

Article

AE-DD: Autoencoder-Driven Dictionary with Matching Pursuit for Joint ECG Denoising, Compression, and Morphology Decomposition

Fars Samann ¹  and Thomas Schanze ^{2,*}

¹ Department of Biomedical Engineering, University of Duhok, Duhok 42001, Kurdistan Region, Iraq; fars.samann@uod.ac

² Life Science Engineering, Technische Hochschule Mittelhessen, 35390 Gießen, Germany

* Correspondence: thomas.schanze@lse.thm.de

Abstract

Background: Electrocardiogram (ECG) signals are crucial for cardiovascular diagnosis, but their analysis face challenges from noise contamination, compression difficulties due to their non-stationary nature, and the inherent complexity of its morphological components, particularly for low-amplitude P- and T-waves obscured by noise. **Methodology:** This study proposes a novel, multi-stage framework for ECG signal denoising, compressing, and component decomposition. The proposed framework leverages the sparsity of ECG signal to denoise and compress these signals using autoencoder-driven dictionary (AE-DD) with matching pursuit. In this work, a data-driven dictionary was developed using a regularized autoencoder. Appropriate trained weights along with matching pursuit were used to compress the denoised ECG segments. This study explored different weight regularization techniques: L1- and L2-regularization. **Results:** The proposed framework achieves remarkable performance in simultaneous ECG denoising, compression, and morphological decomposition. The L1-DAE model delivers superior noise suppression (SNR improvement up to 18.6 dB at -3 dB input SNR) and near-lossless reconstruction ($MSE < 10^{-5}$). The L1-AE dictionary enables high-fidelity compression (CR = 28:1 ratio, $MSE \approx 0.58 \times 10^{-5}$, PRD = 2.1%), outperforming non-regularized models and traditional dictionaries (DCT/wavelets), while its trained weights naturally decompose into interpretable sub-dictionaries for P-wave, QRS complex, and T-wave enabling precise, label-free analysis of ECG components. Moreover, the learned sub-dictionaries naturally decompose into interpretable P-wave, QRS complex, and T-wave components with high accuracy, yielding strong correlation with the original ECG ($r = 0.98$, $r = 0.99$, and $r = 0.95$, respectively) and very low MSE (1.93×10^{-5} , 9.26×10^{-4} , and 3.38×10^{-4} , respectively). **Conclusions:** This study introduces a novel autoencoder-driven framework that simultaneously performs ECG denoising, compression, and morphological decomposition. By leveraging L1-regularized autoencoders with matching pursuit, the method effectively enhances signal quality while enabling direct decomposition of ECG signals into clinically relevant components without additional processing. This unified approach offers significant potential for improving automated ECG analysis and facilitating efficient long-term cardiac monitoring.

Keywords: autoencoder-driven dictionary; matching pursuit; weight regularization; autoencoder; denoising autoencoder



Academic Editor: Juraj Gregaň

Received: 28 July 2025

Revised: 10 September 2025

Accepted: 15 September 2025

Published: 17 September 2025

Citation: Samann, F.; Schanze, T. AE-DD: Autoencoder-Driven Dictionary with Matching Pursuit for Joint ECG Denoising, Compression, and Morphology Decomposition. *AI* **2025**, *6*, 234. <https://doi.org/10.3390/ai6090234>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Electrocardiogram (ECG) is a non-invasive test method widely used for diagnosing cardiovascular diseases. The morphology of the ECG—including the P-wave, QRS complex, and T-wave—is an important metric for assessing a patient's cardiac health. For example, a comprehensive study in 2015 [1] showed that both short and long P-wave durations are associated with an elevated risk of atrial fibrillation (AF). Another study suggested that T-wave alternans (TWA) is a highly sensitive predictor of cardiac death in patients with mild-to-moderate congestive heart failure undergoing cardiac rehabilitation [2]. Moreover, in [3], it was shown that the use of microvolt TWA is more accurate than QRS duration in identifying groups at high and low risk of dying among heart failure patients. Therefore, accurately detecting and estimating the location and duration of P-waves and T-waves is essential for the reliable diagnosis of cardiac conditions. However, these waves typically have low amplitudes and are, therefore, highly susceptible to being obscured by noise or artifacts, particularly during prolonged ECG recordings, such as those obtained through ambulatory cardiac monitoring (ACM). In such settings, various types of physical noise—such as baseline wander, motion artifacts, and electrode displacement—introduced by patient movement during routine daily activities can significantly degrade ECG signal quality [4]. During prolonged ECG monitoring, the volume of data recorded for subsequent analysis is substantial. Consequently, compression techniques are commonly employed to reduce the amount of data that must be stored or transmitted, enabling more efficient handling and analysis of long-term ECG recordings.

The current state-of-the-art approach for denoising biosignals is the denoising autoencoder (DAE). Numerous DAE architectures have been proposed in the literature, and their performance largely depends on two key factors: (1) the type of input segment (whether it is QRS-aligned or a non-aligned ECG segment), and (2) the structure of the DAE (whether it employs fully connected layers or convolutional layers). Most studies utilize long, non-aligned ECG segments in conjunction with deep convolutional denoising autoencoders (CDAEs) to exploit temporal correlations and achieve strong denoising performance [4–7]. In contrast, fully connected DAEs typically require QRS-aligned input segments but can achieve effective denoising with fewer hidden layers [8–11]. A novel DAE architecture, known as the Running DAE model, was proposed in [4], and it effectively leverages the high correlation between successive overlapping segments—commonly referred to as the sliding window—to successfully remove unwanted noise from the signal.

Prolonged ECG recordings, like ACM, enable continuous assessment of heart functions throughout a person's daily activities, including sleep, exercise, and routine tasks. Unlike traditional in-clinic ECG tests that offer only a brief snapshot of cardiac activity, ACM provides a more comprehensive and dynamic view of the heart's condition over an extended period. However, this continuous monitoring generates a large volume of ECG data, posing significant challenges in terms of data storage, transmission, and processing—particularly in resource-constrained environments. Efficient medical data management becomes critical to ensure that, without overburdening the communication infrastructure or computational resources, clinically relevant information is preserved and made accessible for timely diagnosis. Several methods have been proposed in the literature for ECG signal compression, ranging from traditional approaches, such as the direct data compression method [12], compressed sensing [13–16], and support value and tensor decomposition method [17,18], to more advanced techniques based on autoencoders [6]. The effectiveness of sparse coding largely depends on the choice of dictionary used to represent the signal. However, due to the complex and highly variable morphology of ECG signals, capturing their structure with predefined dictionaries is particularly challenging, often limiting the compression performance of classical methods. In [11], a regularized DAE model demonstrated highly

promising results—not only in effectively denoising ECG signals, but also in capturing underlying features that closely resemble key morphological components of the ECG, such as the P-wave, QRS complex, and T-wave.

This study introduces an autoencoder-based dictionary framework for denoising, compressing and decomposing ECG morphology. To generate suitable data-driven dictionary for these tasks, we examined two weight regularization strategies: L1- and L2-regularization. L1-regularization enforces sparsity by forcing many weights to zero, producing representations that closely align with ECG morphologies [11]. In contrast, L2-regularization penalizes large weights more strongly and distributes them more uniformly, resulting in redundant morphological representations [11]. Based on these observations, an autoencoder with L1-weight regularization was adopted to construct distinct dictionaries for the P-wave, QRS complex, and T-wave. When combined with the matching pursuit algorithm, the proposed framework achieves effective and interpretable representation of ECG morphology, thereby enabling automated ECG decomposition without any labeling.

2. Materials and Methods

The proposed framework consists of three stages, as shown in Figure 1: denoising, compression, and decomposition. Once the noisy ECG signal is denoised using a denoising autoencoder, the resulting clean signal is compressed using an autoencoder-driven dictionary D_{ECG} , which is derived from the trained weights of the decoder. The compression is, subsequently, performed using the matching pursuit algorithm. The reconstructed ECG segment is further decomposed into its fundamental morphological components—namely, the P-wave, QRS complex, and T-wave—using structured sparse approximation. This decomposition is achieved by employing three distinct sub-dictionaries, denoted as D_{P-wave} , D_{QRS} , and D_{T-wave} . These sub-dictionaries are formed by thresholding the learned weights based on both the amplitude and temporal location of each atom in relation to the known ECG morphology. The amplitude–time threshold is used to identify and group atoms that significantly contribute to each waveform component, enabling a morphology-aware separation of the ECG signal. This decomposition is also carried out using the matching pursuit algorithm.

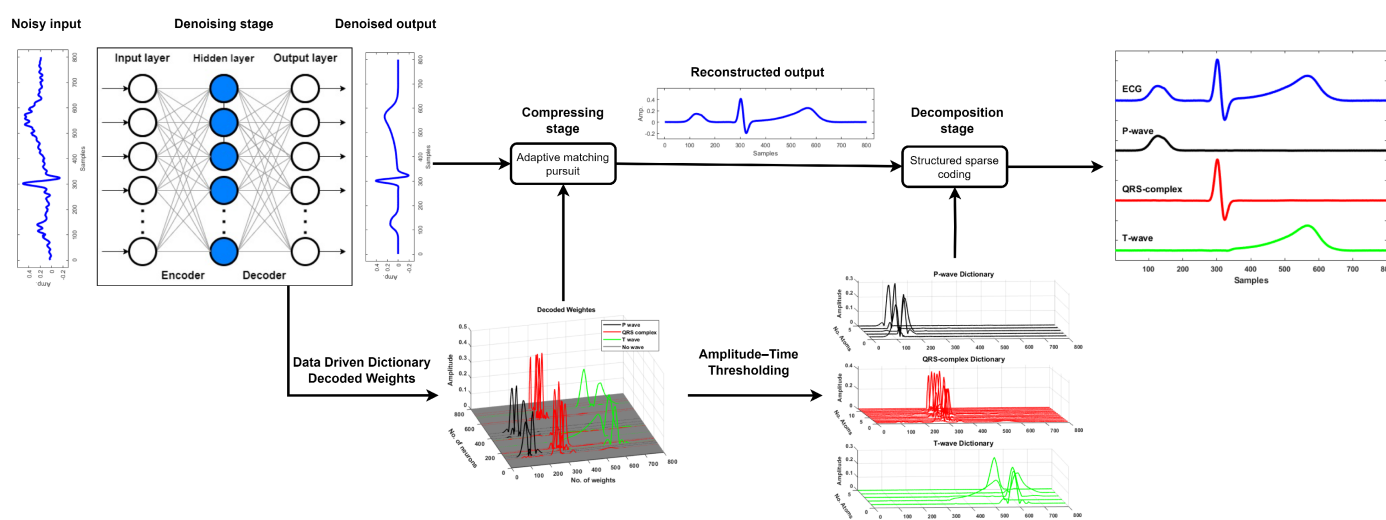


Figure 1. The proposed framework consists of the denoising stage, compressing stage, and decomposition stage. The denoising stage is based on a regularized DAE model, whereas the compressing stage is based on a regularized AE model to generate the autoencoder-driven dictionary.

2.1. Denoising Stage

An autoencoder (AE) is an unsupervised machine learning model commonly used for solving regression problems. A typical AE consists of an input layer, one or more hidden layers, and an output layer, with the input and output layers having the same number of neurons, as shown in Figure 2. To help the AE to learn a compact and meaningful representation of the data, the hidden layer typically has fewer neurons than the input and output layers—a design known as the bottleneck effect. This bottleneck layer compresses the input, encouraging the model to retain only the most relevant features while eliminating redundant information. Such dimensionality reduction improves learning efficiency and generalization [11]. However, choosing the right size for the bottleneck layer is essential to prevent underfitting or overfitting [10]. As a general rule of thumb, the number of hidden neurons is typically set to around half the number of neurons in the input/output layer (see Figure 2). This is a simple AE version, many AE have more hidden layers. Mathematically, the AE encodes the input segment into latent representation; then, decodes it back to produce the desired output, as follows.

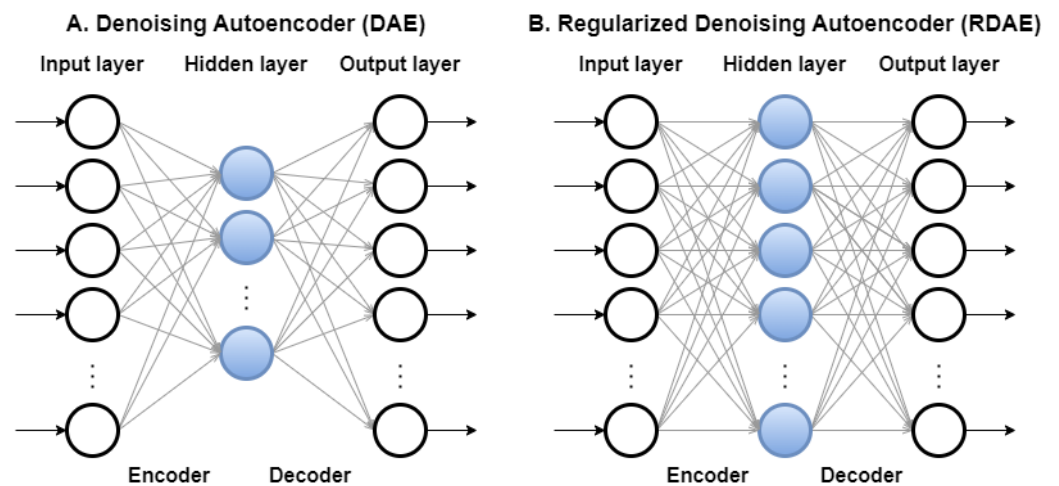


Figure 2. (A) Denoising autoencoder. (B) Regularized denoising autoencoder.

$$\mathbf{z} = f(\mathbf{W}_e \hat{\mathbf{x}} + \mathbf{b}_e), \quad (1)$$

$$\mathbf{y} = f(\mathbf{W}_d \mathbf{z} + \mathbf{b}_d), \quad (2)$$

where

- $\mathbf{x}$: Original segment (clean);
- $\hat{\mathbf{x}}$: Input segment (noisy);
- $\mathbf{z}$: Latent (encoded) representation;
- $\mathbf{y}$: Output segment (denoised);
- $\mathbf{W}_e$: Weight matrix of the encoder;
- $\mathbf{b}_e$: Bias vector of the encoder;
- $\mathbf{W}_d$: Weight matrix of the decoder;
- $\mathbf{b}_d$: Bias vector of the decoder;
- $f(\cdot)$: Activation function (e.g., ReLU).

Note: in the case of autoencoder, the input segment $\hat{\mathbf{x}}$ is equal to the desired segment $\mathbf{x}$. The AE models learn by minimizing the reconstruction error between the desired segments and the output segments, effectively capturing the underlying features of the training data. The only difference between a standard autoencoder (AE) and a denoising autoencoder (DAE) lies in the data used for learning: while the AE uses clean segments as both the input and target, the DAE takes noisy segments as the input and learns to reconstruct the

corresponding clean target segments by minimizing the error between the output and the clean reference.

In this work, the bottleneck effect is introduced by forcing the model to have small or nearly zero weights for the unwanted features of the input segment employing either L2- or L1-weight regularization.

Mathematically, a regularized loss function combines a data fidelity term (loss function) with a regularization term to prevent overfitting and improve generalization. The general form is expressed as follows:

$$\mathcal{L}_{\text{reg}} = \mathcal{L}(\mathbf{x}, \mathbf{y}; \mathbf{w}) + \lambda \mathcal{R}(\mathbf{w}), \quad (3)$$

where $\mathcal{L}(\mathbf{x}, \mathbf{y}; \mathbf{w})$ denotes the primary loss function measuring the discrepancy between the ground truth $\mathbf{x}$, the model prediction $\mathbf{y}$, $\mathcal{R}(\mathbf{w})$ is the regularization term applied to model parameters $\mathbf{w}$, and λ is a regularization parameter.

In this work, we considered two forms of regularized loss functions, where both the loss term and the regularization term adopt the same norm. Specifically, we explored either a fully L2-regularized formulation or a fully L1-regularized formulation. Mathematically, the regularization term added to loss function as a combination of mean square error (MSE) with L2-regularization term or mean absolute error (MAE) with L1-regularization, as follow,

$$\mathcal{L}_{\text{L2-norm}} = \frac{1}{M} \sum_{i=1}^M (y_i - x_i)^2 + \lambda \sum_{i=1}^M w_i^2. \quad (4)$$

$$\mathcal{L}_{\text{L1-norm}} = \frac{1}{M} \sum_{i=1}^M |y_i - x_i| + \lambda \sum_{i=1}^M |w_i|. \quad (5)$$

These formulations illustrate the integration of the loss term with either L1- or L2-regularization, which are commonly used for promoting sparsity or penalizing large weights, respectively.

2.2. Compressing Stage

Sparse modeling is a mathematical representation of a noiseless signal $\mathbf{y} \in \mathbb{R}^{N \times 1}$ by a linear combination of a few basis elements (denoted as atoms), which are extracted from a dictionary

$$\mathbf{D} = [\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3, \dots, \mathbf{d}_M] \in \mathbb{R}^{N \times M}$$

of M atoms (see Figure 3), such that

$$\mathbf{y} = \mathbf{D}\boldsymbol{\alpha}, \quad (6)$$

where $\boldsymbol{\alpha}$ is a sparse coefficient vector. The dictionary $\mathbf{D}$ with $M > N$ is known as an overcomplete dictionary, and when $M = N$, it is referred to as a complete dictionary. In this work, the output of the DAE model is considered as the noiseless signal $\mathbf{y}$, and the dictionary $\mathbf{D}$ is obtained from the trained weights $\mathbf{W}_d$ of AE's decoder. The main goal is to estimate the sparse vector $\hat{\boldsymbol{\alpha}}$, which fulfills the assumption of Equation (6). The sparse modeling problem can be tackled by using error-constraint problem. We propose the mean square error as an error-constraint as follows:

$$\hat{\boldsymbol{\alpha}} = \min_{\boldsymbol{\alpha}} \|\boldsymbol{\alpha}\|_1 \quad \text{s.t.} \quad \frac{1}{N} \|\mathbf{y} - \mathbf{D}\boldsymbol{\alpha}\|_2^2 \leq \epsilon. \quad (7)$$

Here, ϵ represents the threshold for the allowable reconstruction error. In this work, the matching pursuit algorithm is primarily employed due to its simplicity and computational efficiency. Since the focus is on biomedical signals, such as ECG, the error threshold ϵ is set

to be as small as possible (e.g., 10^{-5}) to ensure accurate and low reconstruction error of the compressed signal $\hat{\mathbf{y}} = \mathbf{D}\hat{\boldsymbol{\alpha}}$. Three dictionaries based on the autoencoder-driven dictionary (AE-DD) were used to compress the denoised ECG segments: D_{AE} , D_{L2-AE} , and D_{L1-AE} . As suggested by [4], four predefined dictionaries, e.g., DCT, Sym6, db6, and coif2, were also considered as well-suited dictionaries for representing ECG segments (see Figure 4). We compared the performance of our learned dictionaries against these predefined ones.

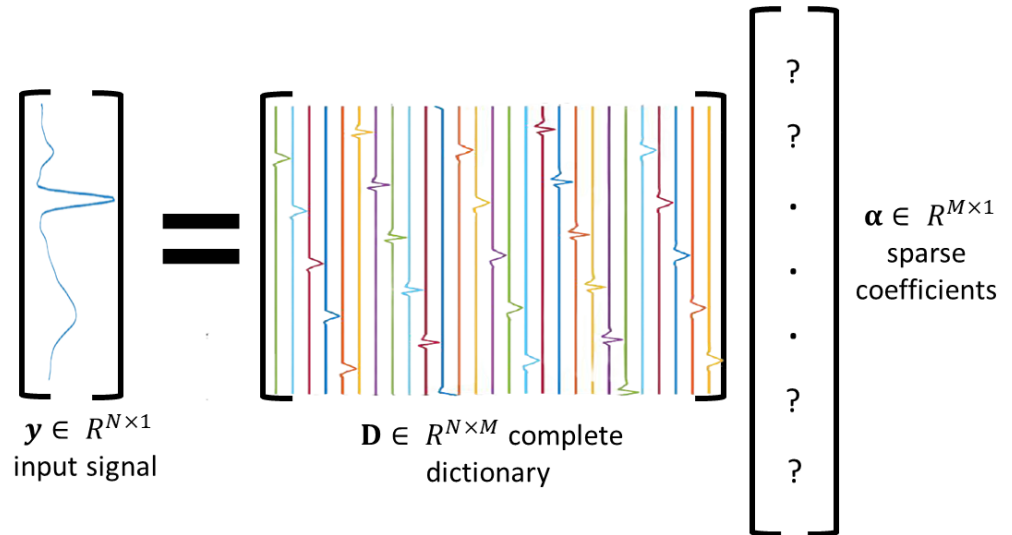


Figure 3. Sparse approximation of a signal using a predefined dictionary. An input signal $\mathbf{y} \in \mathbb{R}^{N \times 1}$ is approximated as a linear combination of atoms from a complete dictionary $\mathbf{D} \in \mathbb{R}^{N \times M}$, where $M = N$. The objective is to determine the sparse coefficient vector $\boldsymbol{\alpha} \in \mathbb{R}^{M \times 1}$, containing mostly zeros, such that $\mathbf{y} \approx \mathbf{D}\boldsymbol{\alpha}$. This sparse representation is widely used in applications, such as signal reconstruction, compression, and denoising.

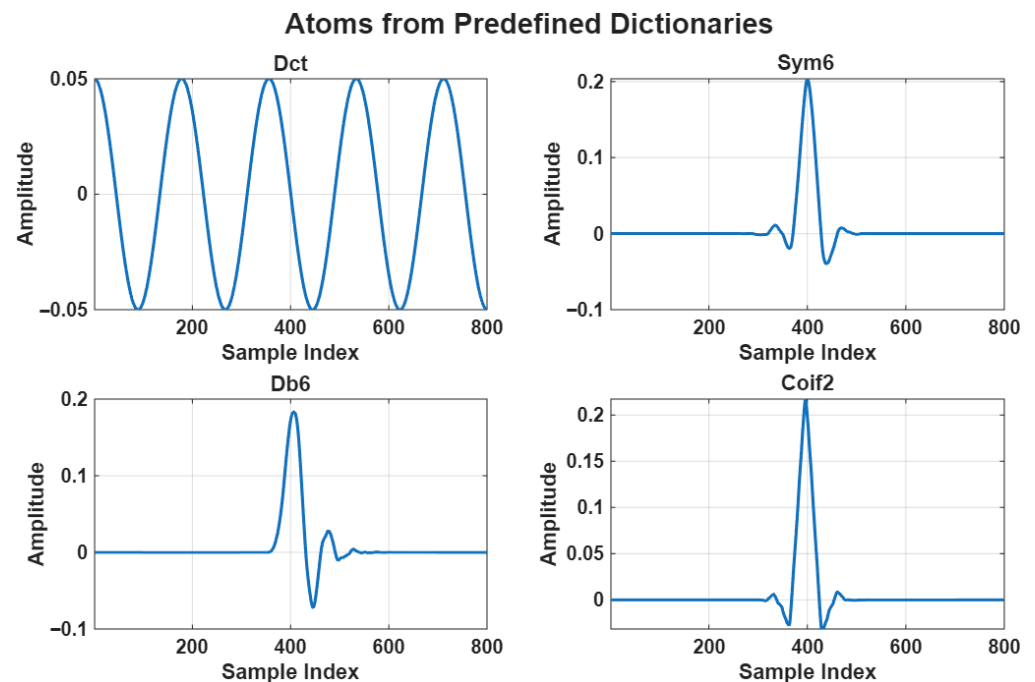


Figure 4. Representative atoms from predefined dictionaries used for signal compression. The plots display typical atoms from the Discrete Cosine Transform (DCT), Symlet 6 (sym6), Daubechies 6 (db6), and Coiflet 2 (coif2) dictionaries.

2.3. Decomposing Stage

In this work, structured sparse coding was employed to decompose the ECG morphology into its fundamental components. This was achieved using the trained weights of an L1-regularized autoencoder, denoted as $\mathbf{D}_{\text{L1-AE}}$. This dictionary was grouped into three subdictionaries corresponding to distinct ECG components:

$$\mathbf{D}_{\text{L1-AE}} = [\mathbf{D}_{\text{P-wave}} \mid \mathbf{D}_{\text{QRS}} \mid \mathbf{D}_{\text{T-wave}}]. \quad (8)$$

The grouping was carried out through an amplitude–time thresholding technique, which identifies and assigns atoms that make significant contributions to each morphological component based on their temporal location and amplitude. Matching pursuit was applied using the full dictionary to estimate a sparse coefficient vector $\hat{\mathbf{a}}$:

$$\hat{\mathbf{a}} = \text{MP}(\hat{\mathbf{y}}, \mathbf{D}_{\text{L1-AE}}), \quad (9)$$

where $\hat{\mathbf{y}}$ denotes the reconstructed ECG segment obtained from the compression stage. To isolate a specific morphological component, a binary mask $\mathbf{h}_i$ (with ones in the positions corresponding to subdictionary i and zeros elsewhere) was used to extract the relevant part of the sparse code. The component reconstruction is then given by the following:

$$\text{Component}_i = \mathbf{D}_{\text{L1-AE}} \cdot (\hat{\mathbf{a}} \odot \mathbf{h}_i), \quad (10)$$

where $i \in \{\text{P-wave}, \text{QRS}, \text{T-wave}\}$, and $\odot$ denotes element-wise multiplication. This allows the reconstruction of each ECG component independently from its corresponding subdictionary.

2.4. Hyperparameters and Computational Environment

All the hyperparameters used for training both the denoising autoencoder (DAE) and the autoencoder (AE) models are summarized in Table 1. The choice of loss function depends on the regularization type: mean absolute error (MAE) was used with L1-regularization, while mean squared error (MSE) was paired with L2-regularization.

Table 1. Hyperparameters of the denoising autoencoder (DAE) and autoencoder (AE).

Parameter	DAE Model	AE Model
Activation Function	ReLU	ReLU
Regularization (Encoder)	$\lambda = 1 \times 10^{-7}$	$\lambda = 1 \times 10^{-5}$
Regularization (Decoder)	$\lambda = 1 \times 10^{-6}$	$\lambda = 1 \times 10^{-5}$
Loss Function	MSE or MAE	MSE or MAE
Optimizer	Adam	Adam
Learning Rate	0.0001	0.001
Epochs	100	50

The models were implemented in Python 3.7 using the TensorFlow 2.11.0 library. All experiments were conducted on a personal laptop equipped with an Intel Core i5 processor and 20 GB of RAM.

2.5. Evaluation Metrics

The performance of the denoising stage was evaluated using signal-to-noise ratio improvement (SNR_{imp}) and mean squared error (MSE), as defined below:

$$\text{SNR}_{\text{imp}} = 10 \log_{10} \left(\frac{\sum_{n=1}^N |\hat{\mathbf{x}}[n] - \mathbf{x}[n]|^2}{\sum_{n=1}^N |\mathbf{y}[n] - \mathbf{x}[n]|^2} \right), \quad (11)$$

$$\text{MSE} = \frac{1}{N} \sum_{n=1}^N (\mathbf{x}[n] - \mathbf{y}[n])^2, \quad (12)$$

where

- $\mathbf{x}$ is the original ECG segment;
- $\hat{\mathbf{x}}$ is the noisy ECG segment;
- $\mathbf{y}$ is the denoised output.

A higher SNR_{imp} indicates better noise reduction performance, while a lower MSE implies a closer match between the denoised and the original ECG signal.

The effectiveness of the compression stage is evaluated using peak signal-to-noise ratio (PSNR), mean squared error (MSE), percentage root-mean-square difference (PRD), and compression ratio (CR), which are defined as follows:

$$\text{MSE} = \frac{1}{N} \sum_{n=1}^N (\mathbf{y}[n] - \hat{\mathbf{y}}[n])^2, \quad (13)$$

$$\text{PSNR} = 10 \log_{10} \left(\frac{\text{MAX}^2}{\text{MSE}} \right), \quad (14)$$

$$\text{PRD} = 100 \times \frac{\|\mathbf{y} - \hat{\mathbf{y}}\|_2}{\|\mathbf{y}\|_2}, \quad (15)$$

$$\text{CR} = \frac{N \cdot b}{K \cdot (b + \lceil \log_2 M \rceil)}, \quad (16)$$

where

- $\hat{\mathbf{y}}$ is the reconstructed ECG segment;
- N is the length (in samples) of the ECG segment;
- K is the number of non-zero coefficients used for reconstruction;
- b is the number of bits per coefficient value (e.g., 16 bits);
- M is the size of the dictionary (i.e., number of atoms);
- MAX is the maximum absolute value in $\mathbf{y}$.

Higher PSNR values and lower MSE and PRD values indicate improved signal fidelity after compression, while a higher CR reflects greater compression efficiency.

The effectiveness of the decomposition stage is assessed using the correlation coefficient (r) and mean squared error (MSE) between the original component $\mathbf{c}$ (e.g., P-wave, QRS complex, or T-wave) and its reconstructed version $\hat{\mathbf{c}}$:

$$r = \frac{\sum_{n=1}^N (\mathbf{c}[n] - \bar{\mathbf{c}})(\hat{\mathbf{c}}[n] - \bar{\hat{\mathbf{c}}})}{\sqrt{\sum_{n=1}^N (\mathbf{c}[n] - \bar{\mathbf{c}})^2} \sqrt{\sum_{n=1}^N (\hat{\mathbf{c}}[n] - \bar{\hat{\mathbf{c}}})^2}}, \quad (17)$$

$$\text{MSE} = \frac{1}{N} \sum_{n=1}^N (\mathbf{c}[n] - \hat{\mathbf{c}}[n])^2, \quad (18)$$

where

- $\mathbf{c}$ is the original ECG component (P, QRS, or T);
- $\hat{\mathbf{c}}$ is the reconstructed ECG component;
- $\bar{\mathbf{c}}$ and $\bar{\hat{\mathbf{c}}}$ are their respective means.

A higher r indicates stronger similarity in morphology, while a lower MSE reflects better reconstruction accuracy of the decomposed component. Note: the original ECG components used for reference were segmented manually.

3. Results and Discussion

The absence of ground truth signals is a crucial problem when assessing a novel denoising or compressing technique. Typically, noise is superimposed on the raw recorded ECG signals. As a result, the evaluation of these denoising techniques is partially acceptable. However, the majority of the advanced DAE models were tested using filtered ECG signals as the ground truth. Therefore, in order to verify the effectiveness of the proposed framework, we considered simulated ECG signals. A simulating model proposed by [19] was used to generate 160 ECG recordings, all representing normal and clean ECG signals with heart rates ranging between 60 and 100 beats per minute. The suggested models were trained using 110 of the 160 simulated ECG recordings that were sampled at a sampling frequency of 1 KHz, and their performance was assessed using the remaining recordings. One hundred QRS aligned ECG segments with length of 800 samples (about duration 0.8 s) were collected from each recording.

To evaluate the denoising stage, both the training and testing datasets were contaminated with a mixture of physiological noise types—namely baseline wander, motion artifacts, and electrode movement—collected from [20]. Various input signal-to-noise ratio (SNR) levels were applied: for the training set, input SNR values of ∞ , 5, 0, and -5 dB were used; and for the testing set, input SNR values of ∞ , 3, 0, and -3 dB were considered. The denoised ECG segments were further compressed by the compressing stage; then, the reconstructed ECG segments were decomposed into their main components: P-wave, QRS complex, and T-wave.

The autoencoder model with and without weight regularization was trained with clean ECG segments of the training set (the trained decoded weights of different AE models are presented in Figure 5). The trained L1-AE model had some weights related to components of ECG, such as P-wave, QRS complex, and T-wave, as shown in Figure 6.

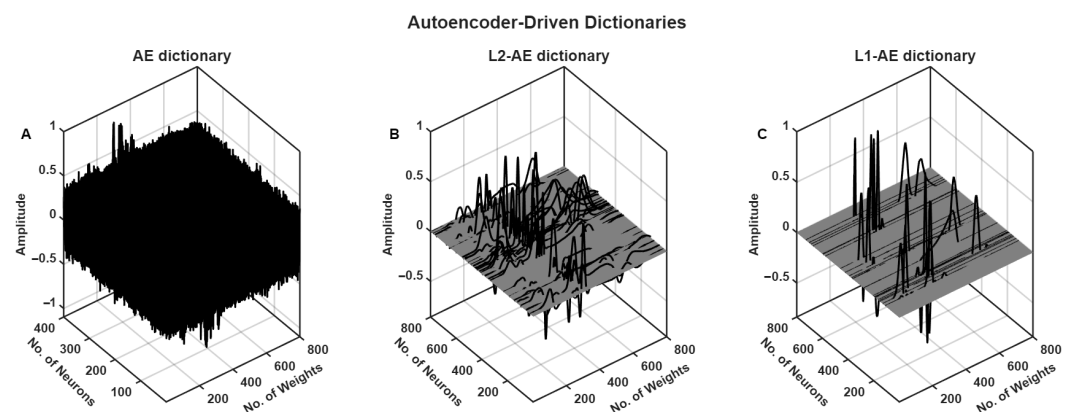


Figure 5. The autoencoder-driven dictionary in the case of (A) AE, (B) L2-AE, and (C) L1-AE.

The performance of the denoising stage was evaluated for the DAE models with and without the weight regularization technique at different input SNR levels. Figure 7 clearly shows that the DAE model with L1-regularization (L1-DAE) achieved better denoising performance compared to the L2-DAE and DAE models, especially in the case of input SNR ∞ , 3, and 0 dB. In the case of input SNR = -3 dB, the L2-DAE model achieved slightly higher $\text{SNR}_{\text{imp}} = 18.75$ dB compared to L1-DAE $\text{SNR}_{\text{imp}} = 18.60$ dB. This slight improvement by the L2-DAE model under the challenging -3 dB SNR conditions may be attributed to the nature of the regularization terms. L2-regularization typically encourages smaller, more evenly distributed weight values, which can help the model better generalize when the input signal is heavily corrupted by noise. Under such low SNR conditions, the L2 norm is more effective in preventing large fluctuations in the learned parameters, leading to a more stable reconstruction despite the high noise levels. On the other hand,

L1-regularization promotes sparsity, which generally improves denoising performance by focusing on dominant features, but this may be less advantageous when the noise is extremely strong and widespread, as in -3 dB SNR. Therefore, the numerical advantage of L2-DAE here reflects its increased robustness to heavy noise, complementing the strengths of L1-DAE observed at a higher input SNR.

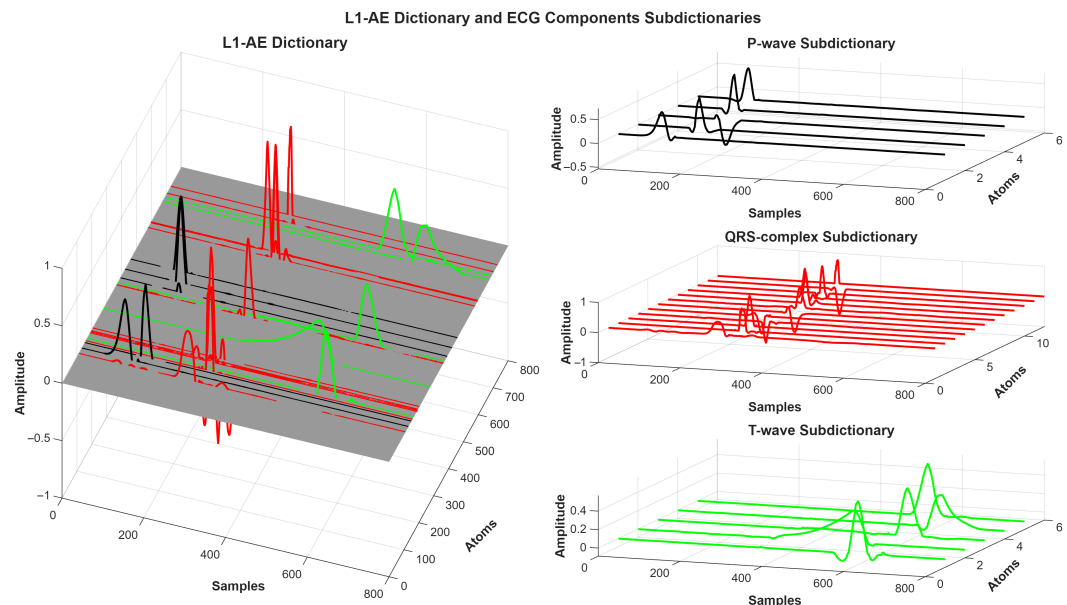


Figure 6. Visualization of the L1-AE dictionary and its decomposition into ECG component subdictionaries. The left panel displays the full L1-AE dictionary, with atoms color-coded by ECG component: black for P-wave, red for QRS complex, and green for T-wave. The right panels show the corresponding subdictionaries for each ECG component, highlighting the distinct waveform patterns learned for P-waves (**top**, black), QRS complexes (**middle**, red), and T-waves (**bottom**, green). Each subdictionary illustrates the characteristic morphology and temporal localization of its respective ECG component.

Figure 8 shows that the L1-AE model, which was trained using weight regularization, effectively compressed the denoised ECG segment with a high compression ratio ($CR = 28$), obtaining the lowest MSE and highest PSNR for all input SNR values. The DAE model, on the other hand, was unable to learn efficient basis functions for representing the ECG signal in the absence of weight regularization (see Figure 5). This highlights the critical role of weight regularization in enabling the model to learn more compact and meaningful representations using fewer atoms. Moreover, the use of the mean squared error (MSE) as an error constraint in Equation (7) ensures the quality of the reconstructed ECG segment following the compression stage, maintaining a reconstruction error below 1×10^{-5} , as clearly demonstrated in Figure 8.

The learned AE-DD dictionary, e.g., L1-AE, was also compared to a number of pre-defined dictionaries, including DCT, Sym6, db6, and coif2, which were found to have a strong correlation with the ECG signals reported in [4]. Table 2 highlights the effectiveness of the proposed method compared to different predefined dictionaries. The L1-regularized autoencoder (L1-AE) dictionary attained the best overall performance, while requiring the fewest atoms at 18, with the highest peak signal-to-noise ratio (PSNR) of 47.84 dB, the lowest mean squared error (MSE) of 0.58×10^{-5} , the lowest PRD of 2.13, and the highest compression ratio of 28. This demonstrates that the L1-AE dictionary provides a more compact and accurate representation of ECG signals, leading to improved compression results compared to traditional dictionaries, such as DCT and various wavelets. Additionally, the proposed compression method demonstrated superior performance compared to

various existing ECG compression techniques, as summarized in Table 3. Specifically, it achieved a notably high compression ratio of 28 while maintaining acceptable (PRD) of 2.13, outperforming most referenced methods in terms of both compression efficiency and signal fidelity.

Table 2. The performance of the compressing stage with respect to different dictionaries in terms of PSNR, MSE, compression ratio (CR), and number of atoms, where input SNR = −3 dB.

Dictionary	PSNR (dB)	MSE	PRD	CR	Atoms
DCT	45.45	0.92×10^{-5}	2.73	9	59
Sym6	45.56	0.89×10^{-5}	2.7	17	28
db6	45.59	0.89×10^{-5}	2.69	16	31
coif2	45.40	0.93×10^{-5}	2.75	17	29
DAE	45.17	0.98×10^{-5}	2.83	4	117
L2-AE	45.56	0.90×10^{-5}	2.70	24	21
L1-AE	47.84	0.58×10^{-5}	2.13	28	18

Table 3. Performance comparison of the different ECG compression techniques taken from [21] (experimental settings, such as dataset, noise characteristics, and compression ratio calculation methods, may differ from the present study).

Compression Technique	CR	PRD
AZTEC [22]	10	28
Improved modified AZTEC [23]	9.91	7.99
FAN/SAPA [24]	3.0	4.0
SLOPE [25]	4.8	7.0
TP [26]	2.5	5.1
SPIHT [27]	8	6.7
Al-Shrouf [28]	11.6	5.3
Perceptual Masks [29]	3.5	1.24
Quantized DCT coefficients [30]	10.9	3
Mukhopadhy [31]	15.72	7.89
Mukhopadhy [32]	7.18	0.023
Proposed method	28	2.13

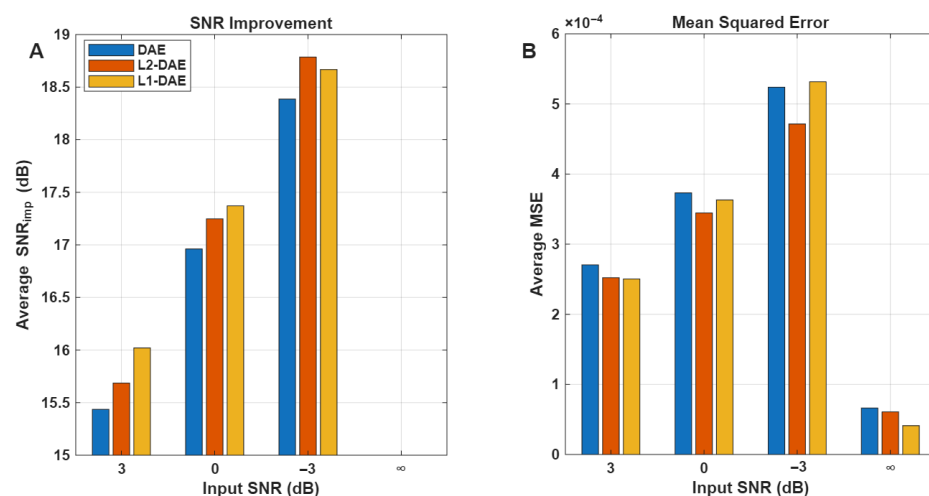


Figure 7. Comparison of the denoising performance of different DAE models, with and without weight regularization, in terms of (A) SNR_{imp} and (B) MSE. Note: SNR_{imp} had zero value because the original and input segment were the same when input SNR = ∞.

Figure 9 presents the complete signal processing pipeline of the proposed framework, demonstrating three key stages: denoising, compression, and morphological decomposition. The time course analysis revealed accurate component decomposition, where the P-waves (visible at Samples 100–200) and T-wave (Samples 450–650) were clearly resolved despite their lower amplitudes relative to the R-peak. Each component was clearly isolated, showing minimal overlap, which verifies the framework's capacity for morphological decomposition. This decomposition enables precise analysis of individual ECG waves, which is particularly valuable in clinical applications such as arrhythmia detection, ischemia monitoring, and cardiac pathology characterization. Overall, the results confirm that the proposed model not only performs robust denoising and compression, but also enables interpretable and clinically meaningful decomposition of the ECG signal.

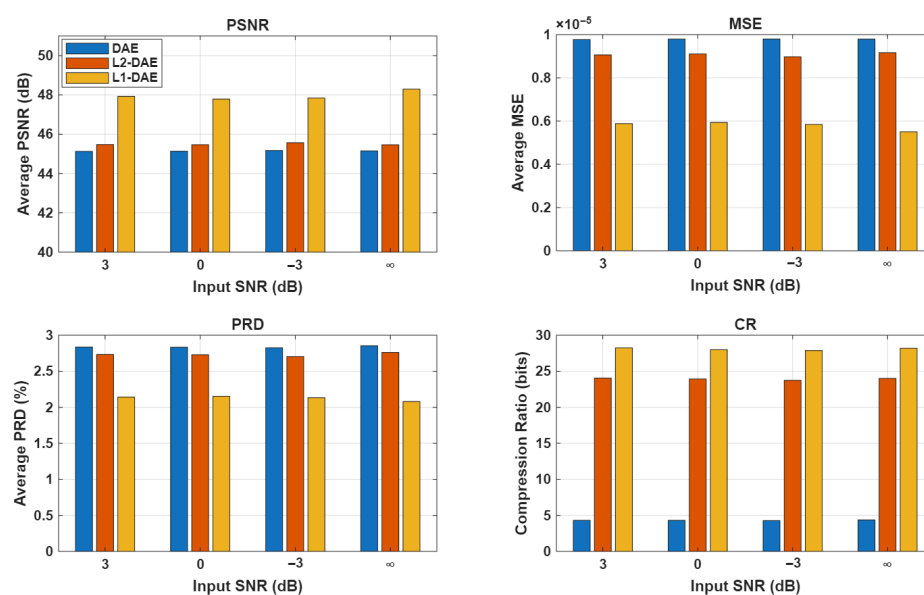


Figure 8. Comparison of the compressing performance using the different trained weight AE models, with and without weight regularization, in terms of PSNR, MSE, PRD and compression ratio (CR).

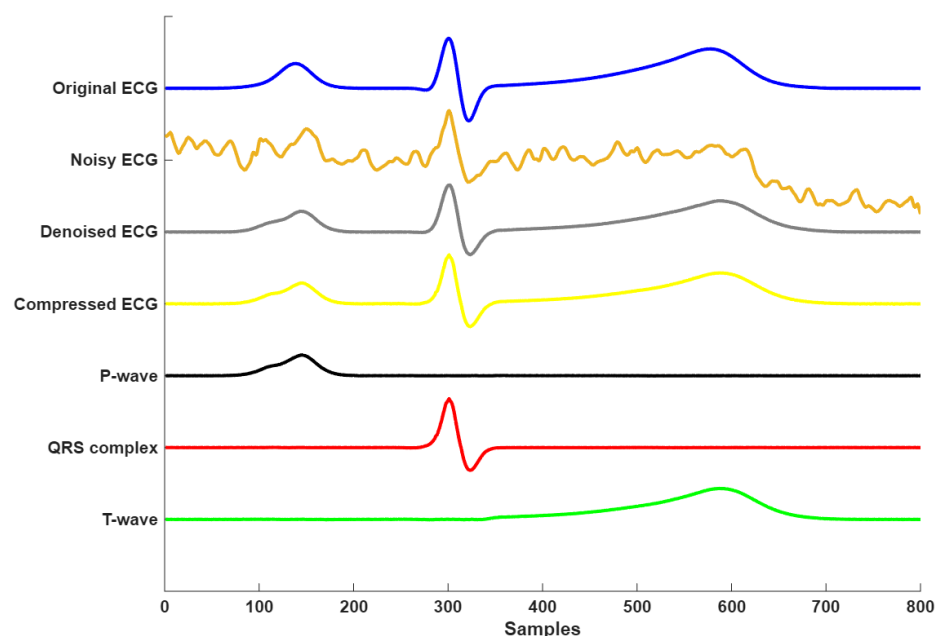


Figure 9. Time course of the ECG segments in samples after denoising (Gray), compressing (Yellow), decomposition to its components, P-wave (Black), QRS complex (Red), and T-wave (Green). Note: L1-DAE was used to denoised the noisy ECG segment of input SNR = −3 dB, and L1-AE was used to compress and decompose.

To assess the generalization performance of the proposed autoencoder-driven dictionary approach, we considered real recorded ECG signals from the St. Petersburg INCART 12-lead Arrhythmia Database [33], focusing on QRS-aligned ECG segments categorized into four AAMI classes: Normal (N), Ventricular ectopic beat (V), Supraventricular ectopic beat (S), and Fusion beat (F). Each segment was 0.78 s in duration (corresponding to 200 samples at a sampling rate of 257 Hz).

All segments underwent preprocessing using a third-order butterworth bandpass filter (0.5–40 Hz) to remove baseline wander and high-frequency noise. An adaptive screening method, as described by [4], was then applied to identify and retain low-noise, outlier-free segments suitable for model evaluation. Following screening, 80% of the segments from each class were used to train the proposed L1-regularized autoencoder (L1-AE) model, while the remaining 20% were reserved for evaluating the model's performance in compressing and decomposing the morphological structure of ECG signals.

Table 4 summarizes the number of segments collected before and after filtering and from the adaptive screening across the four AAMI classes.

Table 4. Number of aligned ECG segments by AAMI class before and after adaptive screening.

AAMI Class	Before Screening	After Screening
N	153,611	57,828
V	20,007	8165
S	1959	940
F	219	96

As illustrated in Figure 10, the compression stage applied to real ECG segments demonstrated promising performance, achieving a high PSNR of approximately 50 dB, a low MSE on the order of 10^{-5} , and a PRD of around 1.5%. However, the primary limitation of the method is its relatively modest compression ratio, which remains around 5:1 across all AAMI classes. This limitation could be attributed not only to residual noise or artifacts, which increase the number of atoms required to accurately represent the signal, but also to other factors, such as ECG waveform diversity and QRS alignment errors. In particular, the dataset exhibits an R-peak jitter of approximately ± 7.69 ms, which may further affect the ability of the method to efficiently compress the signal while preserving its morphological features. Despite the low compression ratio, the proposed AE-DD method demonstrated strong generalization capabilities in accurately decomposing both normal and arrhythmic ECG segments into their fundamental morphological components, such as P-wave, QRS complex, and T-wave, as illustrated in Figure 11.

Table 5 presents the correlation coefficients and mean squared error (MSE) values obtained when comparing the original ECG signal with the reconstructed morphologies. The results demonstrate that the decomposition method preserves the main characteristics of the ECG waveform with high fidelity. In particular, the QRS complex showed the highest similarity, with a correlation of 0.99 and an MSE of 9.26×10^{-4} , reflecting the robustness of the method in capturing the sharp and high-amplitude features of the ECG. The P-wave also exhibited strong agreement between the original and reconstructed signals, achieving a correlation of 0.98 and a very low MSE of 1.93×10^{-5} , indicating accurate preservation of low-amplitude deflections. For the T-wave, a correlation of 0.95 and an MSE of 3.38×10^{-4} were obtained, which, although slightly lower than the other components, still confirmed the reliability of the method in reconstructing repolarization morphology. Overall, these results highlight that the proposed approach achieves accurate ECG morphology reconstruction across all key waveform components.

Table 5. The average correlation coefficient and mean squared error (MSE) for different ECG morphologies.

Morphology	Corr.	MSE
P-wave	0.98	1.93×10^{-5}
QRS complex	0.99	9.26×10^{-4}
T-wave	0.95	3.38×10^{-4}

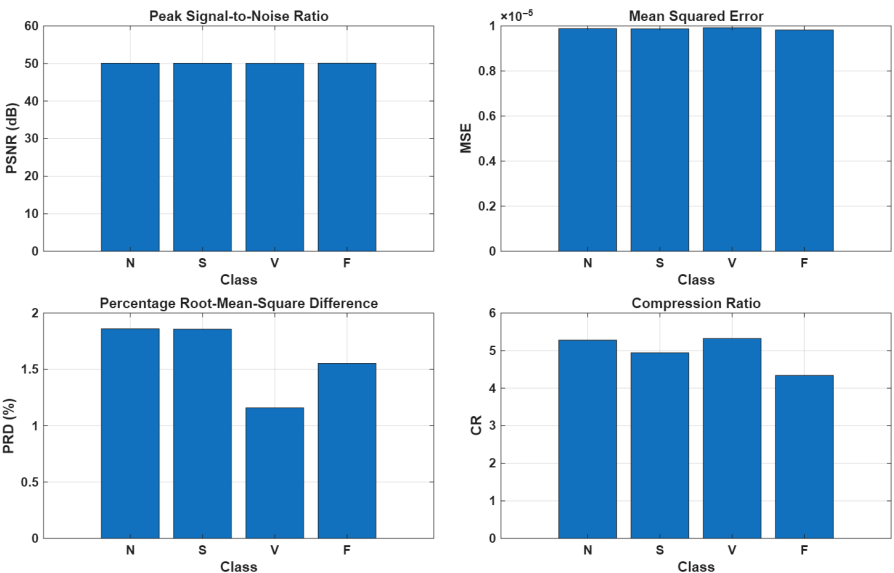


Figure 10. Compression performance across four ECG classes (N, S, V, and F), evaluated using peak signal-to-noise ratio (PSNR), mean squared error (MSE), percentage root-mean-square difference (PRD), and compression ratio (CR).

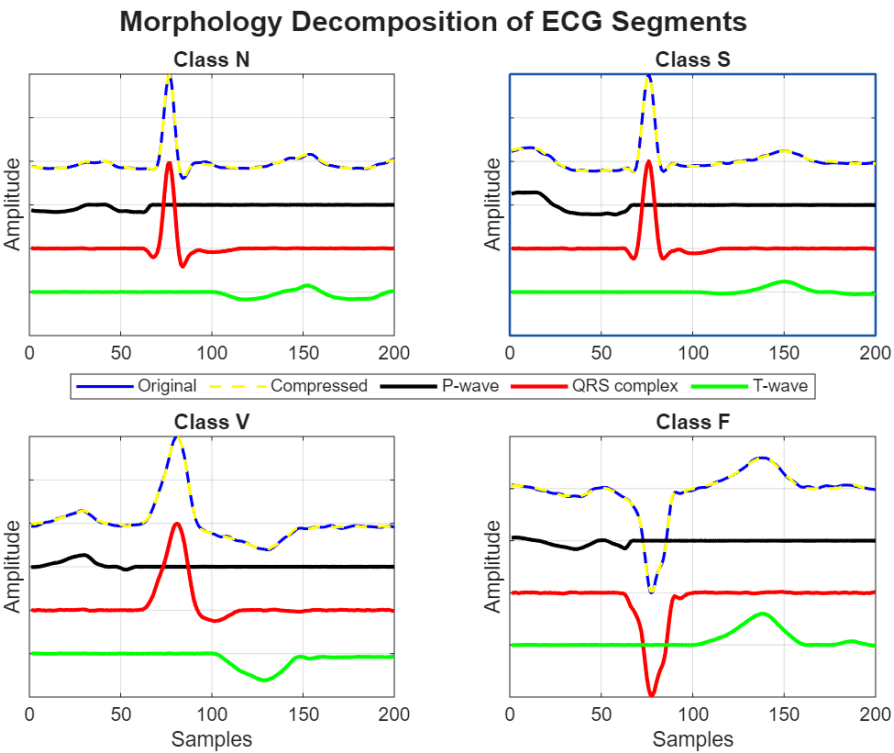


Figure 11. Morphological decomposition of the ECG segments for four classes (N, S, V, and F). Each subplot shows the original and compressed signals, along with the extracted P-wave, QRS complex, and T-wave components, highlighting class-specific waveform variations.

4. Conclusions

This study presents a novel framework based autoencoder-driven dictionary (AE-DD) that simultaneously addresses ECG denoising, compression, and morphological decomposition. By leveraging L1-regularized autoencoders with the matching pursuit algorithm, the method achieves high-fidelity signal compression and the precise decomposition of ECG components. Key contributions include end-to-end high-performance denoising along with ECG morphology decomposition.

In controlled experiments, the L1-AE dictionary achieved a 28:1 compression ratio with near-perfect reconstruction fidelity, which was reflected by an extremely low MSE of 0.58×10^{-5} and PRD of 2.1%—significantly outperforming non-regularized autoencoders and predefined dictionaries. In the case of real ECG segments from the INCART database, the framework continued to demonstrate a good performance, achieving a PSNR of approximately 50 dB, a PRD of 1.5%, and an MSE on the order of 10^{-5} . Although the compression ratio in the case of real signals was more modest (approximately 5:1), this was primarily due to the presence of residual noise and artifacts even after adaptive screening, which required more atoms for accurate representation. However, the decomposition stage showed excellent generalization in decomposing both normal and arrhythmic real ECG segments.

The detection and duration estimation of P- and T-waves remains a significant challenge due to their relatively low amplitudes and the potential overlap with noise and artifacts. To address this challenge, the trained weights of the L1-AE model form basis functions that naturally align with key ECG components, including the P-wave, QRS complex, and T-wave. These learned bases enable not only the automatic identification of P- and T-waves, but also the precise estimation of their durations, offering a robust and interpretable approach for detailed morphological assessment.

Finally, the proposed framework addresses the challenges of prolonged ambulatory ECG monitoring by offering a unified solution for signal enhancement, compression, and clinically interpretable feature extraction. Notably, the morphological decomposition capability is both highly novel and clinically valuable as it enables accurate identification and temporal localization of individual ECG components, particularly the P-wave and T-wave durations. In addition, we recorded initial indications that separate bases are better suited for representing the P-wave, QRS complex, and T-wave when the time interval between these signal components varies. This capability facilitates precise temporal and morphological analysis, essential for diagnosing conditions such as atrial fibrillation and heart failure.

Author Contributions: Conceptualization, F.S. and T.S.; Methodology, F.S. and T.S.; Software, F.S.; Resources, F.S.; validation, F.S. and T.S.; formal analysis, F.S.; investigation, F.S.; resources, F.S.; Writing—original draft, F.S.; Writing—review and editing, T.S.; Supervision, T.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original data presented in the study are openly available in INCART 12-lead Arrhythmia Database, PhysioNet at <https://doi.org/10.13026/C2V88N> (accessed on 27 July 2025).

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Nielsen, J.B.; Kühl, J.T.; Pietersen, A.; Graff, C.; Lind, B.; Struijk, J.J.; Olesen, M.S.; Sinner, M.F.; Bachmann, T.N.; Haunsø, S.; et al. P-wave duration and the risk of atrial fibrillation: Results from the Copenhagen ECG Study. *Heart Rhythm* **2015**, *12*, 1887–1895. [\[CrossRef\]](#) [\[PubMed\]](#)
- Klingenheben, T.; Zabel, M.; D'Agostino, R.; Cohen, R.; Hohnloser, S. Predictive value of T-wave alternans for arrhythmic events in patients with congestive heart failure. *Lancet* **2000**, *356*, 651–652. [\[CrossRef\]](#)
- Bloomfield, D.M.; Steinman, R.C.; Namerow, P.B.; Parides, M.; Davidenko, J.; Kaufman, E.S.; Shinn, T.; Curtis, A.; Fontaine, J.; Holmes, D.; et al. Microvolt T-wave alternans distinguishes between patients likely and patients not likely to benefit from implanted cardiac defibrillator therapy. *Circulation* **2004**, *110*, 1885–1889. [\[CrossRef\]](#) [\[PubMed\]](#)
- Samann, F. Towards Real-Time ECG Signal Denoising Using Sparse and Shallow Running Denoising Autoencoder. Ph.D. Thesis, Technische Hochschule Mittelhessen, Gießen, Germany, 2025. [\[CrossRef\]](#)
- Chiang, H.-T.; Hsieh, Y.-Y.; Fu, S.-W.; Hung, K.-H.; Tsao, Y.; Chien, S.-Y. Noise reduction in ECG signals using fully convolutional denoising autoencoders. *IEEE Access* **2019**, *7*, 60806–60813. [\[CrossRef\]](#)
- Yildirim, O.; Tan, R.S.; Acharya, U.R. An efficient compression of ECG signals using deep convolutional autoencoders. *Cogn. Syst. Res.* **2018**, *52*, 198–211. [\[CrossRef\]](#)
- Qiu, L.; Cai, W.; Zhang, M.; Zhu, W.; Wang, L. Two-stage ECG signal denoising based on deep convolutional network. *Physiol. Meas.* **2021**, *42*, 115002. [\[CrossRef\]](#)
- Xiong, P.; Wang, H.; Liu, M.; Liu, X. Denoising autoencoder for electrocardiogram signal enhancement. *J. Med. Imaging Health Inform.* **2015**, *5*, 1804–1810. [\[CrossRef\]](#)
- Xiong, P.; Wang, H.; Liu, M.; Lin, F.; Hou, Z.; Liu, X. A stacked contractive denoising auto-encoder for ECG signal denoising. *Physiol. Meas.* **2016**, *37*, 2214–2230. [\[CrossRef\]](#)
- Samann, F.; Schanze, T. Multiple ECG segments denoising autoencoder model. *Biomed. Eng. Biomed. Tech.* **2023**, *68*, 275–284. [\[CrossRef\]](#)
- Samann, F.; Schanze, T. Resembling the morphologies of ECG signals using regularized denoising autoencoder. *Passer J. Basic Appl. Sci.* **2024**, *6*, 341–351. [\[CrossRef\]](#)
- Devindi, I.; Liyanage, S.; Jayarathna, T.; Alawatugoda, J.; Ragel, R. A novel ECG compression algorithm using pulse-width modulation integrated quantization for low-power real-time monitoring. *Sci. Rep.* **2024**, *14*, 17162. [\[CrossRef\]](#)
- Dias, F.M.; Khosravy, M.; Cabral, T.W.; Monteiro, H.L.M.; Filho, L.M.d.; Honório, L.d.M.; Naji, R.; Duque, C.A. Compressive sensing of electrocardiogram. In *Compressive Sensing in Healthcare*; Elsevier: Amsterdam, The Netherlands, 2020; pp. 165–184. [\[CrossRef\]](#)
- Fira, M.; Goras, L. Basis pursuit for ECG compression. In Proceedings of the 2009 International Symposium on Signals, Circuits and Systems, Iasi, Romania, 9–10 July 2009; IEEE: Iasi, Romania, 2009; pp. 1–4. [\[CrossRef\]](#)
- Luengo, D.; Meltzer, D.; Trigano, T. An efficient method to learn overcomplete multi-scale dictionaries of ECG signals. *Appl. Sci.* **2018**, *8*, 2569. [\[CrossRef\]](#)
- Samann, F.; Schanze, T. Finding an optimal dictionary of different wavelet types using sparse modeling to denoise ECG signal. *Curr. Dir. Biomed. Eng.* **2021**, *7*, 125–128. [\[CrossRef\]](#)
- Schanze, T. Compression and noise reduction of biomedical signals by singular value decomposition. *IFAC-PapersOnLine* **2018**, *51*, 361–366. [\[CrossRef\]](#)
- Schanze, T. Tensor decomposition-based compression and noise reduction of multichannel ECG signals. *Passer J. Basic Appl. Sci.* **2024**, *6*, 326–340. [\[CrossRef\]](#)
- Sološenko, A.; Petrėnas, A.; Paliakaitė, B.; Marozas, V.; Sörnmo, L. Model for simulating ECG and PPG signals with arrhythmia episodes. *PhysioNet* **2022**. [\[CrossRef\]](#)
- Moody, G.B.; Muldrow, W.; Mark, R.G. The MIT-BIH noise stress test database. *PhysioNet* **1992**. [\[CrossRef\]](#)
- Kumar, H.; Niwaria, K.; Chourasia, B. A review on various types of ECG data compression techniques. *Int. Res. J. Eng. Technol.* **2020**, *7*, 1496–1499.
- Cox, J.R.; Nolle, F.M.; Fozzard, H.A.; Oliver, G.C. AZTEC, a preprocessing program for real-time ECG rhythm analysis. *IEEE Trans. Biomed. Eng.* **1968**, *15*, 128–129. [\[CrossRef\]](#) [\[PubMed\]](#)
- Kumar, V.; Saxena, S.C.; Giri, V.K.; Singh, D. Improved modified AZTEC technique for ECG data compression: Effect of length of parabolic filter on reconstructed signal. *Comput. Electr. Eng.* **2005**, *31*, 334–344. [\[CrossRef\]](#)
- Barr, R.C.; Blanchard, S.M.; Dipersio, D.A. SAPA-2 is the fan. *IEEE Trans. Biomed. Eng.* **1985**, *32*, 337. [\[CrossRef\]](#)
- Tai, S.C. SLOPE—A real-time ECG data compressor. *Int. J. Bio-Med. Comput.* **1991**, *29*, 175–179. [\[CrossRef\]](#)
- Mueller, W.C. Arrhythmia detection program for an ambulatory ECG monitor. *Biomed. Sci. Instrum.* **1978**, *14*, 81–85. [\[PubMed\]](#)
- Lu, Z.; Kim, D.Y.; Pearlman, W.A. Wavelet compression of ECG signals by the set partitioning in hierarchical trees algorithm (SPIHT). *IEEE Trans. Biomed. Eng.* **2000**, *47*, 849–856.

28. Al-Shrouf, A.; Abo-Zahhad, M.; Ahmed, S.M. A novel compression algorithm for electrocardiogram signals based on the linear prediction of the wavelet coefficients. *Digit. Signal Process.* **2003**, *13*, 604–622. [[CrossRef](#)]
29. Duarte, R.C.M.; Matos, F.M.; Batista, L.V. Near-lossless compression of ECG signals using perceptual masks in the DCT domain. *IFMBE Proc.* **2007**, *18*, 229–231. [[CrossRef](#)]
30. Batista, L.V.; Melcher, E.U.; Carvalho, L.C. Compression of ECG signals by optimized quantization of discrete cosine transform coefficients. *Med. Eng. Phys.* **2001**, *23*, 127–134. [[CrossRef](#)]
31. Mukhopadhyay, S.K.; Mitra, S.; Mitra, M. An ECG signal compression technique using ASCII character encoding. *Measurement* **2012**, *45*, 1651–1660. [[CrossRef](#)]
32. Mukhopadhyay, S.K.; Mitra, S.; Mitra, M. A lossless ECG data compression technique using ASCII character encoding. *Comput. Electr. Eng.* **2011**, *37*, 486–497. [[CrossRef](#)]
33. Tihonenko, V.; Khaustov, A.; Ivanov, S.; Rivin, A. St.-Petersburg Institute of Cardiological Technics 12-Lead Arrhythmia Database. *PhysioNet* **2007**. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.