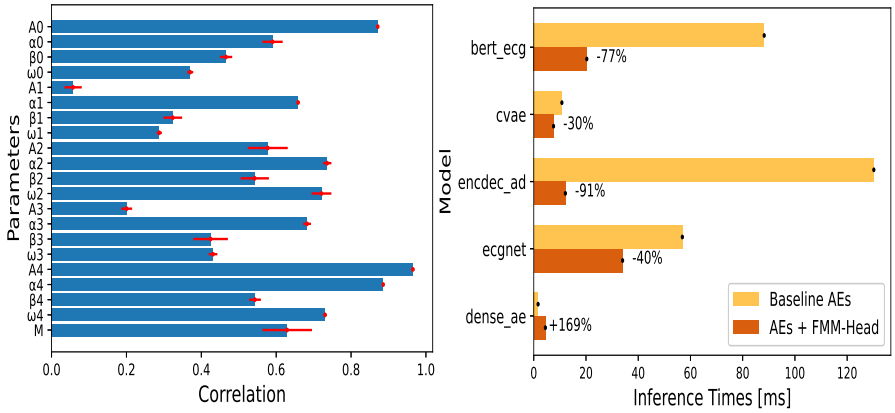
Table 1: Highest AUROC values for best learning rate and gains compared to baselines. Our FMM-Head enhanced models produce consistently better results compared to the baselines when the warm-up regression is employed (*i.e.,* for Shaoxing and PTB-XL)

	Shaoxing		**PTB-XL**		**ECG5000**	
Model	**AUROC**	**Gain**	**AUROC**	**Gain**	**AUROC**	**Gain**
ecgnet	0.575		0.661		0.993	
fmm_ecgnet	0.659	**0.084**	0.731	**0.070**	0.988	*–0.005*
encdec_ad	0.461		0.384		0.982	
fmm_encdec_ad	0.617	**0.156**	0.695	**0.311**	0.988	**0.006**
bert_ecg	0.545		0.536		0.971	
fmm_bert_ecg	0.650	**0.105**	0.697	**0.161**	0.989	**0.017**
dense_ae	0.653		0.691		0.992	
fmm_dense_ae	0.699	**0.046**	0.698	**0.007**	0.990	*–0.002*
cvae	0.749		0.693		0.992	
fmm_cae	0.729	*–0.020*	0.727	**0.034**	0.990	*–0.002*
ecg_adgan	0.587		0.662		0.927	
diffusion_ae	0.491		0.464		0.919	

correlated (*i.e.,* value greater than 0.4). Noteworthy is that the P and T waves manifest in general high correlation to the optimized ones. This is due to the two peaks being spaced apart by the QRS complex, thus making them more straightforward to extract. In contrast, the waves belonging to the QRS complex are close, thus enabling the same reconstructed time series through possibly non-ideal wave combinations. For instance, the same ECG pattern could be obtained through the sum of three wide extracted waves instead of three narrow peaks. Therefore, the inference of the QRS complex with a NN is less correlated to the original when compared to the P and T wave. The results for PTB-XL are similar to those for Shaoxing, except for the inference time gains for the LSTM-based models, which are lower due to the smaller input size.

Figure 4b shows the decrease in inference time compared to baseline models for the Shaoxing dataset. Compared to the original FMM optimization problem, which takes in average $38\,\mathrm{s}$, AEs with FMM-Head reduces the computation time by more than four orders of magnitude. Compared to the baseline AEs, the benefits of FMM-Head varies from -30% to -91%, with the exception of DenseAE. The main reason behind this is that although FMM-Head is small, it is still more complex than most state-of-the-art layers. For DenseAE, replacing the fully connected 3-layer decoder with FMM-Head does *not* reduce the inference time but actually increases it by $+169\%$. However, dense, convolutional, and LSTM layers have been extensively researched and optimized in the last decades; hence, we argue that FMM-Head execution times could be further improved by optimizing the code.

By replacing standard decoders with FMM-Head, the model size is nearly halved for the non-LSTM evaluated models, greatly reducing the storage requirements on mobile and wearable ECG monitoring devices. Figure 5a shows the reduction in model size obtained by replacing each baseline's decoders with FMM-Head. In most cases, the file size of the obtained models is nearly halved. The large decrease in model size on BertECG is due to a reduction in the encoder, which we experimentally showed



(a) Correlation between FMM coefficients for Shaoxing

(b) Inference Times for Shaoxing

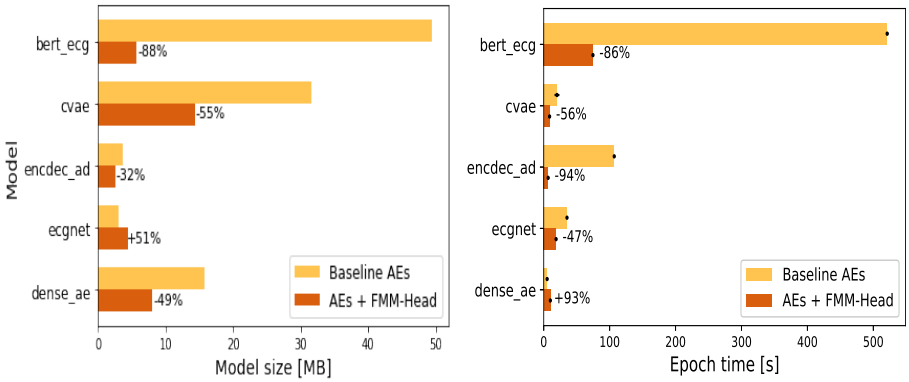
Fig. 4: Despite not being built for FMM coefficient extraction, FMM-Head produces FMM parameters that are highly correlated with those extracted by the original method (see FMM-DenseAE in *(a)*). *(b)* Compared to the baselines, our NN-based models provide lower inference, except for models with low-complexity such as DenseAE.

did not produce a relevant impact on the performance of the correspondent model integrated with FMM-Head. Noteworthy is that LSTM-based models such as ECG-NET and EncDecAD gain the least benefits from FMM-Head in terms of model size. This is due to the inherent design of LSTM layers, which favor minimizing model size over training and inference time; hence, the gains of FMM-Head are less prominent. In particular, the size of ECG-NET is 51% higher when FMM-Head is employed since the long inputs of Shaoxing increase the size of the pooling layer and consequently increase the size of the first fully connected layer.

Figure 5b shows the epoch duration of each model on the Shaoxing dataset. As expected, the reduced model size enables faster training than the baselines. Similarly to Figure 4b, the only exception is DenseAE, which trades off the model size with the additional complexity of FMM-Head, thus showing an increase in the epoch duration.

Warm-up regression is essential to maximize the benefits of the FMM-Head. Column 3 in Table 1 shows the performance of different models on the ECG5000 dataset. Whereas for EncDecAd and BertECG the AUROC is slightly better, for the other baselines we notice a decrease instead. We claim that this is due to the fact that we did not perform any warm-up regression on ECG5000, thus not exploiting the prior knowledge of the ECG shape. Without warm-up, the waves are not constrained in a meaningful position and the increase in AUROC is consequently less prominent.

Compared to state-of-the-art GAN and diffusion-based models, FMM-Head enhanced AEs increase anomaly detection performance. The last two rows of Table 1 show the AUROC of baseline ECG-ADGAN [21] and DiffusionAE [20]. Whereas the two models produce acceptable performance on the small ECG5000 dataset, they strug-



(a) Model sizes for Shaoxing dataset (b) Epoch duration for Shaoxing dataset

Fig. 5: *(a)* FMM-Head has negligible size compared to most decoders, and thus one can considerably reduce the amount of file storage and memory needed to compute training and inference for neural network models. *(b)* shows the epoch duration for different models during training.

gle to perform well in more complex scenarios, such as in PTB-XL and Shaoxing. In particular, ECG-ADGAN shows the well-known mode collapse issue [28]: the generator produces outputs that belong to a single mode, thus being unsuitable for the normal samples with high complexity and variability in the bigger datasets. Moreover, ECG-ADGAN shows limited robustness to hyper-parameters, such as the learning rate, as shown in the appendix. DiffusionAE has been proposed to tackle anomaly detection on time series with anomalies in seasonality and trends, thus different inputs compared to ECG shapes during a single heartbeat, on which it performs poorly.

Compared to [31], FMM-Head can be applied to multiple baseline NNs for anomaly detection and not just classification. [31] provides a way to estimate 12-leads FMM coefficients by means of a custom NN and use them for ECG classification with Support Vector Machine (SVM) or logistic regression. In contrast, FMM-Head can be applied to multiple AEs for anomaly detection instead of classification with a single model. In [31] the waves are centered into the correct position by adding a loss regularization term, which is based on 100 random samples, whose coefficients are obtained through the non-ECG specific code from [23]. We instead perform warm-up on *all* the coefficients, obtained from the ECG-specific R code from [24]. Moreover, we explicitly consider α and β as circular variables, thus correctly estimating their values and correlations.

5 Conclusion and Future Works

We introduced a novel way to insert *a priori* knowledge into AEs for anomaly detection applied to ECG data. Our FMM-Head increased the AUROC of five baselines and reduced model size and inference time, thus allowing to perform real-time, explainable anomaly detection for continuously monitored patients. As future work, we will vary

the regression loss function to provide more precise FMM coefficients and better exploit parallel operations in TensorFlow to decrease the inference time.

Acknowledgements This work was supported by the Swedish Research Council (project "Scalable Federated Learning", no. 2021-04610) and by the King Abdullah University of Science and Technology (KAUST) Office of Research Administration (ORA) under Award No. ORA-CRG2021-4699.

References

1. Alamr, A., Artoli, A.: Unsupervised transformer-based anomaly detection in ECG signals. Algorithms **16**(3) (2023). https://doi.org/10.3390/a16030152, https://www.mdpi.com/1999-4893/16/3/152
2. AlGhatrif, M., Lindsay, J.: A brief review: history to understand fundamentals of electrocardiography. Journal of community hospital internal medicine perspectives **2**(1), 14383 (2012). https://doi.org/10.3402/jchimp.v2i1.14383
3. Atwood, J., Towsley, D.: Diffusion-convolutional neural networks. In: Proceedings of the 30th International Conference on Neural Information Processing Systems. p. 2001–2009. NIPS'16, Curran Associates Inc., Red Hook, NY, USA (2016). https://doi.org/10.5555/3157096.3157320
4. Bortolan, G., Christov, I., Simova, I.: Potential of rule-based methods and deep learning architectures for ecg diagnostics. Diagnostics **11**, 1678 (2021). https://doi.org/10.3390/diagnostics11091678
5. Chalapathy, R., Chawla, S.: Deep learning for anomaly detection: A survey. CoRR **abs/1901.03407** (2019), http://arxiv.org/abs/1901.03407
6. Chauhan, S., Vig, L.: Anomaly detection in ECG time signals via deep long short-term memory networks. In: 2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA). pp. 1–7 (2015). https://doi.org/10.1109/DSAA.2015.7344872
7. Chen, Y., Keogh, E.: ECG5000 dataset (2000), http://www.timeseriesclassification.com/description.php?Dataset=ECG5000
8. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial networks. Commun. ACM **63**(11), 139–144 (oct 2020). https://doi.org/10.1145/3422622
9. Hinton, G.E., Salakhutdinov, R.R.: Reducing the dimensionality of data with neural networks. Science **313**(5786), 504–507 (2006). https://doi.org/10.1126/science.1127647
10. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Computation **9**(8), 1735–1780 (1997). https://doi.org/10.1162/neco.1997.9.8.1735
11. Jammalamadaka, R., Sarma, Y.R.: Correlation coefficient for angular variables. In: Statistical Theory and Data Analysis II: Proceedings of the Second Pacific Area Statistical Conference. pp. 349–364. Pacific Area Statistical Conference, North-Holland, Amsterdam; New York City (1988), https://archive.org
12. Jang, J.H., Kim, T.Y., Lim, H.S., Yoon, D.: Unsupervised feature learning for electrocardiogram data using the convolutional variational autoencoder. PLOS ONE **16**(12), 1–16 (12 2021). https://doi.org/10.1371/journal.pone.0260612
13. Kaplan Berkaya, S., Uysal, A.K., Sora Gunal, E., Ergin, S., Gunal, S., Gulmezoglu, M.B.: A survey on ECG analysis. Biomedical Signal Processing and Control **43**, 216–235 (2018). https://doi.org/10.1016/j.bspc.2018.03.003
14. Kingma, D.P., Welling, M.: Auto-encoding variational Bayes (2022). https://doi.org/10.48550/arXiv.1312.6114

15. Lee, A.: Circular data. WIREs Computational Statistics **2**(4), 477–486 (2010). https://doi.org/10.1002/wics.98
16. Liu, P., Sun, X., Han, Y., He, Z., Zhang, W., Wu, C.: Arrhythmia classification of LSTM autoencoder based on time series anomaly detection. Biomedical Signal Processing and Control **71**, 103228 (2022). https://doi.org/10.1016/j.bspc.2021.103228
17. Malhotra, P., Ramakrishnan, A., Anand, G., Vig, L., Agarwal, P., Shroff, G.: LSTM-based encoder-decoder for multi-sensor anomaly detection. CoRR **abs/1607.00148** (2016), http://arxiv.org/abs/1607.00148
18. Nigusse, A.B., Mengistie, D.A., Malengier, B., Tseghai, G.B., Langenhove, L.V.: Wearable smart textiles for long-term electrocardiography monitoring—a review. Sensors **21**(12) (2021). https://doi.org/10.3390/s21124174, https://www.mdpi.com/1424-8220/21/12/4174
19. Pan, J., Tompkins, W.J.: A real-time QRS detection algorithm. IEEE Trans. on Biomedical Engineering **BME-32**(3), 230–236 (1985). https://doi.org/10.1109/TBME.1985.325532
20. Pintilie, I., Manolache, A., Brad, F.: Time series anomaly detection using diffusion-based models. In: 2023 IEEE International Conference on Data Mining Workshops (ICDMW). pp. 570–578 (2023). https://doi.org/10.1109/ICDMW60847.2023.00080
21. Qin, J., Gao, F., Wang, Z., Wong, D.C., Zhao, Z., Relton, S.D., Fang, H.: A novel temporal generative adversarial network for electrocardiography anomaly detection. Artificial Intelligence in Medicine **136**, 102489 (2023). https://doi.org/10.1016/j.artmed.2023.102489
22. Roy, M., Majumder, S., Halder, A., Biswas, U.: ECG-NET: A deep LSTM autoencoder for detecting anomalous ECG. Engineering Applications of Artificial Intelligence **124**, 106484 (2023). https://doi.org/10.1016/j.engappai.2023.106484
23. Rueda, C., Larriba, Y., Peddada, S.D.: Frequency Modulated Möbius Model accurately predicts rhythmic signals in biological and physical sciences. Scientific Reports **9**, 18701 (2019). https://doi.org/10.1038/s41598-019-54569-1
24. Rueda, C., Rodríguez-Collado, A., Fernández, I., Canedo, C., Ugarte, M.D., Larriba, Y.: A unique cardiac electrocardiographic 3D model. toward interpretable AI diagnosis. iScience **25**(12), 105617 (2022). https://doi.org/10.1016/j.isci.2022.105617
25. Sampson, M.: Continuous ECG monitoring in hospital: part 1, indications. British Journal of Cardiac Nursing **13**(2), 80–85 (2018). https://doi.org/10.12968/bjca.2018.13.2.80
26. Sampson, M.: Continuous ECG monitoring in hospital: part 2, practical issues. British Journal of Cardiac Nursing **13**(3), 128–134 (2018). https://doi.org/10.12968/bjca.2018.13.3.128
27. Shan, L., Li, Y., Jiang, H., Zhou, P., Niu, J., Liu, R., Wei, Y., Peng, J., Yu, H., Sha, X., Chang, S.: Abnormal ecg detection based on an adversarial autoencoder. Frontiers in Physiology **13**, 961724 (2022). https://doi.org/10.3389/fphys.2022.961724, https://doi.org/10.3389/fphys.2022.961724
28. Srivastava, A., Valkov, L., Russell, C., Gutmann, M.U., Sutton, C.: Veegan: reducing mode collapse in gans using implicit variational learning. In: Proceedings of the 31st International Conference on Neural Information Processing Systems. p. 3310–3320. NIPS'17, Curran Associates Inc., Red Hook, NY, USA (2017). https://doi.org/10.5555/3294996.3295090
29. Verardo, G.: Fmm-head. https://github.com/giacomoverardo/FMM-Head (2023)
30. Wagner, P., Strodthoff, N., Bousseljot, R.D., Kreiseler, D., Lunze, F.I., Samek, W., Schaeffter, T.: PTB-XL, a large publicly available electrocardiography dataset. Scientific Data **7**(1), 154 (2020). https://doi.org/10.1038/s41597-020-0495-6
31. Yang, Y., Rocher, M., Moceri, P., Sermesant, M.: Explainable electrocardiogram analysis with wave decomposition: Application to myocardial infarction detection. Statistical Atlases and Computational Models of the Heart. Regular and CMRxMotion Challenge Papers pp. 221–232 (2022). https://doi.org/10.1007/978-3-031-23443-9_21
32. Zheng, J., Zhang, J., Danioko, S., Yao, H., Guo, H., Rakovski, C.: A 12-lead electrocardiogram database for arrhythmia research covering more than 10,000 patients. Scientific Data **7**(1), 48 (2020). https://doi.org/10.1038/s41597-020-0386-x