technologies

MDPI

Article

Performance of an Efficient Hybrid Dilated–Long Short-Term Memory with Residual Learning for High-Fidelity Electrocardiogram Denoising Signal

Suchada Sitjongsataporn [1,†], Pipat Sakarin [1,†] and Theerayod Wiangtong [2,*,†]

[1] Department of Electronic Engineering, School of Electrical and Electronic Engineering (SEE), Faculty of Engineering and Technology, Mahanakorn University of Technology, Nongchok, Bangkok 10530, Thailand; ssuchada@mut.ac.th (S.S.); pipat.skr@gmail.com (P.S.)
[2] Department of Electrical Engineering, School of Engineering, King Mongkut's Institute of Technology Ladkrabang, 1 Chalong Krung 1 Alley, Lat Krabang, Bangkok 10520, Thailand
* Correspondence: theerayod.wi@kmitl.ac.th
† These authors contributed equally to this work.

Abstract

Addressing the critical challenge of signal degradation in biosensor cardiac monitoring, this paper introduces an efficient hybrid dilated–long short-term memory (LSTM) with residual learning (HDLR), which is a novel architecture engineered for a high-fidelity electrocardiogram (ECG) denoising signal. The proposed HDLR model synergistically integrates with dilated convolutions to expand the receptive field for multi-scale feature extraction. This is an LSTM-based backbone used to resolve long term temporal dependencies with residual learning paths to stabilize gradient flow and accelerate convergence. The proposed HDLR architecture integrates three core functional components with dilated convolutional layers utilized for local temporal feature extraction, where varying dilation rates expand the receptive field to capture both local waveform patterns and broader morphological structures without increasing computational complexity. Experimental results demonstrate a significant leap in performance, with the HDLR model achieving a mean squared error (MSE) of 0.002176, a signal-to-noise ratio (SNR) of 14.4420 dB, and a Matthews correlation coefficient (MCC) of 0.9822. Beyond quantitative metrics, the proposed HDLR architecture exhibits exceptional robustness in preserving cardiac morphology, specifically the P-wave and QRS complex of the ECG signal under stochastic noise conditions. These findings underscore the HDLR model's potential as a backbone for next generation, real time diagnostic systems in intelligent healthcare.

Keywords: long short-term memory (LSTM); gated recurrent unit (GRU); dilated convolution; residual learning; electrocardiogram (ECG) signal

Check for updates

Academic Editors: Hongying Meng, Everardo Inzunza-González and Fabrizio Stasolla

Received: 11 June 2026
Revised: 16 July 2026
Accepted: 20 July 2026
Published: 22 July 2026

Copyright: © 2026 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.

1. Introduction

Electrocardiogram (ECG) signals are highly susceptible to various types of noise, ranging from high-frequency muscle artifacts to low-frequency baseline wander. Therefore, effective denoising models are crucial for extracting clean signals from contaminated data, ensuring high-fidelity waveforms for clinical analysis. Precise interpretation of these waveforms is vital for detecting life-threatening cardiac conditions, as even minor signal distortions can lead to misdiagnosis. Consequently, the development of the advanced noise-filtering architectures to preserve the delicate morphological

characteristics of the ECG amidst severe noise interference is a top priority in modern cardiovascular engineering.

In the context of deep learning models, the authors in [1] proposed a hybrid approach combining convolutional neural networks (CNNs) and long short-term memory (LSTM), integrated with standard wavelet-based techniques to enhance ECG noise reduction. Similarly, the Grunet model was introduced in [2], utilizing a CNN architecture to extract complex features from ECG signals followed by a gated recurrent unit (GRU) for temporal analysis, an approach holding particular significance in sports medicine. GRUs [3] serve as an alternative to LSTMs, employing fewer parameters while maintaining comparable performance across various prediction fields. In [4], a Kalman-iterative shrinkage and thresholding algorithm for ECG noise reduction was presented, and its performance was analyzed in conjunction with deep learning. Furthermore, variations of recurrent networks have been widely adopted; for instance, the authors in [5] applied LSTM to enhance wind power forecasting for wind speed calibration, while the Bidirectional LSTM (BiLSTM) introduced in [6] extends LSTM capabilities by processing data in both forward and backward directions. In [7], a modified BiLSTM-based model was proposed for epilepsy seizure detection, with simulation results demonstrating high effectiveness in capturing seizures. Beyond recurrent architectures, a multiscale extreme learning machine was introduced in [8] to classify cardiovascular diseases, and a modified CNN model was utilized in [9] to detect arrhythmias that disrupt the circulatory system by leveraging ECG images.

Meanwhile, a denoising autoencoder based on a fully convolutional network was introduced in [10] to eliminate noise from ECG signals. The model was trained using the MIT-BIH arrhythmia database and evaluated against various noise levels from the MIT-BIH noise stress test database. While these studies demonstrate significant advancements in ECG denoising, there remains a critical need for an architecture capable of achieving high-fidelity signal restoration and real-time computational efficiency.

In order to improve a multi-scale CNN with a larger receptive field in [11,12], the authors in [11] have presented the non-zero elements of the convolutional based on the gradient method. While the authors in [12] have combined the dilated convolutional layer with the same size of convolutional layer without redundant information. In [13], the hierarchical dilated deep learning based on an exponential smoothing model has been applied with multilayer LSTM for electric load forecasting. Simulation results depicted that dilated convolution based on LSTM can capture the long-term relationships with efficient training. In [14], the concept of residual connections has been confirmed that can accelerate and improve significantly the training speed for inception architecture in deep convolutional model.

Therefore, this paper proposes the efficient hybrid dilated–LSTM with residual learning (HDLR) model designed to reduce noise in ECG signals acquired from cardiac biosensors. As presented in Figure 1, the framework leverages the combined strengths of multi-scale local temporal feature extraction and deep temporal sequence modeling to achieve high-fidelity ECG signal restoration under stochastic noise conditions. The noisy input signal first undergoes a specialized dilated convolution stage, which utilizes multiple dilation levels to effectively expand the receptive field and perform multi-scale feature extraction. For temporal modeling, an LSTM Backbone is integrated to resolve long-term temporal dependencies while simultaneously stabilizing gradient flow and accelerating model convergence during training.

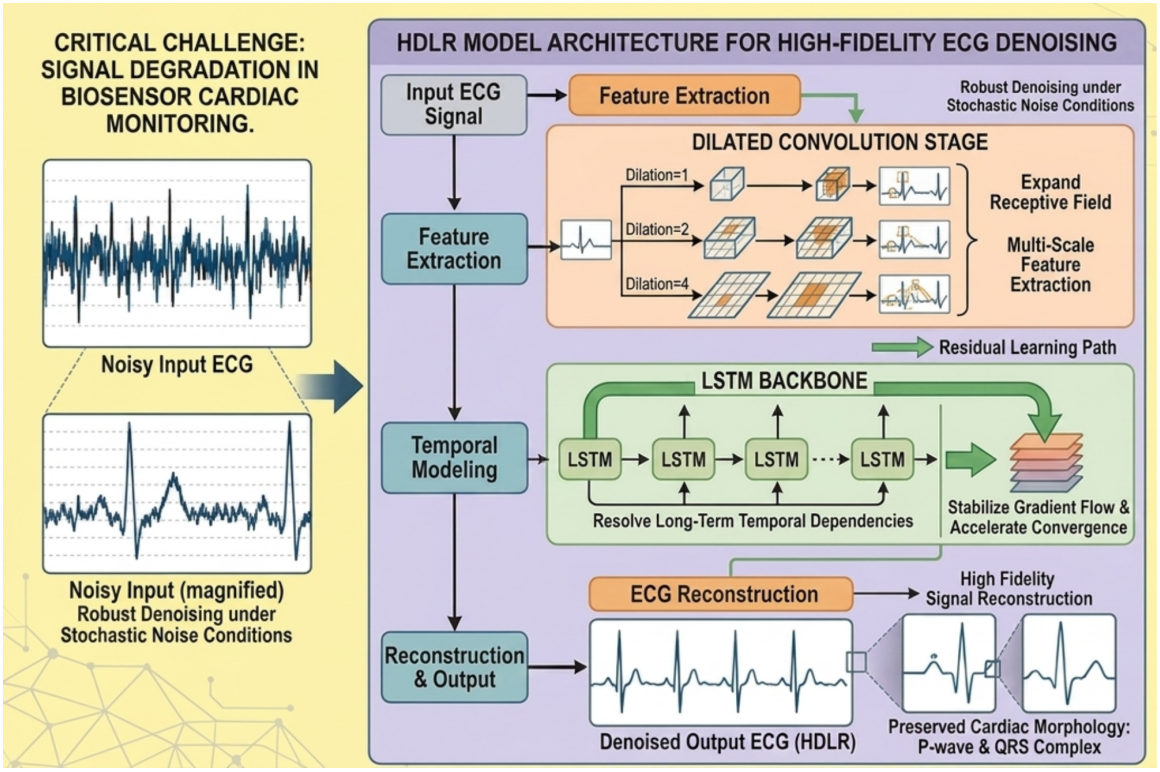


Figure 1. Overview of proposed HDLR framework.

Furthermore, the architecture incorporates a dedicated residual learning path that runs parallel to these main processing stages. This path ensures that delicate morphological features and critical high-frequency components of the ECG, which might otherwise be lost during deep feature transformations, are effectively retained and mapped to the output. By synergizing these local and temporal elements, the proposed HDLR model achieves high-fidelity ECG reconstruction that successfully preserves the structural integrity of the P-wave and QRS complex necessary for accurate clinical diagnosis.

2. Proposed Efficient Hybrid Dilated–LSTM with Residual Learning (HDLR)

An efficient HDLR is a model designed for restoring raw ECG signals contaminated by diverse noise sources as presented in Figure 2. Proposed HDLR architecture integrates three core functional components dilated convolutional layers utilized for local temporal feature extraction, where varying dilation rates expand the receptive field to capture both local waveform patterns and broader morphological structures without increasing computational complexity. LSTM layers are responsible for modeling temporal dependencies that allow the network to learn the sequential nature and rhythmic variations of the cardiac cycle and a residual layer incorporated for identity preservation.

The proposed hybrid HDLR architecture shown in Figure 2 is designed with the principle of efficient local and temporal integration to achieve high-accuracy, real-time ECG signal restoration. The process begins with a noisy ECG signal being fed into the system. The internal structure is clearly divided into two main blocks.

The first block acts as a hybrid local temporal feature extraction block focusing on multi-dimensional feature extraction and fusion. The raw signal is first fed into a local feature extractor, the Standard Conv1D layer, with 64 filters, a kernel size of 7, and a dilation rate of 1. This layer detects short-range morphological features with steep slopes and high frequencies, such as the peaks of QRS-complex ECG waveforms.

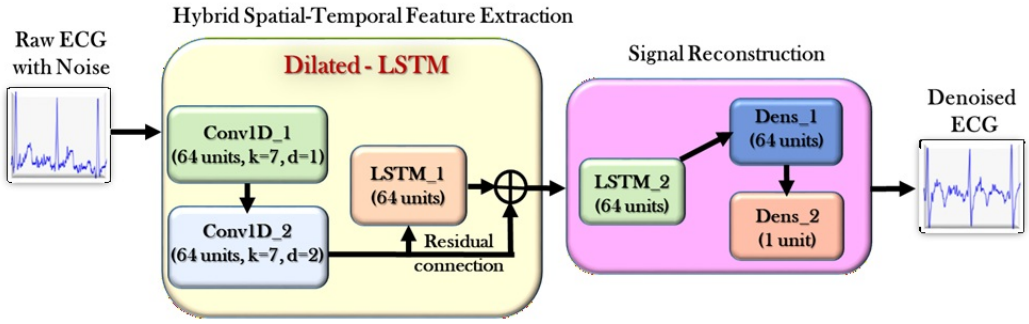


Figure 2. Proposed hybrid dilated–LSTM with residual learning (HDLR) architecture layout.

The initial local information is then driven into the subsequent dilated Conv1D layer, which uses the same number of filters and kernel size but with an increased dilation rate of 2. This mechanism expands the receptive field of the model by resulting in the network seeing a signal window that is almost twice as wide while maintaining the same number of parameters. This allows for a more accurate extraction of contextual morphological features at low frequencies such as the baseline structures of P-waves and T-waves of ECG signal. This local signal is then split into two parallel paths. The primary path enters the deep temporal dimension via the 64-unit LSTM Layer 1 to recognize sequential continuity, direction, and time-series trend behaviors. The secondary path acts as a layer-crossing shortcut via a residual connection, passing the original local temporal features from the dilated Conv1D layer around to merge with the temporal features emitted from LSTM Layer 1 at the element-wise addition processor. This two-dimensional information fusion called *Feature Fusion* is the core of the residual learning mechanism, preserving the original geometric features from being lost in the deep layers and enabling faster network convergence.

The second block of the architecture is called the *Signal Reconstruction Block*, which receives the fused hybrid information and processes it for signal recovery and reconstruction. All features are fed into the 64-unit global LSTM Layer 2, which aggregates and summarizes the contextual overview of the time series from all time points into a single vector. This temporal summary vector is then passed through the first 64-unit dense layer for decoding, where advanced feature matching is performed. Finally, it is sent to the final output layer, the 1-unit dense output layer, which, utilizing a hyperbolic tangent (tanh) activation function, converts the advanced information back into individual time-point ECG amplitude values. As a result, the proposed HDLR model delivers a clean, highly accurate restored ECG output, ready for real-time medical vital sign monitoring applications.

2.1. Dilated–LSTM Model

Following [12], dilated convolution functions by introducing a spacing dilated factor between kernel weights. This mechanism effectively broadens the receptive field without augmenting the parameter count by enabling the model to extract extensive local features from ECG signals, while preserving high computational speed.

If x_i is an input sequence and the convolutional kernel w_k of length k is the dilated convolution, they can be defined as

$$z_i = \sum_{k=0}^{K-1} x_{(i+l \cdot k)} \cdot w_k \,, \tag{1}$$

where z_i is the output of the dilated convolution operation and l is the dilation rate representing the spacing between kernel weight elements w_k. A dilated convolution with a kernel size of 3×3 and $l = 2$ is depicted with a receptive field, 5×5, in Figure 3.

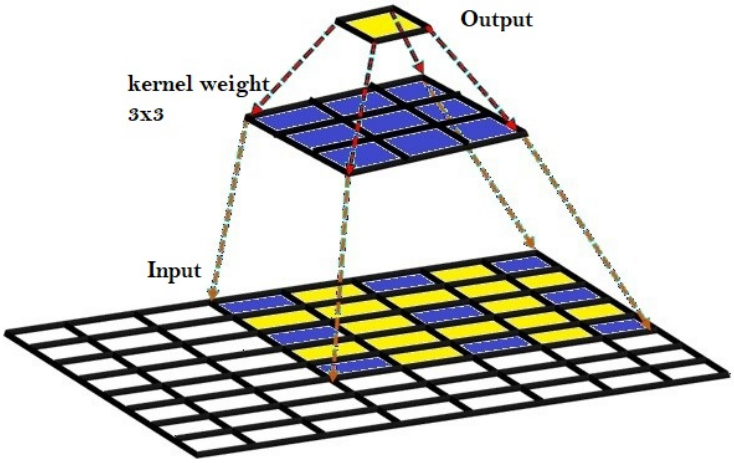


Figure 3. Dilated convolution 3×3, $l = 2$.

Temporal depth residual learning refers to adaptive deep Residual Networks (ResNet) to model the time-series data or video frame sequences by expanding network depth along the temporal dimension. Skip connections enable the network to learn temporal changes over multiple frames without suffering from vanishing gradients.

Distinguished by ability to resolve the vanishing gradient dilemma, LSTM [7] serves as an advanced sequence processor with autonomous memory control. According to the layout in Figure 4, the forget gate serves as the primary regulator by systematically eliminating redundant information to optimize the model's internal state for high-precision forecasting.

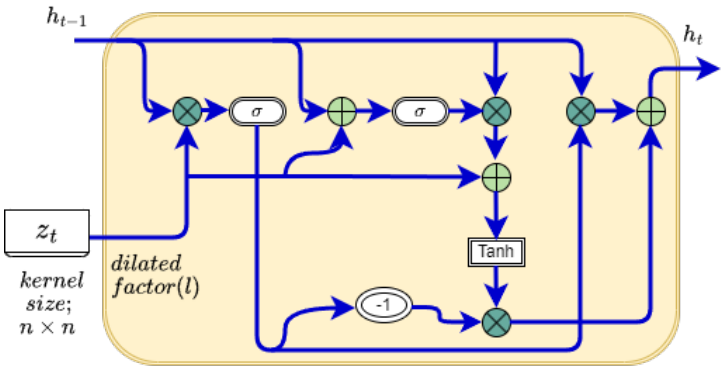


Figure 4. Dilated–LSTM structure.

The memory management within the dilated–LSTM is governed by a series of gating mechanisms. Initially, the forget gate f_t dictates which information from the previous cell state should be discarded:

$$f_t = \sigma\Big(W_f \cdot [h_{t-1}, z_t] + b_f\Big) \,. \tag{2}$$

where σ represents the sigmoid activation function, while W_f and b_f denote the weight matrix and bias vector associated with the forget gate. The terms h_{t-1} and x_t refer to the hidden state from the preceding step and the current input, respectively. To determine which new data will be integrated into the long-term memory, the input gate i_t and the candidate cell state $\tilde{C}_t$ are computed by

$$i_t = \sigma(W_i \cdot [h_{t-1}, z_t] + b_i) \,, \tag{3}$$

$$\tilde{C}_t = \sigma(W_c \cdot [h_{t-1}, z_t] + b_c) \,, \tag{4}$$

where i_t acts as a filter for the candidate state $\tilde{C}_t$, with b_i and b_c representing their respective bias vectors. The updated cell state (C_t) is then synthesized by combining the preserved past memory with the newly filtered information:

$$C_t = (f_t * C_{t-1}) + (i_t * \tilde{C}t) \, , \tag{5}$$

where $*$ signifies element-wise multiplication as the Hadamard product. Finally, the output gate Θ_t regulates the information to be propagated into the next hidden state h_t, formulated as follows:

$$\Theta_t = \sigma(W_o \cdot [h_{t-1}, z_t] + b_o) \, . \tag{6}$$

The hidden state serving as the short-term memory and the model's actual output is updated by

$$h_t = \Theta_t * \tanh(C_t) \, , \tag{7}$$

where W_o and b_o are the parameters for the output gate and tanh ensures the cell state values are scaled between -1 and 1.

2.2. Residual Connection

In the development of sophisticated deep learning models, simply increasing the number of layers often leads to a paradoxical decline in accuracy, a phenomenon known as the degradation problem. Unlike the vanishing gradient issue, which is largely addressed by normalized initialization, degradation suggests that deeper systems are inherently more difficult to optimize.

To overcome this, the residual connection was introduced. The core idea is to transition from learning an unreferenced mapping to learning a residual mapping. Instead of expecting a stack of layers to directly approximate a target function $\mathcal{H}(x)$, we reconfigure the architecture to focus on the residual, defined as $\mathcal{F}(x) := \mathcal{H}(x) - x$. This approach simplifies the optimization process if the ideal transformation for a given stage is merely an identity. So, this network can easily achieve this by suppressing the weights within the residual function $\mathcal{F}(x)$ toward zero by allowing the original signal to propagate unchanged. The interaction within a residual building block can be formulated through a feedforward operation with an integrated shortcut. Let x and y represent the input and output vectors of the considered layers, respectively. This operation is defined as

$$y = \mathcal{F}(x, \{W_i\}) + x_i \, , \tag{8}$$

where x is the original input features entering the residual unit. $\mathcal{F}(x, \{W_i\})$ is the residual mapping to be learned through the weighted layers. x_i is the identity shortcut, which performs a direct bypass of the nonlinear transformations. y is the synthesized output, which combines the refined residual features with the preserved input context.

3. Performance Metrics of Proposed HDLR Model

These six metrics evaluate different dimensions of model quality across the proposed HDLR framework for ECG denoising signal as follows.

3.1. Mean Squared Error

The mean squared error (MSE) measures the average squared deviation between the original clean signal and the restored denoising signal by

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (x_i - \hat{x}_i)^2 \tag{9}$$

where x_i represents the original clean signal sample, $\hat{x}_i$ denotes the restored or predicted signal sample by the network model, and N is the total number of data points or signal samples.

3.2. Root Mean Squared Error

The root mean squared error ($RMSE$) penalizes larger errors more heavily and presents the error metric in the same operational unit as the target signal by

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \hat{x}_i)^2} \tag{10}$$

3.3. R-Squared Score

R-squared (R^2) score indicates the proportion of variance in the pristine signal that is predictable from the model's output framework as given by

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(x_i - \hat{x}_i)^2}{\sum_{i=1}^{N}(x_i - \bar{x})^2} \tag{11}$$

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{12}$$

where $\bar{x}$ represents the empirical mean value of the original clean signal sequence.

3.4. Post-Filtering Signal-to-Noise Ratio

The post-filtering signal-to-noise ratio (SNR_{post_dB}) evaluates the signal power relative to the residual noise floor remaining after the deep neural network enhancement process, as expressed by

$$SNR_{\text{post_dB}} = 10\ \log_{10}\left(\frac{\sum_{i=1}^{N} x_i^2}{\sum_{i=1}^{N}(x_i - \hat{x}_i)^2}\right) \tag{13}$$

3.5. Delta Signal-to-Noise Ratio

The Delta SNR (ΔSNR_{dB}) quantifies the direct metric improvement achieved by the denoising paradigm, reflecting the net gain between post-filtering and pre-filtering signal environments, as given by

$$\Delta SNR_{\text{dB}} = SNR_{\text{post_dB}} - SNR_{\text{pre_dB}} \tag{14}$$

where the pre-filtering SNR ($SNR_{\text{pre_dB}}$) representing the initialized noise state of corrupted input signal is formulated by

$$SNR_{\text{pre_dB}} = 10\ \log_{10}\left(\frac{\sum_{i=1}^{N} x_i^2}{\sum_{i=1}^{N}(x_i - \tilde{x}_i)^2}\right) \tag{15}$$

where $\tilde{x}_i$ denotes the raw noisy signal sample recorded prior to model optimization.

3.6. Matthews Correlation Coefficient

For detection tasks, the Matthews correlation coefficient (MCC) is adopted due to its strict mathematical reliability on highly imbalanced datasets, as calulated by

$$MCC = \frac{(TP \cdot TN) - (FP \cdot FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{16}$$

where true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) are obtained from the categorical confusion matrix evaluation.

4. Experimental Results

The workflow of proposed HDLR framework following in Algorithm 1 begins by loading the noisy ECG dataset [15–17] and extracting the raw signal into the variable $full_signal$. A Butterworth bandpass filter is then defined with a low-cut frequency of 0.5 Hz, a high-cut frequency of 45.0 Hz, a sampling rate (fs) of 250.0 Hz, and a filter order of 5. In order to extract the training dataset, the first 200,000 data points of the raw signal are sliced into *train_raw*, which is subsequently passed through the defined Butterworth filter to generate the ground-truth target signal (*train_clean*).

Then, two separate instances of MinMaxScaler are initialized to fit and transform both the raw and clean signals independently by scaling their amplitude values into the range of $[-1, 1]$. The normalized signals are restructured into overlapping sequences using a sliding window method with a sequence length (*seq_len*) of 100-time points in order to capture consecutive temporal dependencies. This dataset is partitioned into training and validation sets ($X_train, X_val, Y_train, Y_val$) as a 10% validation ratio. For evaluation, an unknown testing block from index 200,000 to 205,000 is extracted and normalized through the raw signal sliced into the inference sequences (X_test_inf), and its corresponding ground truth is filtered and prepared as *test_truth* to evaluate the model performance.

When the configured data is fed into the network, the noisy ECG sequence with a shape of $(100, 1)$ enters the first block designated as the hybrid local temporal feature extraction block. The process initiates with short-range local feature extraction through a standard Conv1D layer featuring 64 filters, a kernel size of 7, and a dilation rate of 1 for capturing the high-frequency morphological transitions such as the sharp slopes of QRS-complex waveforms followed by a batch normalization (BN) layer and a LeakyReLU activation function. The extracted local features are then driven into a subsequent dilated Conv1D layer with the same filter configurations, while an increased dilation rate of 2 is expanding the receptive field to capture macro-level and low-frequency contextual structures as the baseline geometries of P-waves and T-waves without the multiplying parameters.

After another cycle of BN and LeakyReLU, the local tensor splits into two parallel streams. The primary path transitions into the deep temporal dimension via a 64-unit LSTM Layer 1 configured with `return_sequences=True` to map sequential continuity and directional tendencies, while the secondary path functions as a layer-crossing shortcut via a residual connection mechanism. In the next phase, the fused hybrid features enter the second main division called the signal reconstruction block. The unified features are fed into a 64-unit Global LSTM Layer 2 that can aggregate and summarize the overall contextual time-series behavior across all time points into a singular compressed context vector. This summarized temporal vector is passed through a 64-unit dense layer for high-level feature matching and latent decoding. The output is finally forwarded to a 1-unit dense output layer running a hyperbolic tangent (tanh) activation function, which maps the advanced representations back into individual time-point ECG amplitude values.

Following the proposed HDLR architecture, the network is compiled using the Adam optimizer with a learning rate of 0.0001 and a gradient clipping norm of 0.5, alongside a mean squared error (MSE) for loss function. The training execution runs on X_train and Y_train governed by EarlyStopping, ModelCheckpoint, and ReduceLROnPlateau callbacks to prevent overfitting and ensure optimal weight convergence. Once training is complete, the model performs inference on the unknown X_test_inf sequences to generate normalized predictions (*predicted_scaled*). These predictions are processed through the inverse transformation of the clean signal to reverse the normalization by restoring the

voltage amplitudes to their original physical scale as *denoised_ECG_signal*. The final denoised signal is returned for real-time medical vital sign monitoring applications.

Algorithm 1: Proposed HDLR framework for ECG denoising signal

Input: Raw ECG signal with Noise
Output: Denoised ECG signal
Initial of variables: Set Sequence length $seq_len = 100$, train batch size = 512 inference batch size = 1024, and max epochs = 60.

1 **Load** ECG with noise dataset and extract raw signal: $full_signal$
2 **Define** $butter_bandpass_filter$ with $lowcut = 0.5$, $highcut = 45.0$, $fs = 250.0$, and $order = 5$
3 **Split** $full_signal$ to extract training raw data:
$train_raw = full_signal[0 : 200000]$
4 **Apply** filtering: $train_clean = butter_bandpass_filter(train_raw)$
5 **Initialize** MinMaxScalers (sc_raw, sc_clean) and fit-transform signals into range $[-1, 1]$:
$raw_scaled = sc_raw.fit_transform(train_raw)$
$clean_scaled = sc_clean.fit_transform(train_clean)$
6 **Generate** overlapping sequences (X, Y) from scaled data with window size seq_len
7 **Partition** data into training and validation sets ($X_train, X_val, Y_train, Y_val$) with a 10% validation ratio
8 **Extract** unseen test data: $test_raw_unseen = full_signal[200000 : 205000]$
9 **Prepare** ground truth:
$test_truth = butter_bandpass_filter(test_raw_unseen)[seq_len :]$
10 **Transform** and slice test data into inference sequences X_test_inf using sc_raw
11 **Construct Proposed HDLR Architecture using Two-Block Integration:**
- **Input Layer:** ECG with noise sequence shape $(100, 1)$
- **Hybrid Local Temporal Feature Extraction Block:**
  (1) *Standard Local Extraction:*
  $\mathbf{X}_{\text{conv1}} = \text{Standard Conv1D}(\text{filters} = 64, \text{kernel} = 7, \text{dilation} = 1)(\text{Input})$
  $\rightarrow$ Batch Normalization $\rightarrow$ LeakyReLU
  (2) *Contextual Local Extraction:*
  $\mathbf{X}_{\text{conv2}} = \text{Dilated Conv1D}(\text{filters} = 64, \text{kernel} = 7, \text{dilation} = 2)(\mathbf{X}_{\text{conv1}})$
  $\rightarrow$ Batch Normalization $\rightarrow$ LeakyReLU
  (3) *Deep Temporal Dimension (Primary Path):*
  $\mathbf{X}_{\text{lstm1}} = \text{LSTM Layer 1}(\text{units} = 64, \text{return_seq} = \text{True})(\mathbf{X}_{\text{conv2}})$
  (4) *Residual Connection (Secondary Path):*
  Shortcut path routing $\mathbf{X}_{\text{conv2}}$ around $\mathbf{X}_{\text{lstm1}}$
  (5) *Feature Fusion:*
  $\mathbf{X}_{\text{fusion}} = \mathbf{X}_{\text{lstm1}} + \mathbf{X}_{\text{conv2}}$ (Element-wise Addition)
- **Signal Reconstruction Block:**
  (1) *Global Temporal Integration:*
  $\mathbf{X}_{\text{global}} = \text{Global LSTM Layer 2}(\text{units} = 64)(\mathbf{X}_{\text{fusion}})$
  (2) *Decoding Phase:*
  $\mathbf{X}_{\text{dense}} = \text{Dense}(64)(\mathbf{X}_{\text{global}})$
  (3) *Signal Restoration Output:*
  $\hat{\mathbf{Y}} = \text{Dense}(1, \text{activation} = \text{`tanh}')(\mathbf{X}_{\text{dense}})$
12 **Compile** model using Adam optimizer ($lr = 0.0001$, $clipnorm = 0.5$) and MSE loss
13 **Execute training** on X_train, Y_train with EarlyStopping, ModelCheckpoint, and ReduceLROnPlateau
14 **Perform inference** to reconstruct signal:
$predicted_scaled = Model.predict(X_test_inf)$
15 **Denormalize** output to original scale:
$denoised_ECG_signal = sc_clean.inverse_transform(predicted_scaled)$
16 **Return** $denoised_ECG_signal$

The training loss over 60 epochs is depicted in Figure 5. Proposed HDLR model is represented by the blue line, while the FCN-DAE [18], Conv-Transformer [19], LSTM, BiLSTM, GRU, and BiGRU models serve as baselines, alongside three ablation configurations designed in order to isolate the contributions of its architectural components. It is evident that the proposed HDLR model achieves the lowest convergence loss significantly by outperforming the other models. Although the proposed HDLR model initially starts with a higher validation loss due to its structural complexity, it demonstrates a superior learning rate by converging more rapidly than its counterparts. This consistent performance and lower loss floor validate the effectiveness of integrating dilated–LSTM method with residual connections for complex ECG signal processing tasks.

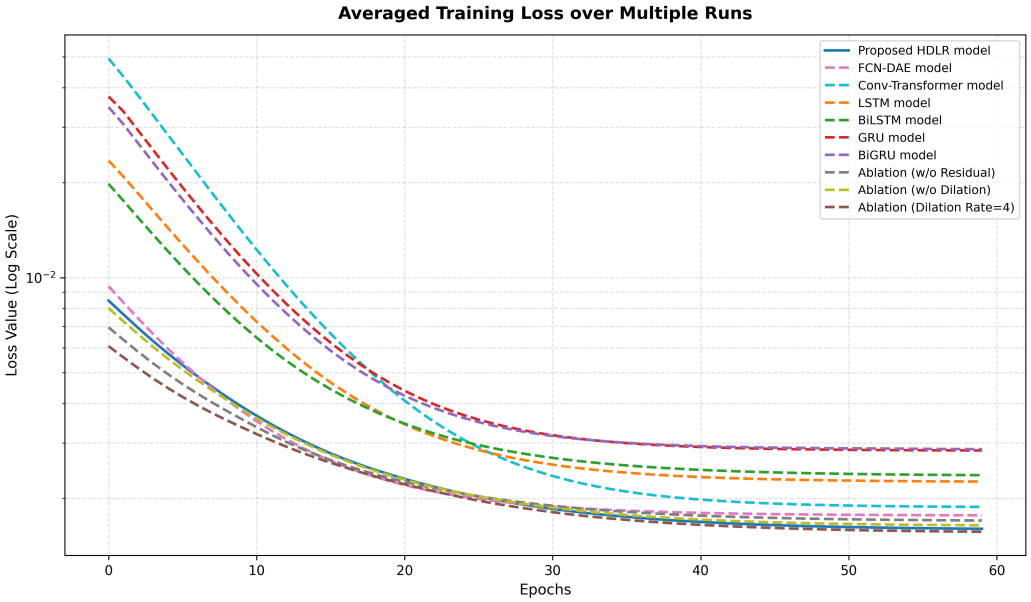


Figure 5. Averaged training loss over multiple runs of proposed HDLR model compared with different models.

The training and learning dynamics captured across the performance curves provide crucial insights into the optimization behavior, convergence speed, and generalization capabilities of the evaluated models over 60 epochs. Analysis of the convergence characteristics via the training and validation loss trajectories are illustrated in Figure 5 and Figure 6, respectively. It reveals that the proposed HDLR model exhibits an exceptionally superior optimization profile, specifically the proposed HDLR framework on the training loss curve in Figure 5, alongside its ablation variants, which initiates from a relatively lower error boundary compared to standard recurrent architectures by undergoing a continuous monotonic decline toward a final steady state with a training loss of approximately at 0.0016. This learning process is significantly validated by the validation loss trajectory in Figure 6.

Although the proposed HDLR model starts with a higher validation loss at epoch 0 due to its complex and deep architectural capacity, it demonstrates a remarkably steep and rapid convergence path by continuously optimizing past at epoch 30. It asymptotically approaches a minimum validation floor of approximately 0.0015 without any signs of divergence or fluctuations by confirming that the neural network does not suffer from overfitting despite its deep capacity. In contrast, while the traditional baseline models, such as FCN-DAE, Conv-Transformer, LSTM, and BiLSTM, exhibit smooth learning paths in the initial stage, their training and validation loss trajectories level off prematurely at a

higher error floor are reaching a plateau between 0.0020 and 0.0022, which highlights their structural limitations in capturing long-term multi-scale dependencies.

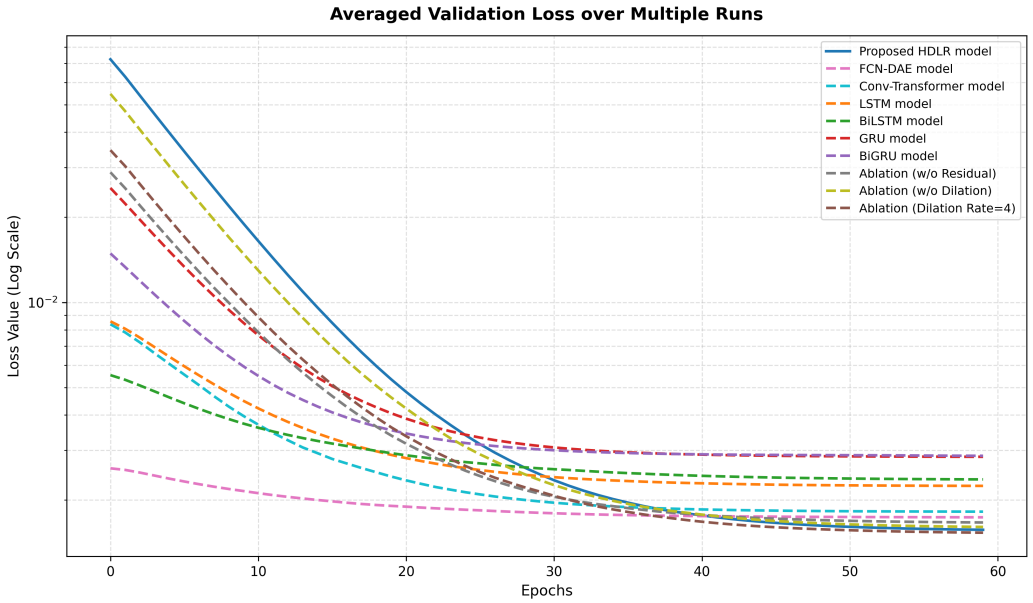


Figure 6. Averaged validation loss over multiple runs of proposed HDLR model compared with different models.

Furthermore, these architectural constraints are most evident in the GRU and BiGRU models, whose loss curves remain trapped at the highest error bound of approximately 0.0028 throughout half of the training process. Regarding the ablation configurations, namely Ablation (w/o Residual), Ablation (w/o Dilation), and Ablation (Dilation Rate = 4), they closely cluster around the final convergence zone of the proposed HDLR model. However, the intact and complete HDLR framework maintains the most balanced and stable downward trajectory across both training and validation spaces by confirming that the integration of dilated structures with residual connections successfully optimizes parameter management and temporal data flow to minimize errors with maximum efficiency.

Parallel to the loss curve convergence, the trends of the restoration accuracy curves emphasize the robustness of proposed HDLR model. On the training accuracy graph in Figure 7, the proposed HDLR framework, along with its ablation variants and the FCN-DAE baseline that demonstrates a stable upward trajectory into an asymptotic plateau that reaches an outstanding accuracy ceiling of approximately 97.1%. This learning behavior is further elucidated by the validation accuracy graph in Figure 8. Due to its structural complexity, the proposed HDLR model represented by the blue line that initially starts from a lower validation accuracy bound of approximately 78.0% at epoch 0.

However, it exhibits a remarkably steep and rapid learning acceleration, overcoming its initial lag within the first 15 epochs. It continuously optimizes to join the asymptotically approaching a stable validation accuracy floor above 96.4% without any performance degradation. This positive parallelism between the training and validation trajectories provides empirical visual evidence proving that the embedded residual connections and dilated convolutions successfully preserve and transmit multi-scale local temporal features into deep structural layers effectively.

Conversely, the accuracy curves of the traditional baseline models reflect systematic structural limitations. Although the BiLSTM and LSTM methods show the smooth movement and stable convergence at steady state, they encounter visual saturation by turning horizontally early and remaining stagnant below the HDLR curve. This creates a distinct

visual performance gap in both the training in Figure 7 and validation in Figure 8, where the BiLSTM and LSTM curves achieve a peak horizontal convergence of only 96.3% and 95.5%, respectively.

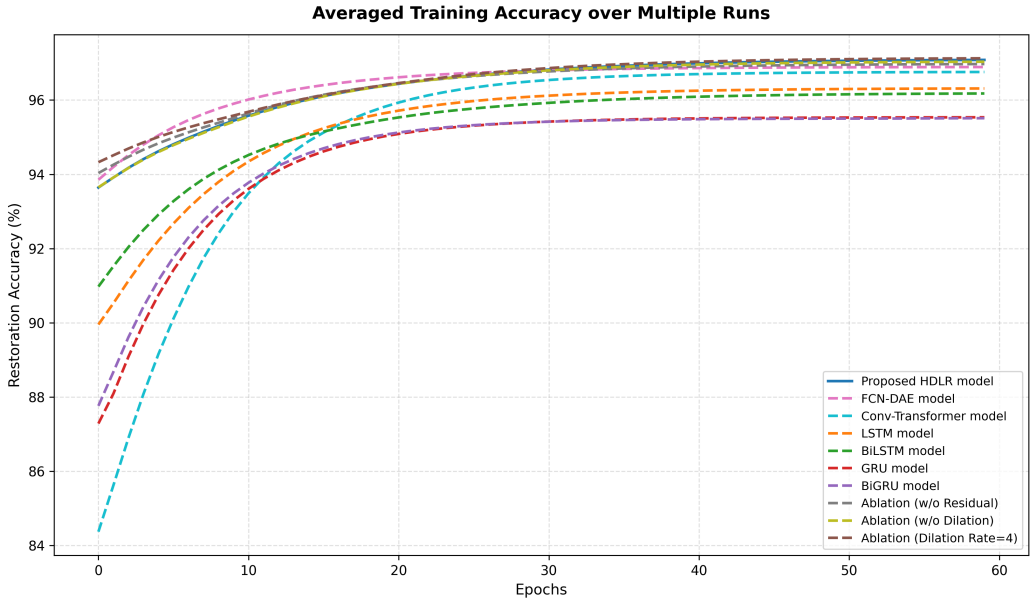


Figure 7. Averaged training accuracy over multiple runs of proposed HDLR model compared with different models.

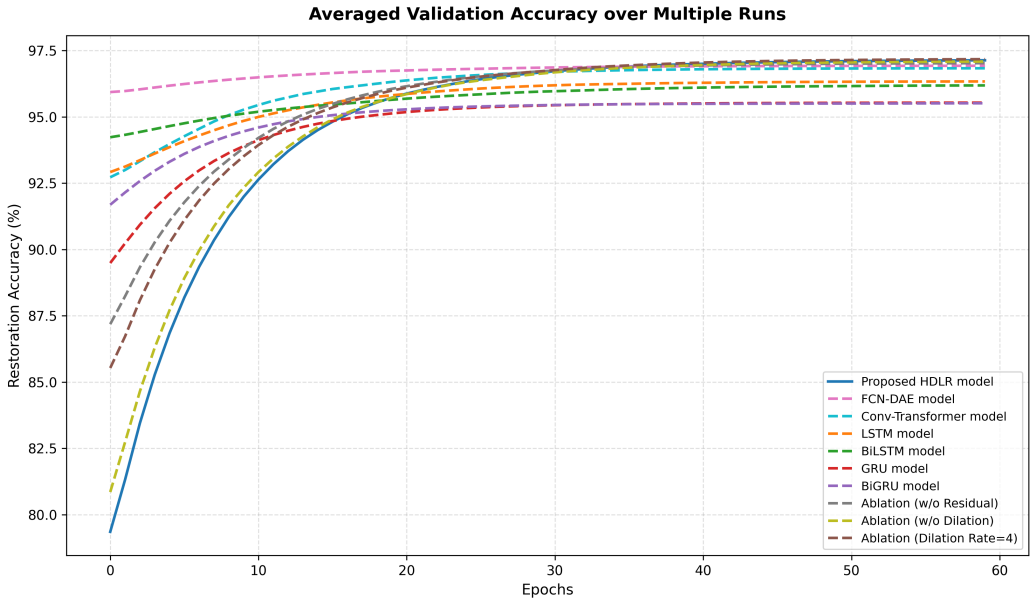


Figure 8. Averaged validation accuracy over multiple runs of proposed HDLR model compared with different models.

Furthermore, the GRU-based networks reflect these limitations even more pronouncedly through their learning rates. In particular, the BiGRU model tracks on the validation accuracy graph in Figure 8 demonstrate a severe performance dip in the earliest epochs by forming a deep concave valley before gradually recovering to converge at a flat level below 95.5%. The systemic behavioral patterns and clear stratification of the curves across all four plots provide an explicit visual proof of the performance margins, highlighted by their respective final endpoint values.

Table 1 presents a statistical evaluation of various deep learning architectures for ECG signal denoising and restoration over five independent runs. The performance of each model is quantified using six primary metrics: MSE, Root MSE (RMSE), coefficient of determination (R^2 score), post-denoising signal-to-noise ratio (SNR_{Post}) [20], SNR improvement (ΔSNR), and Matthew's correlation coefficient (MCC). All results are reported as mean $\pm$ standard deviation (SD) to demonstrate the stability and robustness of the models.

Table 1. Summary of statistical performance in form of (Mean $\pm$ SD).

Model	MSE	RMSE	R^2 Score	SNR_{Post} (dB)	ΔSNR (dB)	MCC
Proposed HDLR model	0.002176 ± 0.000074	0.046640 ± 0.000788	0.9640 ± 0.0012	14.4420 ± 0.1466	15.2787 ± 0.1466	0.9822 ± 0.0007
FCN-DAE model [18]	0.002395 ± 0.000131	0.048924 ± 0.001356	0.9604 ± 0.0022	14.0284 ± 0.2429	14.8650 ± 0.2429	0.9802 ± 0.0011
Conv-Transformer model [19]	0.002498 ± 0.000170	0.049958 ± 0.001741	0.9587 ± 0.0028	13.8484 ± 0.3098	14.6851 ± 0.3098	0.9793 ± 0.0015
LSTM model [7]	0.003124 ± 0.000111	0.055886 ± 0.000998	0.9483 ± 0.0018	12.8713 ± 0.1556	13.7079 ± 0.1556	0.9741 ± 0.0010
BiLSTM model [7]	0.002913 ± 0.000072	0.053968 ± 0.000670	0.9518 ± 0.0012	13.1739 ± 0.1077	14.0106 ± 0.1077	0.9758 ± 0.0007
GRU model [21]	0.003568 ± 0.000114	0.059728 ± 0.000958	0.9410 ± 0.0019	12.2935 ± 0.1399	13.1302 ± 0.1399	0.9703 ± 0.0009
BiGRU model [21]	0.003692 ± 0.000034	0.060762 ± 0.000277	0.9390 ± 0.0006	12.1436 ± 0.0395	12.9802 ± 0.0395	0.9693 ± 0.0003
Ablation (w/o Residual)	0.002139 ± 0.000046	0.046247 ± 0.000500	0.9646 ± 0.0008	14.5149 ± 0.0943	15.3516 ± 0.0943	0.9825 ± 0.0003
Ablation (w/o Dilation)	0.002191 ± 0.000078	0.046802 ± 0.000828	0.9638 ± 0.0013	14.4120 ± 0.1536	15.2487 ± 0.1536	0.9822 ± 0.0007
Ablation (Dilation Rate = 4)	0.002154 ± 0.000059	0.046406 ± 0.000636	0.9644 ± 0.0010	14.4854 ± 0.1182	15.3221 ± 0.1182	0.9824 ± 0.0006

The experimental results indicate that the proposed HDLR model exhibits a clear superiority over the baseline architectures as the FCN-DAE model yields an MSE of 0.002395 ± 0.000131, while the Conv-Transformer model obtains an MSE of 0.002498 ± 0.000170. In contrast, the proposed HDLR model significantly reduces the signal distortion by achieving the lowest MSE of 0.002176 ± 0.000074 and an RMSE of 0.046640 ± 0.000788. Compared to the conventional recurrent neural network (RNN) baselines, the proposed HDLR model demonstrates a superior restoration capability. Among the RNN baselines, the BiGRU model [21] exhibits the worst performance with the highest error rate (MSE: 0.003692 ± 0.000034), followed by the GRU model [21] (MSE: 0.003568 ± 0.000114), the LSTM model [7] (MSE: 0.003124 ± 0.000111), and the BiLSTM model [7] (MSE: 0.002913 ± 0.000072). Furthermore, regarding noise suppression, the proposed HDLR model achieves an outstanding SNR_{Post} of 14.4420 ± 0.1466 dB and a ΔSNR of 15.2787 ± 0.1466 dB, outperforming the FCN-DAE and Conv-Transformer models by approximately 0.41 dB and 0.59 dB, respectively. Crucially, the exceptionally high MCC [22] of 0.9822 ± 0.0007 and R^2 score of 0.9640 ± 0.0012 confirm that the proposed network preserves the ECG morphological features such as the QRS complex and R-peaks, without introducing artificial distortions.

The ablation study in Table 1 reveals the specific contributions of individual network components. By removing the dilation mechanism (Ablation w/o Dilation) leads to a noticeable decline in filtering performance with the MSE rising to 0.002191 ± 0.000078. This highlights of dilated convolutions in expanding the receptive field to capture broader temporal contexts. Under this simulation setup, both Ablation (w/o Residual) and Ablation (Dilation Rate = 4) yield excellent and highly competitive statistical results with MSE values of 0.002139 ± 0.000046 and 0.002154 ± 0.000059, respectively.

To assess these models on the resource-constrained mobile health devices and edge computing nodes, Table 2 evaluates their computational complexity and hardware efficiency. The evaluation is based on several key metrics as follows: the total number of model parameters (comprising all weights and biases), the parameter memory footprint (Param Memory) required for storage on system RAM or VRAM, and the activation memory (Activation Mem) at a batch size of 1 (BS = 1), which denotes the memory allocated for storing intermediate feature maps during the forward pass.

The combination of these two memory values defines the total runtime memory usage (Memory Usage), representing the minimum system resources required to execute the model. Additionally, the models are compared based on their inference time (Inf Time) per segment and training time (Train Time) per epoch, both reported along with their standard deviations to reflect operational stability and consistency.

Table 2. Computational complexity and hardware resource efficiency evaluation of proposed HDLR model compared with baseline networks.

Model Name	Parameters	Param Memory (MB)	Activation Mem (MB)	Memory Usage (MB)	Inf Time (ms)	Train Time (min)
Proposed HDLR model	100,033	0.3816	0.1958	0.5774	0.0873 ± 0.0296	11.4358 ± 0.2934
FCN-DAE model [18]	246,241	0.9393	0.0857	1.0250	0.2225 ± 0.1699	2.5031 ± 0.0225
Conv-Transformer model [19]	505,537	1.9285	0.3176	2.2461	0.9267 ± 0.4745	9.5628 ± 0.1746
LSTM model [7]	29,345	0.1119	0.0245	0.1364	0.0499 ± 0.0024	7.7239 ± 0.1393
BiLSTM model [7]	54,433	0.2076	0.0489	0.2565	0.0834 ± 0.0294	11.7299 ± 0.3141
GRU model [21]	22,305	0.0851	0.0245	0.1096	0.0501 ± 0.0114	7.7241 ± 0.3092
BiGRU model [21]	41,313	0.1576	0.0489	0.2065	0.0650 ± 0.0034	10.7259 ± 1.1857
Ablation (w/o Residual)	100,033	0.3816	0.1714	0.5530	0.1102 ± 0.0341	11.6103 ± 0.3540
Ablation (w/o Dilation)	100,033	0.3816	0.1958	0.5774	0.0751 ± 0.0145	11.6928 ± 0.1932
Ablation (Dilation Rate = 4)	100,033	0.3816	0.1958	0.5774	0.0776 ± 0.0137	11.5502 ± 0.1952

Based on these metrics, the proposed HDLR model stands out for its structural compactness. The network contains only 100,033 parameters, requiring a minimal parameter memory footprint of 0.3816 MB. At a batch size of 1, its total runtime memory usage is a mere 0.5774 MB. This total memory is 1.77-fold smaller than that of the FCN-DAE model (1.0250 MB total memory) and 3.89-fold smaller than the Conv-Transformer model (2.2461 MB total memory).

In terms of network parameters, the proposed HDLR model is 2.46 times and 5.05 times smaller than FCN-DAE (246,241 parameters) and Conv-Transformer (505,537 parameters), respectively. Although the RNN baselines like LSTM and GRU models demand less memory, they fail to achieve acceptable denoising accuracy, as demonstrated in Table 1. Crucial for real-time clinical deployment, the proposed HDLR model achieves an ultra-fast average inference time of 0.0873 ± 0.0296 ms per segment. Given an ECG sampling rate of 250 Hz, where new data points arrive every 4 ms, the execution time of the proposed model represents a minute fraction of this real-time constraint, thereby guaranteeing seamless, latency-free processing.

Moreover, hardware comparisons across the ablation variants uncover an essential architectural insight: the proposed HDLR model operates significantly faster than the Ablation (w/o Residual) variant (0.1102 ± 0.0341 ms). This empirically proves that the integration of shortcut residual connections not only enhances statistical gradient flow but also optimizes the runtime computational graph during inference.

Finally, regarding the cost–benefit analysis of the dilation mechanism, both Ablation (w/o Dilation) and Ablation (Dilation Rate = 4) share the exact same parameter count (100,033) and memory usage (0.5774 MB) as the main HDLR architecture. This validates that the proposed dilated convolution expands contextual awareness and improves signal filtering without imposing any computational overhead or hardware resource penalties.

For performance evaluation, ECG signals from record 100 of the standard MIT-BIH arrhythmia database [15] were resampled to 250 Hz to serve as the test dataset. This evaluation was conducted under an environment to eliminate inverse scaling bias by yielding the mean and standard deviation metrics, as presented in Table 3. The experimental results demonstrate that the proposed HDLR model significantly outperforms all other comparative models across all evaluated dimensions. Specifically, it achieves the lowest cumulative error with a MSE of 0.001458 and reconstructs the morphological structure of the

ECG signals with the highest accuracy, as indicated by an R^2 score of 0.8772. Furthermore, the proposed HDLR model delivers the most substantial signal quality improvement by yielding an output SNR improvement (Δ SNR) of 8.8582 dB. Additionally, the proposed HDLR model exhibits outstanding stability and robustness demonstrated by having the narrowest standard deviation among all evaluated architectures.

Table 3. Denoising performance evaluation using MIT-BIH database (record 100).

Model	MIT_MSE ($\times 10^{-3}$)	MIT_RMSE	MIT_R^2	MIT_Δ SNR (dB)
Proposed HDLR model	**1.458 ± 0.153**	**0.038151 ± 0.001988**	**0.8772 ± 0.0129**	**8.8582 ± 0.4505**
Ablation (Dilation Rate = 4)	1.554 ± 0.184	0.039387 ± 0.002340	0.8691 ± 0.0155	8.5830 ± 0.5401
Ablation (w/o Residual)	1.673 ± 0.229	0.040859 ± 0.002821	0.8590 ± 0.0193	8.2435 ± 0.6724
Ablation (w/o Dilation)	1.905 ± 0.354	0.043537 ± 0.004068	0.8395 ± 0.0298	7.6896 ± 0.9524
BiLSTM model [7]	2.354 ± 0.384	0.048406 ± 0.003949	0.8018 ± 0.0334	7.0255 ± 0.6401
BiGRU model [21]	2.502 ± 0.354	0.049925 ± 0.003504	0.7893 ± 0.0307	6.7909 ± 0.5901
FCN-DAE model [18]	2.673 ± 0.329	0.051638 ± 0.003238	0.7748 ± 0.0277	6.2328 ± 0.5561
Conv-Transformer model [19]	2.695 ± 0.285	0.051865 ± 0.002713	0.7730 ± 0.0240	6.1909 ± 0.4476
GRU model [21]	2.768 ± 0.540	0.052402 ± 0.005041	0.7669 ± 0.0456	6.1360 ± 0.8385
LSTM model [7]	2.905 ± 0.654	0.053655 ± 0.006272	0.7553 ± 0.0551	5.9296 ± 1.0524

The ablation study results that strongly support the architectural design choices of proposed HDLR model. In particular, disabling the receptive field expansion capability by removing the dilation rate in the convolutional layers (*Ablation (w/o Dilation)*) leads to the most performance drop among the ablation configurations. In this case, the R^2 score decreases to 0.8395 and the Δ SNR drops to 7.6896 dB which confirms that dilated convolutions play a critical role in capturing the global characteristics of ECG waveforms. However, excessively expanding the receptive field, as seen in the *Ablation (Dilation Rate = 4)* configuration, also degrades the performance slightly compared to the main model. Meanwhile, omitting the shortcut connections in the *Ablation (w/o Residual)* configuration reduces the denoising quality that the residual connection is essential for facilitating stable information flow and gradient propagation across the processing layers.

Compared to the standard RNN baselines, BiLSTM and BiGRU models outperform their unidirectional counterparts as LSTM and GRU models, but still fall short of the proposed HDLR model and all ablation variants. This gap is primarily due to the inherent limitations of traditional recurrent architectures in extracting localized structural and frequency features, where models integrated with convolutional layers hold a distinct advantage. Similarly, baseline models including the FCN-DAE and the Conv-Transformer models that achieve moderate performance with R^2 scores of approximately 0.77 and Δ SNR values of approximately 6.2 dB.

These results suggest that while the attention mechanism in the Conv-Transformer model is highly effective at capturing long-range context and the sample-level ECG denoising relies heavily on a balanced combination of localized local features and temporal sequences. The proposed HDLR model addresses this requirement optimally, proving to be the most effective approach for this task.

The methodology commences with the acquisition of a comprehensive arrhythmia dataset [16,17] comprising 200,000 ECG segments to encompass a diverse range of cardiac morphologies. Prior to the learning phase, rigorous data sanitization is performed using a Butterworth bandpass filter [23] to eliminate extraneous noise by addressing low frequency baseline wander and high frequency powerline interference as presented in Figure 9. This ensures the preservation of essential clinical components critical for diagnosis. The filtered signals are subsequently partitioned into three subsets 80% (approximately

160,000 samples) for training, 10% for validation, and 10% as an independent test set for final performance evaluation.

Figure 10 shows the high-precision signal recovery capability of the proposed HDLR model compared to the conventional Butterworth filter. The top waveform shows the reconstructed signal generated by the HDLR model, while the bottom waveform displays the reference signal.

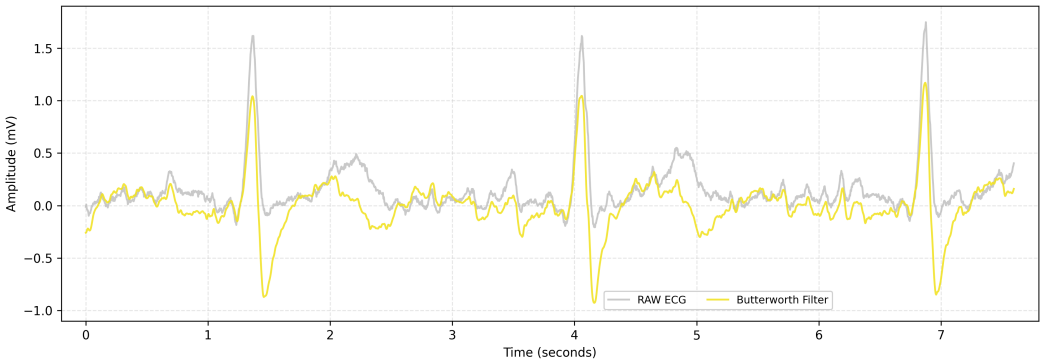


Figure 9. Denoising using Butterworth filter compared with raw ECG signal.

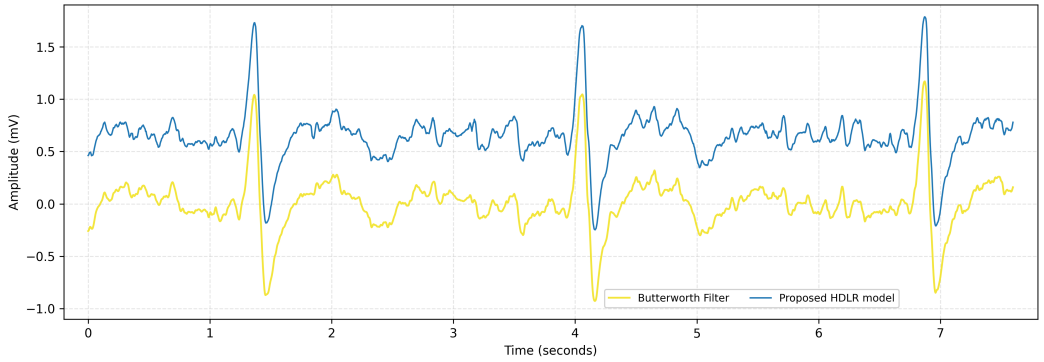


Figure 10. Comparison of reconstructed ECG waveforms between proposed HDLR model and Butterworth filter output.

Visual inspection reveals that the proposed model successfully recovers key morphological features of the ECG signal, including the P-wave, QRS complex, and T-wave [24] with high accuracy. Importantly, the proposed HDLR model demonstrates the ability to remove complex noise and artifacts while maintaining the structural integrity of the cardiac cycle, resulting in outcomes highly consistent with conventional signal processing methods.

A comparative analysis between the proposed HDLR model's output and raw ECG data is presented in Figure 11. The upper blue waveform showcases the successful denoising and signal restoration achieved by the proposed HDLR architecture, while the lower gray waveform represents the original data heavily obscured by noise. The results indicate that the proposed HDLR model can autonomously filter out significant interference, yielding a high-fidelity output that facilitates more precise clinical observation.

A visualization of the reconstructed ECG signals from multiple denoising approaches is presented in Figure 12 by showcasing the superior performance of the proposed HDLR model represented by the dark blue line. The middle gray waveform exhibits the original raw ECG data heavily obscured by noise, serving as the common input. The yellow line represents the Butterworth bandpass filter output used as the standard reference for successful denoising.

In the comparison among deep learning architectures in Figure 12, the proposed HDLR model generates a purified waveform that is most congruent with the Butterworth reference,

demonstrating optimal noise suppression and feature preservation. Although baseline models such as LSTM, BiLSTM, GRU, and BiGRU, along with competitive architectures like the FCN-DAE and Conv-Transformer models effectively reduce a significant portion of the raw interference, they still exhibit varying degrees of signal distortion, phase shifts, and baseline fluctuations.

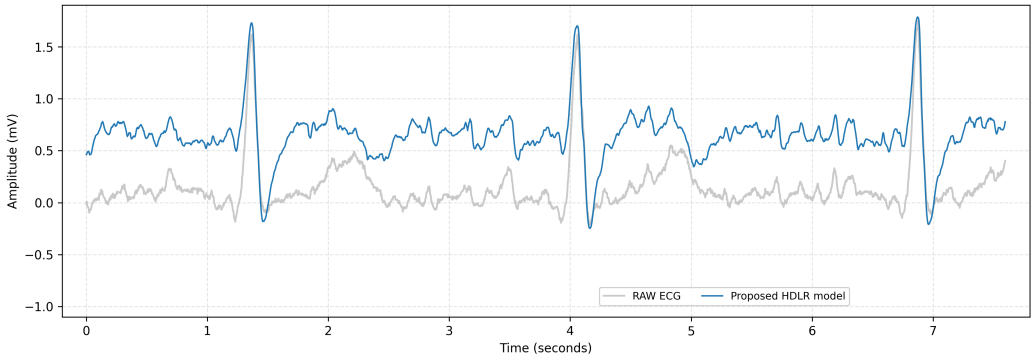


Figure 11. Comparison of reconstructed ECG waveforms between proposed HDLR model and raw ECG signal.

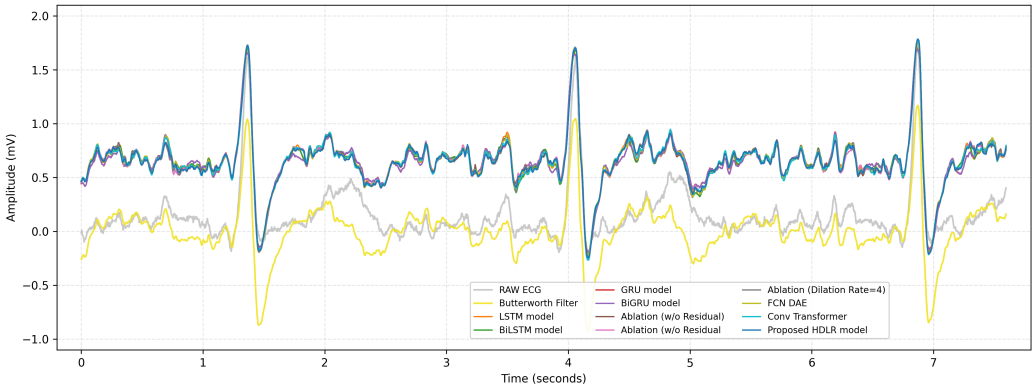


Figure 12. Comparison of reconstructed ECG waveforms between all baseline models and Butterworth filter output.

Furthermore, the results of the ablation studies, specifically the variants without residual connections (*Ablation w/o Residual*), without dilation (*Ablation w/o Dilation*), and with an altered dilation rate (*Ablation Dilation Rate* = *4*), clearly underperform compared to the complete architecture. This underscores the necessity of each integrated component in the proposed framework. Importantly, the proposed HDLR model maintains higher fidelity in recovering the critical P-wave and T-wave components, which are essential for clinical diagnosis, thereby confirming its robustness in handling real-world and non-sanitized ECG data.

5. Conclusions

This paper presents an efficient HDLR model for ECG noise reduction from cardiac biosensors. The proposed HDLR architecture integrates three core functional components: a dilated convolutional layer for multi-scale local feature extraction with an expanded receptive field, an LSTM layer for capturing long-term temporal dependencies, and a residual connection to mitigate the vanishing gradient problem while accelerating convergence. The ECG dataset was partitioned into 80% for training, 10% for validation, and 10% for testing.

Performance evaluation reveals that the proposed HDLR model achieves an outstanding MSE of 0.002176, a RMSE of 0.046640, a coefficient of determination (R^2 score) of

0.9640, a post-denoising signal-to-noise ratio (SNR) of 14.4420 dB, and a MCC of 0.9822, significantly outperforming the baseline LSTM, BiLSTM, GRU, BiGRU, FCN-DAE, and Conv-Transformer models. Furthermore, the ablation study highlights the significance of each architectural component; while simpler ablation configurations, such as those without dilated layers or residual connections, yield closely competitive reconstruction metrics, they suffer from degraded training stability in loss trajectories. The proposed HDLR framework consistently exhibits the most stable convergence behavior and robust generalization.

Beyond quantitative superiority, the proposed HDLR model demonstrates exceptional morphological preservation of critical ECG features, such as the P-wave and QRS complex, even under heavy and stochastic interference. These results suggest that the proposed hybrid HDLR model provides a robust and reliable solution for real-time cardiac monitoring systems, potentially enhancing the accuracy of automated arrhythmia detection and clinical diagnosis in wearable healthcare devices.

Author Contributions: Conceptualization, S.S., P.S. and T.W.; methodology, S.S., P.S. and T.W.; formal analysis, S.S., P.S. and T.W.; writing—original draft preparation, S.S., P.S. and T.W.; writing—review and editing, S.S., P.S. and T.W.; visualization, S.S. and P.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 data presented in this study are available in [15].

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

Abbreviations

The following abbreviations are used in this manuscript:

BiGRU	Bidirectional Gated Recurrent Unit
CNN	Convolutional Neural Network
ECG	Electrocardiogram
GRU	Gated Recurrent Unit
HDLR	Hybrid Dilated–Long Short-Term Memory with Residual Learning
LSTM	Long Short-Term Memory
MCC	Matthews Correlation Coefficient
MSE	Mean Squared Error
RMSE	Root Mean Squared Error
SNR	Signal-to-Noise Ratio

References

1. Arsene, C.T.; Hankins, R.; Yin, H. Deep learning models for denoising ECG signals. In Proceedings of the IEEE European Signal Processing Conference (EUSIPCO), A Coruña, Spain, 2–6 September 2019; pp. 1–5.
2. Ji, W.; Zhu, D. ECG classification exercise health analysis algorithm based on GRU and convolutional neural network. *IEEE Access* **2024**, *12*, 59842–59850. [CrossRef]
3. Ma, P.; Jiao, Z.; Wu, Z. Robust time-series InSAR deformation monitoring by integrating variational mode decomposition and gated recurrent units. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2024**, *18*, 3208–3221.
4. de Vries, I.R.; de Jong, A.M.; van der Hout-van der Jagt, M.B.; van Laar, J.O.E.H.; Vullings, R. Deep learning for sparse domain kalman filtering with applications on ecg denoising and motility estimation. *IEEE Trans. Biomed. Eng.* **2024**, *71*, 2321–2329. [CrossRef] [PubMed]
5. Wang, D.; Xu, M.; Guangming, Z.; Chen, Y. Enhancing wind power forecasting accuracy through LSTM with adaptive wind speed calibration (C-LSTM). *Sci. Rep.* **2025**, *15*, *5352*. [CrossRef] [PubMed]

6. Lynn, H.M.; Pan, S.B.; Kim, P. A Deep Bidirectional GRU Network Model for Biometric Electrocardiogram Classification Based on Recurrent Neural Networks. *IEEE Access* **2019**, *7*, 145395–145405. [CrossRef]
7. Cao, X.; Zheng, S.; Zhang, J.; Chen, W.; Du, G. A hybrid CNN-Bi-LSTM model with feature fusion for accurate epilepsy seizure detection. *BMC Med. Inform. Decis. Mak.* **2025**, *25*, *1*. [CrossRef] [PubMed]
8. Prathom, C.; Sato, R.; Watanabe, S.; Kondo, S.; Sato, K.; Okada, Y. A Lightweight Multiscale Deep Learning Framework for Automated Cardiovascular Disease Classification from Standard 12-Lead ECG Images. *Technologies* **2026**, *14*, *326*. [CrossRef]
9. Prathom, C.; Fukuda, R.; Yokoyanagi, Y.; Okada, Y. Preprocessing-Free Convolutional Neural Network Model for Arrhythmia Classification Using ECG Images. *Technologies* **2025**, *13*, *128*. [CrossRef]
10. 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]
11. Hassani, I.K.; Pellegrini, T.; Masquelier, T. Dilated convolution with learnable spacings. *arXiv* **2021**, arXiv:2112.03740.
12. He, J.; Wu, P.; Tong, Y.; Zhang, X.; Lei, M.; Gao, J. Bearing Fault Diagnosis via Improved One-Dimensional Multi-Scale Dilated CNN. *Sensors* **2021**, *21*, 7319. [CrossRef] [PubMed]
13. Dudek, G.; Pelka, P.; Smyl, S. A Hybrid Residual Dilated LSTM and Exponential Smoothing Model for Midterm Electric Load Forecasting. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**, *33*, 2879–2891. [CrossRef] [PubMed]
14. Szegedy, C.; Ioffe, S.; Vanhoucke, V.; Alemi, A. Inception-v4, inception-resnet and the impact of residual connections on learning. In Proceedings of the AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 4–9 February 2017; Volume 31, p. 1.
15. Moody, G.; Mark, R. MIT-BIH Arrhythmia Database. Available online: https://physionet.org/content/mitdb/1.0.0/ (accessed on 15 July 2026).
16. Goldberger, A.; Amaral, L.; Glass, L.; Hausdorff, J.; Ivanov, P.C.; Mark, R.; Mietus, J.E.; Moody, G.B.; Peng, C.K.; Stanley, H.E. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation* **2000**, *101*, e215–e220. [CrossRef] [PubMed]
17. Moody, G.B.; Mark, R.G. The impact of the MIT-BIH Arrhythmia Database. *IEEE Eng. Med. Biol.* **2001**, *20*, 45–50. [CrossRef]
18. Maurya, J.P.; Manoria, M.; Joshi, S. A proposed deep learning model for multichannel ECG noise reduction. *Discov. Artif. Intell.* **2025**, *5*, *65*. [CrossRef]
19. Ding, Y.; Dang, H.; Luo, J.; Zhuo, X.; Huang, N.; Xiao, J.; Lv, Z. Exploring a Novel Conv-Transformer Network for Multi-Modality Heart Segmentation. *Iradiology* **2026**, *4*, 13–22.
20. Guer, I.; Ubeyli, E.D. Feature saliency using signal-to-noise ratios in automated diagnostic systems developed for ECG beats. *Expert Syst. Appl.* **2005**, *28*, 295–304.
21. Wang, X.; Nourmohammadi, S. A novel framework for sentiment classification employing Bi-GRU optimized by enhanced human evolutionary optimization algorithm. *Sci. Rep.* **2025**, *15*, 17038. [CrossRef] [PubMed]
22. Chicco, D.; Jurman, G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genom.* **2020**, *21*, 6. [CrossRef]
23. Hussin, S.F.; Birasamy, G.; Hamid, Z. Design of butterworth band-pass filter. *Politek. Kolej Komuniti J. Eng. Technol.* **2016**, *1*, 32–46.
24. Wijaya, C.; Crispin, A.R.; Harahap, M.; Christnatalis, C.; Turnip, M.; Turnip, A. Abnormalities state detection from P-wave, QRS complex, and T-wave in noisy ECG. *J. Phys. Conf. Ser.* **2019**, *1230*, 12–15. [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.