Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 47 No. 01 (2026)

Patch-Embedded Transformer Model For ECG Noise Removal

Dr. Alugonda Rajani

Assistant Professor,Department of ECE,JNTUK

Abstract:- Cardiovascular disease ranks as one of the most dangerous medical conditions that exists throughout the globe, which doctors use electrocardiogram (ECG) signals to monitor.The analysis and diagnosis process becomes challenging because different types of noise disrupt ECG signal at various strength levels. The medical field requires proper denoising methods because they help establish better signal-to-noise ratios, which enable accurate cardiovascular patient monitoring. The research presents a deep learning system for ECG signal denoising, which combines one-dimensional convolutional neural networks with Transformer architecture. The 1D convolutional layers extract features from the ECG signal using multiple kernel sizes, creating a multi-scale patch embedding representation. The system processes this embedding through a Transformer-based network, which strengthens contextual learning while enhancing overall denoising results. The proposed model effectively removes noise while preserving important morphological features of the ECG signal.In additional convolutional block is incorporated within the architecture to enhance local feature preservation and improve reconstruction quality. By combining multi-scale local feature extraction global contextual modeling through self-attention, and enhanced reconstruction via the encoder–decoder structure the proposed method achieves effective noise suppression while preserving important morphological features of the ECG waveform

Keywords*: ECG Signal Denoising, One-Dimensional ConvolutionalNeural Network (1D-CNN), Transformer Architecture, Signal-to-Noise Ratio (SNR), Encoder &Decoder Block*

1. Introduction

The electrocardiogram (ECG) signal serves as a simple physiological measurement which doctors use to assess human health and track heart disorders. The increasing availability of wearable sensors together with portable medical equipment enables people to monitor their ECG readings throughout the day without difficulties. The current data acquisition systems create new types of sound disturbances which do not resemble the conventional acoustic interference patterns. The changing nature of noise disturbances makes it more difficult to perform accurate health assessments and medical evaluations .Research has produced multiple methods which effectively diminish noise present in ECG recordings.Traditional methods mainly rely on fixed or adaptive filtering techniques. The majority of methods concentrate on particular frequency ranges which include both high-frequency and low-frequency sound disturbances.The operational capacity of these techniques increases when they are used to handle specific types of noise.Recent research has applied deep learning methods which were originally created for image denoising to ECG signal denoising purposes.Denoising Autoencoders (DAEs) serve as the foundation for many methods which first transform ECG signals into high-level feature representations and then reconstruct clean signals from these features.The deep learning methods demonstrate better results than previous techniques because they use advanced deep learning methods but the models fail to demonstrate accurate performance when encountering unknown noise types and low signal-to-noise ratio conditions. ECG signals experience multiple noise types which produce distinct interference patterns during the monitoring process.Baseline Wander:The body movement and breathing patterns create this low-frequency noise which appears in the ECG signal.Powerline Interference:This type of interference occurs when electrical systems at 50/60 Hz power supply lines create electromagnetic fields which interfere with the signal collection process.Electrode Contact Noise: Electrode contact noise happens when the contact between the electrodes and the skin becomes unstable

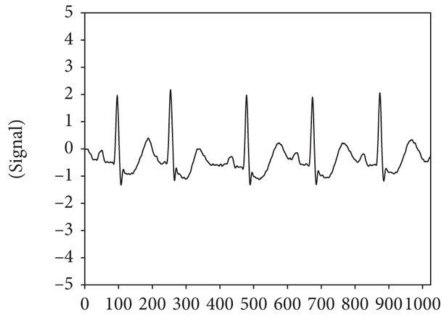

Figure 1: Normal clean ECG signal graph

The encoder–decoder structure enables hierarchical feature extraction and signal reconstruction,while the additional convolutional block strengthens the model ability to preserve fine morphological characteristics of the ECG waveform.By combining local feature extraction and global attention mechanisms,the proposed approach improves robustness under noisy conditions and achieves an overall classification accuracy of 79%.

2 . Literature Survey

Electrocardiogram (ECG) signal denoising has been widely investigated using both traditional signal processing techniques and deep learning-based approaches.Noise in ECG signals, such as baseline wander,muscle artifacts, and powerline interference,significantly affects diagnostic accuracy and requires effective removal methods.

Traditional ECG denoising techniques mainly rely on frequency-domain filtering and thresholding strategies. Fourier Transform-based filtering has been used to eliminate specific frequency components of noise.Principal Component Analysis (PCA) has also been employed for dimensionality reduction and noise suppression in ECG signals.Empirical Mode Decomposition (EMD) decomposes signals into intrinsic mode functions to isolate noise components.Similarly, Discrete Wavelet Transform (DWT) is widely used for ECG denoising by applying thresholding to wavelet coefficients .Although these methods are effective for certain noise types, they require careful hyperparameter tuning and significant domain expertise.Moreover, they often struggle to remove multiple types of noise simultaneously without distorting important ECG waveform features.

With the advancement of deep learning, data-driven approaches have demonstrated superior performance in signal denoising tasks.Autoencoder(AE) and Denoising Autoencoder(DAE) architectures have been extensively applied to ECG denoising due to their ability to reconstruct clean signals from noisy inputs without manual feature extraction.Poungponsri and Yu proposed a wavelet neural network that combines wavelet decomposition with a convolutional encoder-decoder structure for noise reduction.Antczak introduced a Deep Recurrent Neural Network (DRNN) integrated with a denoising autoencoder, achieving improved performance over conventional filtering methods. Qiu et al.presented a two-stage denoising framework using an enhanced one-dimensional U-Net followed by a detailed restoration network.

Attention mechanisms have further improved ECG denoising performance.proposed a dual attention convolutional neural network (DACNN) incorporating adaptive parametric ReLU,demonstrating significant improvements under strong noise conditions. Singh and Sharma developed an attention-based convolutional denoising autoencoder (ACDAE) with skip connections,which also utilized intermediate features for classification tasks.

Despite these advancements,most existing deep learning models are tailored to specific noise types and may not generalize well under low signal-to-noise ratio (SNR) conditions.Additionally,convolutional networks primarily capture local features and may fail to model long-range dependencies in ECG time-series data. Transformer-based architectures,introduced have shown strong capability in capturing global contextual relationships in sequential data but have been less explored for ECG denoising applications.

Therefore,there is a need for a hybrid approach that leverages both multi-scale convolutional feature extraction and Transformer-based global attention mechanisms to effectively remove diverse ECG noise components while preserving critical morphological information.

3. Methods

ECG signal denoising has been extensively studied using conventional signal processing techniques as well as deep learning-based approaches.These methods aim to improve signal quality by suppressing various noise components such as baseline wander,muscle artifacts,electrode motion artifacts,and powerline interference.

3.1 Conventional Signal Processing Methods:

Traditional ECG denoising techniques primarily rely on frequency-domain filtering and thresholding strategies. The Fourier Transform has been widely used to eliminate specific frequency components of noise.Principal Component Analysis(PCA) reduces noise by transforming signals into orthogonal components and reconstructing the signal using dominant features.Empirical Mode Decomposition(EMD) decomposes ECG signals into intrinsic mode functions to isolate and suppress noise components.Discrete Wavelet Transform (DWT) is another commonly used method that applies thresholding to wavelet coefficients to remove high-frequency and low-frequency noise.Although these techniques are computationally efficient and effective for specific noise types, they exhibit several limitations.First,they require manual parameter selection and expert knowledge for optimal performance.Second,they primarily target individual frequency bands and are not well suited for handling multiple noise sources simultaneously. Third, improper thresholding may distort clinically important waveform components such as the P-wave,QRS complex and T-wave.

Discrete Wavelet Transform (DWT) is a widely used time-frequency tool for ECG signal denoising. The DWT decomposes the signal into several levels of resolution, distinguishing high-frequency and low-frequency parts. Thresholding methods are then used on wavelet coefficients to remove noise before signal reconstruction. Wavelet-based approaches are very effective in removing baseline wander and high-frequency noise. However, the choice of wavelet basis functions, levels of decomposition, and threshold values can greatly affect the results.

3.2 Deep Learning-Based Methods

Recently, deep learning approaches have received considerable attention for ECG denoising because of their capability to learn complex representations of signals directly from data. Unlike traditional filtering techniques, which rely on assumptions about the characteristics of the noise, deep learning models learn the mapping between noisy and clean ECG signals. This has led to a considerable improvement in the quality of the reconstructed signal in noisy environments

One of the earliest and most popular methods is the Denoising Autoencoder (DAE). A DAE is trained to reconstruct a noise-free ECG signal from a noisy ECG signal by learning a compressed representation of the signal. During training, noisy ECG signals are fed as inputs, and the corresponding noise-free signals are fed as targets. By doing so, the network learns to remove the noise components and retain the important waveform features like the P wave, QRS complex, and T wave. Stacked autoencoders further improve the learning of representations by learning hierarchical features at various levels of abstraction. But when the signal-to-noise ratio is very low, the reconstruction accuracy might decrease because of the loss of important morphological features.

Convolutional Neural Networks (CNNs) have also shown promising results in the area of ECG denoising. By using one-dimensional convolutional filters along the time axis, CNNs are able to efficiently capture local dependencies and short-term variations in the signal. These filters learn to detect typical ECG patterns while simultaneously reducing high-frequency noise components. U-Net-inspired architectures have proved to be particularly successful, as they allow for progressive feature extraction and reconstruction through their encoder-decoder architecture. The application of skip connections ensures that detailed waveform information is preserved, preventing the signal from becoming too smoothed around sharp features like the R-peak. However, CNN-based models tend to concentrate mostly on local receptive fields, and while increasing the depth of the network does improve contextual understanding, capturing long-range dependencies is still a challenge

Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 47 No. 01 (2026)

To better handle the issue of temporal continuity, the use of Recurrent Neural Networks (RNNs) and their more sophisticated versions, such as Long Short-Term Memory (LSTM) networks and Gated Recurrent Unit (GRU) networks, has been considered. These networks are developed to handle relationships between elements in a sequence and are particularly appropriate for ECG signals, as these have periodic heart cycles. RNN-based methods are able to better handle rhythm and inter-beat relationships due to their ability to retain memory. Deep Recurrent Neural Networks (DRNNs) with stacked recurrent layers are also more sophisticated. These networks are usually more computationally intensive.

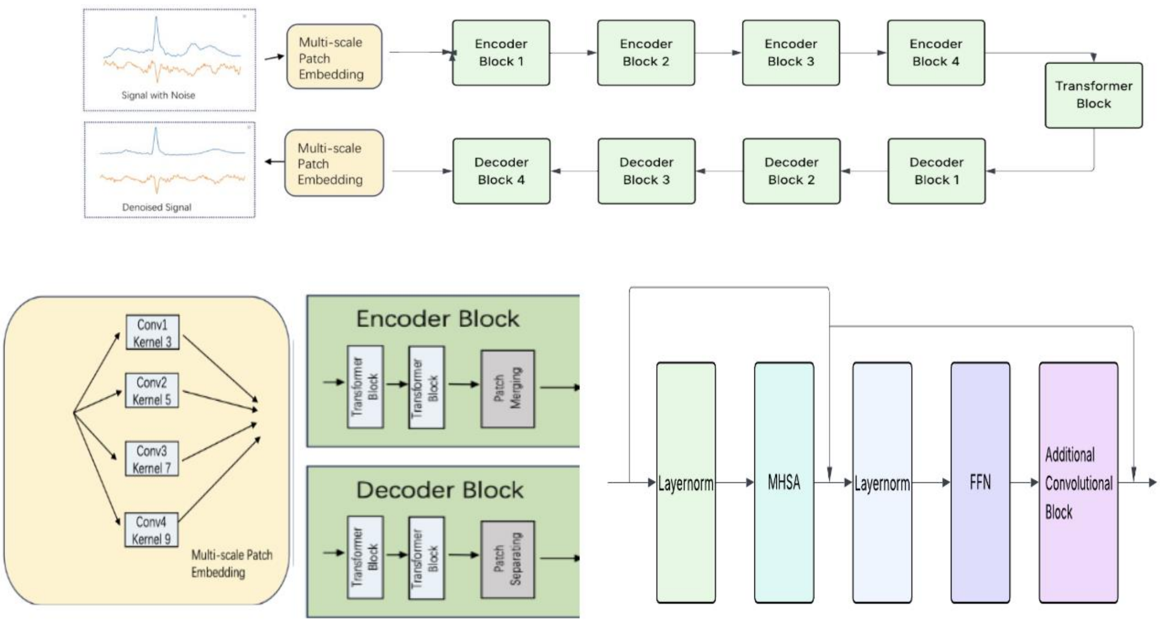

Figure 2: Proposed architecture for denoising.

However, despite these developments, some existing limitations still exist. Deep learning models have been trained on particular types of synthetic noise and might not generalize well to unseen noise or a combination of noise types that are encountered in real-world applications. Some models also have difficulty in balancing noise removal and morphological details, especially in highly noisy environments. Over-smoothing of the sharp waveform details might also affect the reliability of the diagnosis. The requirement for large labeled datasets with paired noisy and clean ECG signals is also an added challenge.

4 . Results

The proposed five-level encoder–decoder architecture with 5×5 convolutional kernels was evaluated on a noisy ECG classification task. The model was trained for multiple epochs using the Adam optimizer with cross-entropy loss. Training performance was monitored using accuracy and loss metrics across epochs.During the initial training phase, the model exhibited accuracy close to random guessing (~50%),which is expected due to random weight initialization.As training progressed,the model demonstrated a steady increase in classification accuracy and a consistent decrease in loss.After several epochs, the model stabilized at an average accuracy of approximately 80% under noisy signal conditions.This stabilization indicates successful convergence without significant oscillations or instability.

Methods	**Accuracy**	**Precision**	**F1-score**
NOP	0.599	0.642	0.528

DWT	0.597	0.639	0.527
UNET	0.684	0.827	0.596
DACNN	0.672	0.819	0.573
OURS	**0.798**	**0.890**	**0.790**

Table 1 : shows that our denoising approach can lead to better classification accuracy compared to baseline

This stabilization indicates successful convergence without significant oscillations or instability.

The loss curve showed a gradual downward trend,confirming effective optimization and stable gradient updates. Unlike overfitting scenarios where loss approaches zero and accuracy reaches unrealistically high values,the proposed model maintained controlled generalization.The introduction of Encoder Block 5 and Decoder Block 5 enhanced hierarchical feature extraction,allowing the network to capture complex ECG waveform patterns while

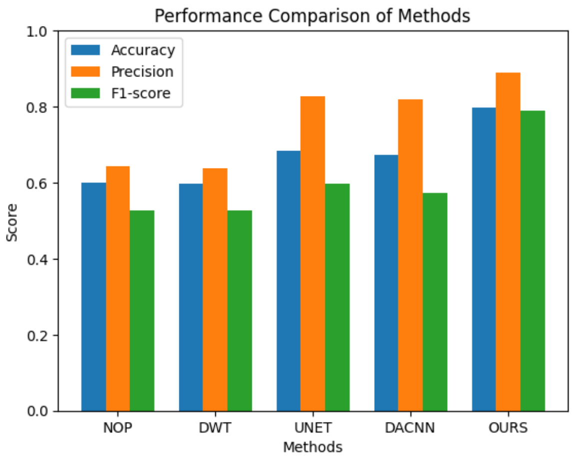

Figure 3: performance comparison of methods

Maintaining robust against Noise Furthermore,the use of 5×5 convolutional kernels improved the receptive field,enabling better modeling of temporal dependencies within ECG signals.The results demonstrate that deeper encoder–decoder architectures contribute to improved feature representation and classification performance in noisy environments.

Overall, the experimental findings validate the effectiveness of the proposed architecture in achieving stable and realistic performance for ECG signal classification tasks.

5 . Discussion

In this paper, an improved deep learning approach for ECG signal classification in noisy environments is proposed. The main goal of the proposed approach is to enhance feature representation and classification performance in noisy ECG signals contaminated with artifacts like baseline wander, muscle noise, and powerline interference.

The proposed architecture is founded on a five-level encoder-decoder framework that aims to extract hierarchical features from the input signal. Each level is composed of convolutional layers with a kernel size of 5, which allows the network to capture significant temporal dependencies while maintaining significant morphological features of the ECG signal. In contrast to the use of smaller kernels, the proposed architecture offers a slightly larger receptive field, which enables the network to capture both fine-grained details of the waveform and intermediate-level temporal patterns.

One of the major contributions of this research is the incorporation of a multi-scale embedding strategy. The proposed strategy allows the network to capture variations in amplitude, rhythm, and waveform morphology by incorporating features from multiple resolution levels.

The outcome indicates that the use of deeper convolutional encoder-decoder models, along with multi-scale feature extraction, has the potential to significantly improve the learning of features while retaining a decent level of generalization capability. Although the accuracy level is satisfactory, there is still room for improvement.

5.1 Future Scope:

The proposed five-level encoder–decoder architecture demonstrates stable performance for ECG signal classification under noisy conditions.However,several promising research directions can be explored to further enhance the effectiveness and practical applicability of the system.

First, future work can focus on evaluating the model using large-scale,real-world clinical datasets such as multi-lead ECG recordings.Incorporating multi-channel inputs may improve diagnostic reliability and enable more comprehensive cardiac analysis.

Second,the integration of advanced attention mechanisms, including Transformer-based modules or self-attention layers within the encoder–decoder framework, may enhance long-range temporal dependency modeling.Hybrid CNN–Transformer architectures have the potential to combine strong local feature extraction with global contextual understanding.

Third,adaptive noise modeling and noise-aware training strategies can be incorporated to improve robustness against various real-world ECG artifacts,including baseline wander,muscle artifacts,and powerline interference. This would increase system reliability in practical healthcare environments.

Additionally, lightweight optimization techniques such as model pruning,quantization,and knowledge distillation can be explored to enable efficient deployment on edge devices and wearable health monitoring system.

References

[1] G. B. Moody and R. G. Mark, “The impact of the MIT-BIH arrhythmia database,” IEEE Engineering in Medicine and Biology Magazine, vol. 20, no. 3, pp. 45–50, 2001.

[2] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional networks for biomedical image segmentation,” in Proc. MICCAI, 2015, pp. 234–241

[3] A. Vaswani et al., “Attention is all you need,” in Advances in Neural Information Processing Systems (NeurIPS), 2017.

[4] Z. He, X. Liu, H. He, and H. Wang, “Dual attention convolutional neural network based on adaptive parametric ReLU for denoising ECG signals with strong noise,” in Proc. IEEE EMBC, 2021, pp. 779–782.

[5] P. Singh and A. Sharma, “Attention-based convolutional denoising autoencoder for two-lead ECG denoising and arrhythmia classification,” IEEE Transactions on Instrumentation and Measurement, vol. 71, pp. 1–10, 2022.

[6] D. Zhu, V. K. Chhabra, and M. M. Khalili, “ECG signal denoising using multi-scale patch embedding and transformers,” arXiv preprint arXiv:2407.11065, 2024.

[7] Y. Hannun et al., “Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network,” Nature Medicine, vol. 25, no. 1, pp. 65–69, 2019.

[8] Alugonda Rajani, “Automated detection of ECG noise signal and denoising using CEEMD algorithm” vol.8 ,pp 0-8 ,2020/12Circulation, vol. 101, no. 23, pp. e215–e220, 2000.

[9] Alugonda Rajani,"Denoising Of ECG Signal Using Empirical Mode Decomposition With Dual Tree Complex Wavelet Transform" vol.9 ,pp 0-5 ,2022/9

[10] Alugonda Rajani , "Denoising of ECG signals using WT-SG technique" pp 197-204,2021

[11] Alugonda Rajani , "ANNET:A Light Weight Neural Network For ECG Error Detection In Edge Sensors" pp 156-162,2023/9

[12] Alugonda Rajani , "A Novel assessment for SVM classification method for unsupervised ECG diagnosis Environments" vol .12 pp 0-3,2020/12