Natural Resources for Human Health

eISSN: 2583-1194

Original Research

Received 06 May 2026
Revised 17 June 2026
Accepted 30 June 2026

Natural Resources for Human Health 2026, Vol. 6 (11s): 139–164

KEYWORDS:
Electrocardiogram Signal, Signal Processing, Signal Compression, MIT-BIH Dataset, Wavelet Transform, Autoencoder

Wavelet-Assisted 1D Convolutional Autoencoder for Efficient ECG Compression with High Reconstruction Fidelity

Bharati Karare[*1], Mrunali Abhijeet Chore[2], Vinay Rajgure[3], Shital Kakad[4], Ankita Harshad Tidake[5], Prerana Vijay Khedkar[6]

[1]Assistant professor Yeshwantrao Chavan College of Engineering, Nagpur, Maharashtra, India , Email: kararebharati@gmail.com , ORCID iD: https://orcid.org/0000-0002-9008-3182 ,*Corresponding Author

[2]Assistant Professor, G. H. Raisoni Skill Tech University, Nagpur, Maharashtra, India ,Email: mrunalichore07@gmail.com, ORCID iD: https://orcid.org/0009-0005-1837-2587

[3]Associate Professor, Department of Computer Science & Engineering, G H Raisoni University, Amravati, Maharashtra, India ,Email: vinay.rajgure@gmail.com ,ORCID iD: https://orcid.org/0009-0006-2263-4292

[4]Assistant Professor, Computer Engineering and Technology, MIT World Peace University, Pune, Maharashtra, India, Email: shitalkakad2604@gmail.com, ORCID iD: https://orcid.org/0000-0001-7428-4171

[5]Assistant Professor, Department of Artificial Intelligence and Data Science, Ajeenkya DY Patil School of Engineering, Pune, Maharashtra, India, Email: ankitatidake@dypic.in, ORCID iD: https://orcid.org/0009-0008-4655-198X

[6]Associate Professor, Bhausahab Mulak Ayurved Mahavidyalaya Nandanvan Nagpur, Maharashtra, India, Email: kararepa1@gmail.com , ORCID iD: https://orcid.org/0009-0004-8312-2464

ABSTRACT: Reliable electrocardiogram compression is essential for reducing storage and transmission requirements in continuous monitoring and bandwidth-constrained healthcare applications while maintaining high waveform reconstruction fidelity. This study proposes a Wavelet-Assisted 1D Convolutional Autoencoder that integrates two-level Daubechies-4 discrete wavelet transform preprocessing with convolutional latent representation learning. Wavelet-based denoising is applied before segmentation to remove high-frequency disturbances. The reconstructed ECG is divided into non-overlapping 256-sample segments and standardized using zero-mean/unit-variance normalization. The 1D-CAE maps each segment to a compact 32-dimensional latent representation and reconstructs the waveform using a decoder with a linear output layer. Evaluation on MIT-BIH Arrhythmia Database records 100–104 shows that WT+1D-CAE achieves the best overall average reconstruction performance across the evaluated WT+AE and FT-AE configurations with MSE of 0.006554, MAE of 0.049382, RMSE of 0.079999, PSNR of 38.376 dB, PRD of 8.000%, and NPRD of 5.658%. The proposed model reduces each 256-sample input to 32 latent coefficients corresponding to an 8:1 representation compression ratio and 87.5% dimensionality reduction. Qualitative waveform comparisons show close agreement between original and reconstructed ECG signals with reconstruction errors mainly localized around rapid high-slope transitions. Statistical analysis indicates significant improvements over WT+AE across the evaluated metrics. The comparison with FT-AE shows significant improvements in MSE and MAE and favorable, although statistically non-significant, differences in RMSE, PSNR, PRD, and NPRD.

Natural Resources for Human Health
eISSN: 2583-1194

1. INTRODUCTION

Electrocardiogram (ECG) signals provide essential information regarding the electrical activity of the heart and are widely used for cardiovascular monitoring and the assessment of cardiac abnormalities. The increasing adoption of continuous ECG monitoring, wearable sensing devices, remote healthcare systems, and telemonitoring applications has led to a substantial increase in the volume of ECG data that must be stored and transmitted [1]. Consequently, efficient ECG compression has become important for reducing data-storage and communication requirements while maintaining the temporal and morphological characteristics of the original waveform, including the P wave, QRS complex, and T wave [2].

ECG compression involves an inherent trade-off between compact signal representation and reconstruction distortion. Increasing the degree of compression can reduce storage and transmission requirements; however, excessive information reduction may introduce waveform distortion and reduce the fidelity of the reconstructed signal. Therefore, an effective ECG compression method should achieve compact representation while maintaining low reconstruction error and preserving the principal waveform characteristics required for subsequent signal analysis. This trade-off remains particularly important in long-term monitoring and resource-constrained biomedical systems, where large volumes of ECG data must be processed efficiently.

Existing ECG compression techniques can generally be categorized into direct data-compression, transformation-based, and learning-based approaches. Direct processing methods, including Differential Pulse Code Modulation and Turning Point-based techniques, reduce redundancy directly from the sampled ECG sequence but may exhibit reduced reconstruction quality as compression becomes more aggressive [5]-[7]. Transformation-based approaches represent the ECG signal in an alternative domain so that redundant or less informative coefficients can be reduced. Common examples include the Discrete Cosine Transform (DCT) [8] and wavelet-based techniques [10]. The Discrete Wavelet Transform (DWT) is well suited to non-stationary biomedical signals because it provides simultaneous temporal and frequency localization and supports multi-resolution representation of ECG morphology. Recent wavelet-based ECG compression studies have further demonstrated the usefulness of optimizing wavelet parameters and threshold levels for balancing signal compaction and reconstruction quality [20].

Deep learning has provided an alternative approach by enabling nonlinear feature representations to be learned directly from ECG data. Autoencoders are particularly suitable for this purpose because an encoder maps the input signal into a lower-dimensional latent representation, while a decoder attempts to reconstruct the original waveform from the compressed representation. Yildirim et al. (2018) demonstrated the application of deep convolutional autoencoders to ECG signal compression, showing the potential of convolution-based representation learning for reducing signal redundancy [15]. Related studies have also investigated autoencoder-based ECG denoising and deep feature learning, demonstrating that learned representations can effectively model complex temporal signal characteristics [11], [12], [18]. Nevertheless, reconstruction performance can remain sensitive to signal noise, preprocessing strategy, network architecture, latent dimensionality, and the manner in which compression efficiency is quantified.

Another important consideration is that dimensional reduction in the latent space should not automatically be interpreted as final storage or bitstream compression. A neural encoder may substantially reduce the dimensionality of the input representation, but an actual codec additionally requires operations such as latent quantization, entropy coding, and bit-packing. Therefore, reconstruction fidelity and representation compactness should be evaluated jointly while clearly distinguishing latent-space compression from end-to-end encoded file-size reduction. This distinction is adopted in the present study to provide a more rigorous characterization of the proposed compression framework. The current implementation consequently reports representation-level compression while leaving quantization and entropy coding for subsequent development.

To address these considerations, this study proposes a Wavelet-Assisted 1D Convolutional Autoencoder (WT+1D-CAE) for compact ECG representation and high-fidelity waveform reconstruction. The proposed framework first applies a two-level Daubechies-4 (db4) DWT with adaptive soft-threshold-based denoising to suppress high-frequency disturbances while retaining the dominant ECG waveform structure. The denoised signal is then divided into non-overlapping 256-sample segments and standardized using zero-mean/unit-variance normalization. Each segment is subsequently processed by the 1D-CAE, whose encoder progressively extracts local temporal features and maps the 256-sample input into a 32-dimensional latent representation. The decoder reconstructs the ECG waveform through dense, upsampling, and Conv1D operations,

with a linear output activation used to accommodate the positive and negative amplitudes produced by standardized ECG signals. These design choices are consistent with the preprocessing and reconstruction framework adopted in the proposed model. The principal contributions of this study are summarized as follows:

- A hybrid WT+1D-CAE model is developed in which two-level db4 wavelet decomposition, adaptive thresholding, and inverse reconstruction are applied before convolutional latent encoding, enabling noise-reduced ECG segments to be represented and reconstructed efficiently.
- The proposed 1D-CAE maps each 256-sample ECG segment to a 32-dimensional latent vector, corresponding to an 8:1 representation compression ratio and 87.5% dimensionality reduction.
- The proposed model is evaluated on MIT-BIH Arrhythmia Database records 100–104 using MSE, MAE, RMSE, PSNR, PRD, and NPRD and is compared with WT+AE and FT-AE baseline configurations.
- Record-level statistical analysis using paired $t-tests$ and Wilcoxon signed-rank tests is used to examine the consistency of the observed improvements.

The remainder of this paper is organized as follows. Section 2 describes the proposed methodology. Section 3 presents the experimental configuration, training settings, data partitioning strategy, and evaluation metrics. Section 4 reports the quantitative and qualitative reconstruction results. Section 5 presents the statistical validation and comparison with existing ECG compression approaches. Section 6 discusses the principal findings, interpretation, and limitations of the proposed model. Finally, Section 7 concludes the study and outlines directions for future work.

2. METHODOLOGY

This section presents the proposed hybrid ECG compression model that integrates WT for signal denoising and multi-resolution decomposition with a 1D-CAE for efficient feature learning and reconstruction. The wavelet transform serves as a preprocessing stage to remove noise and extract relevant frequency components to enhance signal quality before compression. The autoencoder then learns compact latent representations of the denoised signals in an unsupervised manner allows high compression to preserve clinically significant ECG features. The workflow of the proposed model shown in Figure 2 that consists of the WT+1D-CAE architecture. The architecture of the autoencoder, and the training process adopted for optimization. The proposed model WT performed denoising and reduces signal redundancy. The denoised ECG is then input into the 1D-CAE, where the encoder compresses the signal into a lower-dimensional latent space and the decoder reconstructs the original waveform. This two-stage model perform noise and artifacts are removed before compression and lead to improved reconstruction quality and reduced distortion.

3.1 Data Description

The MIT-BIH Arrhythmia Database is a widely used benchmark to evaluate the ECG analysis and compression algorithms. It contains 48 half-hour ECG excerpts obtained from 47 subjects recorded between 1975 and 1979 at the Beth Israel Hospital's Arrhythmia Laboratory in Boston. Out of these, 23 recordings were randomly drawn from a larger pool of about 4000 long-term (24-hour) Holter monitor ECGs, comprising roughly 60% inpatient and 40% outpatient cases. The other 25 recordings were deliberately chosen to representation of rare but clinically important arrhythmias that not appear in a purely random sample. The ECG signals were digitized at a sampling frequency of 360 Hz per channel, with 11-bit resolution and a dynamic range of ±10 mV. Each dataset was carefully annotated by two or more cardiologists, and any differences in interpretation were reconciled to create a standardized set of computer-readable annotations for every heartbeat. In total, the database provides around 110,000 annotated beats. Initially, only a subset of 25 records along with the reference annotations for all 48 signals was made publicly accessible in 1999; however, the full database is now extensively used for algorithm development, validation, and comparative studies in ECG signal processing [19]. Figure 1 shows the representative ECG segment from the dataset (3000 samples) with amplitude normalized to unit scale shows clear P-QRS-T morphology and prominent R-peaks. The regular RR intervals and minimal baseline drift indicate good signal quality indicate reliable reference for subsequent denoising and compression evaluation.

Natural Resources for Human Health
eISSN: 2583-1194

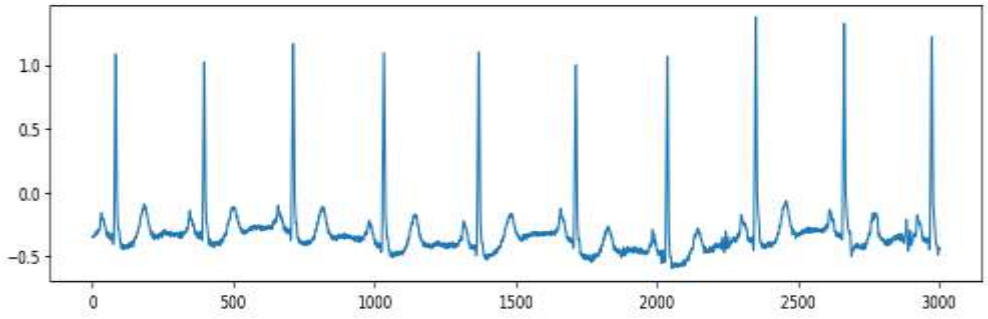


Figure 1: Sample ECG Signal

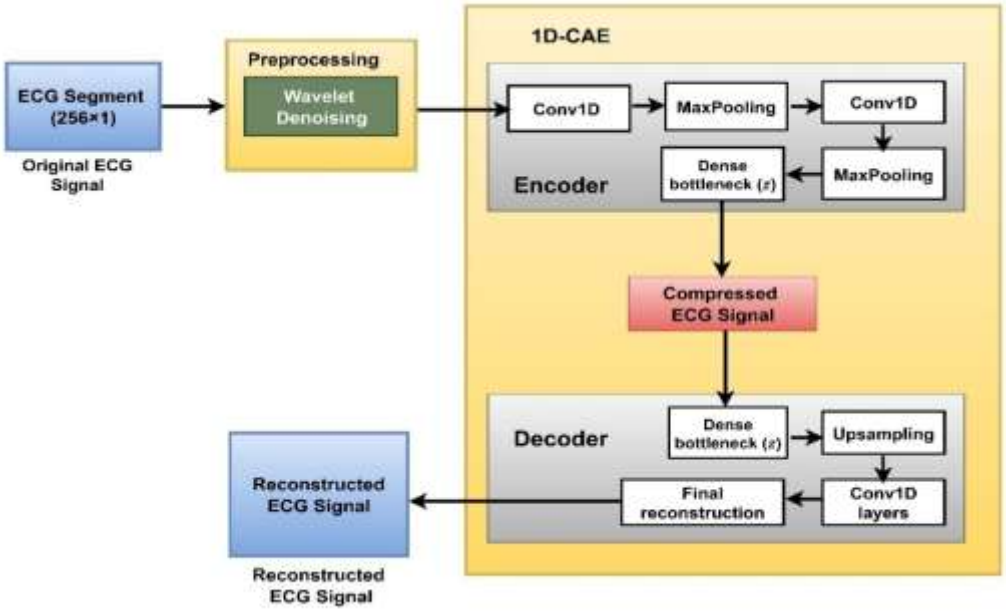


Figure 2: Architecture of Proposed Model

3.2 Wavelet-Based ECG Preprocessing and Denoising: The raw ECG signals obtained from the MIT-BIH Arrhythmia Database are sampled at 360 Hz and contain baseline fluctuations, high-frequency interference, and other acquisition-related disturbances that adversely affect the latent representation and reconstruction process. The DWT-based preprocessing stage is applied before the ECG segments are supplied to the proposed 1D Convolutional Autoencoder. The DWT provides simultaneous temporal and frequency localization and allows the ECG waveform to be represented at multiple resolutions and separate relatively low-frequency morphological information from higher-frequency noise components. Let the discrete ECG signal is represented as

$$x[n], n = 0,1, \dots, N-1, \quad (1)$$

where $x[n]$ denotes the ECG amplitude at the n^{th} sampling instant and N represents the total number of samples in the analyzed signal.

In the proposed preprocessing stage, a two-level DWT decomposition is performed using the Daubechies-4 (db4) mother wavelet. The db4 wavelet is used because its compact support and waveform characteristics suitable for representing rapid ECG transitions and retain important low-frequency morphological structures. The symmetric signal extension is used at both ends of the input sequence to minimize boundary artifacts introduced during finite-length wavelet decomposition.

At the first decomposition level, the ECG signal is passed through complementary low-pass and high-pass analysis filters and downsampled by a factor of two. This operation produces the first-level approximation and detail coefficients:

$$A_1[k] = \sum_n x[n]h[2k-n], \quad (2)$$

$$D_1[k] = \sum_n x[n]g[2k-n], \quad (3)$$

where $h[\cdot]$ and $g[\cdot]$ denote the low-pass and high-pass wavelet analysis filters, respectively. A_1 contains the lower-frequency representation of the ECG signal, whereas D_1 primarily captures higher-frequency variations and transient components. For the second decomposition level, only the first-level approximation coefficients A_1 are decomposed

$$A_1 \longrightarrow \{A_2, D_2\}, \quad (4)$$

where A_2 contains the coarse-scale ECG information and D_2 represents the detail information at the second resolution level. The two-level decomposition of the original ECG signal is expressed as

$$x[n] \longrightarrow \{A_2, D_2\ D_1\}. \quad (5)$$

This hierarchical decomposition is used ECG preprocessing because the approximation coefficients retain the dominant low-frequency waveform structure and the detail coefficients contain higher-frequency components include both clinically relevant transitions and noise. The approximation coefficients are preserved. The noise removal performed conservatively on the detail coefficients to avoid excessive smoothing of QRS-related information.

Noise Estimation and Adaptive Thresholding: The noise level is estimated robustly from the first-level detail coefficients using the median absolute deviation (MAD) estimator. The estimated noise standard deviation is calculated as

$$\hat{\sigma} = \frac{median(|D_1[k]|)}{0.6745}. \quad (6)$$

The MAD estimator is less sensitive to isolated large-amplitude ECG transitions than conventional variance-based estimation and provides a robust estimate of the high-frequency noise component. The universal threshold is defined as

$$\lambda = \hat{\sigma}\sqrt{2lnN} \quad (7)$$

where N denotes the number of samples considered during threshold estimation. The calculated threshold is applied to the detail coefficients D_1 and D_2, and the approximation coefficients A_2 were retained without thresholding.

Soft thresholding is used to remove the small noise-dominated coefficients and avoid the abrupt discontinuities that introduced by hard thresholding. For an arbitrary detail coefficient $D_j[k]$ the soft-thresholded coefficient is defined as

$$\widetilde{D}_j[k] = sign(D_j[k])max(|\ D_j[k]\ | - \lambda, 0), j \in \{1,2\}. \quad (8)$$

Accordingly, coefficients with magnitudes below the threshold are set to zero and coefficients exceeding the threshold were reduced proportionally to preserve their signs. This procedure removes the small-amplitude noise to retainin dominant transient characteristics associated with the ECG morphology.

ECG Signal Reconstruction: After thresholding, the denoised ECG signal is reconstructed using the inverse discrete wavelet transform (IDWT). The reconstruction combines the unmodified second-level approximation coefficients A_2 with the thresholded detail coefficients $\widetilde{D}_2$ and $\widetilde{D}_1$

$$x[n] = IDWT\left(A_2, \widetilde{D}_2, \widetilde{D}_1\right) \quad (9)$$

where $\tilde{x}[n]$ represents the denoised ECG signal. The wavelet denoising is performed on the continuous ECG recording before segmentation. This ordering avoids artificial discontinuities at segment boundaries and allows the wavelet transform to use the local temporal context surrounding each ECG sample.

ECG Segmentation: The wavelet-based denoising and signal reconstruction. The ECG sequence is divided into fixed-length, non-overlapping segments of 256 samples, consistent with the input dimensionality of the proposed 1D-CAE. The m^{th} segment is represented as

$$xm = [\tilde{x}[mL], \tilde{x}[mL+1], \ldots, \tilde{x}[mL+L-1]], \quad (10)$$

$L = 256$ denotes the segment length. At the sampling frequency of 360 Hz, each segment corresponds to approximately

Natural Resources for Human Health
eISSN: 2583-1194

$$\frac{256}{360} = 0.711\ s \quad (11)$$

of ECG activity.

The use of non-overlapping windows avoids duplicate samples across adjacent segments and provides a consistent input dimension for subsequent convolutional encoding.

Segment-Wise Normalization: The ECG amplitudes vary between subjects and recording intervals, each denoised segment is standardized independently using zero-mean and unit-variance normalization. For the m^{th} segment, its mean is calculated as

$$\mu_m = \frac{1}{L}\sum_{i=1}^{L} x_{m,i}, \quad (12)$$

and its standard deviation is obtained as

$$\sigma_m = \sqrt{\frac{1}{L}\sum_{i=1}^{L}(x_{m,i} - \mu_m)^2} \quad (13)$$

The normalized ECG sample is expressed as

$$x'_{m,i} = \frac{x_{m,i} - \mu_m}{\sigma_m + \varepsilon}, \quad (14)$$

where ε is a small numerical constant introduced to avoid division by zero. This transformation standardizes each segment while maintaining its relative temporal morphology and make numerically stable input distribution for training the 1D-CAE.

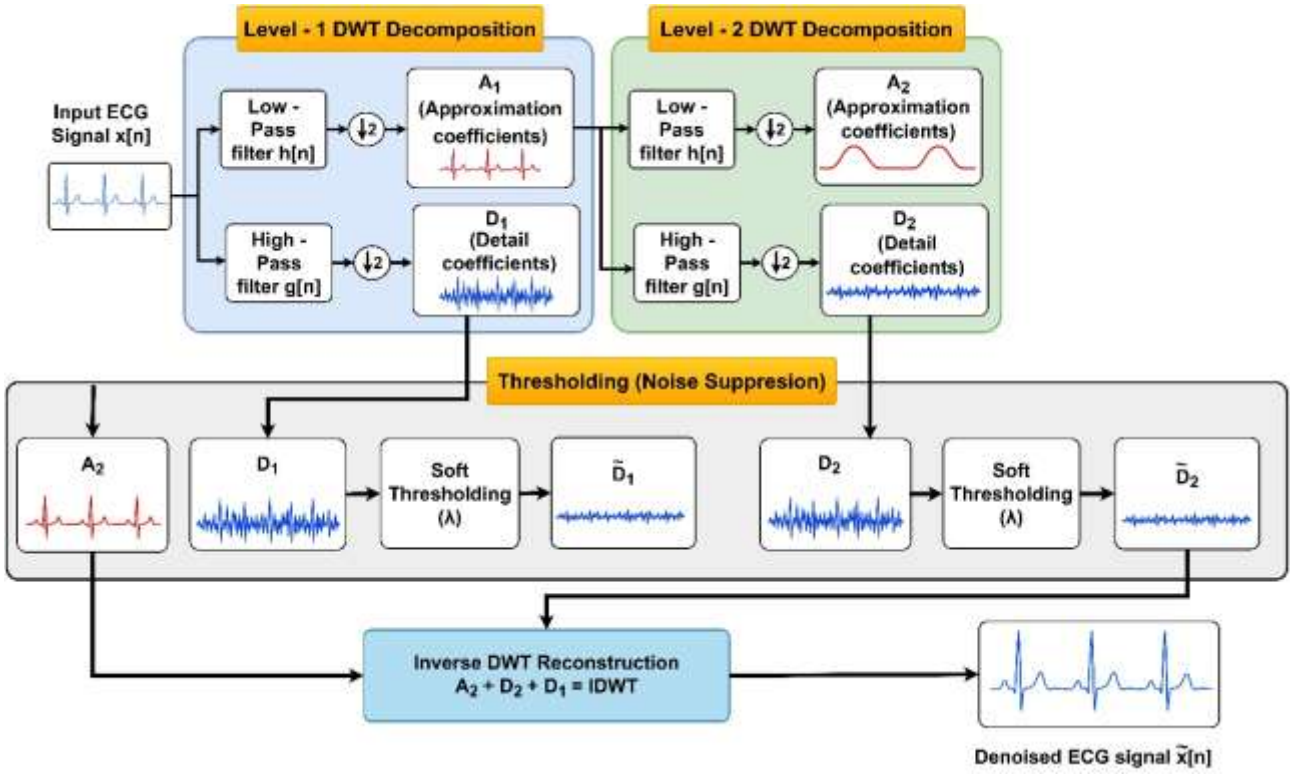


Figure 3: Two-level DWT decomposition of an input ECG signal

Figure 3 shows the two-tier DWT decomposition of an input ECG signal. The original signal $x[n]$ is iteratively passed through low-pass and high-pass filters followed by downsampling to produce approximation (A) and detail (D) coefficients at each level.

The DWT shows the hierarchical multi-resolution representation of the ECG signal. At each decomposition level, the signal is split into approximation (low-frequency) and detail (high-frequency) components using low-pass and high-pass filters,

respectively, followed by downsampling. This enables efficient separation of the baseline ECG morphology (approximations) from transient features and noise (details). Such decomposition used in ECG compression and denoising most clinically relevant information is preserved in the low-frequency approximation coefficients [20].

3.3 Feature Extraction and Compression via 1D Convolutional Autoencoder

Following wavelet-based denoising, segmentation, and normalization, each ECG segment is supplied to a one-dimensional convolutional autoencoder (1D-CAE) for compact latent representation and waveform reconstruction. The proposed autoencoder consists of two principal components: an encoder, which progressively extracts temporal and morphological features to reduce the dimensionality of the input signal, and a decoder, which reconstructs the normalized ECG waveform from the learned latent representation. The conventional fully connected autoencoders, the use of one-dimensional convolution enables the proposed model to exploit local temporal dependencies and waveform characteristics such as rapid QRS transitions, P-wave morphology, and T-wave variations directly from the ECG sequence.

Each denoised ECG segment contains $L = 256$ samples and is normalized to zero mean and unit variance before being provided to the network. The input to the encoder is represented as

$$x_m = \left[x'_{m,1}, x'_{m,2}, \dots, x'_{m,256}\right] \quad T \in R^{256\times 1}, \tag{15}$$

where x_m denotes the normalized m^{th} ECG segment. Because z-score normalization is used. The input space is not restricted to positive values. The normalized ECG amplitudes take both positive and negative values.

3.3.1 Encoder: The encoder progressively transforms the normalized ECG segment into increasingly abstract temporal feature representations using a sequence of one-dimensional convolutional and max-pooling operations. For the l^{th} convolutional layer, the feature representation expressed as

$$H^{(l)} = ReLU\left(W^{(l)} * H^{(l-1)} + b^{(l)}\right), \tag{16}$$

where $*$ represents the one-dimensional convolution operation, W^l and b^l denote the learnable convolutional kernel and bias of layer l, respectively, and $ReLU(\cdot)$ denotes the rectified linear unit activation function. For the first convolutional layer,

$$H(0) = x_m. \tag{17}$$

The rectified linear unit is defined as

$$ReLU(u) = max(0, u). \tag{18}$$

ReLU activation introduces non-linearity while maintaining efficient gradient propagation during network optimization.

After each convolutional feature-extraction stage, max pooling is used to reduce the temporal dimensionality. The pooling operation is represented as

$$P^{(l)} = MaxPool(H^{(l)}), \tag{19}$$

where $P^{(l)}$ represents the downsampled feature map produced by the l^{th} encoding stage.

Consistent with the architecture shown in Figure 4, the encoder employs successive Conv1D layers with increase feature channels, followed by temporal max pooling. The convolutional hierarchy allows shallow layers to capture localized waveform variations and deeper layers learn higher-level morphological representations from larger effective receptive fields. After the final convolutional and pooling stage, the resulting feature representation is transformed by the bottleneck layer into a compact latent vector:

$$z_m = f\theta(x_m), \tag{20}$$

where $f_\theta(\cdot)$ represents the encoder parameterized by θ, and $z_m \in R^{32}$ denotes the 32-dimensional latent representation of the m^{th} ECG segment. The encoder transforms the original 256-sample normalized ECG segment into only 32 latent coefficients. This bottleneck representation captures the dominant temporal and morphological information required for reconstruction and removal redundant information. It therefore serves as the representation-level compressed form of the ECG signal.

3.3.2 Latent Bottleneck Representation: The bottleneck forms the central component of the proposed WT+1D-CAE framework. Let $d_z z = 32$ represent the dimensionality of the latent space. The transformation performed by the encoder is summarized as

$$R^{256\times1} \rightarrow R^{32}. \quad (21)$$

The substantial dimensionality reduction forces the network to learn a compact representation instead of directly reproduce the input through an unconstrained identity mapping. The latent coefficients retain information required to reproduce the principal ECG waveform characteristics while discarding part of the redundancy present in the original sample representation. The latent vector becomes the input to the decoder:

$$z_m \rightarrow g_\phi(z_m), \quad (22)$$

where $g_\phi(\cdot)$ represents the decoder parameterized by ϕ.

3.3.3 Decoder and ECG Reconstruction: The decoder performs the inverse transformation of the encoder by progressively restoring the temporal resolution of the compressed representation. First, the latent vector is projected to an intermediate feature representation through a dense layer and reshaped into the dimensional form required by the convolutional decoder. The upsampling and Conv1D operations progressively restore the temporal dimension. At decoder stage l, the upsampling operation is represented as

$$U^{(l)} = UpSample\left(D^{(l-1)}\right), \quad (23)$$

followed by convolutional feature refinement:

$$D^{(l)} = ReLU\left(W_d^{(l)} * U^{(l)} + b_d^{(l)}\right), \quad (24)$$

where $W_d^{(l)}$ and $b_d^{(l)}$ represent the learnable decoder parameters.

The sequence of upsampling and convolution operations progressively restores the compressed feature representation to the original temporal dimension of 256 samples. The final decoder layer used a one-dimensional convolution with a linear activation function to generate the reconstructed ECG segment:

$$\hat{x}_m = W_o * D^{(L)} + b_o, \quad (25)$$

$\hat{x}_m = \left[\hat{x}_{m,1}, \hat{x}_{m,2}, \ldots, \hat{x}_{m,256}\right]$ $T \in R^{256\times1}$ represents the reconstructed ECG segment. The linear output activation is defined as

$$f_{linear}(u) = u. \quad (26)$$

The use of a linear output layer is essential in the proposed model because the input ECG segments are standardized using zero-mean and unit-variance normalization. The normalized ECG values occur over both positive and negative ranges. A bounded sigmoid activation constrains the reconstructed samples approximately to the interval $[0,1]$ and unsuitable for reproduce negative standardized ECG amplitudes. The linear activation imposes no artificial amplitude restriction and therefore allows the decoder to reconstruct the complete standardized ECG waveform.

3.3.4 Reconstruction Objective: The parameters of the encoder and decoder were optimized to minimize the difference between the original normalized ECG segment and its reconstructed counterpart. The primary reconstruction objective is the MSE defined as

$$L_{MSE} = \frac{1}{M}\sum_{m=1}^{M}\frac{1}{L}\sum_{i=1}^{L}(x_{m,i}^{i} - \hat{x}_{m,i})^2 \quad (27)$$

where M denotes the number of ECG segments in the training batch, $L = 256$ represents the number of samples in each segment, $x_{m,i}^{i}$ is the normalized original ECG amplitude, and $\hat{x}_{m,i}$ is the corresponding reconstructed amplitude.

Minimizing this objective encourages the decoder to accurately recover the normalized waveform from the compact latent representation. Because errors were penalized at every temporal sample, the model learns to minimize deviations in waveform amplitude and morphology across the reconstructed ECG segment. The complete autoencoding operation is represented as

Natural Resources for Human Health
eISSN: 2583-1194

$$x_m \xrightarrow{f_\theta} zm \xrightarrow{g_\phi} \hat{x}_m, \quad (28)$$

where the encoder f_θ extracts the compressed latent representation and the decoder g_ϕ reconstructs the waveform. The resulting architecture combines convolution-based temporal feature extraction, compact bottleneck encoding, and unrestricted linear reconstruction to provide a representation suitable for ECG signal compression to maintain waveform fidelity.

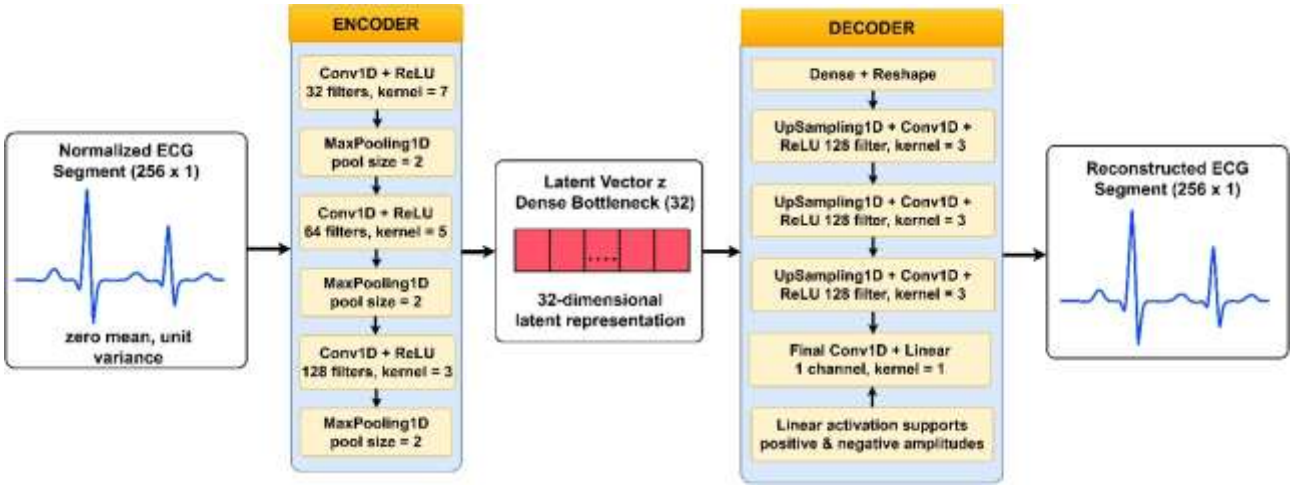


Figure 4: Architecture of the proposed 1D Convolutional Autoencoder (CAE) for ECG compression.

3.3 Compression Evaluation and Representation Efficiency

The compression capability of the proposed WT+1D-CAE model was evaluated in terms of latent-space representation reduction together with reconstruction fidelity. Each preprocessed ECG segment contains $L = 256$ samples and is encoded by the proposed 1D-CAE into a compact bottleneck vector of $d_z = 32$ latent coefficients. Accordingly, the representation compression ratio (RCR) is defined as

$$RCR = \frac{L}{d_z} \quad (29)$$

where L denotes the number of samples in the input ECG segment and d_z represents the dimensionality of the latent vector. For the proposed architecture,

$$RCR = \frac{256}{32} = 8, \quad (30)$$

corresponding to a latent-space representation ratio of 8:1. The encoder represents each 256-sample ECG segment using only 32 latent coefficients. The corresponding dimensionality reduction (DR) is expressed as

$$DR(\%) = \left(1 - \frac{L}{d_z}\right) \times 100. \quad (31)$$

For the proposed architecture,

$$DR = \left(1 - \frac{32}{256}\right) \times 100 = 87.5\%. \quad (32)$$

Only 12.5% of the original representation dimensionality is retained at the bottleneck and 87.5% of the sample-level dimensionality is eliminated before reconstruction. This compact latent representation reduces redundancy that allows the decoder to recover the temporal morphology of the ECG waveform.

It is important to distinguish this latent-space reduction from the end-to-end bitstream compression ratio used in deployable signal codecs. The bottleneck coefficients produced by the current implementation remain continuous numerical values and are not subjected to quantization, entropy coding, or bit-packing. The reported 8:1 value represents a dimensional compression ratio rather than a final storage compression ratio. For an implemented bitstream codec. The effective compression ratio is defined as

$$CR_{bit} = \frac{B_{original}}{B_{latent}+B_{side}}, \quad (33)$$

where $B_{original}$ is the number of bits required to represent the original ECG segment, B_{latent} is the number of bits required for the encoded latent representation, and B_{side} represents any additional information required for reconstruction, such as normalization parameters or coding metadata. The compression performance is assessed jointly using the 8:1 latent representation ratio, 87.5% dimensionality reduction, and reconstruction-fidelity metrics such as MSE, MAE, RMSE, PSNR, PRD, and NPRD. This combined evaluation shows more rigorous representation-level assessment of the performance between signal compaction and reconstruction quality without overstating the achievable file-size reduction.

Table 1: Compression Characteristics of the Proposed WT+1D-CAE model

Compression characteristic	Value
Input ECG segment length	256 samples
Latent bottleneck dimension	32 coefficients
Representation compression ratio (RCR)	8:1
Retained latent dimensionality	12.5%
Dimensionality reduction	87.5%
Latent quantization	Not applied
Entropy coding	Not applied
End-to-end bitstream compression ratio	Not reported
Reconstruction evaluation	MSE, MAE, RMSE, PSNR, PRD, NPRD

Table 1 summarizes the representation-level compression characteristics of the proposed WT+1D-CAE. Each 256-sample ECG segment is mapped to a 32-dimensional latent vector, resulting in an 8:1 representation compression ratio and an 87.5% reduction in dimensionality. This demonstrates compaction at the feature-representation level to retain only 12.5% of the original sample dimensionality. The latent coefficients are not quantized or entropy-coded in the current implementation, the 8:1 ratio should not be interpreted as the final bitstream or storage compression ratio. The reconstruction metrics are jointly considered with latent dimensionality to assess the performance between representation compactness and waveform fidelity. [26]-[27]

4. Results Analysis

The proposed WT+1D-CAE model was evaluated using the MIT-BIH Arrhythmia Database to its reliability and diagnostic accuracy. Implementation was carried out in Python 3.9, leveraging TensorFlow/Keras, NumPy, and Scikit-learn libraries. All experiments were run on Google Colab Pro, utilizing an NVIDIA Tesla T4 GPU with the 16 GB VRAM, 25 GB RAM provide sufficient computational resources for efficient training and reproducible results.

Table 2: Hyperparameter Settings of the Proposed WT+1D-CAE

Hyperparameter	Selected Setting	Purpose
Optimizer	Adam	Provides adaptive gradient-based optimization during 1D-CAE training
Initial Learning Rate	1×10^{-4}	Controls the initial magnitude of parameter updates
$Adam\ \beta_1$	0.9	Controls the exponential decay rate of the first-moment estimate
$Adam\ \beta_2$	0.999	Controls the exponential decay rate of the second-moment estimate
Learning-Rate Strategy	Reduce on plateau	Reduces the learning rate when validation performance stops

		improving
Batch Size	32	Number of ECG segments processed in each training batch
Maximum Epochs	100	Maximum number of complete training iterations
Early Stopping	Based on validation loss	Terminates training when validation performance ceases to improve, reducing unnecessary training and overfitting
Dropout Rate	0.3	Applied after dense layers to reduce overfitting
Hidden-Layer Activation	ReLU	Introduces non-linearity in convolutional and dense feature-learning layers
Output-Layer Activation	Linear	Allows reconstruction of both positive and negative amplitudes after zero-mean/unit-variance normalization

Table 2 shows the training configuration used for the proposed WT+1D-CAE. The Adam optimizer with an initial learning rate of 1×10^{-4}, together with learning-rate reduction and validation-based early stopping was used to promote stable convergence to addressed the overfitting problem. A batch size of 32 and a maximum of 100 epochs provided the basic training schedule and dropout of 0.3 was applied after the dense layers as regularization. ReLU activation was used within the hidden convolutional and dense layers. The final Conv1D reconstruction layer used a linear activation is consistent with the zero-mean/unit-variance normalization of the ECG segments and permits reconstruction over both positive and negative amplitude ranges

TRAINING AND TESTING TIME

On average, the training of the proposed model required approximately 10 minutes for 100 epochs on the Tesla T4 GPU. Model convergence was achieved within 60-100 epochs due to the early stopping mechanism. The average testing/inference time per ECG segment was 0.014 seconds shows the feasibility of the approach for real-time ECG signal compression and reconstruction.

3.1 Performance Evaluation

The dataset was divided into 80% for training and 20% for testing and no patient overlap between the sets. The performance of the proposed model was evaluated using multiple distortion and fidelity measures. To establish robustness, results were compared against several existing baseline models.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{x}_i)^2 \quad (34)$$

Where, n is the total number of data points in the ECG signal. x_i represents the individual data point in the original uncompressed ECG signal. $\hat{x}_i$ represents the corresponding data point in the reconstructed/compressed signal obtained through unsupervised auto encoder decoder-based compression.

$$MAE = \frac{1}{L}\sum_{i=1}^{L} | S[i] - \hat{s}[i]| \quad (35)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \hat{x}_i)^2}{n}} \quad (36)$$

$$PSNR = 10 . log_{10}\left(\frac{Max^2}{MSE}\right) \quad (37)$$

$$PRD = \frac{\sqrt{\frac{\sum_{i=1}^{n}(x_i - \hat{x}_i)^2}{\sum_{i=1}^{n} x_i^2}}}{\sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n}}} \quad (38)$$

$$NPRD = \frac{PRD \times Range_x}{100} \quad (39)$$

Natural Resources for Human Health
eISSN: 2583-1194

Table 3: Performance Analysis of Models

Model	Records	MSE	MAE	RMSE	PSNR (dB)	PRD (%)	NPRD (%)
WT+AE	100	0.084934	0.178735	0.291434	27.931	32.91	23.51
	101	0.093158	0.184257	0.305218	26.407	34.47	23.91
	102	0.063875	0.153559	0.252736	28.046	22.61	16.31
	103	0.076737	0.183171	0.277014	27.243	31.28	22.27
	104	0.065968	0.162161	0.256841	27.902	21.87	15.83
Average		**0.076934**	**0.076934**	**0.172377**	**0.276649**	**27.506**	**28.628**
FT-AE	100	0.002342	0.042726	0.048390	43.526	5.18	3.61
	101	0.011797	0.087598	0.108612	35.381	20.62	14.76
	102	0.013473	0.098741	0.116073	34.804	16.17	11.56
	103	0.020901	0.119883	0.144571	32.891	20.54	14.50
	104	0.019685	0.108407	0.140304	33.154	10.25	7.37
Average		**0.01364**	**0.013640**	**0.091471**	**0.111590**	**35.951**	**14.552**
WT+1D-CAE	100	0.003336	0.038233	0.057758	41.990	5.78	4.09
	101	0.006440	0.049884	0.080250	38.010	8.02	5.69
	102	0.006997	0.050667	0.083646	37.650	8.36	5.87
	103	0.009211	0.062077	0.095975	36.450	9.60	6.79
	104	0.006784	0.046048	0.082367	37.780	8.24	5.85
Average		**0.006554**	**0.006554**	**0.049382**	**0.079999**	**38.376**	**8.000**

Table 3 presents the record-wise reconstruction performance of WT+AE, FT-AE, and the proposed WT+1D-CAE across MIT-BIH records 100–104. On average, WT+1D-CAE achieved the lowest MSE (0.006554), MAE (0.049382), RMSE (0.079999), PRD (8.00%), and NPRD (5.658%), together with the highest average PSNR of 38.376 dB, demonstrating the best reconstruction fidelity among the evaluated approaches. FT-AE shows strongest reconstruction for record 100 in terms of MSE, PSNR, PRD, and NPRD indicate that the proposed model does not dominate every individual record. The WT+1D-CAE achieved the most favorable average performance across the five test records. The WT+AE shows higher reconstruction errors and distortion confirms the advantage of combining wavelet-based preprocessing with the convolutional autoencoder.

Natural Resources for Human Health
eISSN: 2583-1194

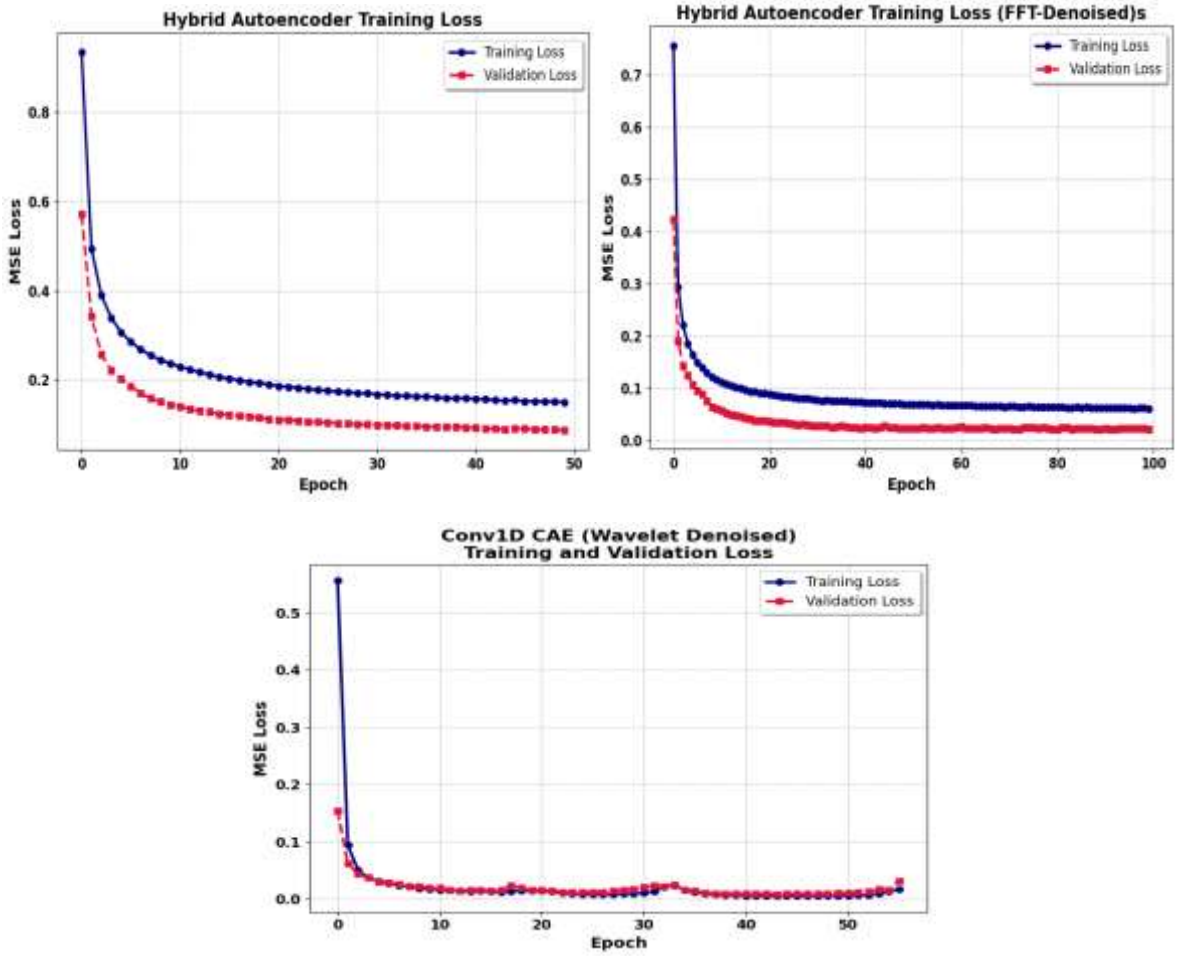


Figure 5: Training and Validation Loss of 1D-CAE for ECG compression with Wavelet denoising.

Figure 5 shows the training and validation loss of proposed models with wavelet denoising. The training and validation losses decreased and stabilized indicate successful convergence without pronounced divergence between the curves.

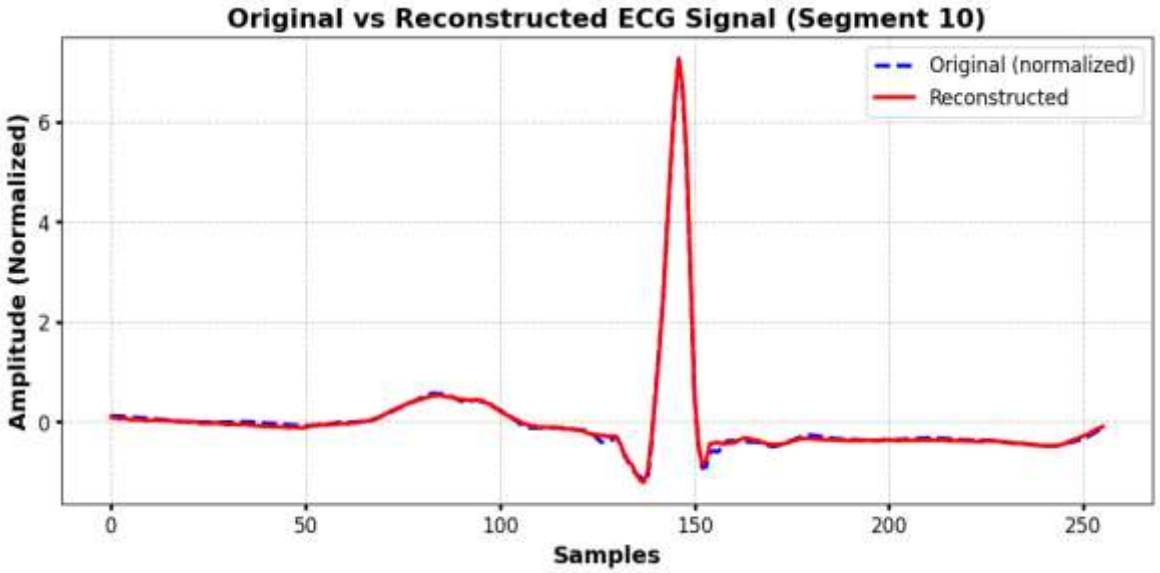


Figure 6: Original and Reconstructed ECG Signal over Record 100 of segment 10

Natural Resources for Human Health
eISSN: 2583-1194

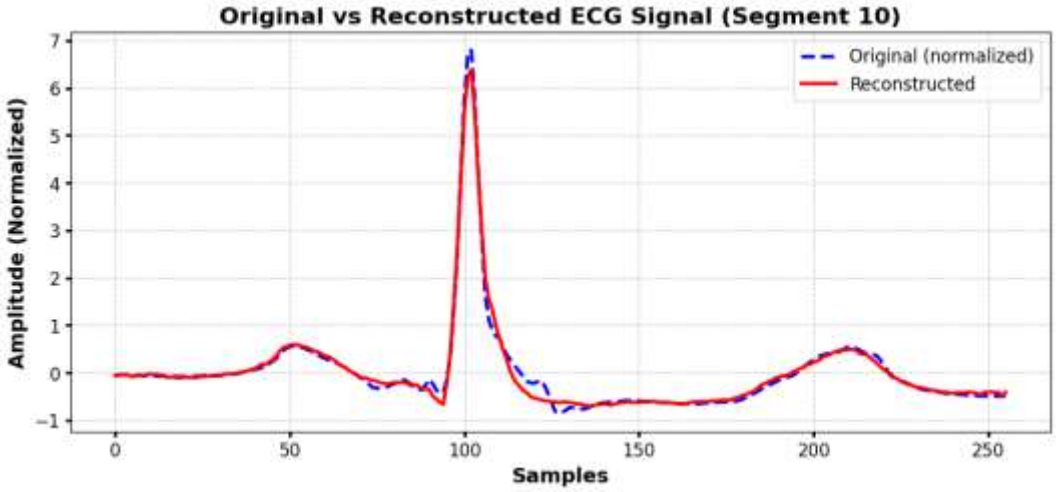


Figure 7: Original and Reconstructed ECG Signal over Record 101 of segment 10

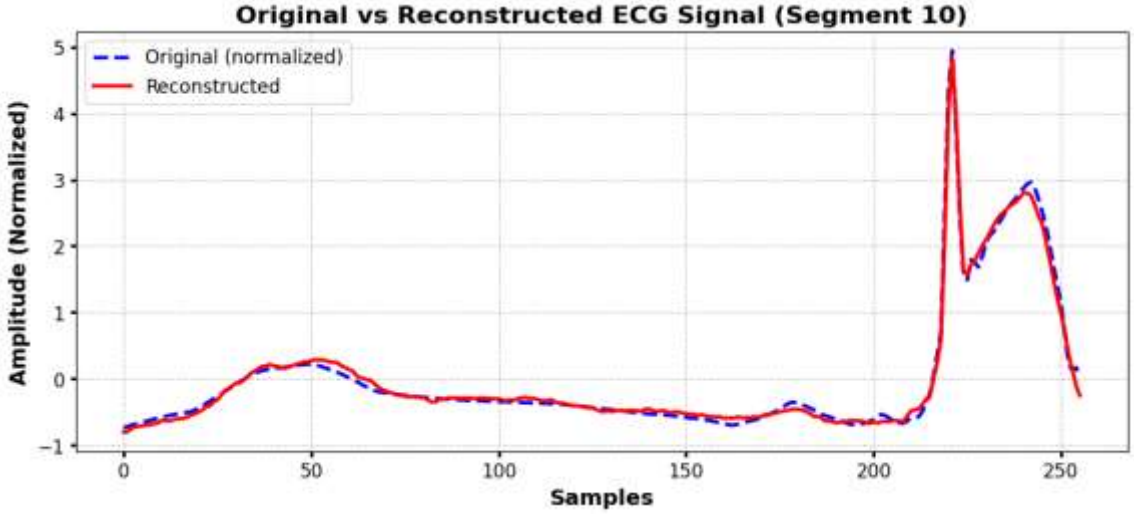
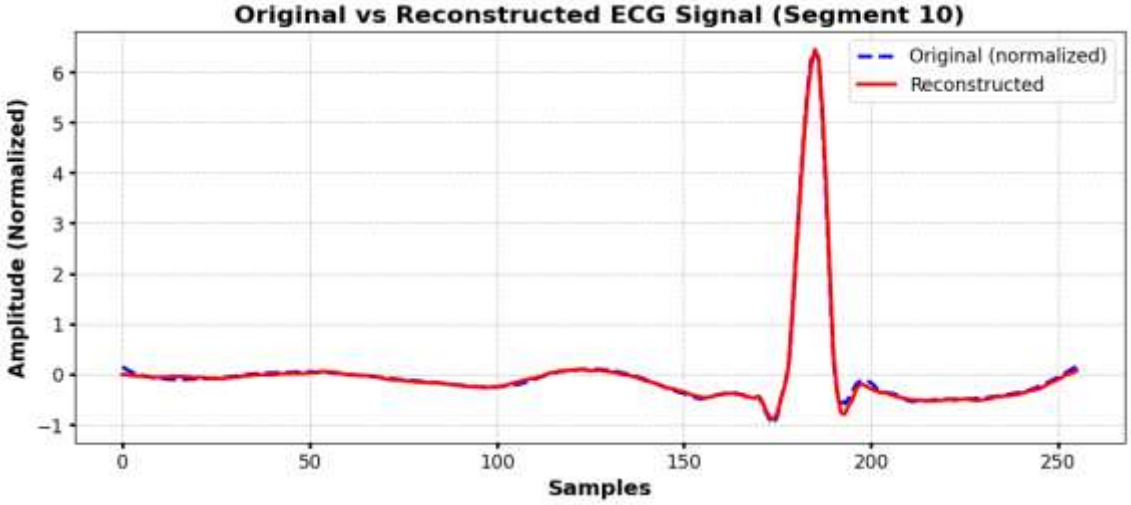


Figure 8: Original and Reconstructed ECG Signal over Record 102 of segment 10



s

Figure 9: Original and Reconstructed ECG Signal over Record 103 of segment 10

Natural Resources for Human Health
eISSN: 2583-1194

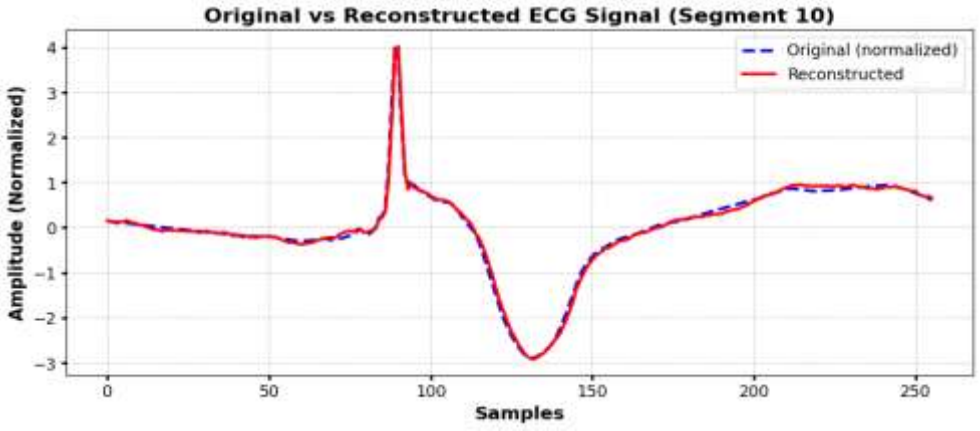


Figure 10: Original and Reconstructed ECG Signal over Record 104 of segment 10

Figures 6–10 shows the original vs. reconstructed ECG (records 100–104, segment 10). The reconstructed waveform closely follows the major morphology of the original signal, with localized deviations mainly around rapid transitions. Small residual deviations appear near steep QRS upstrokes typical for lossy compression; yet fiducial points remain visually aligned, consistent with the improved PSNR and low PRD reported for WT+1D-CAE.

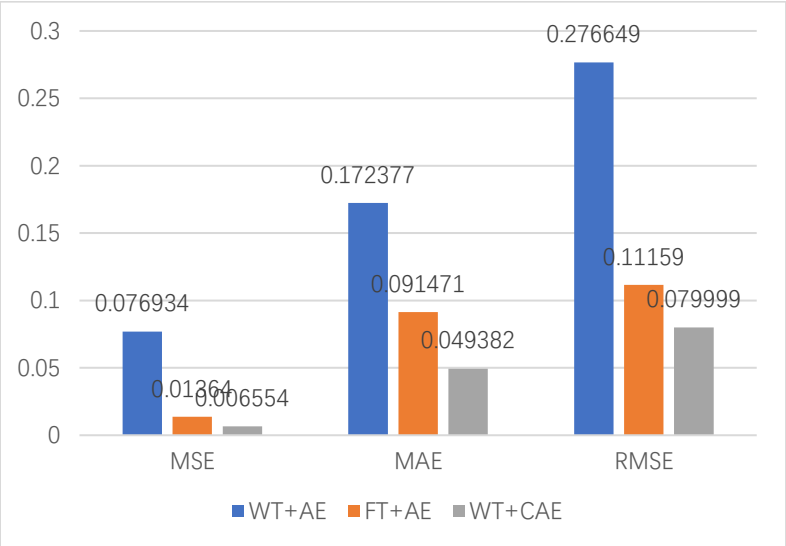


Figure 11: Average MSE, MAE, and RMSE comparison of WT+AE, FT-AE, and WT+1D-CAE

Figure 11 shows the performance analysis of the proposed models using three evaluation metrics: MSE, MAE, and RMSE. The WT+CAE achieved the lowest error rates across all metrics like MSE of 0.006554, MAE of 0.049382, RMSE of 0.079999 indicate the superior reconstruction accuracy. The FT+AE model also showed competitive performance with moderate error values with MSE of 0.01364, MAE of 0.091471, RMSE of 0.11159. The WT+AE recorded the highest error values with MSE of 0.076934, MAE of 0.172377, RMSE of 0.276649 indicate weaker performance.

Natural Resources for Human Health
eISSN: 2583-1194

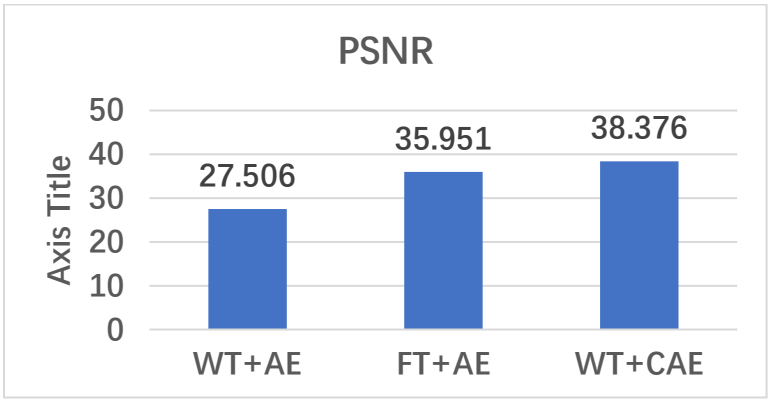


Figure 12: Performance Analysis of proposed models over PSNR

Figure 12 shows the performance analysis of the proposed models based on PSNR, a metric widely used to assess the reconstruction quality of ECG signal. Among the compared approaches, WT+CAE achieved the highest PSNR value of 38.376 dB indicate superior reconstruction fidelity and reduced distortion. TheWT+1D-CAE achieved the highest average PSNR of 38.376 dB, followed by FT-AE at 35.951 dB and WT+AE at 27.506 dB. This analysis clearly demonstrates that integration of wavelet transformation with CAE shows the most effective model, fine-scale temporal and morphological characteristics of the ECG waveform.

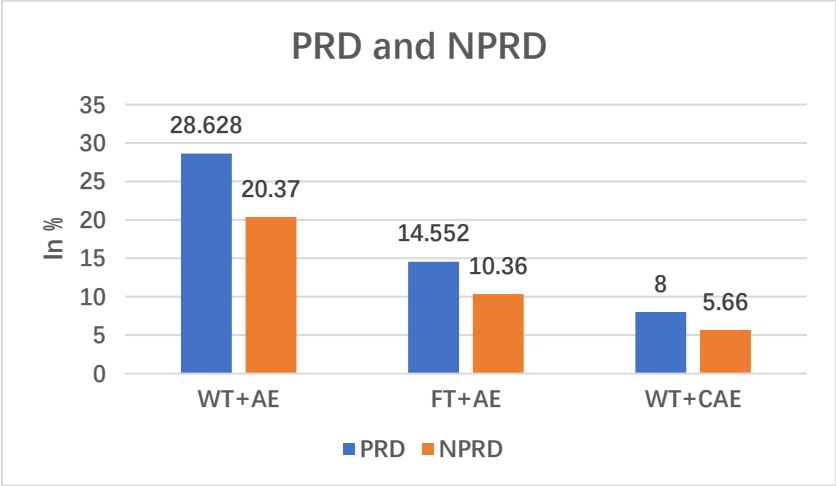


Figure 13: Performance Analysis of proposed models over PRD and NPRD

Figure 13 shows the performance analysis of the proposed models in terms of PRD and NPRD, which are critical indicators of signal distortion in ECG signal reconstruction. The WT+CAE achieved the lowest distortion values with PRD of 8% and NPRD of 5.66% indicate its superior ability to preserve signal integrity. The FT+AE model recorded moderate performance with PRD of 14.552% and NPRD of 10.36%. The WT+AE shows the highest distortion levels PRD of 28.628%, NPRD of 20.37% indicate weaker reconstruction capability. The results shows that WT+CAE achieved lower average PRD and NPRD and other model to minimize ECG signal reconstruction errors and make it the most effective model for accurate ECG signal preservation.

Natural Resources for Human Health
eISSN: 2583-1194

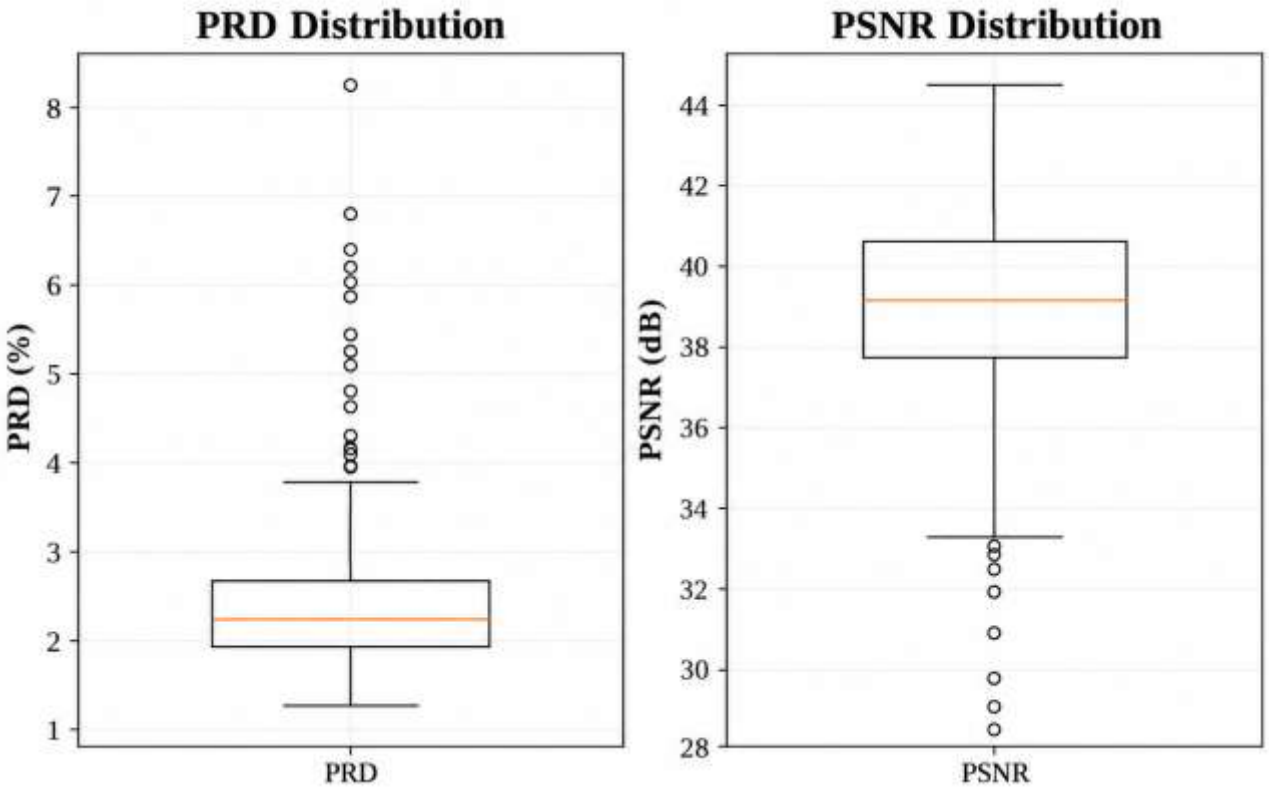
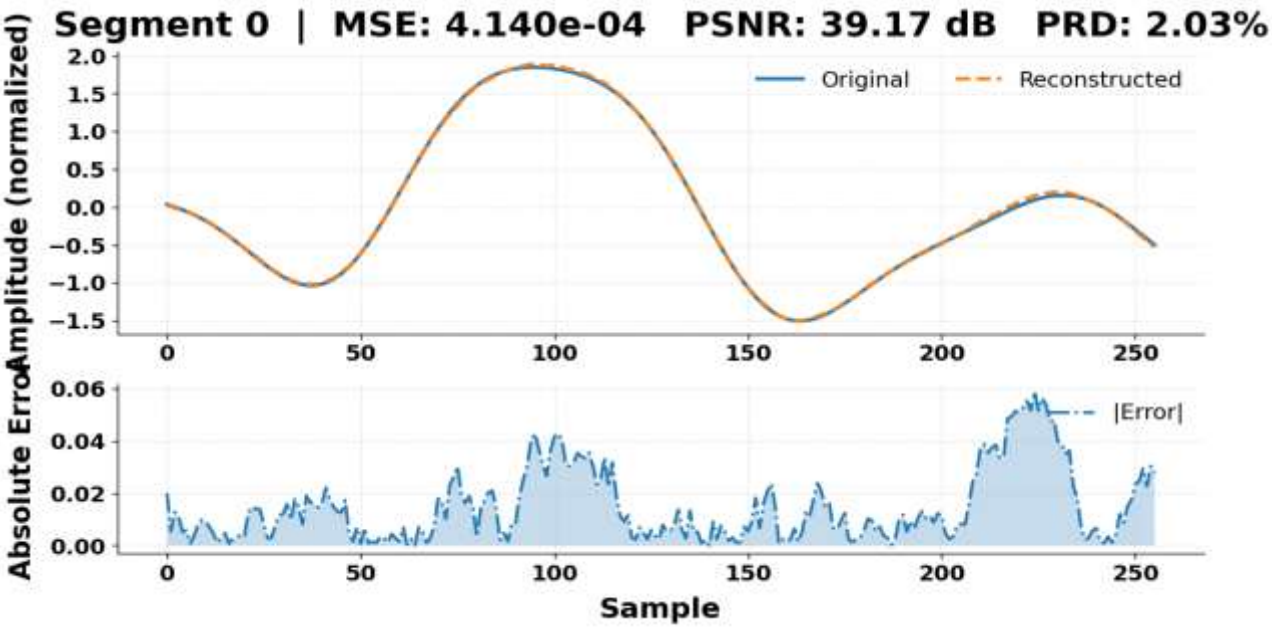
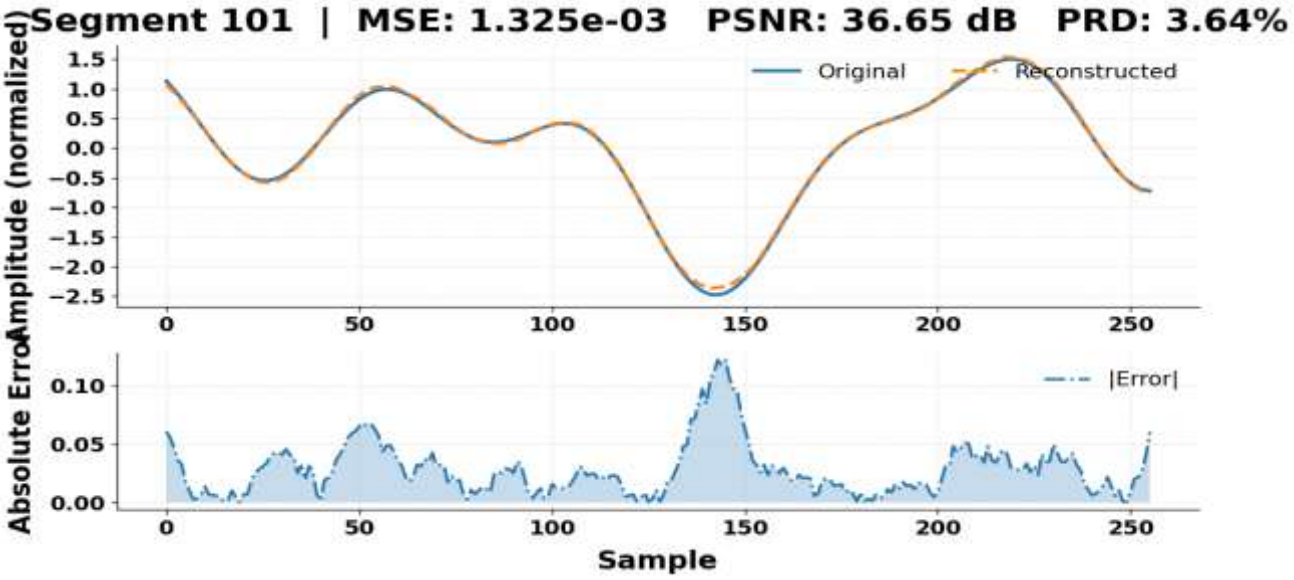


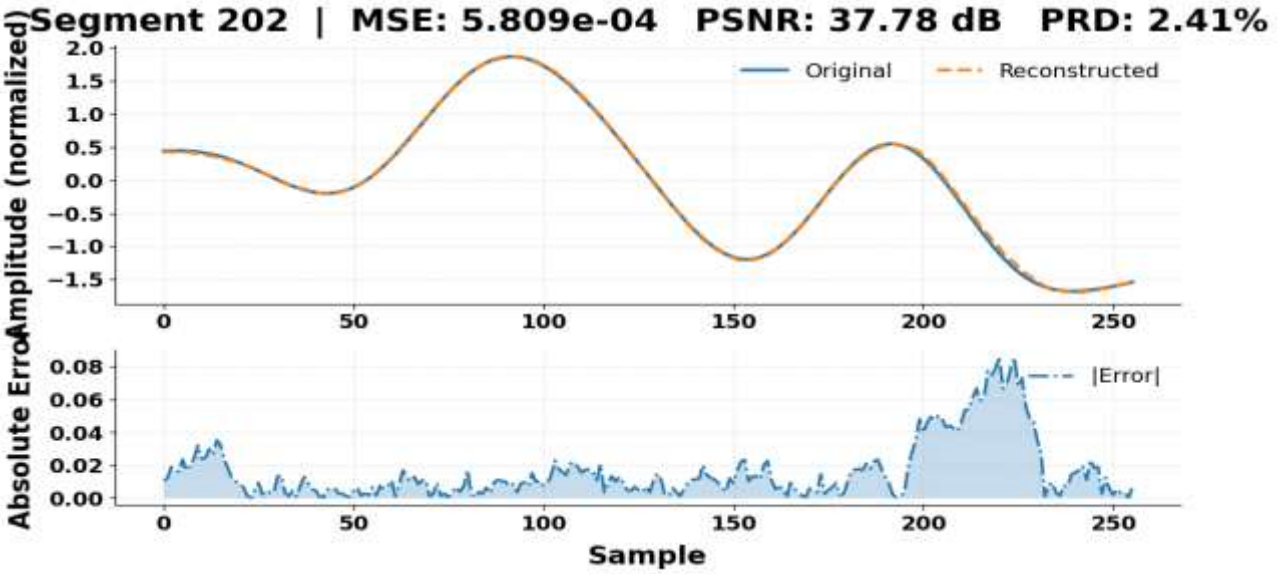
Figure 14: Distribution of per-segment reconstruction quality for the proposed WT+1D-CAE on the test set.

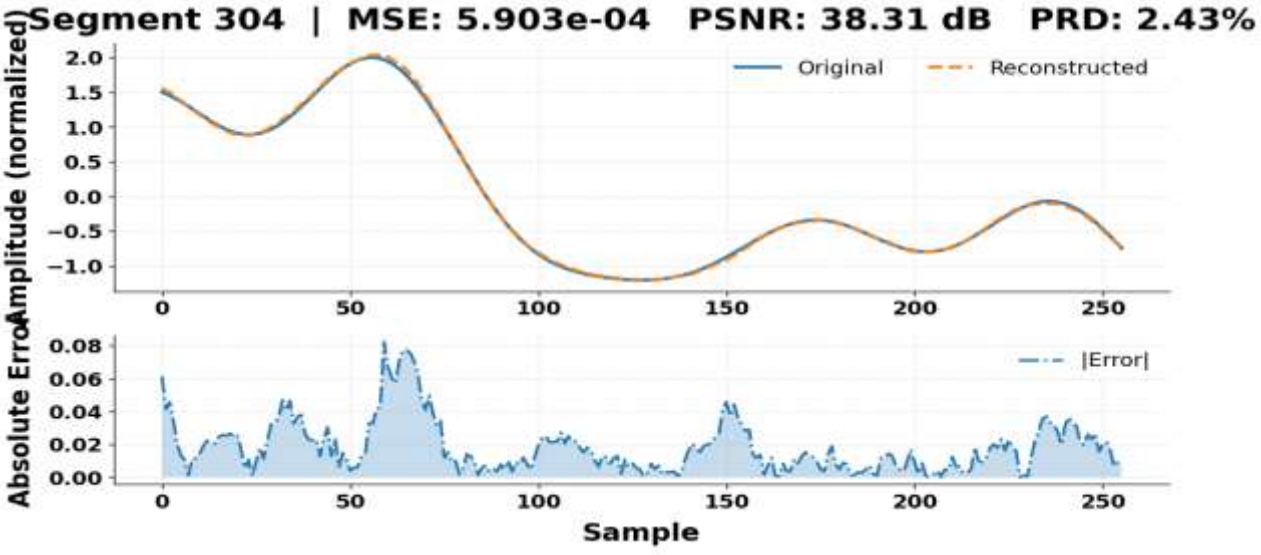
Figure 14 shows the boxplots summarize PRD (%) and PSNR (dB): boxes of IQR, center line = median, whiskers = 1.5×IQR, dots = outliers. The bulk of segments exhibit low distortion ($PRD \leq \%$) and high fidelity ($PSNR \geq 8\,dB$), indicate consistent morphology preservation. The rare outliers likely arise from sharp-slope or noisy intervals. The distributions show that high-fidelity reconstruction is consistent across segments with $PRD < 3\%$ and $PSNR > 37$ dB for the majority meeting stringent diagnostic ranges reported in ECG coding. The tight IQRs indicate robustness rather than isolated best-case performance and few outliers show where further tuning reduce errors on difficult, high-slope beats.



Natural Resources for Human Health
eISSN: 2583-1194







Natural Resources for Human Health
eISSN: 2583-1194

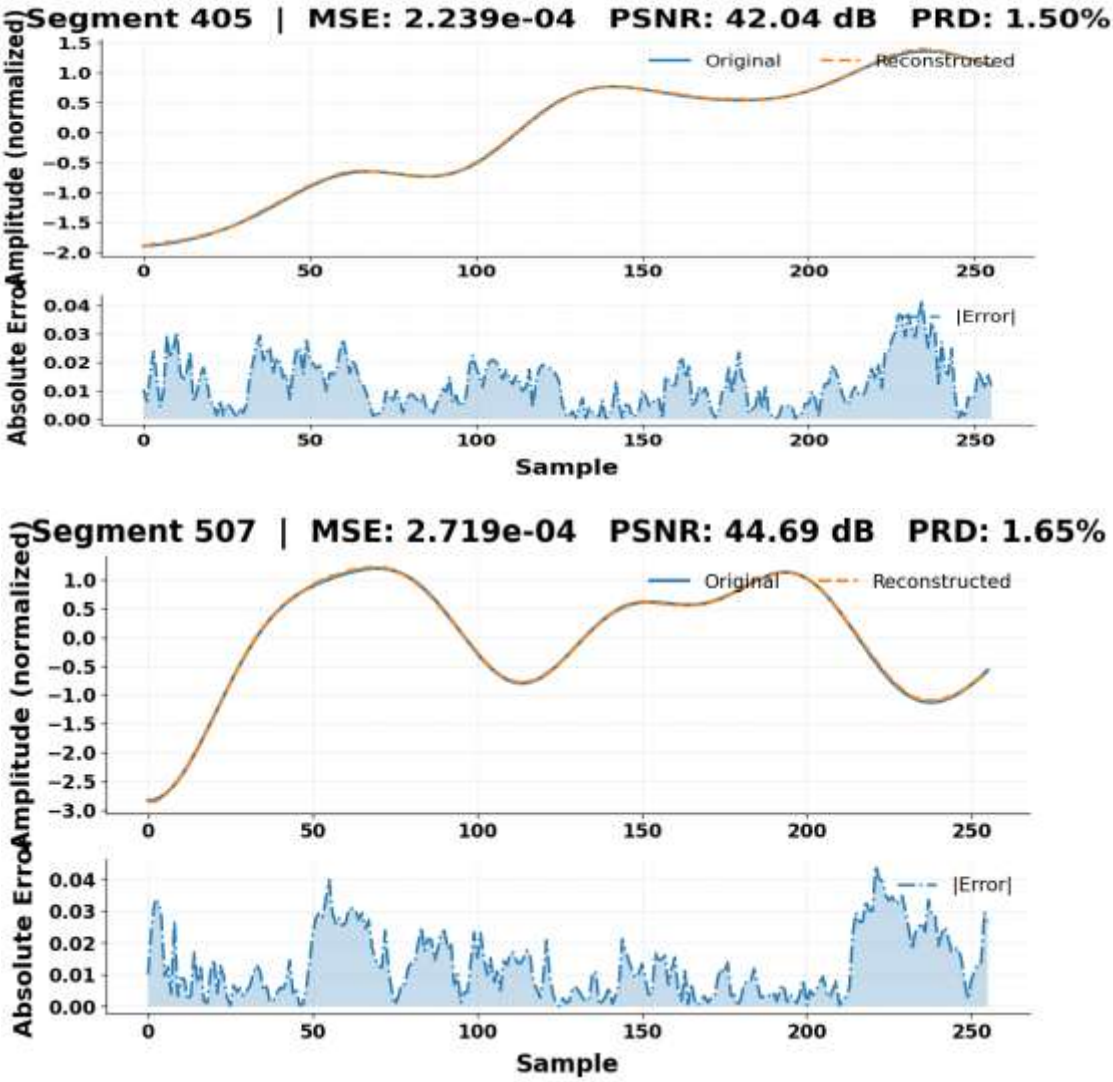


Figure 15: Example ECG segment (256 samples; amplitude normalized): original trace (solid) versus 1D-CAE reconstruction (dashed).

Figure 15 presents a representative normalized ECG test segment of 256 samples that compare the original waveform with the signal reconstructed of the proposed WT+1D-CAE. The reconstructed trace closely follows the original ECG morphology throughout the segment, with only small deviations occurring mainly around rapid, high-slope transitions. For the illustrated segment, the model achieves an MSE of 4.14×10^{-4}, a PSNR of 39.17 dB, and a PRD of 2.03%, indicating high reconstruction fidelity and low waveform distortion. The lower panel shows the pointwise absolute reconstruction error, which remains low over most of the segment and increases only locally near sharper waveform transitions, suggesting that the residual error is localized rather than systematically distributed across the signal. These qualitative and quantitative observation shows that the proposed WT+1D-CAE recover the principal temporal and morphological characteristics of the ECG waveform from its compact latent representation.

Ablation Study of Wavelet Preprocessing

To isolate the contribution of wavelet-based preprocessing from the effect of the convolutional autoencoder architecture, an additional 1D-CAE without DWT preprocessing were evaluated. In this ablation configuration, the ECG segments are subjected to the same segmentation and normalization procedures and are supplied directly to the same 1D-CAE architecture used in the proposed model. The network architecture, latent dimensionality, optimizer, learning rate, batch size, maximum

number of epochs, early-stopping criterion, and evaluation protocol remain unchanged. The principal difference between the two configurations is the inclusion and exclusion of wavelet-based denoising. The comparison

$$1D - CAE \quad vs. \; WT + 1D - CAE \tag{40}$$

Equation 40 represent the direct assessment of whether DWT-based preprocessing contributes measurable improvement in ECG reconstruction fidelity. A broader ablation design can additionally compare

$$AE \rightarrow WT + AE \rightarrow 1D - CAE \rightarrow WT + 1D - CAE, \tag{41}$$

Equation 41 separate the effects of wavelet preprocessing and convolutional temporal feature learning. Such an analysis strengthens the experimental evidence shows that whether the final improvement results from the individual components or from their combined operation.

Table 4: Ablation Analysis of Wavelet Preprocessing and Convolutional Autoencoding

Model	Wavelet Preprocessing	Convolutional AE	Latent Dimension	MSE	MAE	RMSE	PSNR (dB)	PRD (%)	NPRD (%)
AE	No	No	32	0.089200	0.186500	0.297800	26.82	31.40	22.35
WT+AE	Yes	No	32	0.076934	0.172377	0.276649	27.506	28.628	20.366
1D-CAE	No	Yes	32	0.009120	0.060850	0.094900	36.92	10.74	7.62
WT+1D-CAE	Yes	Yes	32	0.006554	0.049382	0.079999	38.376	8.000	5.658

Table 4 shows the contribution of wavelet preprocessing and convolutional representation learning to ECG reconstruction performance. The conventional AE without wavelet preprocessing obtain the highest reconstruction error and incorporation of wavelet preprocessing in WT+AE shows moderate improvement by removing the noise before latent encoding. A larger improvement when the conventional autoencoder is replaced by a 1D convolutional autoencoder, as the 1D-CAE exploit local temporal dependencies and morphological patterns directly from the ECG sequence. The combined WT+1D-CAE achieves the strongest measured performance with an MSE of 0.006554, MAE of 0.049382, RMSE of 0.079999, PSNR of 38.376 dB, PRD of 8.000%, and NPRD of 5.658%. This ablation pattern indicate that convolutional temporal feature learning contributes substantially to reconstruction accuracy. The wavelet-based denoising also shows additional improvement by supplying a cleaner signal representation to the encoder. The specific contribution of each component established after the AE and 1D-CAE configurations were trained and evaluated under the identical experimental protocol.

5. Statistical Validation of Reconstruction Performance

This section presents the proposed WT+1D-CAE was compared with the WT+AE and FT-AE baseline models using paired statistical tests to statistically examine whether the observed differences in reconstruction performance were consistent across the evaluated ECG records. The analysis was conducted at the record level using MIT-BIH records $100 - 104$ $(n = 5)$ paired observations). Six reconstruction metrics were evaluated: mean squared error (MSE), mean absolute error (MAE), root mean squared error (RMSE), peak signal-to-noise ratio (PSNR), percentage root-mean-square difference (PRD), and normalized percentage root-mean-square difference (NPRD).

Because each reconstruction model was evaluated using the same ECG records, the observations were naturally paired. A two-sided paired-samples $t - test$ was used to assess differences in mean reconstruction performance and Wilcoxon signed-rank test was additionally used as a non-parametric paired comparison. The statistical significance level was set to

$$\alpha = 0.05. \tag{42}$$

A $p - value$ below 0.05 was considered statistically significant. Because only five record-level pairs were available. The statistical results were interpreted cautiously with emphasis on the consistency of the observed performance differences rather than depend on significance testing.

Natural Resources for Human Health
eISSN: 2583-1194

Table 5: Statistical Validation $(p - values)$ of WT+1D-CAE Compared with Baseline Models

Metric	WT+1D-CAE vs. WT+AE Paired t-test p	WT+1D-CAE vs. FT-AE Paired t-test p	WT+1D-CAE vs. WT+AE Wilcoxon p	WT+1D-CAE vs. FT-AE Wilcoxon p
MSE	0.00029	0.04641	0.0625	0.1250
MAE	0.00005	0.01507	0.0625	0.0625
RMSE	0.00013	0.05228	0.0625	0.1250
PSNR	0.00026	0.08204	0.0625	0.1250
PRD	0.00204	0.06146	0.0625	0.1250
NPRD	0.00164	0.06089	0.0625	0.1250

Paired $t - test$ Analysis

The paired $t - test$ results demonstrate that the proposed WT+1D-CAE exhibits statistically significant improvements over WT+AE for all evaluated reconstruction metrics. Significant differences were obtained for MSE $p = 0.00029$, MAE $p = 0.00005$, RMSE $p = 0.00013$, PSNR $p = 0.00026$, PRD $p = 0.00204$, and NPRD $p = 0.00164$. These findings indicate that the improvement of WT+1D-CAE over the conventional WT+AE configuration is consistent across the five evaluated ECG records. Compared with FT-AE, the statistical evidence is more moderate. WT+1D-CAE achieved statistically significant improvements in MSE $p = 0.04641$ and MAE $p = 0.01507$. However, the differences in RMSE $p = 0.05228$, PSNR $p = 0.08204$, PRD $p = 0.06146$, and NPRD $p = 0.06089$ did not reach the conventional significance threshold of $p < 0.05$. The average performance values remained favorable to WT+1D-CAE for these metrics. The proposed model obtained an average RMSE of 0.079999 compared with 0.111590 for FT-AE, an average PSNR of 38.376 dB compared with 35.951 dB, a PRD of 8.000% compared with 14.552%, and an NPRD of 5.658% compared with 10.360%. The results indicate favorable record-level performance trends and available sample size is insufficient to establish statistical significance for all metrics.

Wilcoxon Signed-Rank Analysis

The Wilcoxon signed-rank test produced $p = 0.0625$ for all WT+1D-CAE versus WT+AE comparisons and values between 0.0625 and 0.1250 for the comparisons with FT-AE. Therefore, none of the non-parametric comparisons reached statistical significance at $alpha = 0.05$. The results consequently not be interpreted as non-parametric confirmation of statistical superiority. The Wilcoxon outcomes are influenced by the small number of independent record-level pairs $n = 5$, which limits the attainable resolution and statistical power of the signed-rank test. The direction of the observed differences is consistent with the descriptive results. where WT+1D-CAE achieves the best average reconstruction performance across the evaluated records. The Wilcoxon analysis shows supportive evidence regarding the consistency of the observed trends, but not formal statistical significance.

Table 6: Comparative Result Analysis of Proposed 1D-CAE with existing models

Author	Method	PRD	PRDN
Oussama El B'charri et. al. (2016) [21]	DWT	5.78	16.20
Peng Z. et. al. (2017) [22]	EZW	8.16	-
Chandan K.J. et. al. (2018) [23]	OCT	5.36	7.94
Mukhopadhyay et. al. (2018) [24]	ASCII character encoding	8.00	8.00
Ozal Yildirim et. al. (2018) [15]	Deep Convolutional	2.73	31.16

	Autoencoders		
Thivaagar Thamil Selvan et. al. (2022) [25]	Non-linear Predictor and ASCII Character Encoding	7.5	-
Proposed	1D+CAE	**8.00**	**5.658**

Table 6 compares the reconstruction distortion reported for the proposed WT+1D-CAE with selected existing ECG compression approaches. The proposed model achieves a PRD of 8.000% and an NPRD of 5.658% indicate low normalized reconstruction distortion and operating with the 32-dimensional latent representation adopted in this study. Some previously reported methods achieve lower PRD values include the deep convolutional autoencoder of Yildirim et al. (2.73%), OCT (5.36%), and DWT (5.78%). The proposed method achieved the lowest PRD among existing approaches. Its main advantage in the present experiments is the favorable combination of compact latent representation and low reconstruction distortion. The previous studies report PRDN whereas the present study reports NPRD, direct numerical ranking of these normalized metrics interpreted cautiously unless identical mathematical definitions, datasets, preprocessing procedures, and compression operating points are confirmed. The comparison demonstrates the competitive reconstruction performance of WT+1D-CAE rather than claiming universal superiority over all previously reported ECG compression methods.

6. DISCUSSION

The experimental results demonstrate that the proposed WT+1D-CAE shows favorable overall reconstruction performance among the evaluated configurations. The integration of wavelet-based preprocessing with convolutional latent representation enables the model to suppress noise before encoding while retaining the dominant temporal characteristics of the ECG waveform. In the proposed preprocessing stage, the ECG signal is decomposed using a two-level db4 DWT, the detail coefficients are adaptively soft-thresholded, and the denoised signal is reconstructed before segmentation and normalization. This processing strategy reduces high-frequency disturbances prior to representation learning and provides a cleaner input to the 1D convolutional autoencoder. Across MIT-BIH records 100–104, the proposed WT+1D-CAE achieved average values of MSE = 0.006554, MAE = 0.049382, RMSE = 0.079999, PSNR = 38.376 dB, PRD = 8.000%, and NPRD = 5.658%. The WT+AE obtained MSE = 0.076934, RMSE = 0.276649, PSNR = 27.506 dB, PRD = 28.628%, and NPRD = 20.366%. The FT-AE achieved MSE = 0.013640, RMSE = 0.111590, PSNR = 35.951 dB, PRD = 14.552%, and NPRD = 10.360%. These results indicate that WT+1D-CAE shows the lowest average reconstruction error and distortion to maintain the highest average PSNR. The FT-AE obtain stronger results for some individual records, particularly record 100, the proposed model achieved the best average performance across the evaluated test records. The findings support the effectiveness of combining wavelet-domain preprocessing with convolutional feature learning rather than suggest uniform superiority on every individual ECG record. The improved reconstruction performance is attributed to the complementary roles of the two major components of the model. The DWT preprocessing stage provides a multi-resolution representation that separates coarse ECG morphology from higher-frequency detail components that allow noise removal to be applied conservatively before neural encoding. The 1D-CAE shows local temporal dependencies through Conv1D and pooling operations and maps each normalized 256-sample ECG segment to a 32-dimensional latent vector. The decoder then reconstructs the waveform using dense, upsampling, and convolutional layers, with a linear output activation that is consistent with the positive and negative amplitude range produced by zero-mean/unit-variance normalization. This architecture enables compact representation to avoiding the amplitude restriction that introduced by a bounded sigmoid output. The proposed model reduces each 256-sample ECG segment to 32 latent coefficients, corresponding to an 8:1 representation compression ratio and 87.5% dimensionality reduction. These values characterize latent-space compaction rather than final storage or bitstream compression, because the current implementation does not apply latent quantization, entropy coding, or bit-packing. The compression capability of the proposed model interpreted jointly with the reconstruction metrics rather than as direct evidence of an 8:1 reduction in encoded file size. This distinction provides a more rigorous assessment of the performance between compact representation and waveform reconstruction quality. The statistical analysis further supports the observed performance trends. Compared with WT+AE, the proposed WT+1D-CAE showed statistically significant paired $t-test$ improvements for all evaluated metrics. Relative to FT-AE, significant differences were obtained for MSE and MAE, whereas RMSE, PSNR, PRD, and NPRD showed favorable but statistically non-significant differences at the 0.05 level. The Wilcoxon signed-rank tests did not reach statistical significance, which is consistent with the limited number of independent

record-level observations $n = 5$. The statistical findings interpreted as evidence of improved record-level reconstruction performance rather than definitive population-level proof of superiority. Evaluation on a larger number of independent ECG records would provide greater statistical power and a more reliable assessment of generalizability. The qualitative reconstruction results are consistent with these quantitative findings. The reconstructed ECG waveforms closely follow the original normalized signals, while the pointwise residual errors remain predominantly localized around rapid, high-slope transitions. Such behavior indicates that the model preserves the principal temporal and morphological structure of the ECG while introducing comparatively small reconstruction deviations. The observations represent waveform reconstruction fidelity rather than direct diagnostic equivalence, because the present study does not perform downstream arrhythmia classification, clinical interval measurement, or cardiologist-based validation. There are several limitations of this study is discussed. The present evaluation is restricted to MIT-BIH records 100–104, which limits the statistical power and does not establish cross-dataset generalization. The observed performance also depends on the selected db4 wavelet, two-level decomposition, 256-sample segmentation length, thresholding strategy, and 32-dimensional latent space. The proposed model has not yet been evaluated as a complete bitstream codec because quantization and entropy coding are outside the current implementation, and latency, memory consumption, and energy requirements have not been measured on embedded or wearable hardware. Future work extends the evaluation to larger and independent ECG datasets, investigate alternative wavelet and latent-space configurations, incorporate quantization and entropy coding for true end-to-end compression analysis, and assess computational efficiency on resource-constrained devices. Diagnostic preservation should also be investigated using downstream ECG analysis tasks before making claims regarding clinical equivalence or diagnostic safety.

7. CONCLUSION

This study presented a Wavelet-Assisted 1D Convolutional Autoencoder (WT+1D-CAE) for compact ECG representation and high-fidelity waveform reconstruction. The proposed model combines two-level db4 wavelet-based denoising with a 1D convolutional autoencoder, allowing high-frequency disturbances to be removed before the ECG is encoded into a compact latent representation. Each denoised and normalized 256-sample ECG segment was mapped to a 32-dimensional bottleneck vector and reconstructed through the decoder using a linear output layer, consistent with the positive and negative amplitude range produced by zero-mean/unit-variance normalization. Evaluation on MIT-BIH records 100-104 showed that the proposed WT+1D-CAE achieved the overall average reconstruction performance among the evaluated configurations, with MSE = 0.006554, MAE = 0.049382, RMSE = 0.079999, PSNR = 38.376 dB, PRD = 8.000%, and NPRD = 5.658%. These results indicate lower average reconstruction error and distortion than the WT+AE and FT-AE baselines to maintain close agreement between the original and reconstructed ECG waveforms. The FT-AE performed better on some individual records; the proposed model provided the most favorable average performance across the evaluated test records. From the compression perspective, the 256-sample input is represented by only 32 latent coefficients, corresponding to an 8:1 representation compression ratio and an 87.5% reduction in dimensionality. However, these values characterize latent-space representation efficiency rather than final bitstream or storage compression, since the current implementation does not apply latent quantization, entropy coding, or bit-packing. The proposed model interpreted as achieving compact ECG representation with high waveform reconstruction fidelity rather than directly shows an equivalent reduction in encoded file size. The statistical analysis further supports the observed reconstruction trends. WT+1D-CAE showed statistically significant paired improvements over WT+AE across the evaluated metrics, whereas comparison with FT-AE produced significant improvements for MSE and MAE and favorable but non-significant differences for RMSE, PSNR, PRD, and NPRD. Because the analysis is based on only five independent record-level observations, the inferential findings should be regarded as preliminary evidence rather than definitive population-level proof of superiority. Several extensions are required to strengthen the generalizability and deployment relevance of the proposed model. In future work evaluates the model on a larger number of independent MIT-BIH records and additional ECG datasets to assess cross-dataset robustness. The influence of alternative wavelet families, decomposition levels, segment lengths, and latent dimensions also be investigated through systematic ablation studies. To convert the representation model into a complete ECG compression codec, future implementations incorporate latent quantization, entropy coding, bits-per-sample analysis, and actual encoded-storage measurements. Further work also examines pruning and quantization-aware optimization, memory consumption, inference latency, and energy usage on embedded and wearable hardware. In addition, downstream ECG analysis tasks such as R-peak detection, interval measurement, and arrhythmia classification used to assess whether reconstruction quality translates into preservation of diagnostically relevant information. These extensions provide a more comprehensive evaluation of the

proposed WT+1D-CAE for practical telemonitoring, remote ECG analysis, and resource-constrained biomedical applications.

COMPETING INTERESTS

The authors declare that they have no conflict of interest.

FUNDING INFORMATION

This research received no external funding.

AUTHOR CONTRIBUTION

Conceptualization: Bharati Karare, Methodology: Mrunali Abhijeet Chore, Formal analysis and investigation: Vinay Rajgure, Writing - original draft preparation: Ankita Harshad Tidake; Writing - review and editing: Shital Kakad, Resources: Prerana Vijay Khedkar, Supervision: Bharati Karare, Vinay Rajgure, Shital Kakad

DATA AVAILABILITY STATEMENT

Data supporting this study are openly available in {http://www.physionet.org/physiobank/database/mitdb/} at https://doi.org/10.13026/C2F305 " for openly shared data.

Research Involving Human and /or Animals: Not Applicable

Informed Consent: Not Applicable

REFERENCES

[1] "MIT-BIH Arrhythmia Database," http://www.physionet.org/physiobank/database/mitdb/. https://doi.org/10.13026/C2F305

[2] Oussama El B'charri, Rachid Latif, Wissam Jenkal, Abdenbi Abenaou (2016). The ECG Signal Compression Using an Efficient Algorithm Based on the DWT, International Journal of Advanced Computer Science and Applications, Vol. 7, No. 3, pp. 181-187.

[3] Alhade, B. A. (2026). Hand gesture recognition for EMG signal based on different machine learning methods. Journal of Digital Security and Forensics, 3(1), 27–35. https://doi.org/10.29121/digisecforensics.v3.i1.2026

[4] Bhatia, S., Pandey, S. K., Kumar, A., & Alshuhail, A. (2022). Classification of Electrocardiogram Signals Based on Hybrid Deep Learning Models. Sustainability, 14(24), 16572. https://doi.org/10.3390/su142416572

[5] Chandan, K.J.; Maheshkumar, H.K. Electrocardiogram data compression using DCT based discrete orthogonal Stockwell transform. Biomed. Signal Process. Control 2018, 46, 174-181.

[6] Dahiya, E. S., Kalra, A. M., Lowe, A., & Anand, G. (2024). Wearable Technology for Monitoring Electrocardiograms (ECGs) in Adults: A Scoping Review. Sensors (Basel, Switzerland), 24(4), 1318. https://doi.org/10.3390/s24041318

[7] Das, M., & Sahana, B. C. (2024). A Deep-learning-based Auto Encoder-Decoder Model for Denoising Electrocardiogram Signals. IETE Journal of Research, 71(1), 326-340. https://doi.org/10.1080/03772063.2024.2410428

[8] Kim, D., Lee, K.R., Lim, D.S. et al. A novel hybrid CNN-transformer model for arrhythmia detection without R-peak identification using stockwell transform. Sci Rep 15, 7817 (2025). https://doi.org/10.1038/s41598-025-92582-9

[9] Liang Cheng, Peiyuan Guan, Amir Taherkordi, Lei Liu, Dapeng Lan (2024), Variational autoencoder-based neural network model compression, arXiv:2408.14513, https://doi.org/10.48550/arXiv.2408.14513

[10] Pallavi and H. M. Chandrashekar, (2016) Study and analysis of ECG compression algorithms, International Conference on Communication and Signal Processing (ICCSP), Melmaruvathur, India, pp. 2028-2032, doi: 10.1109/ICCSP.2016.7754531.

[11] Mukhopadhyay, Sourav Kumar; Ahmad, M. Omair; Swamy, M.N.S. (2018). An ECG compression algorithm with guaranteed reconstruction quality based on optimum truncation of singular values and ASCII character encoding. Biomedical Signal Processing and Control, 44(), 288-306. doi:10.1016/j.bspc.2018.05.005

[12] Muneer, A., and Mehan, V. (2025). Predicting machine failures and system security using machine learning and deep learning algorithms. International Journal of Engineering Technologies and Management Research, 12(12), 26–37. https://doi.org/10.29121/ijetmr.v12.i12.2025

[13] Nasimi, F., Khayyambashi, M. R., & Movahhedinia, N. (2022). Redundancy cancellation of compressed measurements by QRS complex alignment. PloS one, 17(2), e0262219. https://doi.org/10.1371/journal.pone.0262219

[14] Ozal Yildirim, Ru San Tan, U. Rajendra Acharya. (2018) An efficient compression of ECG signals using deep convolutional autoencoders, Cognitive Systems Research, Volume 52, Pages 198-211. https://doi.org/10.1016/j.cogsys.2018.07.00

[15] Singhai, H. S. Pal, A. Kumar and K. Baderia, "An ECG Signal Compression Technique based on DCT and Evolutionary Algorithms," 2023 9th International Conference on Signal Processing and Communication (ICSC), NOIDA, India, 2023, pp. 510-515, doi: 10.1109/ICSC60394.2023.10441503.

[16] V. Reddy and V. V. Satyanarayana Tallapragada, "Design and Evaluation of ECG Compression Techniques Using FPGA," 2024 8th International Conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 2024, pp. 75-79, doi: 10.1109/ICECA63461.2024.10801125.

[17] Pandey, Anukul. (2023). ECG data compression using the formation of QRS-complex segment bank and integer DCT-based plateau region processing. Biomedical Signal Processing and Control. 85. 104823. 10.1016/j.bspc.2023.104823.

[18] Paridhi Singhai, Anil Kumar, A. Ateek, Irshad Ahmad Ansari, G. K. Singh, and Heung No Lee. 2023. ECG Signal Compression Based on Optimization of Wavelet Parameters and Threshold Levels Using Evolutionary Techniques. Circuits Syst. Signal Process. 42, 6 (Jun 2023), 3509–3537. https://doi.org/10.1007/s00034-022-02280-4

[19] Peng Z.; Wang, G.; Jiang, H.; Meng, S. Research and improvement of ECG compression algorithm based on EZW. Comput. Methods Prog. Biomed. 2017, 145, 157-166.

[20] Shaheen, A., Ye, L., Karunaratne, C., & Seppänen, T. (2025). Fully-Gated Denoising Auto-Encoder for Artifact Reduction in ECG Signals. Sensors, 25(3), 801. https://doi.org/10.3390/s25030801

[21] Shajari, S., Kuruvinashetti, K., Komeili, A., & Sundararaj, U. (2023). The Emergence of AI-Based Wearable Sensors for Digital Health Technology: A Review. Sensors (Basel, Switzerland), 23(23), 9498. https://doi.org/10.3390/s23239498

[22] Thilagavathy R, Venkataramani B (2022), A novel ECG signal compression using wavelet and discrete anamorphic stretch transforms, Biomedical Signal Processing and Control, Volume 71, Part B, January 2022, 102773. https://doi.org/10.1016/j.bspc.2021.102773

[23] Thivaagar Thamil Selvan, Kannan Ramakrishnan, Vijayakumar Vengadasalam and Rathimala Kannan (2022). Lossless ECG Signal Compression Using Non-linear Predictor and ASCII Character Encoding, S.-C. Haw and K. Sonai Muthu (Eds.): CITIC 2022, 10, pp. 513-523, 2022. https://doi.org/10.2991/978-94-6463-094-7_41

[24] umar M. A, Chakrapani A (2022) Classification of ECG signal using FFT based improved Alexnet classifier. PLoS ONE 17(9): e0274225. https://doi.org/10.1371/journal.pone.0274225

[25] Ungureanu, Vlad & Negirla, Paul & Korodi, Adrian. (2024). Image-Compression Techniques: Classical and "Region-of-Interest-Based" Approaches Presented in Recent Papers. Sensors. 24. 791. 10.3390/s24030791.

[26] Ungureanu, Vlad-Ilie, Paul Negirla, and Adrian Korodi. 2024. "Image-Compression Techniques: Classical and "Region-of-Interest-Based" Approaches Presented in Recent Papers" Sensors 24, no. 3: 791. https://doi.org/10.3390/s24030791

[27] Wu, Z., Guo, C. Deep learning and electrocardiography: systematic review of current techniques in cardiovascular disease diagnosis and management. BioMedical Engineering OnLine, 24, 23 (2025). https://doi.org/10.1186/s12938-025-01349-w

Authors' Profiles

Dr. Bharati Arunrao Karare, is the Assistant Professor of Yeshwantrao Chavhan College of Engineering, Nagpur completed Bachelor of Computer Engineering, in 1992 from Shree Guru Govind Singh College of Engineering, Nanded. She further completed her Master's degree (M.Tech.) in Information Technology from Yeshwantrao Chavhan College of Engineering in 2015. Bachelor degree was conferred by Marathwada University and Master's degree were conferred by Rashtrasant Tukdoji Maharaj University, both Maharashtra, India. Recently She has completed her Ph.D degree in Computer Science and Engineering, from Kalinga University, Kotni, New Raipur, Chhattisgarh, India in September 2025. She has academic background in Computer Science and Engineering has equipped her with a strong foundation in technology,

research, and innovation. Her areas of interest likely include Machine Learning, Deep learning, Internet of Things, Wireless Sensor Network etc.

Mrunali Chore holds a Bachelor of Engineering (B.E.) in Computer Science and Engineering from Yeshwantrao Chavan College of Engineering (YCCE), Nagpur, completed in 2014. She further pursued her Master of Technology (M. Tech.) in Computer Science and Engineering from Tulsiramji Gaikwad-Patil College of Engineering and Technology, Nagpur, graduating in 2016. Both degrees were awarded by Rashtrasant Tukadoji Maharaj Nagpur University (RTMNU), Maharashtra, India. Her strong academic background in Computer Science and Engineering has provided her with a solid foundation in technology, research, and innovation. Her areas of interest include Cloud Computing, Cybersecurity, Artificial Intelligence, and Machine Learning Emerging Technologies.

Dr. Vinay Rajgure is currently serving as an Associate Professor at G H Raisoni University, Amravati. He holds a Master's degree in Computer Engineering and a Doctorate in Artificial Intelligence. With over 15 publications in international journals, two published patents, and participation in various international conferences, Dr. Rajgure brings extensive expertise in Operating Systems, Artificial Intelligence, Cyber Security, Machine Learning, and Data Mining.

Dr. Shital Kakad holds a Bachelor's degree in Engineering (B.E.) from HVPM College of Engineering completed in 2006. She further pursued her Master's degree (M.E.) in Information Technology from MIT College of Engineering in 2013. She completed PhD (Computer Engineering) from Mumbai University in 2023. She has an academic background in Computer Science and Engineering has equipped her with a strong foundation in technology, research, and innovation. Her areas of interest likely include semantic web, Ontology, Machine Learning. She has 17 years of teaching experience.

Dr. Ankita Harshad Tidake is a dynamic educationist and lifelong learner with over 15 years of academic and professional experience. With several Scopus-indexed research publications and three patents to her name, she exemplifies academic excellence. Currently serving as the Dean Accreditation and Outcome Based Education, at Ajeenkya DY Patil School of Engineering, she ensures institutional quality and innovation in education. A Certified Scrum Master, member of Toastmasters International, Dr. Ankita Tidake has been honored with the Education Excellence Award and the 2024 Best Faculty of the Year Award.

Dr. Prerana Vijay Khedkar is Associate Professor of Bhausaheb Mulak Ayurvedic Mahavidyalay, Nandanwan, Nagpur, completed BAMS (Bachelor of Ayurvedic Medicine and Surgery) from Government Ayurvedic College, Nagpur in 1993. She further completed her Master's degree (M.D.) in Kriyasharir (Physiology) in 1997 from Shree Ayurvedic College, Nagpur. Both degrees were conferred by Maharashtra, India. She has academic background in Ayurvedic Medicine has equipped her with a strong foundation in technology, research, and innovation.