A Project Report

On

Anomaly Detection in WBANs

BY

Pratyush Bindal

2022A7PS0119H

Under the supervision of

Dr. Chittranjan Hota

SUBMITTED IN FULFILLMENT OF THE REQUIREMENTS OF CS F376: DESIGN PROJECT

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE PILANI (RAJASTHAN)

HYDERABAD CAMPUS

(April 2025)

ACKNOWLEDGMENTS

I want to express my sincere gratitude to Dr. Chittranjan Hota for his guidance and support during this research.

I would also like to express my sincere appreciation to Ms. Shreea Bose for her valuable time and work on the project, without whom this study would remain incomplete. Her insightful feedback, guidance, and meaningful suggestions proved to be very helpful throughout the project.

Pratyush Bindal
(2022A7PS0119H)

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE PILANI (RAJASTHAN)

HYDERABAD CAMPUS

CERTIFICATE

This is to certify that the project report entitled "**Anomaly Detection in WBANs**" submitted by Mr. Pratyush Bindal (ID No. 2022A7PS0119H) in fulfillment of the requirements of the course CS F376, Design Project Course, embodies the work done by him under my supervision and guidance.

Date: April 24, 2025

(Dr. Chittranjan Hota)

BITS Pilani, Hyderabad Campus

ABSTRACT

Wireless Body Area Networks (WBANs) are critical in health monitoring, leveraging wearable sensors to collect physiological data. However, anomaly detection in WBANs is challenging due to the dynamic nature of human activity, sensor noise, and data variability. In this report, we meticulously compare various Deep Learning (DL) and Reinforcement Learning (RL) methods for anomaly detection in WBANs by integrating LSTM and Transformer-based autoencoders with RL methods like Deep Q-Networks (DQN), and Advantage Actor-Critic (A2C). The proposed framework extracts sequential and periodic dependencies using LSTM and Transformer frameworks and derives anomaly scores through reconstruction errors.

These scores are then used to train a DRL agent, which optimizes anomaly classification decisions based on learned policies. We evaluate our approach on the PAMAP2 dataset and compare the efficacy of different models. Results indicate that the LSTM-Transformer-DQN model achieves superior accuracy, outperforming other architectures in anomaly classification. Integrating DRL with sequence modeling significantly enhances robustness, reducing false positives and improving generalization across various activities. This work underscores the potential of DRL-based frameworks in WBAN anomaly detection and its applicability in real-world healthcare scenarios.

CONTENTS

Introduction

Wireless Body Area Networks (WBANs) have emerged as a transformative technology for continuous health monitoring, enabling real-time collection and analysis of physiological signals. These systems support applications such as chronic disease management, early detection of medical emergencies, and personalized healthcare. However, WBANs are prone to anomalies caused by sensor malfunctions, environmental interference, and irregular user behavior. Detecting these anomalies accurately ensures reliable health monitoring and prevents erroneous clinical decisions.

Traditional anomaly detection methods rely on statistical models, threshold-based techniques, or classical machine learning approaches. These methods often struggle to adapt to physiological data's complex, non-stationary nature. Recent advancements in deep learning have enabled more sophisticated sequence modeling through architectures such as Long Short-Term Memory (LSTM) networks and Transformers. These models effectively capture temporal dependencies and contextual patterns, making them suitable for anomaly detection in WBANs.

This work proposes a novel DRL-based anomaly detection framework that integrates LSTM and Transformer autoencoders with reinforcement learning. We leverage LSTM and Transformer models to capture physiological data's local and global temporal dependencies, enhancing anomaly detection performance. Unlike conventional approaches that rely solely on reconstruction errors, our framework employs a DRL agent to optimize anomaly classification decisions based on learned policies. Additionally, we compare the proposed framework against multiple baseline methods to determine the most effective approach for anomaly detection in WBANs.

By integrating DRL with sequence modeling, our approach improves generalization and reduces false positives, making it a viable solution for real-world healthcare applications.

Overview

The proposed multi-context Deep Reinforcement Learning (DRL) based framework represents a novel approach to detecting anomalies in Wireless Body Area Networks (WBANs). The architecture addresses the unique challenges of WBAN sensor data by integrating multiple temporal contexts through specialized autoencoders utilizing Deep Learning (DL) techniques such as LSTM and transformers and employing Reinforcement Learning (RL) for optimal decision-making. *Fig 0* illustrates the overall process of the proposed framework, which consists of five integrated stages: data pre-processing, sequential context modelling, periodic context modelling, feature fusion and DRL-based decision-making.

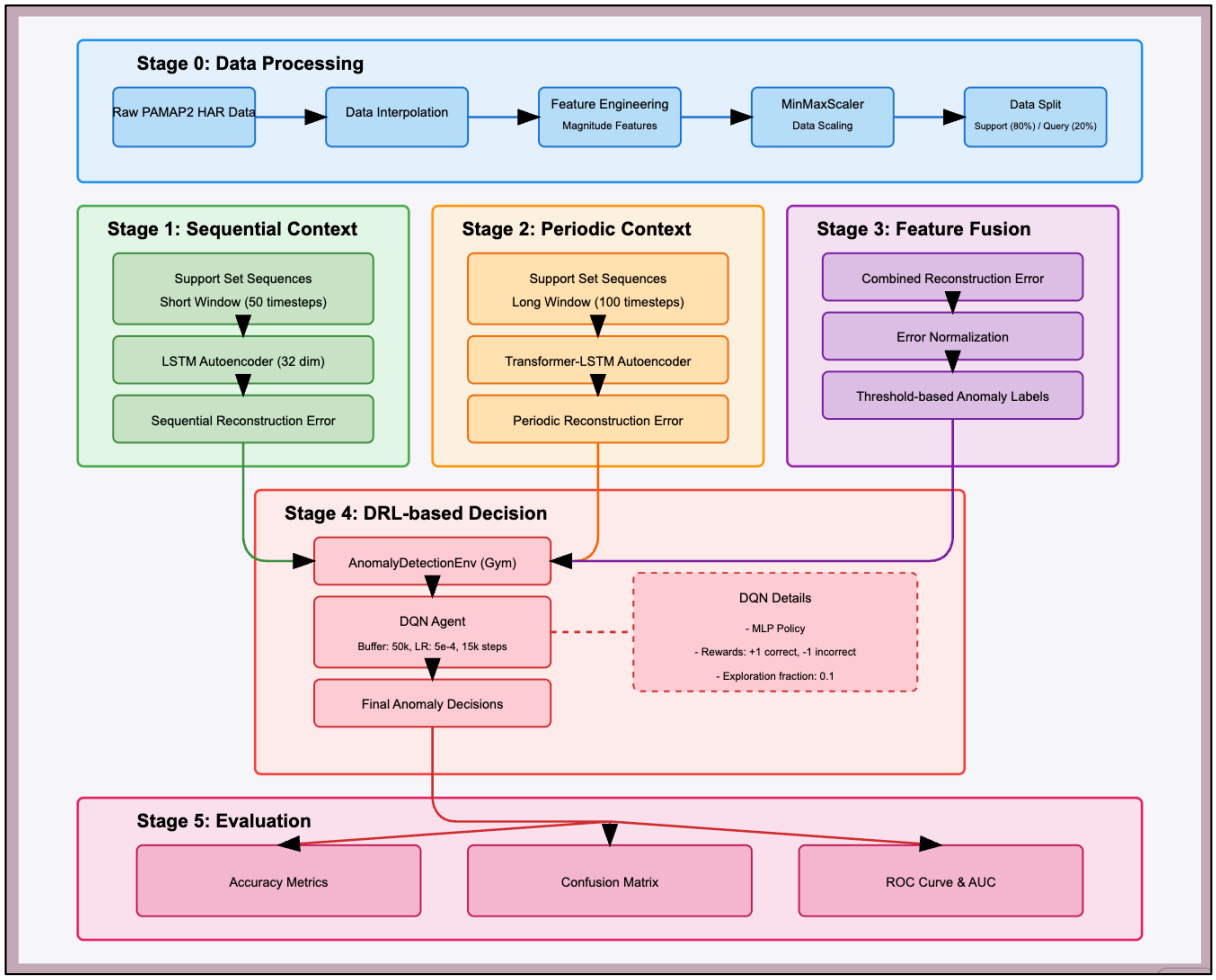


Fig 0. Detailed process overview for anomaly detection in WBANs

Reconstruction-Based Anomaly Detection

We utilize three different types of autoencoder architectures in this work:

- **LSTM Autoencoder** models sequential dependencies using stacked Long Short-Term Memory (LSTM) layers. The encoder compresses the input sequence into a latent representation while the decoder reconstructs the original input. Reconstruction errors,

computed as the mean absolute difference between the input and the output, are a quantitative measure of the anomalous behavior.

- **Transformer Autoencoder** leverages multi-head self-attention mechanisms to capture periodic and global temporal patterns. The encoder-decoder structure incorporates Transformer encoder layers with layer normalization and dropout. Global average pooling is applied in the encoder to obtain a compact representation, and the decoder reconstructs the input sequence. Reconstruction errors are computed similarly to the previous case.
- **CNN Autoencoder** is implemented to capture point anomalies. The architecture comprises convolutional layers with ReLU activations, max-pooling layers for dimensionality reduction, and upsampling and convolution layers with sigmoid activations to reconstruct the original input.

For each autoencoder, a threshold is determined using the 95^{th} percentile of the reconstruction error computed over the training set. Sequences with errors exceeding this threshold are flagged as anomalous. In addition, reconstruction errors from models emphasizing sequential (LSTM) and periodic (Transformer) contexts are aggregated to form a combined anomaly score.

Deep Reinforcement Learning for Anomaly Refinement

We implemented a custom Gym environment where the reconstructed error of each sequence serves as the observation. Binary anomaly labels are derived from the aggregated error using a fixed threshold. Several DRL agents like Deep Q-Networks (DQN), and Advantage Actor- Critic (A2C), are then trained to refine the classification of anomalies. The agents receive a reward signal based on the correctness of their anomaly detection decisions, and the best-performing architecture achieves high accuracy by leveraging the stability and sample efficiency of the underlying DRL algorithm.

Dataset Used and Preprocessing

The methodology consists of four main stages: data preprocessing and feature engineering, sequence generation, reconstruction-based anomaly detection using autoencoders, and DRL-based decision refinement.

Data Preprocessing and Feature Engineering

The experiments utilize the PAMAP2 Monitoring dataset, a widely used benchmark for Human Activity Recognition (HAR) with wearable sensors. The dataset comprises time-series data collected from 9 subjects performing 18 activities, using three Inertial Measurement Units (IMUs) placed on the hand, chest, and ankle, along with a heart rate monitor. Table 1 summarizes the primary columns present in the raw dataset:

Column Name	Description
timestamp	Time (s)
heart_rate	Heart rate (bpm)
IMU_hand_temperature	Hand temperature reading (℃)
IMU_hand_3D_acceleration_1	Hand x-axis acceleration m/s^2 ($\pm$16g scale)
IMU_hand_3D_acceleration_2	Hand y-axis acceleration m/s^2 ($\pm$16g scale)
IMU_hand_3D_acceleration_3	Hand z-axis acceleration m/s^2 ($\pm$16g scale)
IMU_hand_3D_acceleration_4	Hand x-axis acceleration m/s^2 ($\pm$6g scale)
IMU_hand_3D_acceleration_5	Hand y-axis acceleration m/s^2 ($\pm$6g scale)
IMU_hand_3D_acceleration_6	Hand z-axis acceleration m/s^2 ($\pm$6g scale)
IMU_hand_3D_gyroscope_1	Hand x-axis angular velocity (rad/s)
IMU_hand_3D_gyroscope_2	Hand y-axis angular velocity (rad/s)
IMU_hand_3D_gyroscope_3	Hand z-axis angular velocity (rad/s)
IMU_hand_3D_magnetometer_1	Hand x-axis magnetometer reading (μT)
IMU_hand_3D_magnetometer_2	Hand y-axis magnetometer reading (μT)
IMU_hand_3D_magnetometer_3	Hand z-axis magnetometer reading (μT)
IMU_hand_orientation_1	Specific orientation of the hand
IMU_hand_orientation_2	Specific orientation of the hand
IMU_hand_orientation_3	Specific orientation of the hand
IMU_hand_orientation_4	Specific orientation of the hand
IMU_chest_temperature	Chest temperature reading (℃)
IMU_chest_3D_acceleration_1	Chest x-axis acceleration m/s^2 ($\pm$16g scale)
IMU_chest_3D_acceleration_2	Chest y-axis acceleration m/s^2 ($\pm$16g scale)
IMU_chest_3D_acceleration_3	Chest z-axis acceleration m/s^2 ($\pm$16g scale)
IMU_chest_3D_acceleration_4	Chest x-axis acceleration m/s^2 ($\pm$6g scale)
IMU_chest_3D_acceleration_5	Chest y-axis acceleration m/s^2 ($\pm$6g scale)
IMU_chest_3D_acceleration_6	Chest z-axis acceleration m/s^2 ($\pm$6g scale)
IMU_chest_3D_gyroscope_1	Chest x-axis angular velocity (rad/s)
IMU_chest_3D_gyroscope_2	Chest y-axis angular velocity (rad/s)

IMU_chest_3D_gyroscope_3	Chest z-axis angular velocity (rad/s)
IMU_chest_3D_magnetometer_1	Chest x-axis magnetometer reading (μT)
IMU_chest_3D_magnetometer_2	Chest y-axis magnetometer reading (μT)
IMU_chest_3D_magnetometer_3	Chest z-axis magnetometer reading (μT)
IMU_chest_orientation_1	Specific orientation of the chest
IMU_chest_orientation_2	Specific orientation of the chest
IMU_chest_orientation_3	Specific orientation of the chest
IMU_chest_orientation_4	Specific orientation of the chest
IMU_ankle_temperature	Ankle temperature reading (℃)
IMU_ankle_3D_acceleration_1	Ankle x-axis acceleration m/s^2 (±16g scale)
IMU_ankle_3D_acceleration_2	Ankle y-axis acceleration m/s^2 (±16g scale)
IMU_ankle_3D_acceleration_3	Ankle z-axis acceleration m/s^2 (±16g scale)
IMU_ankle_3D_acceleration_4	Ankle x-axis acceleration m/s^2 (±6g scale)
IMU_ankle_3D_acceleration_5	Ankle y-axis acceleration m/s^2 (±6g scale)
IMU_ankle_3D_acceleration_6	Ankle z-axis acceleration m/s^2 (±6g scale)
IMU_ankle_3D_gyroscope_1	Ankle x-axis angular velocity (rad/s)
IMU_ankle_3D_gyroscope_2	Chest y-axis angular velocity (rad/s)
IMU_ankle_3D_gyroscope_3	Chest z-axis angular velocity (rad/s)
IMU_ankle_3D_magnetometer_1	Ankle x-axis magnetometer reading (μT)
IMU_ankle_3D_magnetometer_2	Chest y-axis magnetometer reading (μT)
IMU_ankle_3D_magnetometer_3	Chest z-axis magnetometer reading (μT)
IMU_ankle_orientation_1	Specific orientation of ankle
IMU_ankle_orientation_2	Specific orientation of ankle
IMU_ankle_orientation_3	Specific orientation of ankle
IMU_ankle_orientation_4	Specific orientation of ankle
activityID	Numeric activity label
activity_name	Human-readable activity label

Table 1. Dataset Description

We first import raw sensor data from the PAMAP2 Human Activity Recognition dataset. We use linear interpolation to address missing values, and subsequent removal of any remaining NaNs ensures data integrity. Then, we perform feature engineering on the dataset by computing the Euclidean norm of multi-dimensional inertial measurement unit signals, including acceleration, gyroscope, and magnetometer readings across different body parts. Redundant raw sensor channels are then dropped. Moreover, the resultant feature matrix is normalized using Min–Max scaling to ensure that all features lie within a consistent range, thereby facilitating the convergence of neural network models.

Sequence Generation and Data Splitting:

The preprocessed dataset is partitioned into support (training) and query (testing) sets based on an 80-20 split criterion. Sequential data samples are generated by applying a sliding window approach to form time-series sequences using two sequence lengths: a shorter window (30 - 50 timesteps) for capturing sequential context and a longer window (100 timesteps) to model periodic patterns.

Dataset Visualizations

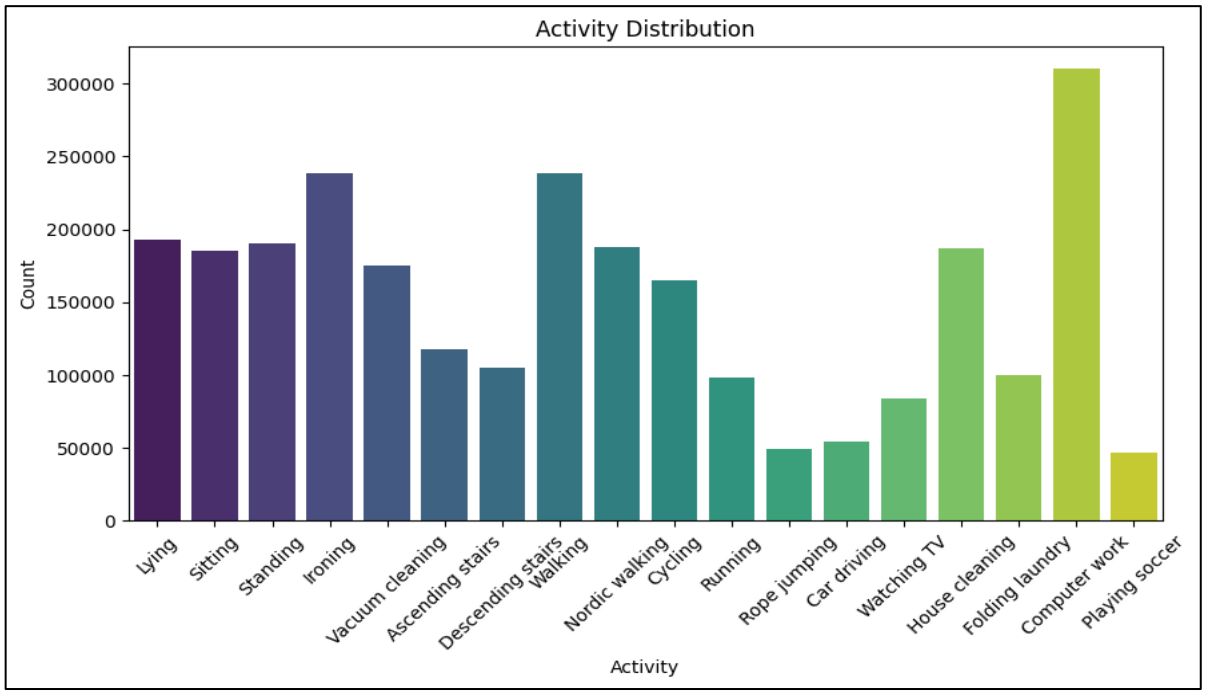


Fig 1. Activity distribution chart

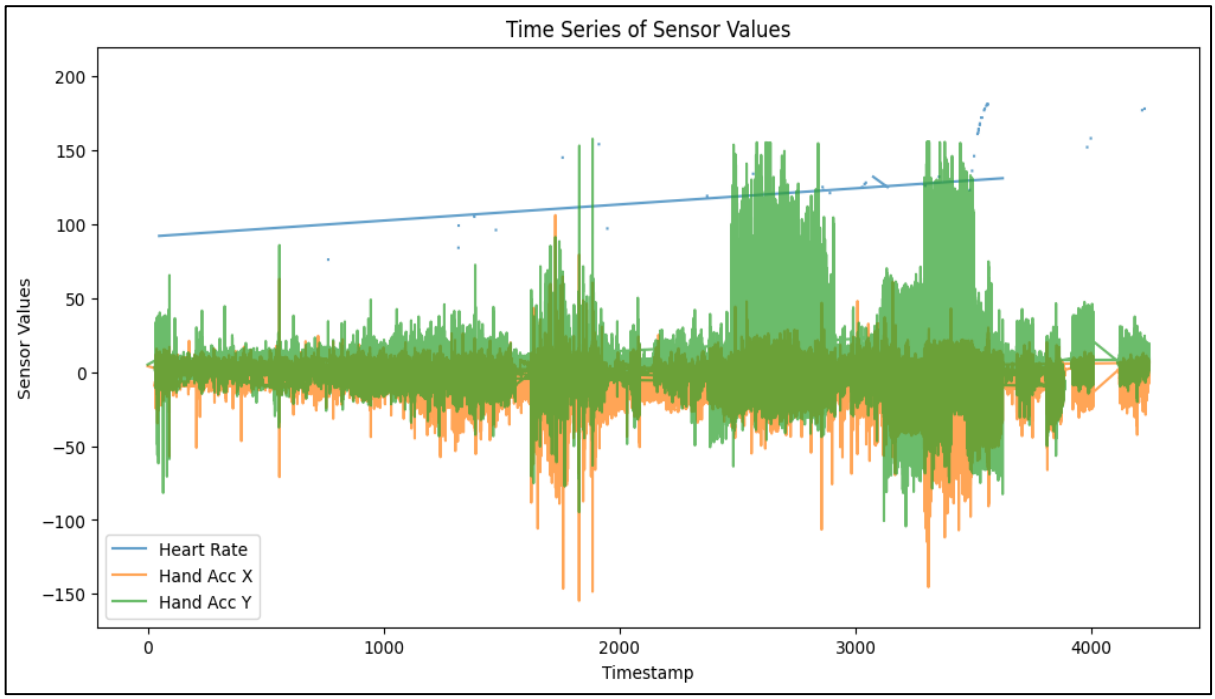


Fig 2. Time series plots for sensor values

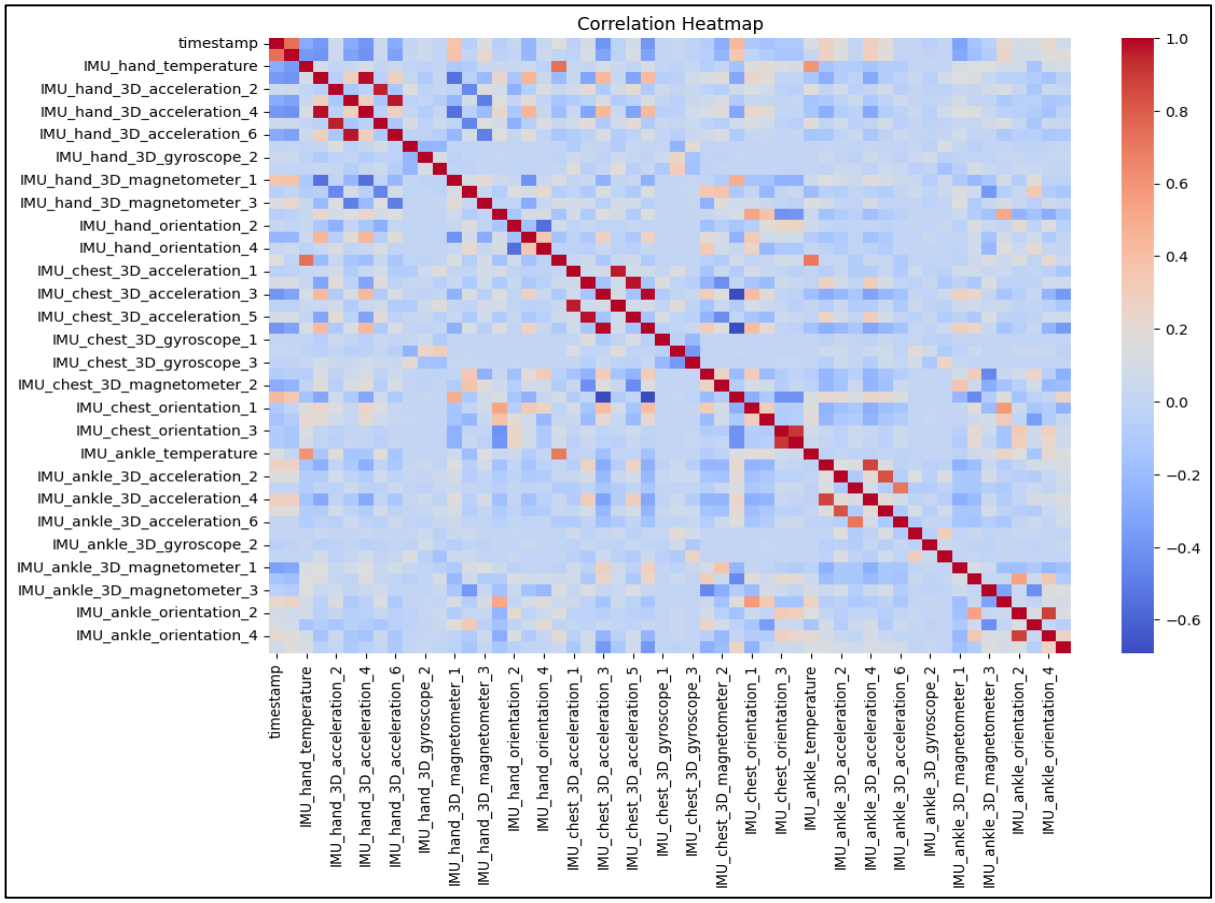


Fig 3. Correlation Heatmap

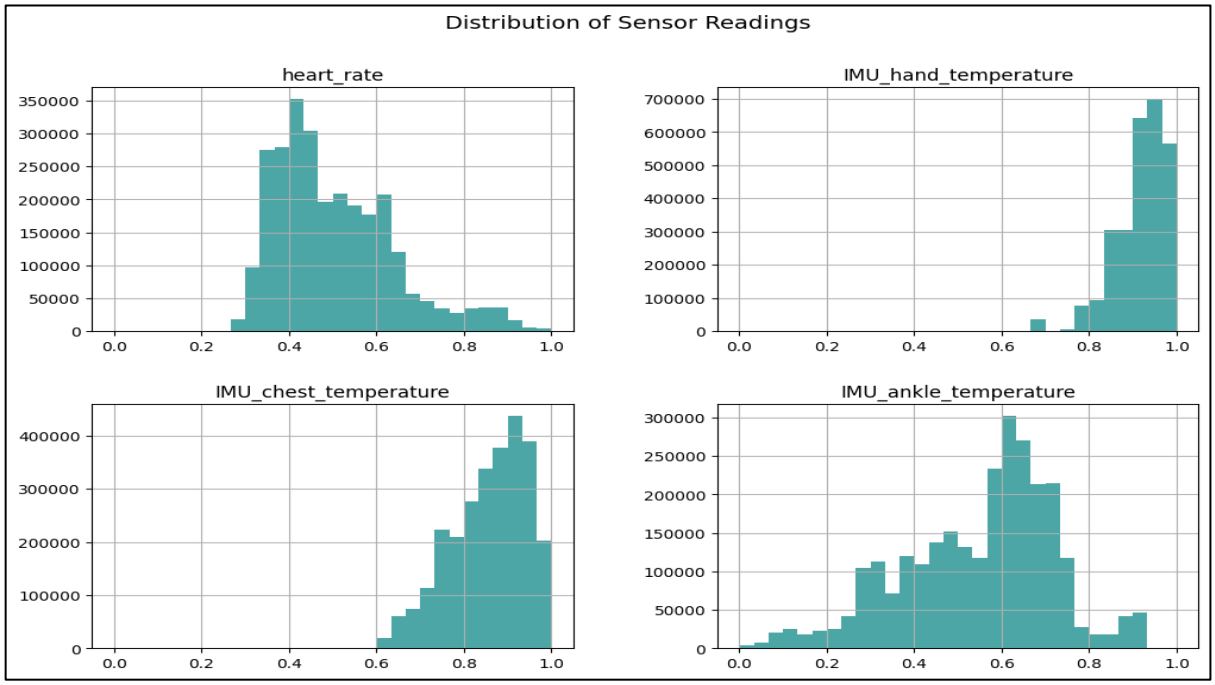


Fig 4. Distribution of sensor readings

Proposed Architecture

As shown in *Fig 5*, the proposed architecture implements a hierarchical, multi-level anomaly detection system for HAR data. We combine deep unsupervised representation-based models such as LSTM and transformer autoencoders with deep reinforcement learning (DRL) frameworks like Deep Q-Network (DQN) to create a robust anomaly detection framework.

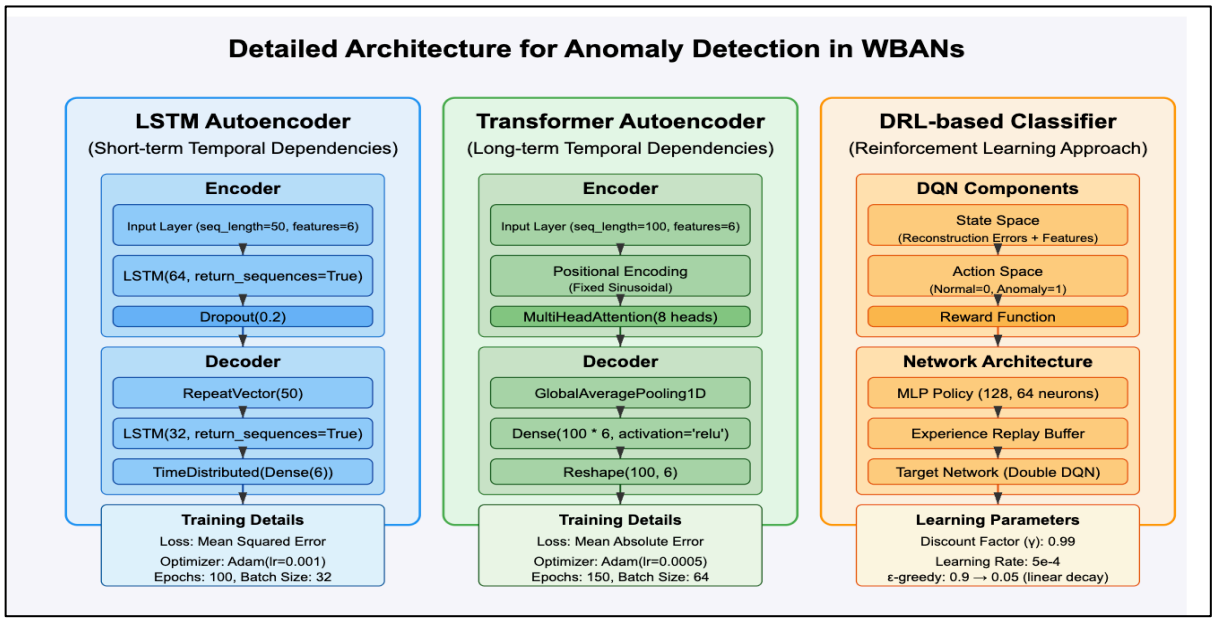


Fig 5. Detailed architecture for anomaly detection in WBANs

Short-term Temporal Dependencies (Sequential Context)

We employ an LSTM-based autoencoder to model short-term temporal dependencies in the sensor data, which is crucial for capturing local patterns and immediate sequential relationships between consecutive sensor readings.

The LSTM autoencoder consists of an encoder-decoder structure designed to reconstruct input sequences, consisting of an input layer which accepts sequences with 50-timestep windows of sensor data. Then, a single LSTM layer, along with the ReLU activation function, is utilised as an encoder, which helps to capture temporal patterns in the input sequence and then compress them into fixed-dimension representation. A dropout layer is also incorporated to reduce the chances of overfitting. Next, a repeat vector layer replicates the encoded representation, which is followed by an LSTM layer acting as a decoder that helps to reconstruct the original sequence.

The LSTM autoencoder is trained on the support set using the mean squared error loss function and Adam optimiser. After training, we apply the model to the query set to compute reconstruction errors that act as anomaly scores. High reconstruction errors indicate that the sequences deviate from standard learned patterns, potentially representing anomalous behaviour.

Long-term Temporal Dependencies (Periodic Context)

We utilise a Transformer-LSTM hybrid autoencoder to capture the sensor data's long-range dependencies and periodic patterns. This component complements the LSTM autoencoder by modelling relationships between distant time points, which is essential for detecting complex anomalies spread over long time horizons.

The Transformer-LSTM autoencoder combines the attention mechanism of transformers with the sequential modelling capability of LSTMs. It consists of an input layer accepting sequences with 100-timestep windows of sensor data to capture extended temporal patterns. A multi-head attention mechanism acts as an encoder which learns the time-distant relationships. Normalised attention outputs are pooled across the temporal dimension to create a global representation. Next, a repeat vector layer followed by an LSTM act as a decoder, which helps to reconstruct the whole sequence, with a final dense layer producing the output.
The Transformer-LSTM autoencoder is trained on more extended sequences from the support set with larger batch sizes to accommodate increased computational requirements.

Like the LSTM autoencoder, we compute the reconstruction errors on the query set as anomaly scores. The attention mechanism enables the model to identify complex patterns that might be missed by the LSTM autoencoder, providing complementary anomaly detection capabilities.

Multi-Context Fusion

We integrate the anomaly scores from both autoencoder models to create a unified representation that leverages their complementary strengths. The fusion approach helps to combine information from different temporal contexts, enabling more robust anomaly detection. The combined errors are then normalised to the [0,1] range to create standardised anomaly scores, and based on the normalised scores, we establish initial anomaly labels using a threshold at the 80^{th} percentile. These labels serve as a baseline for training RL component, further refining anomaly detection decisions.

DRL-based Learning for Decision Optimisation

The final stage of the framework employs DRL to optimise the anomaly detection decision-making process, which offers significant advantages over static thresholds by learning adaptive decision boundaries that aim to maximise detection accuracy. We implemented a custom Open AI Gym environment that encapsulated anomaly detection problems as a sequential decision process in which a Deep Q-Network (DQN) agent learns the optimal policy for anomaly classification based on normalised reconstruction errors. We train the DQN agent for 15,000 timesteps, during which it learns to discriminate between normal and anomalous patterns based on the reward signals. After training, the agent's learned policy makes the final anomaly detection decisions.

The experimental results demonstrate that the proposed framework performs better than traditional threshold-based approaches as the DRL component effectively learns the adaptive decision boundaries, maximising detection accuracy while minimising false alarms.

Comparative Analysis

We analyzed multiple methods based on LSTM, Transformer, and hybrid architectures integrated with various reinforcement learning techniques.

1. **LSTM-FIF (Feedback-based Isolation Forest):** This approach utilizes an LSTM-based autoencoder that learns representative temporal patterns by compressing and accurately representing input sequences. The reconstruction errors are computed and leveraged within a modified Feedback-based Isolation Forest to flag anomalies, with the 95th percentile threshold serving as a reference for anomalous behaviour. We trained for 10 epochs and detected 4980 anomalies in the entire dataset.

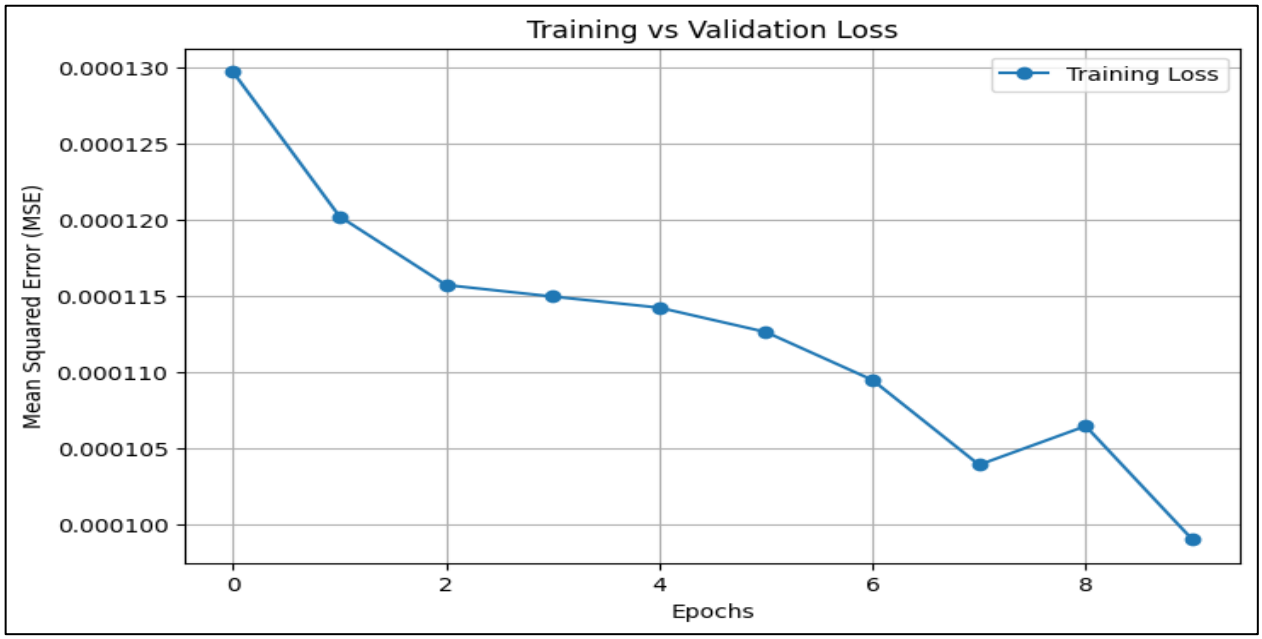


Fig 6. LSTM-FIF Training Loss Curve

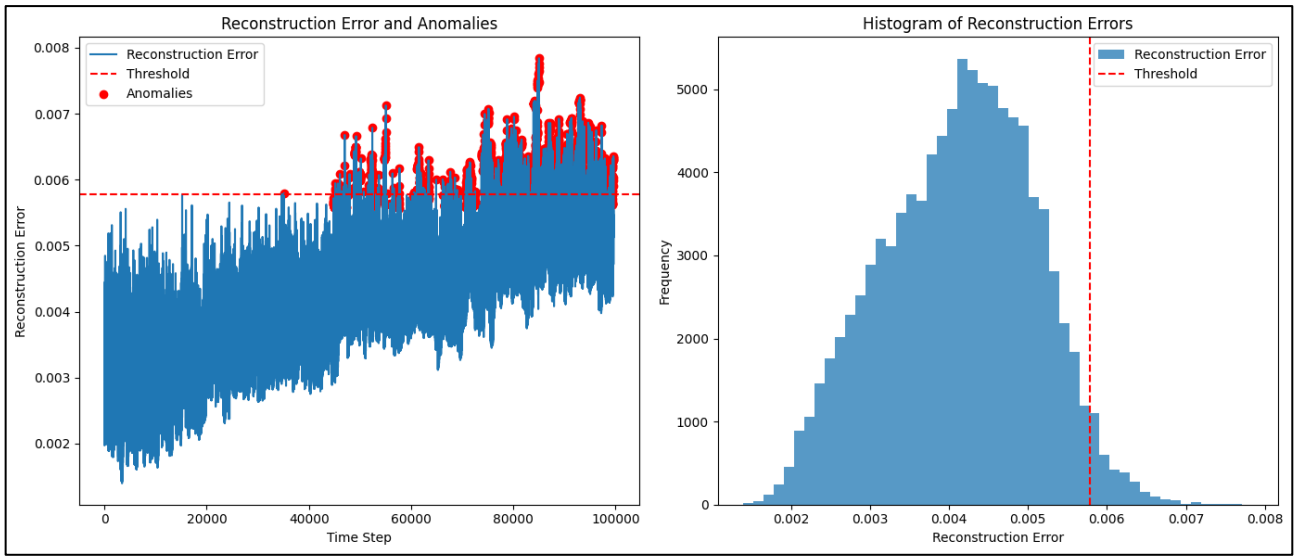


Fig 7. LSTM-FIF Anomaly Plot and Histogram of Reconstruction Errors

2. **LSTM-PPO (Proximal Policy Optimisation):** An LSTM autoencoder, designed with stacked LSTM layers and a corresponding decoder, is trained to reconstruct input sequences. Reconstruction errors are computed as the mean absolute difference between original and reconstructed sequences, and these errors serve to derive unsupervised rewards for further training. A PPO agent, implemented with fully connected layers, is subsequently trained using these rewards to refine anomaly classification. Finally, we evaluate the test set by thresholding the reconstruction errors at the 95th percentile. We trained LSTM model for 20 epochs and PPO algorithm for 10 epochs obtaining 5980 anomalies in entire dataset.

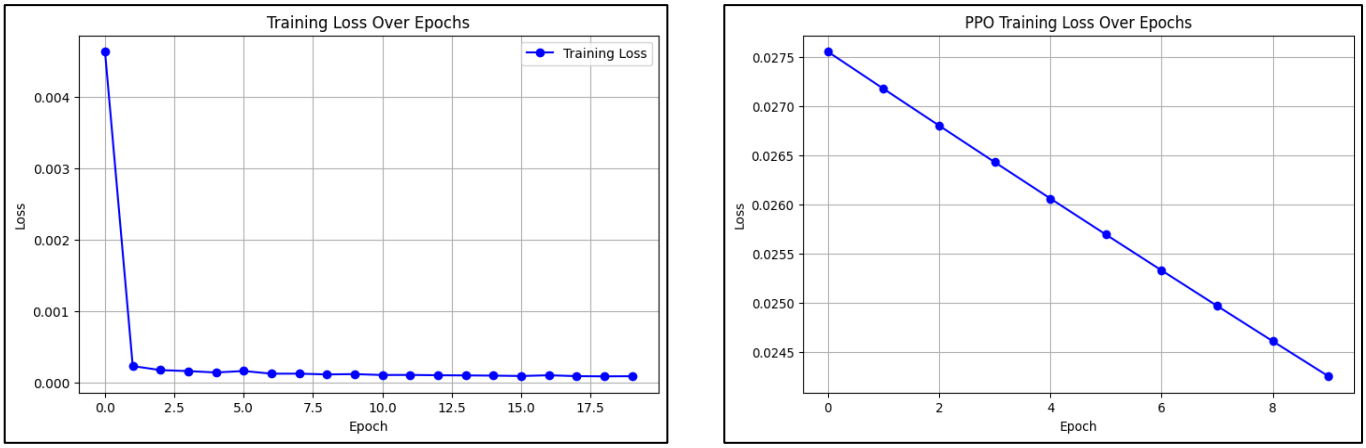


Fig 8. LSTM and PPO Training Loss Curves

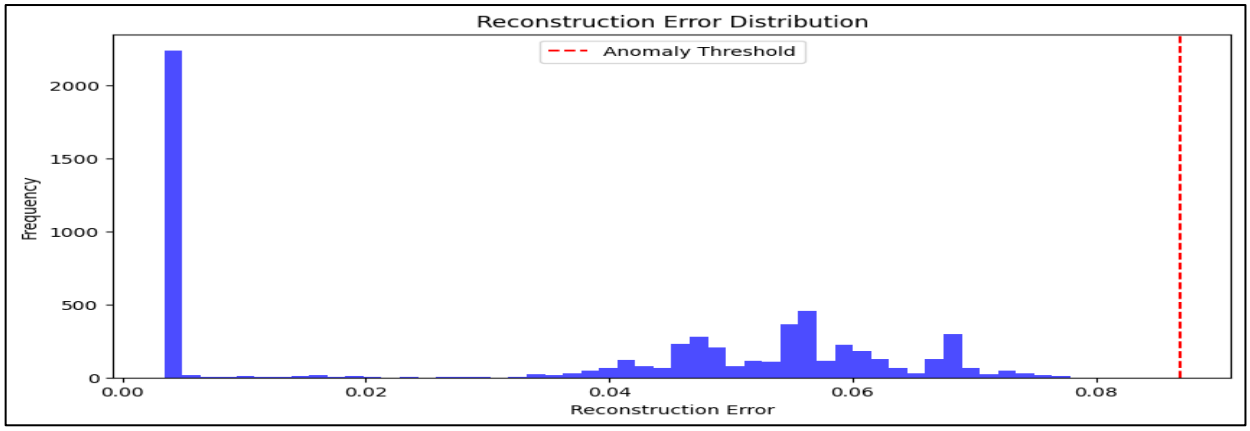


Fig 9. LSTM-PPO Reconstruction Error Distribution

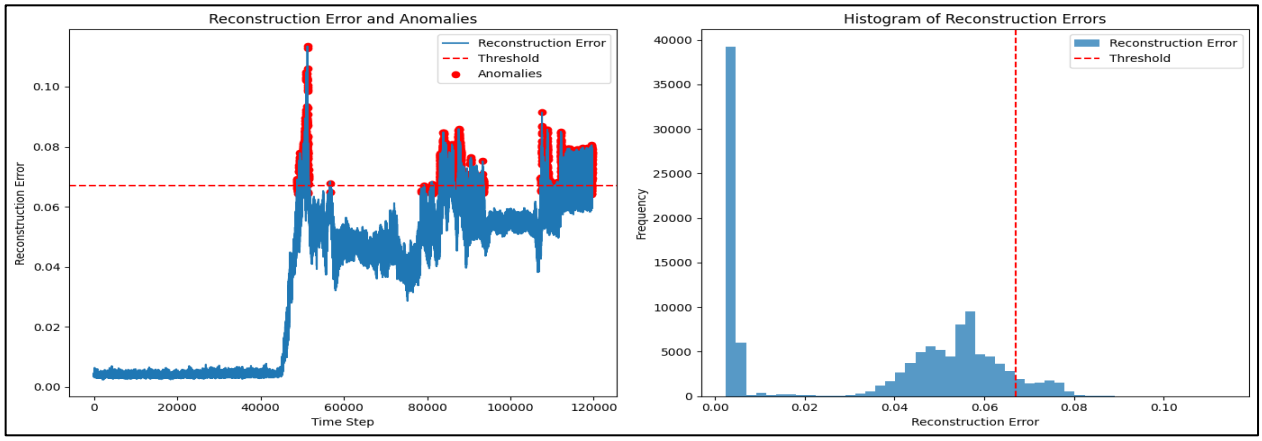


Fig 10. LSTM-PPO Anomaly Plot and Histogram of Reconstruction Errors

3. **Transformer-PPO (Proximal Policy Optimisation):** The autoencoder employs Transformer encoder layers featuring multi-head self-attention, feed-forward networks, layer normalization, and dropout to learn compact representations and reconstruct input sequences. Reconstruction errors, computed as the mean absolute difference between original and reconstructed sequences, are leveraged to derive unsupervised reward signals via 95th-percentile thresholding. A PPO agent, implemented with dense layers, is trained using these rewards to refine the anomaly detection process. We trained transformer architecture for 20 epochs and PPO model for 10 epochs obtaining 5380 anomalies in the entire dataset.

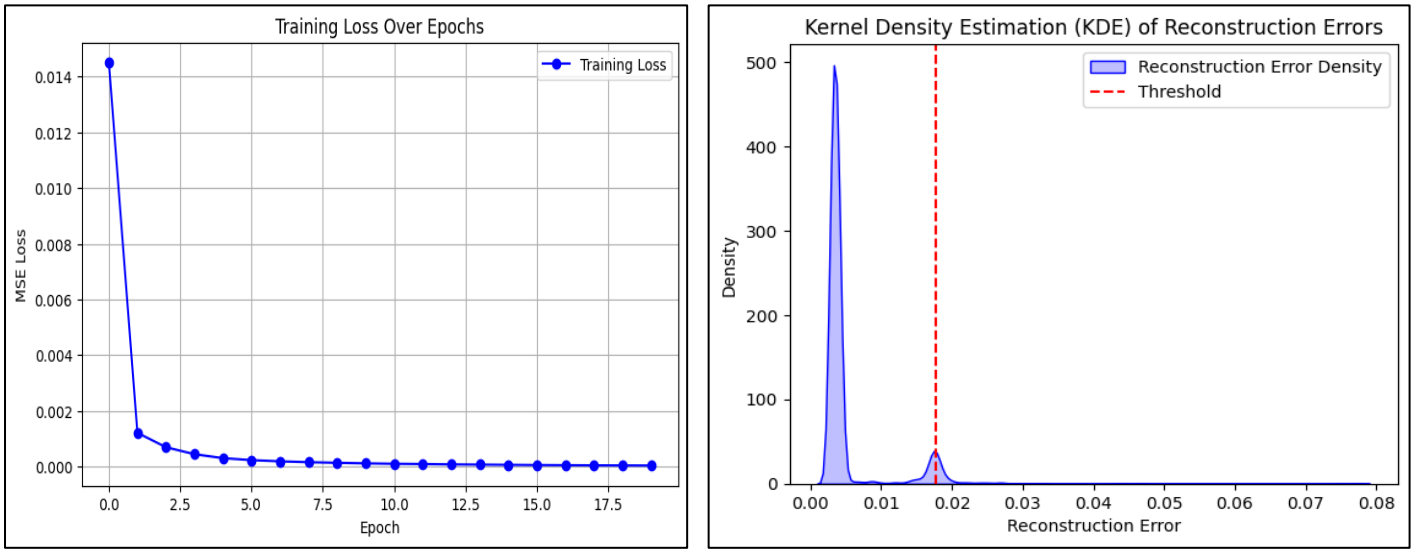


Fig 11. Transformer Training Loss Curve and KDE of Reconstruction Errors

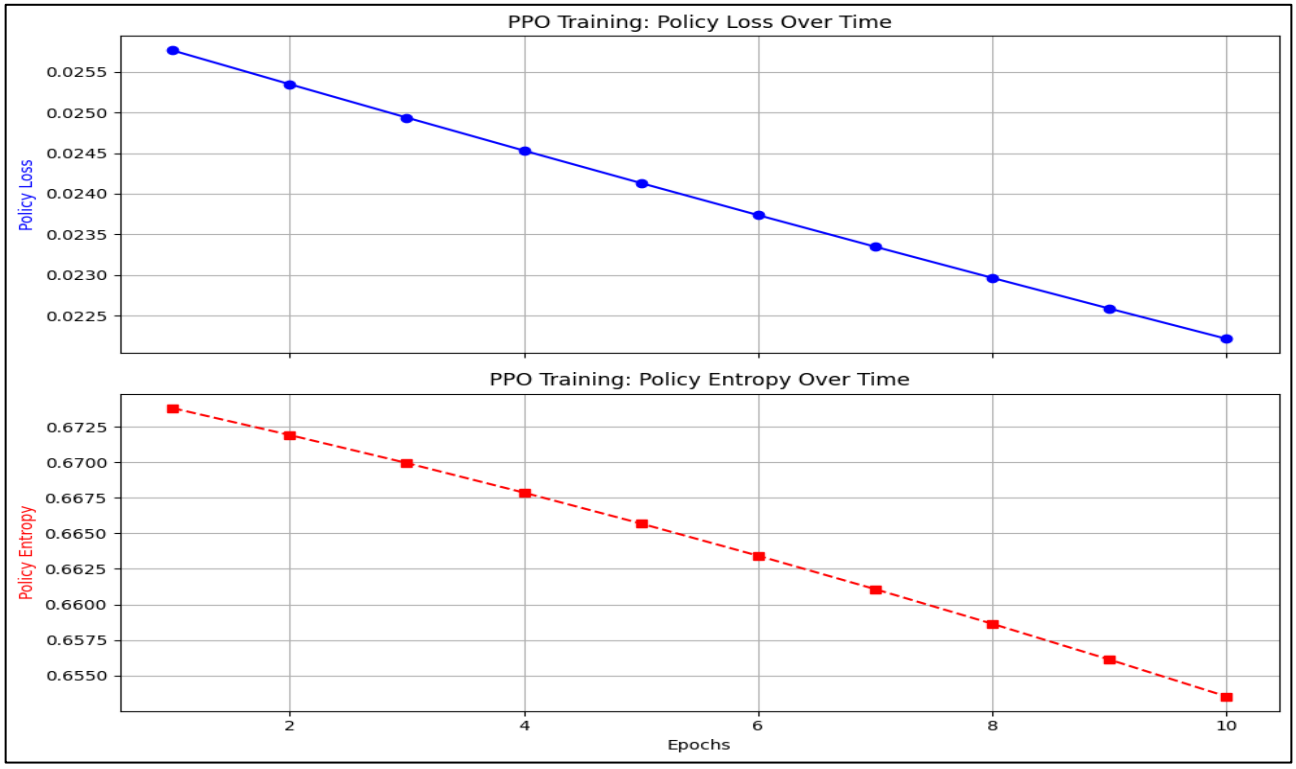


Fig 12. PPO Training Loss Curve and PPO Policy Entropy Plot

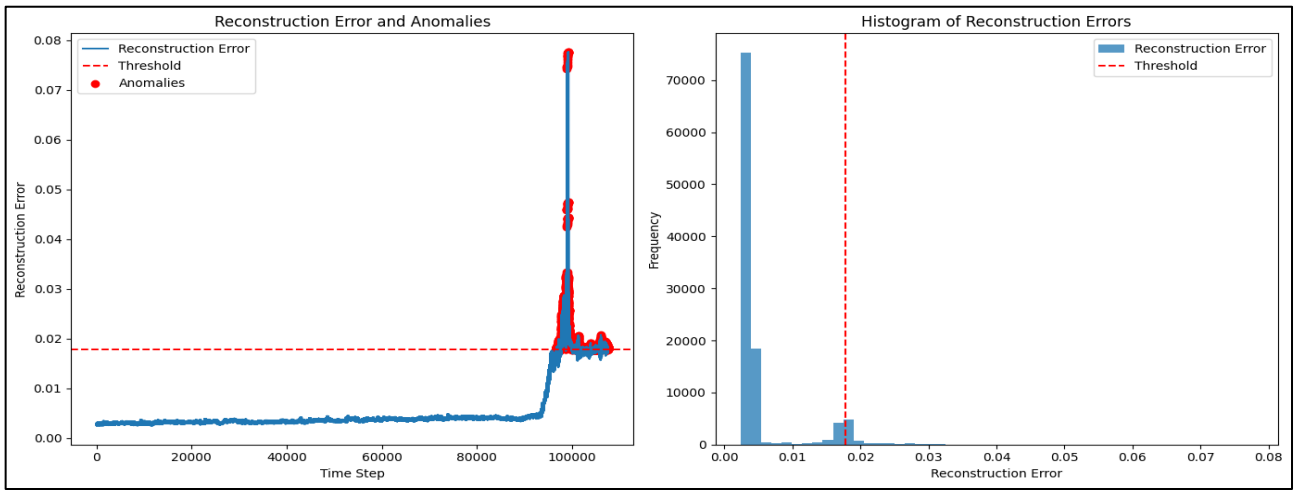


Fig 13. Transformer-PPO Anomaly Plot and Histogram of Reconstruction Errors

The experimental results indicate that three methods- LSTM-PPO, LSTM-FIF, and Transformer-PPO yield varying anomaly counts over the entire dataset. The LSTM-PPO and LSTM-FIF approaches use sequential context, effectively capturing temporal dynamics. However, by focusing solely on sequential information, these methods may overlook anomalies resulting from periodic patterns or mixed contextual features. Conversely, Transformer-PPO emphasizes periodic context, which is beneficial for capturing recurring temporal patterns but may miss anomalies primarily defined by sequential dependencies. These observations suggest that each method, while robust within its designated context, might not fully capture mixed-context anomalies present in complex sensor datasets.

Therefore, a potential future direction would be integrating sequential and periodic features to achieve a more comprehensive anomaly detection framework. Furthermore, due to computational constraints, we restrict further models discussed to the first 25000 entries of the dataset.

4. **CNN AutoEncoder:** This approach uses a CNN autoencoder that comprises convolutional, max-pooling, and upsampling layers to reconstruct input sequences. We show reconstruction errors using the mean absolute difference between original and reconstructed sequences, with anomalies marked as sequences exceeding the 95th percentile threshold. We trained for 20 epochs and detected 249 anomalies.

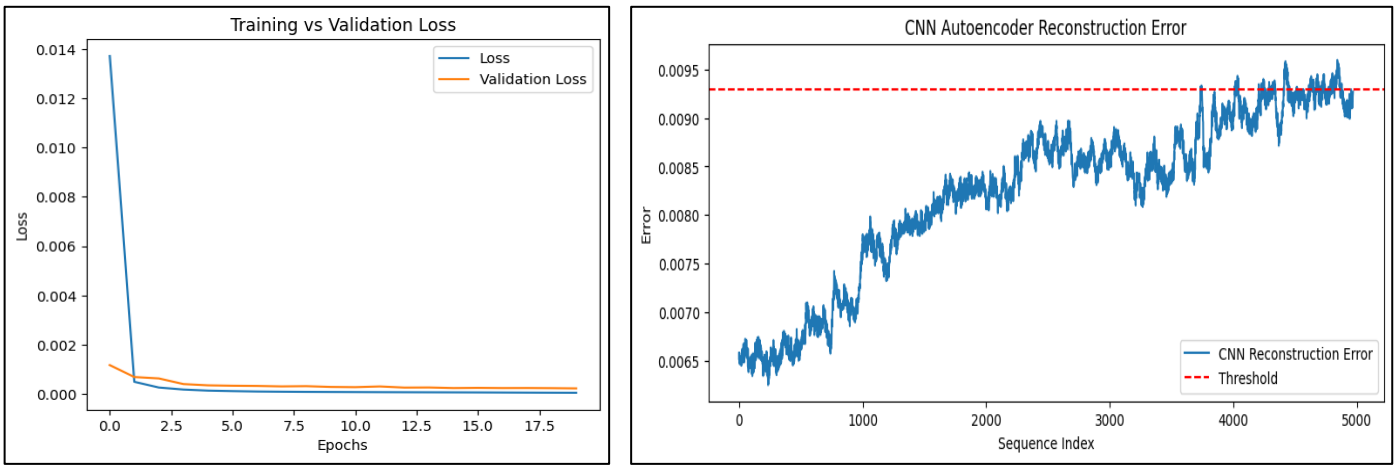


Fig 14. CNN AutoEncoder Training vs. Validation Loss Curve and Reconstruction Error Plot

Experimental results reveal that the CNN-AutoEncoder detects 249 point anomalies. The model's focus on individual data points results in a limited capacity to capture collective anomalies. Hence, while this method effectively identifies isolated outliers, it may overlook anomalies from temporal or contextual dependencies. Hence, we build advanced DRL-based architectures which could be better utilized for complex datasets.

5. **LSTM-Transformer-A2C (Advantage Actor-Critic):** In this model, two autoencoder models are employed: an LSTM autoencoder for capturing short-term sequential context and a Transformer autoencoder for modeling long-term periodic patterns. Reconstruction errors from both models are computed, normalized, and aggregated to generate final anomaly labels. A custom Gym environment is defined using these aggregated errors, and an A2C agent is trained to refine anomaly detection decisions. The performance of the DRL agent is evaluated by comparing its actions with the derived anomaly labels.

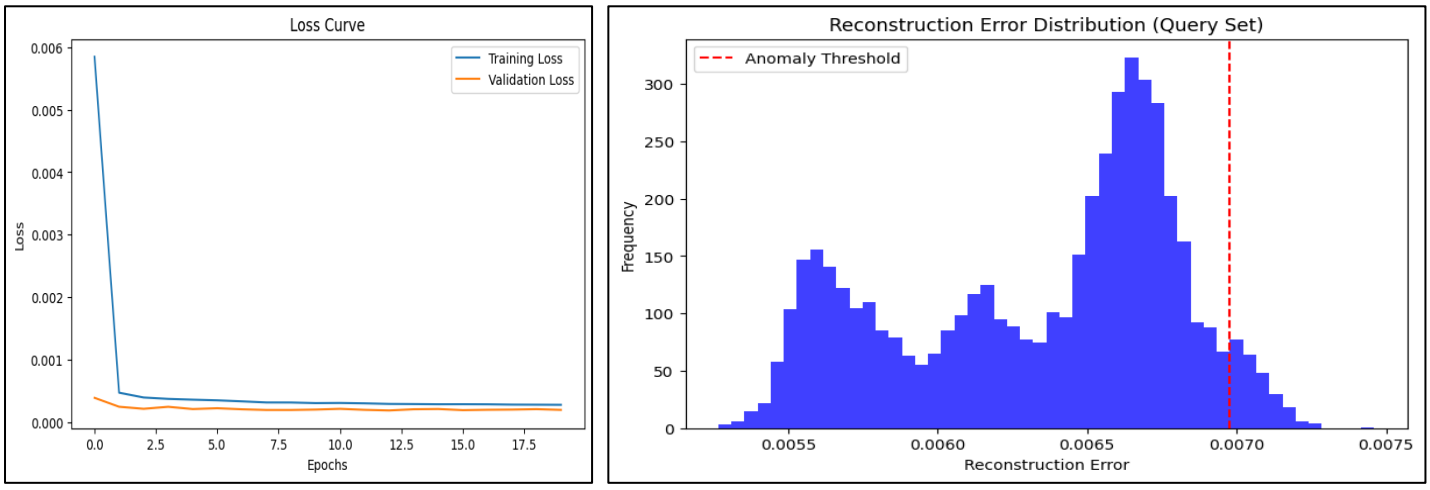


Fig 15. Transformer Training vs. Validation Loss Curve and Reconstruction Error Distribution

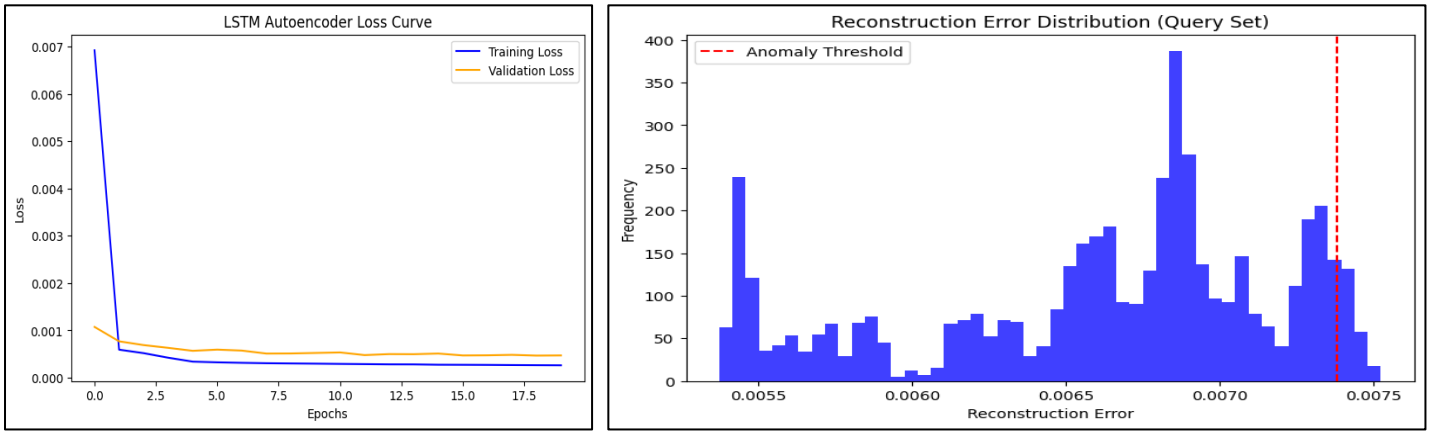


Fig 16. LSTM Training vs. Validation Loss Curve and Reconstruction Error Distribution

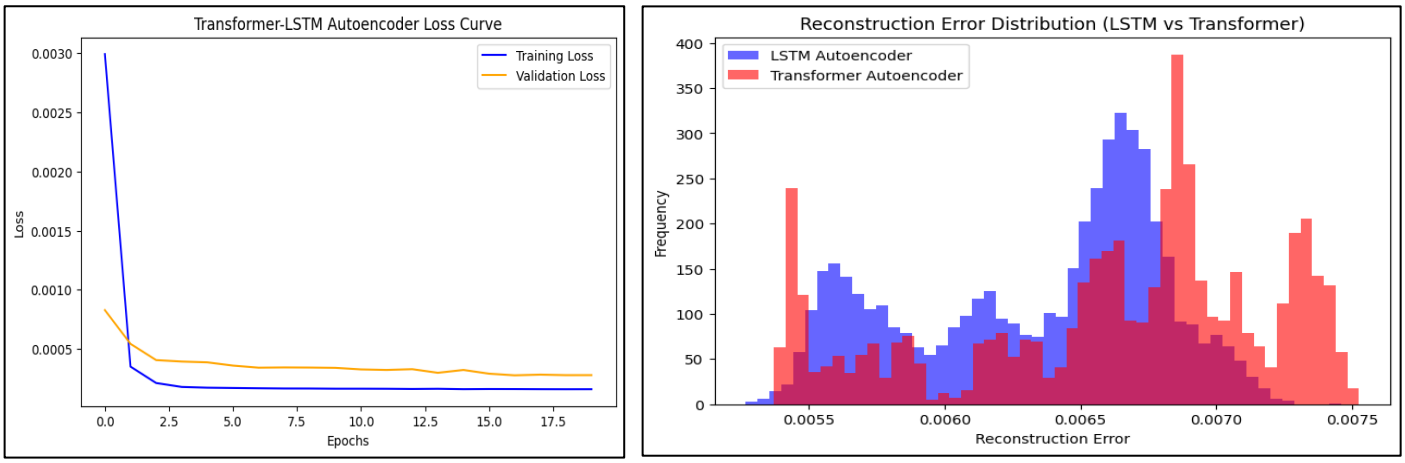


Fig 17. LSTM-Transformer Training vs. Validation Loss Curve and Combined Reconstruction Error Distribution

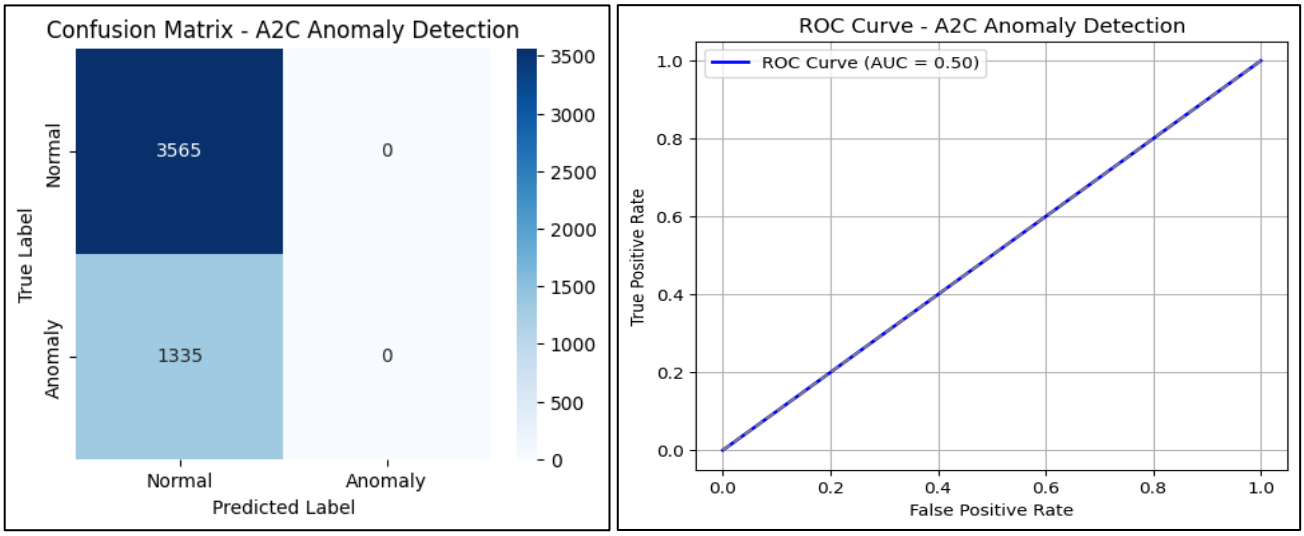
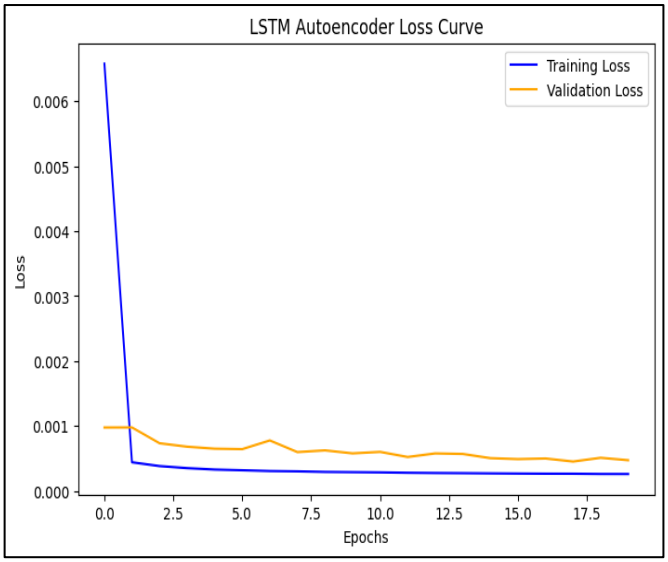


Fig 18. LSTM-Transformer-A2C Confusion Matrix and ROC Curve

The A2C agent, when combined with LSTM and Transformer autoencoder features, yielded 1335 detected anomalies at a 72.76% accuracy. This suboptimal performance is primarily due to the inherent sample inefficiency of the A2C algorithm in complex time-series environments. The diverse sequential and periodic patterns in the data require extensive sampling to converge effectively, and the limited data interactions hinder the agent's ability to discriminate between normal and anomalous events accurately.

6. **LSTM-Transformer-DQN (Deep Q-Networks):** Similar to the previous approach, two autoencoder models are employed: an LSTM autoencoder for capturing sequential context using short windows and a Transformer autoencoder for modeling periodic patterns with longer windows. Reconstruction errors from both models are aggregated and normalized to derive binary anomaly labels. A custom Gym environment is defined based on these aggregated scores, and a DQN agent is subsequently trained to refine anomaly detection decisions. The framework's performance is evaluated by comparing DRL decisions with the aggregated anomaly labels.



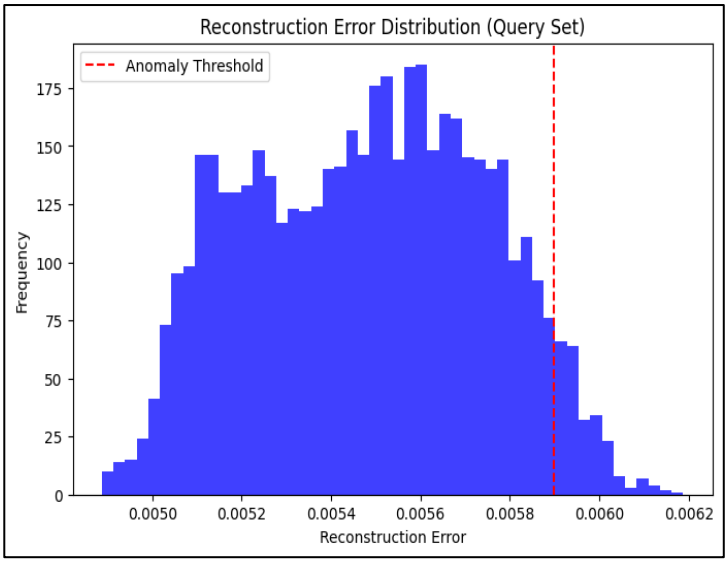
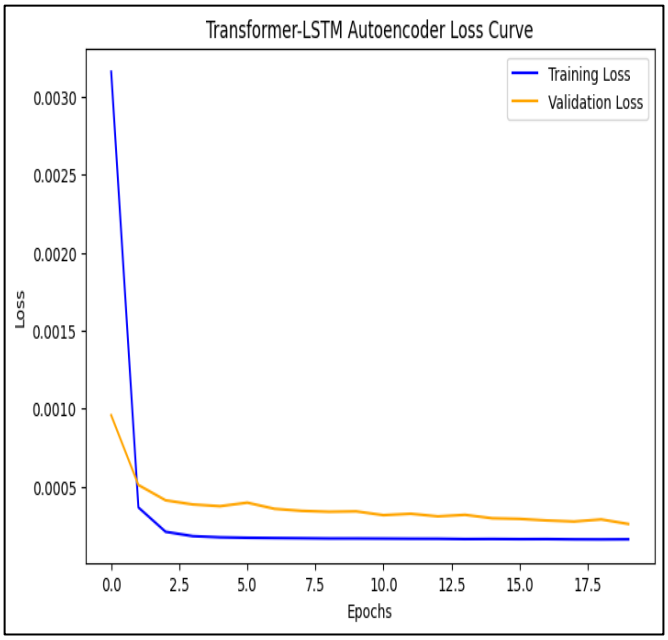


Fig 19. LSTM Training vs. Validation Loss Curve and Reconstruction Error Distribution



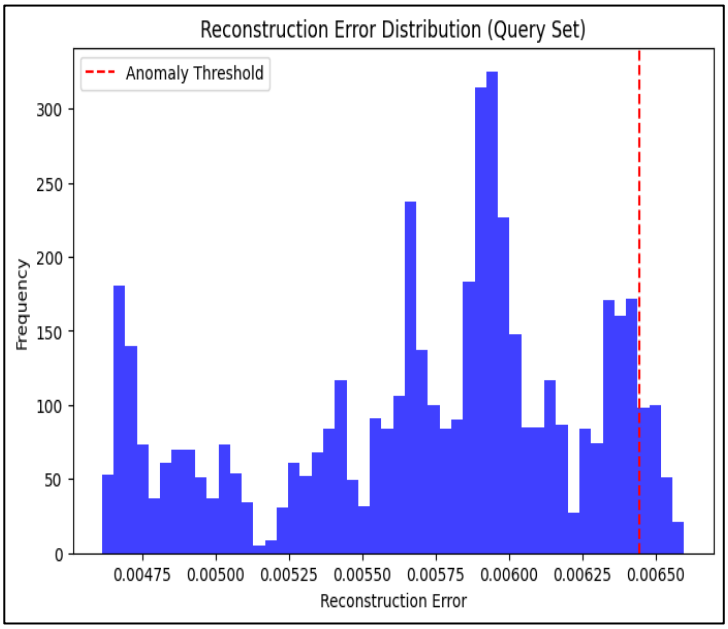


Fig 20. Transformer Training vs. Validation Loss Curve and Reconstruction Error Distribution

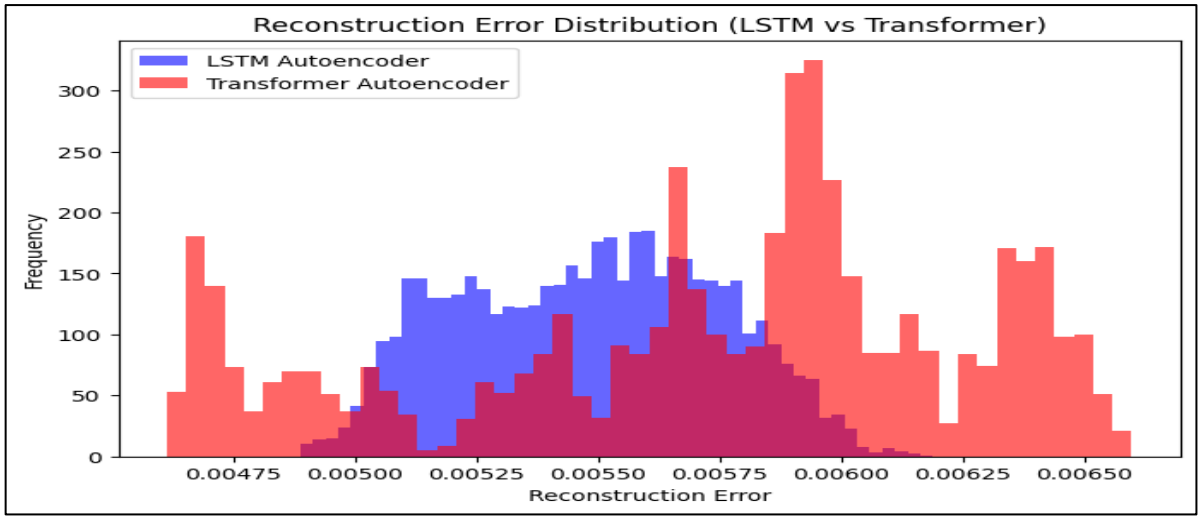


Fig 21. LSTM-Transformer Combined Reconstruction Error Distribution

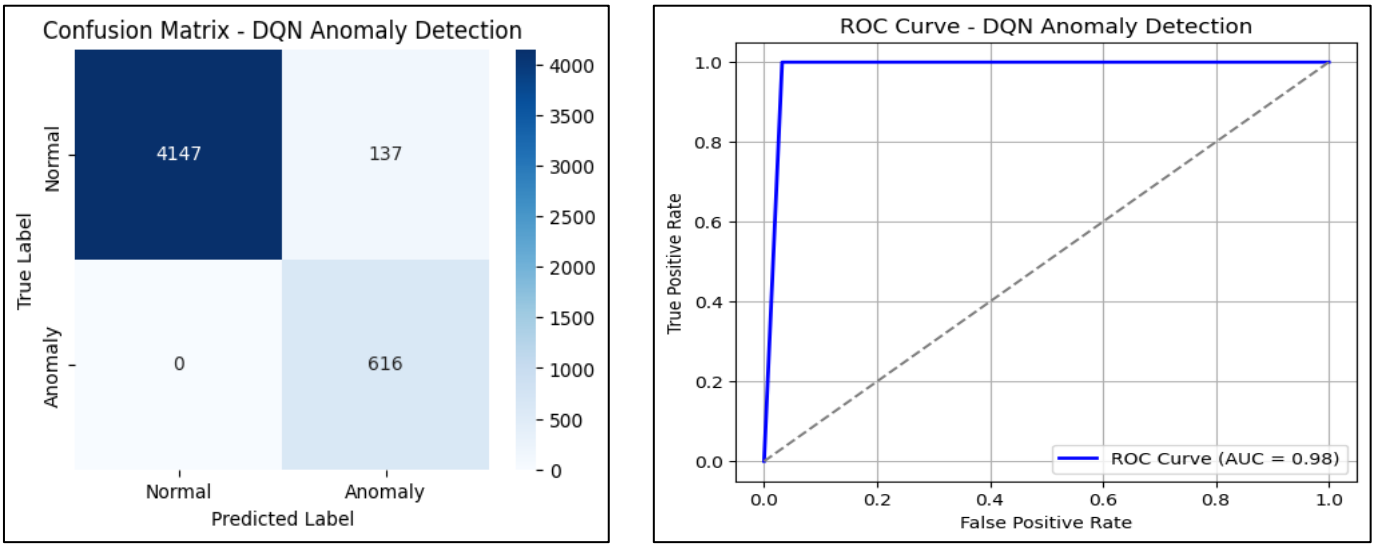


Fig 22. LSTM-Transformer-DQN Confusion Matrix and ROC Curve

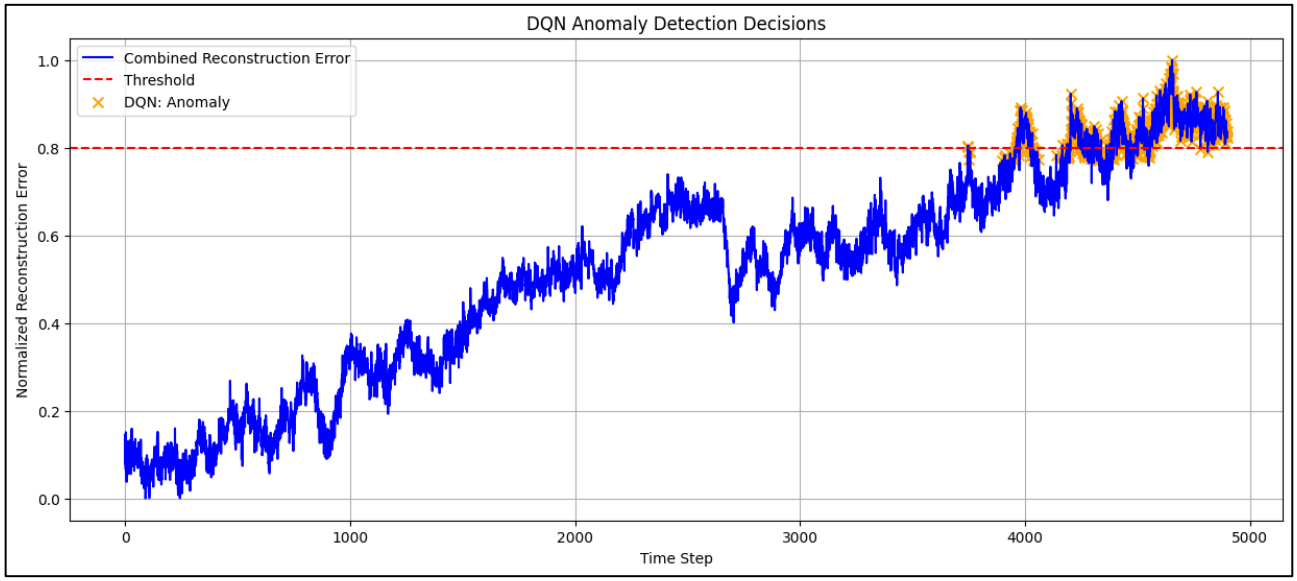


Fig 23. LSTM-Transformer-DQN Anomaly Detection Decisions

The LSTM-Transformer-DQN method achieved 753 detected anomalies with a classification accuracy of 97.20%. This superior performance is attributed to the stability of the DQN algorithm, which is employed as the reinforcement learning component in this hybrid framework. DQN exhibits enhanced sample efficiency and robust convergence properties in complex time-series environments compared to A2C. The hybrid design integrates LSTM and Transformer autoencoders for capturing sequential and periodic patterns, which ensures a comprehensive data representation, while DQN's stable training dynamics facilitate reliable anomaly detection decisions. Consequently, the LSTM-Transformer-DQN approach outperforms alternative methods by achieving a high accuracy rate and a consistent anomaly detection capability.

Conclusion

The increasing reliance on Wireless Body Area Networks (WBANs) for continuous health monitoring necessitates strengthened anomaly detection mechanisms to ensure data security and reliability. This study explored integrating deep learning and reinforcement learning techniques to enhance anomaly detection performance in WBANs. By leveraging LSTM and Transformer-based autoencoders alongside reinforcement learning agents, the proposed framework effectively captured both sequential and periodic dependencies in physiological data.

Experimental evaluations demonstrated that deep reinforcement learning-based models, particularly the LSTM-Transformer-DQN approach, significantly outperformed traditional methods regarding accuracy and robustness. The results underscore the effectiveness of combining sequential modeling with reinforcement learning-driven decision refinement, improving anomaly classification while minimizing false positives. Despite the promising results, specific challenges, such as computational complexity and sample efficiency, remain areas for further investigation.

Future research can explore lightweight architectures to enhance real-time applicability and investigate self-supervised learning methods to reduce reliance on labeled anomaly data. Extending this framework to multi-sensor fusion scenarios could improve adaptability in real-world healthcare applications. The findings of this study highlight the potential of deep reinforcement learning in WBAN anomaly detection, paving the way for more intelligent and adaptive health monitoring systems.

References

- Reiss, A., & Stricker, D. (2012). Introducing a New Benchmarked Dataset for Activity Monitoring. *2012 16th International Symposium on Wearable Computers*, 108-109.

- Pang, G., Van Den Hengel, A., Shen, C., & Cao, L. (2021). Toward Deep Supervised Anomaly Detection. *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining (KDD' 21)*. https://doi.org/10.1145/3447548.3467417\

- *Meta-AAD: Active Anomaly Detection with Deep Reinforcement Learning*. (2020, November 1). IEEE Conference Publication | IEEE Xplore. https://ieeexplore.ieee.org/document/9338270

- Albattah, A., & Rassam, M. A. (2022). A Correlation-Based anomaly detection model for wireless body area networks using convolutional Long Short-Term Memory neural network. *Sensors*, *22*(5), 1951. https://doi.org/10.3390/s22051951

- Dubey, K., & Hota, C. (2024). Anomaly detection in WBANs using CNN-Autoencoders and LSTMs. In *Lecture notes on data engineering and communications technologies* (pp. 187–197). https://doi.org/10.1007/978-3-031-57870-0_17

- *Deep Reinforcement Learning for Anomaly Detection: A Systematic Review*. (2022). IEEE Journals & Magazine | IEEE Xplore. https://ieeexplore.ieee.org/document/9956995

- Rassam, M.A. (2024). Autoencoder-Based Neural Network Model for Anomaly Detection in Wireless Body Area Networks. *IoT*.

- *Reinforcement learning based power control for In-Body sensors in WBANs against jamming*. (2018). IEEE Journals & Magazine | IEEE Xplore. https://ieeexplore.ieee.org/abstract/document/8395379

- Veerabaku, M. G., Nithiyanantham, J., Urooj, S., Quadir, A., MD, Sivaraman, A. K., & Tee, K. F. (2023). Intelligent Bi-LSTM with Architecture Optimization for Heart Disease Prediction in WBAN through Optimal Channel Selection and Feature Selection. *Biomedicines*, *11*(4), 1167. https://doi.org/10.3390/biomedicines11041167

- Rivlin, O., Hazan, T., & Karpas, E. (2020). Generalized Planning With Deep Reinforcement Learning. *CoRR*, *abs/2005.02305*. https://arxiv.org/abs/2005.02305

- *A continuous change detection mechanism to identify anomalies in ECG signals for WBAN-Based healthcare environments*. (2017). IEEE Journals & Magazine | IEEE Xplore. https://ieeexplore.ieee.org/document/7945242