A Reinforcement Learning Framework for Electrocardiogram Baseline Wander Removal

Masoud Nateghi

Abstract—**Baseline wander is a common low-frequency noise in electrocardiogram (ECG) signals, often caused by respiration, body movements, and poor electrode contact. This noise can obscure critical diagnostic features and complicate both manual and automated analyses. Traditional fixed-parameter filtering methods struggle to balance effective noise removal with the preservation of vital signal components. To address this limitation, we propose a reinforcement learning (RL)-based adaptive filtering framework that dynamically adjusts the cutoff frequency of a low-pass filter in response to the characteristics of the noisy ECG signal. The RL agent, utilizing a Long Short-Term Memory (LSTM) network and trained with the Proximal Policy Optimization (PPO) algorithm, was evaluated using the PTB Diagnostic ECG Database and the MIT-BIH Noise Stress Test Database.**

While the framework demonstrated adaptability in selecting cutoff frequencies and reducing baseline wander, its performance in terms of MSE and signal-to-noise ratio (SNR) did not surpass that of traditional fixed-parameter filters. Analysis suggests that the simplicity of the reward function, based solely on the negative MSE, limited the framework's effectiveness in preserving diagnostic features. Despite these challenges, the results highlight the potential of RL-based approaches for adaptive signal filtering. Future work will focus on incorporating more sophisticated reward functions, alternative architectures, and multi-objective optimization to enhance both noise removal and clinical relevance. This study lays the foundation for adaptive, data-driven solutions in ECG signal processing.

I. Introduction

Electrocardiography (ECG) is a cornerstone in cardiac diagnostics, offering crucial insights into heart health through its depiction of electrical activity. However, the utility of ECG signals often suffers due to contamination from various noise sources such as power line interference, muscle artifacts, and motion artifacts. Among these, baseline wander—a low-frequency noise primarily caused by patient respiration, body movements, and poor electrode contact—presents a unique challenge. As shown in Figure 1, baseline wander can obscure diagnostically significant features like P waves, QRS complexes, and T waves, increasing the risk of misdiagnoses when left untreated.

The authors are with the Department of Biomedical Informatics, Emory University (email: masoud.nateghi@emory.edu).

Baseline wander occupies the very low-frequency spectrum, typically below 0.5 Hz, but partially overlaps with vital low-frequency components of ECG signals, such as those found in the ST segment and QT interval. Conventional filtering methods, such as fixed low-pass and high-pass filters, often struggle to balance effective noise removal with the preservation of diagnostic features, potentially leading to distorted clinical markers. This trade-off is evident in Figures 1 and 2, which illustrate that while baseline wander is removed, certain low-frequency components of the signal are also attenuated. These observations highlight the necessity for advanced methods capable of adaptive and context-aware noise removal.

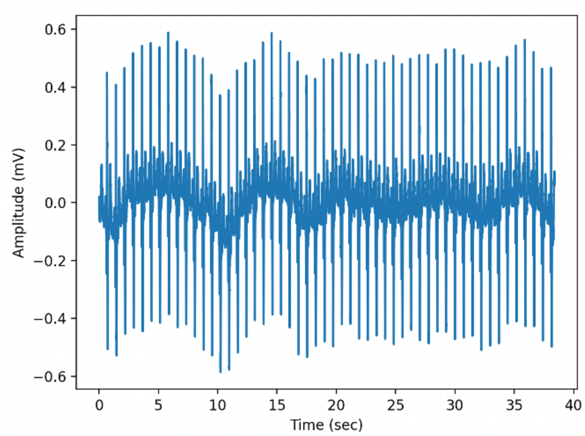

Fig. 1. A sample record of the noisy ECG signal contaminated by baseline wander.

In this paper, we propose a reinforcement learning (RL)-based adaptive filtering framework designed to address the limitations of traditional denoising methods. The RL agent dynamically adjusts filter parameters in real-time, ensuring optimal noise removal while safeguarding critical ECG components. By leveraging model-free RL approaches like Proximal Policy Optimization (PPO) and an LSTM-based agent architecture, this method provides a flexible, data-driven solution to ECG baseline wander removal.

We evaluate the performance of the proposed framework using benchmark datasets, including the PTB Di-

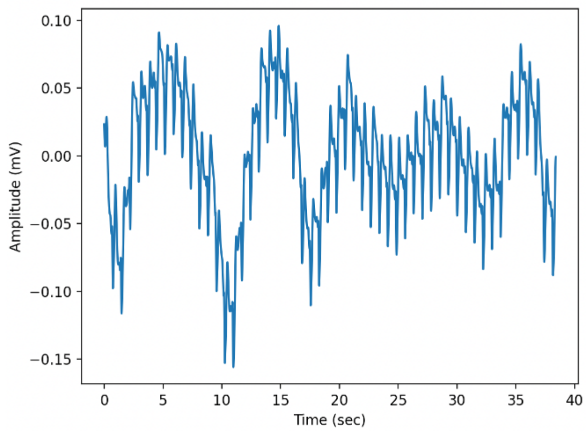

Fig. 2. The noise component obtained after low-pass filtering with a cutoff frequency of 1 Hz. Note that the filtered noise contains components of the ECG signal, highlighting the trade-off between effective noise removal and the preservation of diagnostically important features.

agnostic ECG Database [1] and the MIT-BIH [2] Noise Stress Test Database, which offer a variety of ECG recordings and noise conditions. The results demonstrate the efficacy of our approach in achieving high signal quality and preserving diagnostic integrity.

This work aims to advance the state of ECG signal processing by introducing an adaptive, RL-based paradigm, paving the way for more accurate and reliable cardiac diagnostics in both clinical and remote monitoring settings.

II. Dataset

To evaluate the proposed RL framework for ECG baseline wander removal, two benchmark datasets were utilized: the PTB Diagnostic ECG Database and the MIT-BIH Noise Stress Test Database (NSTDB). These datasets provide diverse ECG signals and noise conditions, enabling a robust evaluation of the proposed approach.

A. PTB Diagnostic ECG Database

The PTB Diagnostic ECG Database is a comprehensive and widely used dataset in the domain of cardiac signal analysis. It contains 549 ECG recordings from 290 subjects, each record spanning over 30 seconds and digitized at a high resolution of 1000 Hz. The dataset includes signals from 15 leads: 12 standard ECG leads and 3 Frank leads, ensuring a detailed representation of cardiac activity.

This dataset is notable for its diverse range of cardiac conditions, including both normal and abnormal cases, such as myocardial infarction, bundle branch block, and arrhythmias. The high sampling rate and rich annotations make it particularly suitable for tasks involving signal denoising and feature preservation. By leveraging this dataset, the proposed approach can be validated across a variety of realistic clinical scenarios.

B. MIT-BIH Noise Stress Test Database

The MIT-BIH Noise Stress Test Database complements the PTB dataset by providing realistic noise profiles for ECG recordings. This database includes three primary noise types: baseline wander, muscle artifact, and electrode motion artifact, all of which are common in real-world ECG data. The noise signals are sampled at 360 Hz and are designed to emulate the conditions under which clinical ECGs are often recorded.

NSTDB is critical for testing the robustness of denoising algorithms under challenging noise conditions. In this study, the baseline wander noise is particularly emphasized, aligning with the focus on removing low-frequency noise while preserving diagnostically relevant features of the ECG signal.

By combining the PTB Diagnostic ECG Database and the NSTDB, the evaluation framework ensures that the proposed RL approach is tested on both diverse cardiac conditions and realistic noise scenarios. This comprehensive assessment underscores the method's potential for clinical applicability and reliability in real-world settings.

III. RL Framework

The design of the RL framework is inspired by the approach proposed in [3], which demonstrates effective adaptation in noisy signal environments. This section details the methodology, including the agent design, states, actions, and reward functions, as illustrated in the agent-environment interaction diagram (Figure 3).

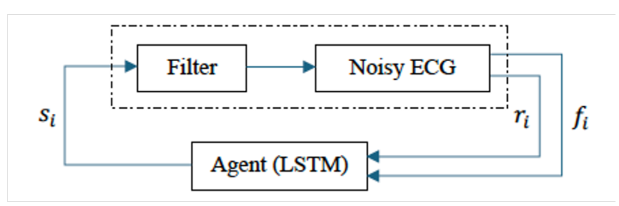

Fig. 3. The reinforcement learning framework for adaptive ECG baseline wander removal. The agent, implemented as a Long Short-Term Memory (LSTM) network, interacts with the environment comprising the noisy ECG signal and the filter. At each segment, the agent observes the noisy signal segment f_i, selects the filter parameter s_i, and receives a reward r_i based on the filtering performance. This interaction allows the agent to iteratively optimize the filter parameters for noise removal.

A. Agent Design

The RL agent is built on a Long Short-Term Memory (LSTM) architecture, enabling it to process sequential ECG data effectively. The temporal nature of the ECG signal makes LSTM an ideal choice for capturing dependencies across successive frames. The agent uses two networks: a policy network, which generates filter parameters, and a value network, which estimates the expected reward for a given state.

The training process leverages the Proximal Policy Optimization (PPO) algorithm [4], which ensures stability and efficiency in policy updates. This algorithm is particularly suited for continuous action spaces, allowing the agent to make fine adjustments to the filter parameters dynamically.

B. States

In the RL framework, the states represent segments of the noisy ECG signal. Each state is defined as a combination of the clean signal and baseline wander noise, represented mathematically as:

$$f_i = g_i + b_i \tag{1}$$

where g_i is the clean ground truth ECG signal, b_i is the baseline wander component, and f_i is the noisy observation at i-th segment. The state transitions occur as the agent processes consecutive segments of the ECG signal.

C. Actions

The action space consists of the filter parameters that the RL agent determines for each segment of the noisy ECG. Specifically, the agent generates a parameter s_i, which defines the cutoff frequency for a simple low-pass filter from the Open Source Electrophysiological Toolbox (OSET) package (lp_filter_zero_phase.m) [5]. This parameter is then used to separate the baseline wander component (b_i) from the clean signal (g_i) as follows:

$$b_i = \text{LPF}(f_i, s_i) \tag{2}$$

$$\hat{g}_i = f_i - b_i \tag{3}$$

The dynamic nature of this action space allows the agent to adaptively tailor the filtering process to the characteristics of each ECG segment. The overall interaction between the agent and the environment is depicted in Figure 4, highlighting the relationship between states, actions, rewards, and the filtering algorithm.

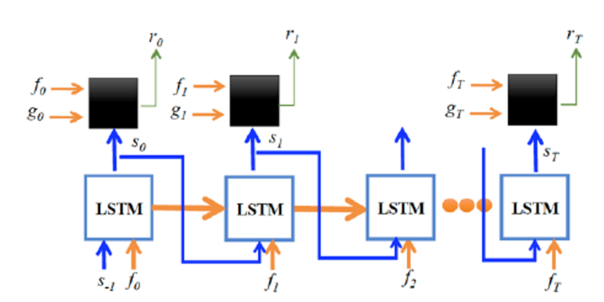

Fig. 4. The reinforcement learning agent architecture for adaptive ECG baseline wander removal [3]. The agent, modeled as a sequence of Long Short-Term Memory (LSTM) units, processes noisy ECG segments (f_i) and outputs filter parameters (s_i) for each segment. The agent receives feedback in the form of rewards (r_i), which are calculated based on the difference between the ground-truth clean signal (g_i) and the denoised signal. This iterative framework allows the agent to optimize filtering decisions over time, preserving important diagnostic features while removing noise.

D. Reward

The reward function is based solely on the negative MSE between the ground-truth clean signal (g_i) and the denoised signal ($\hat{g}_i$) for each segment. This ensures the RL agent prioritizes minimizing the difference between the clean and denoised signals. The reward is calculated as:

$$r_i = -\|g_i - \hat{g}_i\|^2 \tag{4}$$

This simple yet effective formulation drives the agent to focus on achieving the best possible approximation of the clean ECG signal, ensuring that baseline wander is removed while preserving the signal's diagnostic features.

IV. Results

The performance of the RL-based adaptive filtering framework was assessed using key metrics, including signal-to-noise ratio (SNR) and reward values. The results demonstrate the framework's ability to address baseline wander noise while preserving critical diagnostic features of the ECG signal.

Reward values, based on the negative MSE, showed a steady increase over the training episodes, as illustrated in Figure 5. This progression highlights the agent's learning capabilities and its ability to dynamically optimize filter parameters for improved denoising. The rising reward values reflect the adaptability of the RL-based system and its success in managing the complexities of real-time filtering tasks.

The SNR metric validated the quality of the denoised signals. As presented in Figure 6, the proposed method achieved a substantial improvement in SNR from the initial SNR of 10 dB. This improvement underscores the framework's ability to enhance the clarity of ECG signals while safeguarding clinically significant features.

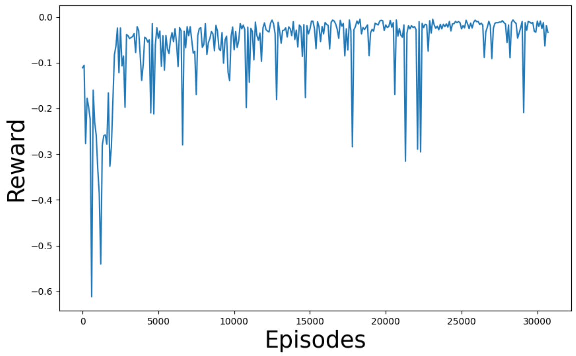

Fig. 5. Average reward accumulated over the course of episodes by agent. This plot shows an overall increase in the reward values which denotes the convergence of the algorithm.

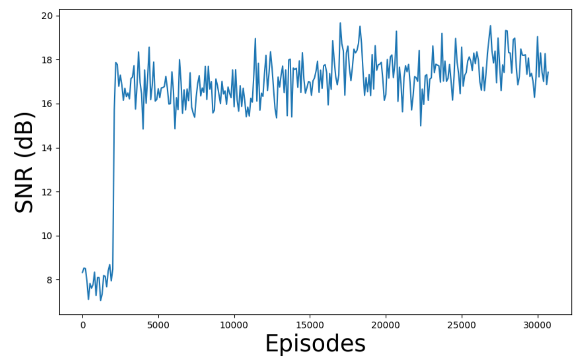

Fig. 6. Average SNR of the signals over the course of training episodes. Note the increasing trend in SNR and the significant improvement from an initial 10 dB SNR of noisy signals.

Additionally, Figure 7 showcases the performance of the framework on a representative ECG record contaminated by baseline wander artifacts. The denoised signal clearly demonstrates the system's robustness in addressing low-frequency noise without distorting diagnostically critical components.

Overall, the results affirm the adaptability of the RL-based framework. The steady improvement in reward values, significant enhancement in SNR, and robust performance on representative ECG records highlight its potential for effective and reliable ECG baseline wander removal in both clinical and remote monitoring contexts.

V. Discussion

The proposed RL-based adaptive filtering framework demonstrated the ability to dynamically adjust filter parameters for baseline wander removal in ECG signals. However, its performance did not surpass that of a traditional fixed-parameter low-pass filter, which achieved

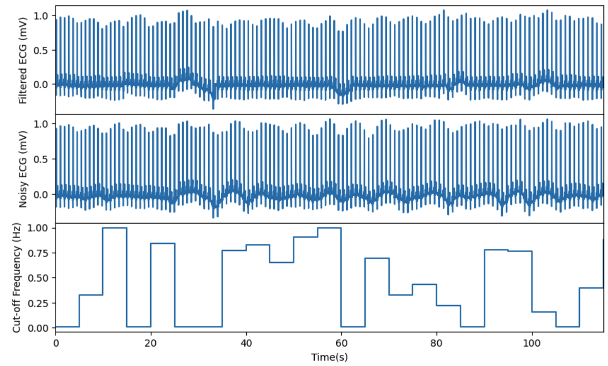

Fig. 7. The top panel shows the filtered ECG signal after baseline wander removal, while the middle panel displays the corresponding noisy ECG signal. The bottom panel illustrates the cutoff frequency selected by the RL agent over the course of the signal. The adaptive behavior of the agent in adjusting the cutoff frequency demonstrates its ability to respond to varying noise levels.

comparable or better results in terms of MSE and SNR. This outcome highlights certain limitations and opportunities for improvement in the current implementation.

One potential reason for the suboptimal performance of the RL framework is the simplicity of the reward function. The reward was based solely on the negative MSE between the clean reference signal and the denoised signal. While effective as a basic metric for signal similarity, this reward function does not account for the preservation of specific diagnostic features of the ECG, such as the ST segment or QT interval, which are critical for clinical applications. A more sophisticated reward function that integrates feature-specific criteria or domain knowledge could guide the RL agent to achieve better results by balancing noise removal with feature preservation.

An interesting observation in the results was the relationship between the cutoff frequency and the noise characteristics of the signal. As shown in Figure 7, the RL agent tended to select a higher cutoff frequency when the error caused by baseline wander was significant, effectively filtering out more noise. Conversely, when the error was small, the agent often chose a cutoff frequency close to zero, resulting in minimal filtering. While this adaptive behavior aligns with the goals of dynamic filtering, it suggests that the agent may benefit from more explicit guidance in selecting cutoff frequencies, potentially through enhanced state representations or reward mechanisms.

The limitations of this study also extend to the architecture of the RL agent and the computational complexity of the training process. The current agent, built on an LSTM network, may lack the capacity to fully capture the nuanced relationships between the signal characteristics and optimal filter parameters. Additionally, the training process, which involves evaluating the agent's performance across numerous noisy signal segments, is computationally intensive. Future implementations could explore more efficient RL algorithms or hybrid models that integrate RL with supervised or unsupervised learning to enhance performance.

VI. Future Works

Several avenues for future work can address the limitations identified in this study. First, designing a more comprehensive reward function that incorporates clinically significant features of the ECG, such as the preservation of wave morphology or the periodicity of the signal, could improve the framework's effectiveness. Second, exploring alternative architectures, such as attention-based models or transformer networks, may allow the agent to better understand the relationships between noise characteristics and optimal filter parameters. Another important direction is to investigate multi-objective optimization strategies that balance noise removal with the preservation of specific diagnostic features.

VII. Conclusion

While the RL-based adaptive filtering framework demonstrates promising adaptive capabilities, its performance highlights the need for more sophisticated reward functions, enhanced architectures, and comprehensive evaluations. Addressing these challenges in future work could establish the framework as a viable alternative to traditional filtering methods for baseline wander removal and other denoising applications in ECG signal processing. By building on these insights, this research lays the groundwork for more effective and clinically relevant adaptive filtering solutions.

Supplementary Materials

The code and data for reproducing the results in this paper can be found and downloaded at the link: https://github.com/MasoudNateghi/BMI-539.

Acknowledgment

The author acknowledges the effort to carefully select and cite relevant references in this work, aiming to provide the necessary context without overwhelming the reader with excessive citations in the introduction. This approach was taken to maintain clarity and focus while presenting the RL framework and its application to ECG baseline wander removal.

References

[1] R.-D. Bousseljot, D. Kreiseler, and A. Schnabel, "The PTB Diagnostic ECG Database," 2004. [Online]. Available: https://physionet.org/content/ptbdb/

[2] G. B. Moody, W. Muldrow, and R. G. Mark, "The MIT-BIH Noise Stress Test Database," 1992. [Online]. Available: https://physionet.org/content/nstdb/

[3] R. Fakoor, X. He, I. Tashev, and S. Zarar, "Reinforcement Learning To Adapt Speech Enhancement to Instantaneous Input Signal Quality," 2017. [Online]. Available: https://arxiv.org/abs/1711.10791

[4] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal Policy Optimization Algorithms," 2017. [Online]. Available: https://arxiv.org/abs/1707.06347

[5] R. Sameni, "OSET: The open-source electrophysiological toolbox," *Version 3.14 [Electronic Resource]*, 2012.