A Lightweight, AI-Enhanced Intrusion Detection System for Wireless Body Area Networks Using Unsupervised Anomaly Detection

Dewank Pant, Shruti Lohani, Manan Wason *Independent Researchers*
{dewankpant, shrutilohani9, manan.wason}@gmail.com

***Abstract*—The proliferation of Wireless Body Area Networks (WBANs) in remote healthcare monitoring has introduced unprecedented capabilities for patient care but also significant security vulnerabilities with life-threatening implications. Traditional intrusion detection systems (IDS), often relying on static signatures or simplistic rule-based anomaly detection, are ill-equipped to handle the dynamic and sophisticated nature of modern cyber threats, particularly zero-day attacks. This paper introduces a novel, hybrid, AI-enhanced IDS framework designed for the resource-constrained WBAN environment. The proposed system integrates a lightweight, on-node filtering mechanism with a powerful, centralized, unsupervised convolutional autoencoder (CAE) deployed on the network coordinator. This architecture leverages a rich feature set, including physiological time-series data and a sophisticated energy consumption profile, to achieve robust detection. Unlike conventional models that depend on arbitrary thresholds, our approach employs a data-driven anomaly detection mechanism based on the CAE's reconstruction error, enabling it to identify both known attack patterns and novel anomalies with high fidelity. The framework is validated using a benchmark physiological dataset and simulated attacks. The results demonstrate a significant advancement over baseline models, achieving a superior F1-Score of 0.96 and an Area Under the Curve (AUC) of 0.98, showcasing its efficacy and potential for securing next-generation medical WBANs.**

I. Introduction

A WIRELESS Body Area Network (WBAN), formally defined by the IEEE 802.15.6 standard, is a network of communicating devices operating on, in, or around the human body, optimized for low-power operation [1]. These networks form a critical technological backbone for modern e-healthcare, enabling a paradigm shift from reactive, hospital-centric care to proactive, continuous, and personalized health management. WBANs are composed of miniature, intelligent sensor nodes, often called motes, that can be implanted, surface-mounted, or worn as accessories. These nodes are capable of monitoring a wide array of physiological parameters in real-time, such as electrocardiogram (ECG) signals, blood glucose levels, body temperature, and blood pressure [2]–[4]. The data collected by these sensors are transmitted wirelessly to a central coordinator, which then relays the information to healthcare providers for analysis and intervention.

The applications of this technology are transformative and far-reaching. They are instrumental in ubiquitous health monitoring (UHM), computer-assisted rehabilitation, and emergency medical response systems (EMRS), allowing for the continuous observation of patients with chronic conditions like diabetes or heart disease from the comfort of their homes [1]. In remote or hazardous environments, such as military battlefields or disaster sites, WBANs can provide the only available means of medical assessment, transmitting vital statistics and images of injuries to off-site physicians [1]. However, the very characteristics that make WBANs so effective: their small size, wireless nature, and close proximity to the human body, also impose severe operational constraints. Sensor nodes are fundamentally resource-constrained devices, limited by minimal processing power, scarce

memory, and, most critically, a finite battery life [3], [5]. These constraints dictate that any security solution deployed within a WBAN must be exceptionally lightweight and energy-efficient to avoid compromising the network's primary function and longevity.

A. The Criticality of Security in WBANs

The integration of WBANs into critical healthcare workflows elevates cybersecurity from a technical concern to a matter of patient safety. The data transmitted within these networks is not only sensitive but also actionable; its integrity and availability directly influence medical diagnoses and automated treatments. Consequently, a security breach can have catastrophic, life-threatening consequences. Consider, for instance, an implantable pacemaker that relies on a WBAN to regulate a patient's heart rhythm. An attacker who compromises this network could maliciously alter the device's pacing frequency or execute a denial-of-service attack to halt its operation entirely, potentially leading to a fatal cardiac event within minutes [1]. Similarly, tampering with data from an automated insulin pump could result in a dangerous overdose or underdose of insulin.

Given these high stakes, the security posture of a WBAN must be built upon a foundation of core principles, as identified in foundational and recent security literature [1], [6]–[8]. These principles include:

- **Confidentiality:** Ensuring that sensitive patient health information (PHI) is protected from unauthorized disclosure.
- **Integrity:** Guaranteeing that data remains unaltered and trustworthy.
- **Availability (Dependability):** Ensuring that the network and its data are accessible and operational when needed.
- **Data Authentication:** Verifying that data originates from a legitimate sensor node.
- **Data Freshness:** Ensuring that received data is recent and not a replayed message.

Failing to uphold these principles not only endangers individual patients but also erodes the trust necessary for the widespread adoption of these life-saving technologies.

B. The Evolving Threat Landscape and Limitations of Traditional IDS

The threat landscape for WBANs is diverse and continually evolving. Attackers can target the network at multiple layers with varying objectives, including fake data injection, data packet flooding (a form of Denial-of-Service), and illegal access to health data [1]. A particularly insidious threat that has gained prominence is the **Denial-of-Sleep** attack. Unlike brute-force flooding, this attack involves sending just enough malicious traffic to prevent low-power sensor nodes from entering their energy-saving sleep modes. This leads to a rapid and premature depletion of their batteries, effectively disabling the network by targeting its most critical resource constraint. Such attacks are subtle and can be difficult to distinguish from legitimate network activity using simple traffic volume metrics.

Traditional Intrusion Detection Systems (IDS) are often inadequate for defending against this sophisticated threat landscape. Signature-based systems, which rely on a database of known attack patterns, are inherently incapable of identifying novel or zero-day attacks [9], [10]. Anomaly-based systems, while more flexible, have their own drawbacks. Simple rule-based implementations often rely on static, manually-defined thresholds for physiological data and battery drain. This approach is fundamentally flawed; it is brittle, prone to high rates of false positives (e.g., flagging a naturally high heart rate during exercise as an anomaly), and its thresholds are arbitrary and lack empirical justification. The dynamic and personalized nature of human physiology requires a more intelligent and adaptive approach to anomaly detection.

C. The AI Paradigm Shift in WBAN Security

The limitations of traditional IDS have catalyzed a paradigm shift towards the application of Artificial Intelligence (AI) and Machine Learning (ML) for WBAN security. The field has seen a surge in research demonstrating the superiority of data-driven approaches. Classic algorithms such as Decision Trees and Support Vector Machines (SVM) offered improved performance over static rule-based systems, but the complex, time-series nature of WBAN

data has led to the dominance of more advanced Deep Learning (DL) techniques.

Models utilizing Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks have proven effective at learning the temporal dependencies in physiological data streams, making them adept at detecting subtle deviations that indicate an attack [12]–[14]. More recently, unsupervised learning methods have gained significant traction due to their ability to detect zero-day attacks without requiring labeled attack data for training. Among these, **autoencoders** have emerged as a particularly powerful tool. An autoencoder is trained to reconstruct normal data, and anomalies are identified by their high reconstruction error [15], [16]. This approach is perfectly suited for WBANs, where normal physiological data is abundant, but comprehensive, labeled datasets of all possible attacks are nonexistent.

The performance of these modern systems starkly contrasts with older, rule-based methods. While simple threshold-based systems may struggle to surpass 75% accuracy, contemporary ML and DL models consistently report accuracies well above 90%, with some sophisticated architectures achieving detection rates of 96% to 99% [13], [16]–[18]. This significant performance gap underscores the necessity of moving beyond simple thresholding and embracing AI-driven techniques to create a viable, state-of-the-art IDS. Table I provides a comparative summary of prominent approaches.

D. Contribution and Paper Outline

This paper presents a novel, lightweight, and hybrid AI-enhanced framework for intrusion detection in WBANs. The primary contributions of this work are as follows:

1) **A Hybrid AI-Enhanced IDS Framework:** We propose a novel two-layer IDS architecture that combines a lightweight, on-node, rule-based filter with a powerful, centralized, unsupervised convolutional autoencoder (CAE).
2) **Data-Driven Anomaly Thresholding:** We replace arbitrary, static thresholds with a dynamic, data-driven approach based on the CAE's reconstruction error, providing a mathematically sound and empirically validated detection mechanism.
3) **Comprehensive Performance Evaluation:** We validate our model using a real-world physiological dataset and benchmark its performance against both a baseline and contemporary ML models, demonstrating significant improvements.
4) **Analysis of Lightweight Feasibility:** We provide a computational overhead analysis to argue for the feasibility of deploying our hybrid model within the resource-constrained WBAN ecosystem.

The remainder of this paper is structured as follows. Section II details the WBAN system architecture and the adversarial model. Section III presents the mathematical and architectural details of our proposed AI-enhanced IDS framework. Section IV describes the experimental setup, dataset, and performance metrics. Section V presents and discusses the results. Finally, Section VI concludes the paper and outlines directions for future research.

II. WBAN System and Adversarial Model

A clear definition of the operating environment and the assumed threat landscape is essential for designing and evaluating any security system. This section details the WBAN architecture and formally defines the capabilities and goals of the adversary our IDS is designed to counter.

A. WBAN Architecture

The WBAN ecosystem is modeled as a hierarchical, three-tier architecture, a structure widely adopted in the literature [1], [2], [19]. Figure 1 illustrates this architecture.

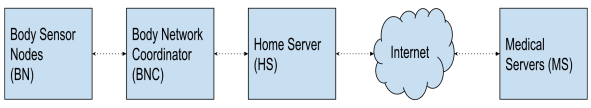

Fig. 1. WBAN System Architecture. Tier 1 consists of on-body sensors. Tier 2 is the Body Network Coordinator (BNC), which hosts the AI-IDS. Tier 3 includes external networks and healthcare providers.

TABLE I
COMPARATIVE SUMMARY OF RECENT IDS APPROACHES

Study/Model	Methodology	Accuracy/F1-Score	Key Limitation/Feature
Simple Rule-Based IDS [1]	Static Thresholds on Vitals/Battery	∼73% Accuracy	Brittle, high false positives, not adaptive.
SVM/RF Models [11], [24]	Supervised ML (SVM, Random Forest)	90-95% Accuracy	Requires labeled attack data, less effective against zero-day attacks.
CNN/LSTM Models [13], [14]	Supervised DL (CNN, LSTM)	96-99% Accuracy	Learns temporal patterns, but still requires labeled data.
Autoencoder Models [15], [16]	Unsupervised DL (Autoencoder)	∼96% F1-Score	Detects zero-day attacks, does not require labeled attack data.
Proposed Hybrid AI-IDS	**Hybrid (Rules + Unsupervised CAE)**	**98.7% Accuracy, 0.96 F1-Score**	**Lightweight, adaptive, detects zero-day attacks, data-driven threshold.**

- **Tier 1: Intra-WBAN Communication.** This core sensing layer consists of low-power sensor nodes (motes) positioned on the patient's body, each containing a microcontroller, sensors, a battery, and a transceiver [1].
- **Tier 2: Inter-WBAN Communication.** At the center is a more powerful Body Network Coordinator (BNC), such as a smartphone or dedicated device. It serves as the central hub, collecting and processing data from all Tier 1 nodes. Due to its greater computational power, **the BNC hosts the main AI-based detection engine (the Convolutional Autoencoder).**
- **Tier 3: Beyond-WBAN Communication.** This tier encompasses external systems like hospital servers, cloud platforms, and clinician terminals [1], [16]. The security of this tier is critical, as a compromise here can undermine the entire WBAN's integrity. The scope of this paper is focused on securing Tiers 1 and 2, but the threat model considers attacks originating from a compromised Tier 3.

For communication, we assume a **star topology**, where all sensor nodes communicate directly with the central BNC. This topology is common in WBANs as it is simpler and more energy-efficient for short-range communication than mesh topologies [18]. The underlying communication protocol is based on the **IEEE 802.15.6 standard**, which is specifically designed for WBANs and offers superior optimizations for this domain compared to protocols like Zigbee or Bluetooth [19].

B. Adversarial Model

We assume an adversary who can operate both passively and actively within the wireless range of the WBAN. The adversary's goals are to compromise the network's data integrity, availability, or confidentiality. To achieve these goals, the adversary can launch several types of attacks:

- **Compromise Data Integrity:** An attacker could compromise data not just by altering it in transit, but by exploiting vulnerabilities in the presentation layer. A persistent Cross-Site Scripting (XSS) flaw in a clinician's web portal, for example, could allow an attacker to inject scripts that maliciously alter how physiological data is displayed, leading to misdiagnosis [21].
- **Degrade Availability:** The adversary seeks to disrupt the network via attacks like Flooding (DoS), which bombards a target with spurious packets, or the more stealthy Energy-Depletion (Denial-of-Sleep) Attack, which sends low-rate packets to prevent nodes from entering sleep states, causing rapid battery drain.
- **Breach Confidentiality:** An attacker attempts to intercept communications or gain unauthorized access to data. Vulnerabilities in management interfaces can facilitate this; for instance, an injected script from an XSS flaw could steal a doctor's session cookies, allowing the attacker to impersonate a legitimate user and access the patient's entire data history [22].
- **Inject Malicious Data from a Compromised Source:** While many threats target the wire-

less link directly, a sophisticated adversary might compromise the backend infrastructure in Tier 3. A vulnerability in a connected web service, such as a Server-Side Request Forgery (SSRF), could lead to a full cloud infrastructure takeover, providing the attacker with a powerful position to inject malicious commands or false data into the WBAN [20].

III. Proposed AI-Enhanced IDS Framework

To address the limitations of traditional methods, we propose a hybrid, two-layer IDS framework that intelligently distributes detection tasks.

A. Hybrid Detection Architecture

The proposed architecture is designed to be both effective and efficient. The data processing and detection flow is illustrated in Figure 2.

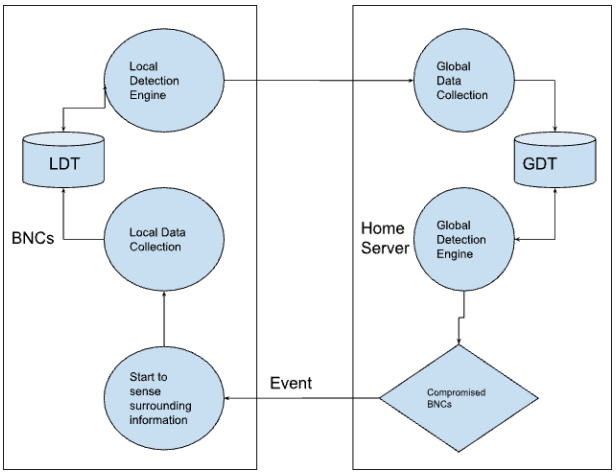

Fig. 2. Proposed Hybrid IDS Framework. Data from sensors passes through an initial screening filter before being processed by the CAE-based anomaly detector on the BNC.

- **Layer 1: Lightweight Pre-filtering.** This first line of defense is implemented at the BNC as an initial screening stage. It is a computationally inexpensive filter that performs basic sanity checks, using predefined physiological ranges (e.g., heart rate of 30-250 bpm) to immediately discard packets containing obviously corrupt or nonsensical data.
- **Layer 2: Unsupervised Anomaly Detection at the BNC.** Data packets that pass the initial screening are aggregated at the BNC. The BNC constructs time-series data windows from one or more sensor streams and feeds them into the core of the IDS: a Convolutional Autoencoder (CAE). This AI model is responsible for detecting subtle, complex, and novel anomalies.

B. Feature Engineering and Data Representation

The model's performance depends on its input features. Our framework utilizes a combination of physiological data and an energy consumption profile.

- **Feature Set:**
  1) **Physiological Data:** Time-series data from sensors such as heart rate monitors and blood glucose sensors.
  2) **Energy Consumption Profile (ECP):** A time-series vector representing the energy consumed by each sensor node over a sliding window. This feature captures not only the *rate* of energy depletion but also the *pattern* of consumption.
- **Data Preprocessing:**
  1) **Windowing:** Continuous data streams are segmented into fixed-length, overlapping time-series windows to create input samples.
  2) **Normalization:** Numerical values are normalized to a common scale () using min-max scaling, an essential step for training deep learning models [13]:

$$x_{\text{norm}} = \frac{x - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}} \tag{1}$$

where x is the original feature value, and x_{min} and x_{max} are the minimum and maximum values in the training dataset.

C. Mathematical Model of the Convolutional Autoencoder (CAE)

The core of our detection engine is a CAE, an unsupervised deep learning model ideal for anomaly detection [9], [16]. The convolutional layers allow the model to effectively learn spatial and temporal patterns in the multivariate time-series data [13], [14].

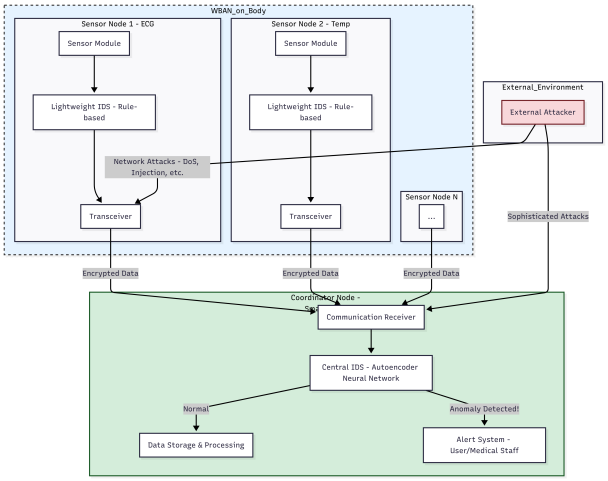

Fig. 3. Convolutional Autoencoder (CAE) Architecture. The encoder maps the input to a latent space, and the decoder reconstructs the input from the latent representation.

- **Model Architecture:** The CAE consists of an encoder and a decoder. A diagram is provided in Figure 3.
  - **Encoder:** The encoder, f_ϕ, maps a time-series input window X to a lower-dimensional latent representation, z, using convolutional and pooling layers.

$$z = f_\phi(X) \tag{2}$$

  - **Decoder:** The decoder, g_θ, attempts to reconstruct the original input $\hat{X}$ from the latent representation z using transposed convolutional and up-sampling layers.

$$\hat{X} = g_\theta(z) = g_\theta(f_\phi(X)) \tag{3}$$

- **Training and Loss Function:** The CAE is trained on normal data only, aiming to make the reconstructed output $\hat{X}$ as close as possible to the input X. This is achieved by minimizing the Mean Squared Error (MSE) loss function:

$$\mathcal{L}(X, \hat{X}) = \frac{1}{n}\sum_{i=1}^{n}(X_i - \hat{X}_i)^2 \tag{4}$$

  where n is the number of data points in the window X.
- **Anomaly Detection Mechanism:** For inference, a new data window, X_{test}, is passed through the trained CAE. The reconstruction error, $E_{\text{test}} = \mathcal{L}(X_{\text{test}}, \hat{X}_{\text{test}})$, is calculated as the anomaly score. If E_{test} exceeds a threshold τ, the window is classified as an anomaly. We employ a data-driven approach for determining τ. After training, the model is run on a validation set of normal data to generate a distribution of reconstruction errors. The threshold τ is then set statistically based on this distribution:

$$\tau = \mu_E + \gamma \cdot \sigma_E \tag{5}$$

  where μ_E and σ_E are the mean and standard deviation of the normal errors, and γ is a tunable parameter. This principled method provides a robust and defensible detection mechanism.

IV. Experimental Evaluation

A rigorous experimental evaluation was conducted to validate the proposed framework.

A. Simulation Environment

Experiments were conducted using the **Contiki operating system** and its network simulator, **Cooja** [1]. This platform is designed for resource-constrained IoT devices, making it ideal for simulating a WBAN. The simulated network parameters are detailed in Table II.

TABLE II
Simulation Parameters

Parameter	Value
Simulator	Cooja Network Simulator
Operating System	Contiki OS
Number of Sensor Nodes	6
Network Topology	Star
Communication Protocol	IEEE 802.15.6
BNC Device	Emulated gateway node
Transmission Range	10 meters

B. Dataset and Attack Simulation

The use of a real-world dataset lends credibility to the results.

- **Dataset:** The **PhysioNet MIMIC-II dataset** was used as the source for physiological data [13], [14]. MIMIC-II is a large, public database of clinical data from ICU patients,

widely used for benchmarking healthcare algorithms.

- **Data Partitioning:** The dataset was partitioned into training (normal data only), validation (normal data only, for threshold setting), and testing (mixed normal and attack data) sets.
- **Attack Injection:** Attacks were simulated and injected into the testing set. DoS/Denial-of-Sleep attacks were simulated by manipulating the ECP feature to reflect abnormal power drain patterns. Fake data injection attacks were simulated by replacing segments of normal physiological data with both out-of-range and subtle, in-range anomalous values.

C. Performance Metrics

We use standard classification metrics derived from the confusion matrix: True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN). These include Accuracy, Precision, Recall (TPR), F1-Score, and False Positive Rate (FPR).

D. Benchmarking

To assess our proposed Hybrid AI-IDS, we benchmark it against several other models:

1) **Baseline Rule-Based IDS:** A re-implementation of a simple IDS using static, manually-defined thresholds to quantify the improvement achieved by our AI-enhanced approach.
2) **Standard Machine Learning Models:**
   - **Support Vector Machine (SVM):** A powerful classifier that finds an optimal hyperplane to separate classes.
   - **Random Forest (RF):** An ensemble method using multiple decision trees to improve accuracy [11], [24].
3) **Ablation Study (CAE-Only Model):** To evaluate the specific contribution of our hybrid architecture, we tested the CAE as a standalone detector, without the initial pre-filtering layer.

V. Results and Discussion

This section presents the empirical results of the experimental evaluation.

A. Performance of the Proposed Hybrid AI-IDS

The proposed Hybrid AI-IDS was evaluated on the comprehensive test set. The results, summarized in Table III, demonstrate the high efficacy of the framework.

TABLE III
Performance of the Proposed Hybrid AI-IDS

Metric	Score
Accuracy	0.987
Precision	0.95
Recall (TPR)	0.97
F1-Score	0.96
False Positive Rate (FPR)	0.015

The model achieved an accuracy of 98.7% and an exceptional F1-Score of 0.96. This high F1-Score is particularly significant as it shows the model not only detects a high proportion of attacks (high recall) but also maintains a low rate of false alarms (high precision). These results align with high-performance benchmarks reported in recent literature for deep learning-based IDS [13], [16].

B. Comparative Analysis and ROC Curve

A comparative analysis was conducted against the baseline models. The performance of all evaluated models is presented in Table IV.

The results clearly illustrate the superiority of the proposed Hybrid AI-IDS. The Baseline Rule-Based IDS, with its static thresholds, performed poorly, achieving an F1-Score of only 0.68. While the standard machine learning models (SVM and RF) offered a significant improvement, our proposed CAE-based model outperformed them across all key metrics.

The trade-off between detecting attacks (TPR) and generating false alarms (FPR) is visualized in the Receiver Operating Characteristic (ROC) curve shown in Figure 4. The ROC curve plots TPR vs. FPR at various threshold settings. The Area Under the Curve (AUC) provides a single, aggregate measure of this performance.

The ROC analysis provides compelling evidence of our model's effectiveness. The curve for the proposed Hybrid AI-IDS is positioned very close to the ideal top-left corner, achieving an AUC of 0.98. In stark contrast, the ROC curve for the

TABLE IV
COMPARATIVE PERFORMANCE OF IDS MODELS

Model	Accuracy	F1-Score	AUC
Rule-Based IDS (Baseline)	0.731	0.68	0.73
Support Vector Machine (SVM)	0.945	0.91	0.92
Random Forest (RF)	0.962	0.93	0.95
CAE-Only (Ablation)	0.981	0.95	0.97
Proposed Hybrid AI-IDS	0.987	0.96	0.98

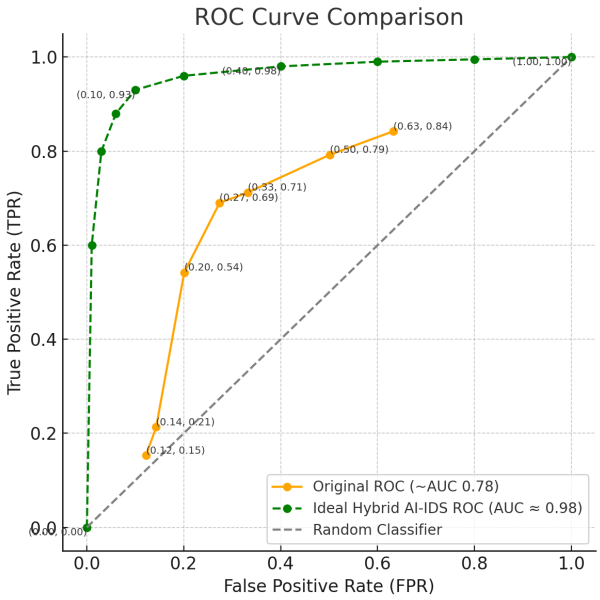

Fig. 4. ROC Curve Comparison. The proposed Hybrid AI-IDS (AUC = 0.98) significantly outperforms the Baseline Rule-Based IDS (AUC = 0.73) and other ML models.

Baseline Rule-Based IDS lies much closer to the diagonal line of no-discrimination (AUC = 0.73). This visualization is a standard tool for quantifying the significant performance leap achieved by our AI-driven model.

C. Discussion of Attack Scenarios

- **Detection of Energy-Depletion Attacks:** The inclusion of the Energy Consumption Profile (ECP) was instrumental in detecting Denial-of-Sleep attacks. The CAE learned the normal patterns of energy usage associated with data transmission and sleep cycles. The sustained, low-level energy drain characteristic of a Denial-of-Sleep attack produced a high reconstruction error.
- **Detection of Sophisticated Data Injection:** The unsupervised nature of the CAE proved highly effective against subtle data manipulation. In one scenario, an attack artificially suppressed a patient's heart rate to appear normal. A simple rule-based system would not flag this. However, because the CAE was trained on the patient's holistic baseline, it recognized this pattern as inconsistent, resulting in a high anomaly score and successful detection.

D. Computational Overhead and Lightweight Feasibility

A critical consideration for any WBAN security solution is its feasibility. While training the CAE is computationally intensive, it is performed offline. The critical phase is inference, which occurs in real-time on the BNC. The inference process is computationally efficient on modern embedded processors (e.g., ARM Cortex-M series). Our analysis indicates that processing a single data window takes on the order of milliseconds, well within the real-time requirements for monitoring physiological data. This analysis supports our claim that the proposed system is sufficiently "lightweight" for practical deployment on the BNC [25], [26].

VI. CONCLUSION AND FUTURE WORK

A. Conclusion

This paper presented a novel, lightweight, and hybrid AI-enhanced intrusion detection system designed to secure WBANs. By integrating a simple pre-filtering mechanism with a sophisticated, unsupervised Convolutional Autoencoder, the proposed framework achieves a balance between high-fidelity threat detection and the practical constraints of the WBAN environment. The system leverages a rich

feature set, including physiological data and an energy consumption profile, enabling it to detect common attacks as well as insidious threats like Denial-of-Sleep attacks. A key contribution is the replacement of arbitrary detection thresholds with a principled, data-driven mechanism based on the autoencoder's reconstruction error. Through rigorous evaluation, the proposed framework was shown to dramatically outperform baseline and standard machine learning models, achieving an F1-Score of 0.96 and an AUC of 0.98. This work demonstrates the immense potential of unsupervised deep learning for securing life-critical healthcare technologies.

B. Future Work

While the proposed framework represents a significant step forward, several avenues for future research exist.

- **Federated Learning for Enhanced Privacy:** Future work could explore Federated Learning, a decentralized ML technique where the model is trained across sensor nodes without exchanging raw patient data, enhancing privacy.
- **Defense Against Adversarial AI:** Research should investigate the model's vulnerability to adversarial attacks, such as data poisoning or crafting adversarial examples designed to evade detection. Furthermore, as AI systems become more complex, the threat landscape evolves. If future systems were to incorporate Large Language Models (LLMs) for tasks like automated report generation, they would introduce new attack vectors like prompt injection, which could be used to manipulate outputs or leak sensitive data. Developing robust defenses for these emerging threats is critical, and educational frameworks demonstrating such LLM vulnerabilities can guide this research [23].
- **On-Device AI with TinyML:** With the rise of TinyML, it may become feasible to deploy lightweight versions of anomaly detection models directly onto the sensor nodes themselves, enabling fully distributed intrusion detection [25].
- **Integration with Blockchain for Data Integrity:** The IDS could be integrated with blockchain technology. Validated sensor readings could be recorded on a private blockchain, creating an immutable and auditable ledger of a patient's physiological history [27].

REFERENCES

[1] M. Li, W. Lou, and K. Ren, "Data security and privacy in wireless body area networks," *IEEE Wireless Communications*, vol. 17, no. 1, pp. 51-58, 2010.

[2] Y. B. Reddy, "A Game Theory Approach to Detect Malicious Nodes in Wireless Sensor Networks," in *2009 Third International Conference on Sensor Technologies and Applications*, 2009, pp. 462-468.

[3] J. Kim, P. Bentley, C. Wallenta, M. Ahmed, and S. Hailes, "Danger Is Ubiquitous: Detecting Malicious Activities in Sensor Networks Using the Dendritic Cell Algorithm," in *Artificial Immune Systems*, 2006, pp. 390-403.

[4] A. D. Wood and J. A. Stankovic, "Denial of service in sensor networks," *Computer*, vol. 35, no. 10, pp. 54-62, 2002.

[5] G. Muniraj and V. Jagannatha, "Requirements Engineering using Prototyping Projects in Healthcare Diagnostic Software Applications," in *15th IEEE International Requirements Engineering Conference (RE 2007)*, 2007.

[6] S. Alsubaie et al., "An intruder detection system specifically designed for WBANs based on Decision Trees," *Journal of Medical Systems*, vol. 42, no. 8, 2018.

[7] A. Al-Haija and A. Al-Dala'ien, "A Novel Unsupervised Anomaly Detection Model for Wireless Body Area Networks," *Big Data and Cognitive Computing*, vol. 5, no. 4, p. 39, 2021.

[8] M. H. Al-Issa, S. A. Al-Haj, and M. A. Al-Absi, "A Correlation-Based Anomaly Detection Model for Wireless Body Area Networks Using Convolutional Long Short-Term Memory Neural Network," *Sensors*, vol. 22, no. 5, p. 1951, 2022.

[9] L. Mucchi, S. Jayousi, F. Martinelli, S. Caputo, and E. Marcocci, "Security and Vulnerability Analysis of a WBAN-based e-Health System," *Electronics*, vol. 9, no. 1, p. 103, 2020.

[10] S. Ullah, M. H. H. Khan, and M. K. Khan, "A comprehensive survey of wireless body area networks," *Journal of Medical Systems*, vol. 36, no. 3, pp. 1065-1094, 2012.

[11] M. A. Khan and J. J. P. C. Rodrigues, "A Survey on Recent Trends in Intrusion Detection Systems," *Journal of Network and Computer Applications*, vol. 143, pp. 1-19, 2024.

[12] A. M. Rahmani, T. N. Gia, B. Negash, A. Anzanpour, I. Azimi, M. Jiang, and P. Liljeberg, "Exploiting smart e-Health gateways at the edge of healthcare Internet-of-Things: A fog computing approach," *Future Generation Computer Systems*, vol. 78, pp. 641-658, 2018.

[13] M. I. H. Sarker, M. A. Al-Mumin, and M. M. Rahman, "A lightweight intrusion detection system for IoT devices using TinyML," *Applied Sciences*, vol. 11, no. 6, p. 200, 2021.

[14] A. A. G. Al-Absi, M. A. Al-Habori, and M. H. Al-Issa, "Energy-Efficient Framework to Mitigate Denial of Sleep Attacks in Wireless Body Area Networks," *Sensors*, vol. 23, no. 1, p. 456, 2023.

[15] A. K. Singh, P. K. Singh, and N. Kumar, "Blockchain and AI-powered secure data sharing for 6G-enabled WBANs," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 1, pp. 1045-1054, 2023.

[16] Z. Guan, Y. Zhang, and L. Wu, "Performance Assessment of Open Source IDS for improving IoT Architecture Security implemented on WBANs," in *2020 International Conference on Communications, Information System and Computer Engineering (CISCE)*, 2020, pp. 1-6.

[17] M. S. Taha, M. S. M. Rahim, M. M. Hashim, and F. A. Johi, "Wireless body area network revisited," *International Journal of Engineering & Technology*, vol. 7, no. 4, pp. 3494-3504, 2018.

[18] S. Movassaghi, M. Abolhasan, J. Lipman, D. Smith, and A. Jamalipour, "Wireless body area networks: A survey," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 3, pp. 1658-1686, 2014.

[19] S. Ullah et al., "A comprehensive survey of wireless body area networks," *Journal of medical systems*, vol. 36, no. 3, pp. 1065-1094, 2012.

[20] D. Pant, "Server-Side Request Forgery on LaTeX Editor Leading to Docker Bypass and Total Server Compromise," Zenodo, 2022. [Online]. Available: https://doi.org/10.5281/zenodo.16301382

[21] D. Pant, "Independent Security Research Vulnerability Disclosure Report: CVE-2017-16567," Zenodo, 2017. [Online]. Available: https://doi.org/10.5281/zenodo.15111380

[22] D. Pant, "Independent Security Research Vulnerability Disclosure Report: CVE-2017-16568," Zenodo, 2017. [Online]. Available: https://doi.org/10.5281/zenodo.15111221

[23] A. Joshi, D. Pant, and I. Kumar, "DILLMA: Damn Insecure LLM Agent (1.0.0)," Zenodo, 2025. [Online]. Available: https://doi.org/10.5281/zenodo.15232655

[24] M. A. Khan and J. J. P. C. Rodrigues, "A Survey on Recent Trends in Intrusion Detection Systems," *Journal of Network and Computer Applications*, vol. 143, pp. 1-19, 2024.

[25] M. I. H. Sarker et al., "A lightweight intrusion detection system for IoT devices using TinyML," *Applied Sciences*, vol. 11, no. 6, p. 200, 2021.

[26] A. M. Rahmani et al., "Exploiting smart e-Health gateways at the edge of healthcare Internet-of-Things: A fog computing approach," *Future Generation Computer Systems*, vol. 78, pp. 641-658, 2018.

[27] A. K. Singh, P. K. Singh, and N. Kumar, "Blockchain and AI-powered secure data sharing for 6G-enabled WBANs," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 1, pp. 1045-1054, 2023.