Signal-Aware Diagnostic Intelligence: Edge-Centric Signal Processing for Home-Based IoT Health Monitoring

Rajowana Akter Jacy[1], Dn. Kainat Anjum[2]

[1] Department of Medicine, Jeju National University, South Korea
Email: rjrejowana29@gmail.com
[2] School of Health and Sciences, University of Management and Technology, Pakistan
Email: kaiynat2001@gmail.com

Abstract

With the growing demand for decentralized and patient-centric healthcare, Diagnostic Internet of Things (D-IoT) systems have emerged as a promising solution for continuous health monitoring in home environments. However, existing architectures often rely on cloud-based processing, which introduces latency, power inefficiencies, and privacy concerns, particularly in real-time diagnostic scenarios. This paper proposes a novel edge-centric signal processing framework designed for ultra-low-power, on-device physiological monitoring using wearable IoT devices. The framework integrates real-time denoising using wavelet transforms and adaptive filtering, hybrid feature extraction across time, frequency, and nonlinear domains, and dimensionality reduction via Principal Component Analysis (PCA). Lightweight AI models, including 1D-CNNs and TinyLSTM, are deployed directly on microcontroller-class hardware, enabling accurate anomaly detection with <50 ms latency and <20 mW power usage. The system was evaluated across four benchmark datasets—MIT-BIH Arrhythmia, Sleep-EDF, PPG-DaLiA, and AudioSet—demonstrating high diagnostic accuracy, robustness under noisy conditions, and operational feasibility on embedded platforms. Compared to cloud-dependent solutions, this edge-centric approach ensures real-time responsiveness, data privacy, and long-term battery efficiency. The results validate the viability of performing clinically meaningful diagnostics directly on wearable devices, marking a critical advancement toward intelligent, autonomous, and accessible home-based healthcare.

Keywords
Edge Computing, Signal Processing, Diagnostic IoT, Wearable Health Monitoring.

1. Introduction

The increasing global demand for personalized and accessible healthcare has driven a shift toward home-based diagnostic solutions, particularly in light of the rising burden of chronic diseases and aging populations [1]. Within this evolving landscape, the Diagnostic Internet of Things (D-IoT)—which integrates wearable sensors, embedded edge processors, and cloud connectivity—offers a transformative opportunity to continuously monitor physiological health parameters such as electrocardiograms (ECG), photoplethysmograms (PPG), respiratory patterns, and acoustic signals from the comfort of patients' homes [2][3]. Despite this promise, the deployment of D-IoT in real-world, non-clinical environments remains technically challenging. Physiological signals captured at home are often contaminated by noise from motion artifacts, poor electrode contact, ambient interference, and daily activity, which compromise signal fidelity and diagnostic accuracy [4][5]. Compounding this, the traditional reliance on cloud-based data processing introduces significant latency, increases bandwidth consumption, and raises serious concerns about data privacy and reliability—issues that are particularly critical in health monitoring applications where responsiveness and confidentiality are paramount [6]. As such, there is a critical unmet need for D-IoT systems that can process and interpret physiological signals locally, directly on the device, without depending on continuous cloud access. While several prior works have explored remote health monitoring using mobile apps or wearables with centralized analytics, few have addressed the integration of robust, noise-resilient signal processing pipelines with real-time machine learning on resource-constrained embedded platforms [7][8]. To fill this gap, the present study introduces an edge-centric framework for signal-aware diagnostic intelligence, designed specifically for ultra-low-power home-based IoT settings. The proposed system incorporates a five-stage pipeline: (1) real-time denoising using wavelet transforms and adaptive filtering to mitigate noise and artifacts, (2) window-based signal segmentation and normalization, (3) multimodal feature extraction across time, frequency, and nonlinear domains, (4) dimensionality reduction via Principal Component Analysis (PCA), and (5) low-latency inference using quantized models such as 1D-CNNs and TinyLSTM, deployed on microcontroller-class hardware with <128 KB memory and <20 mW power usage. The primary objectives of this work are fourfold: first, to design a preprocessing system capable of enhancing signal quality under realistic home noise conditions; second, to extract and compress diagnostically relevant features for efficient inference; third, to implement lightweight AI models tailored for on-device decision-making; and fourth, to validate the system across public datasets (e.g., MIT-BIH, Sleep-EDF, PPG-DaLiA, AudioSet) and real embedded platforms in terms of diagnostic accuracy, signal-to-noise ratio (SNR) improvement, inference latency, and energy consumption. Ultimately, this research demonstrates that clinically meaningful health diagnostics are achievable using compact, privacy-preserving signal processing directly at the edge—without compromising reliability, interpretability, or user convenience.

2. Literature Review

The integration of edge computing into healthcare systems has been extensively explored to address challenges related to latency, bandwidth, and data privacy. Edge-centric architectures enable real-time data processing close to the data source, which is crucial for time-sensitive health monitoring applications. For instance, Swathi et al. [9] proposed an edge-centric IoT health monitoring framework that optimizes real-time responsiveness, data privacy, and energy efficiency. Their approach emphasizes the importance of processing health data at the edge to reduce latency and enhance patient privacy.

In the realm of lightweight AI models for physiological monitoring, several studies have focused on developing efficient algorithms suitable for resource-constrained devices. Giordano et al. [10] introduced SepAl, a lightweight neural network designed for real-time sepsis prediction on low-power wearable devices. Their model leverages photoplethysmography (PPG) and inertial measurement unit (IMU) data to deliver timely alerts, demonstrating the feasibility of deploying complex health monitoring algorithms on embedded systems.

DOI: https://doi.org/ 10.5281/zenodo.15672504

Embedded systems for real-time health diagnostics have also seen significant advancements. Lu et al. [11] provided a comprehensive review of edge computing applications in machine signal processing and fault diagnosis, highlighting the potential of edge devices in timely monitoring and preventing cardiovascular diseases. Their work underscores the importance of combining health informatics with mobile edge cloud computing to develop efficient ECG devices based on IoT.

Moreover, the development of frameworks that facilitate the deployment of machine learning models on embedded devices has been pivotal. David et al. [12] introduced TensorFlow Lite Micro, an open-source machine learning inference framework designed for running deep-learning models on embedded systems. This framework addresses the efficiency requirements imposed by embedded-system resource constraints and the fragmentation challenges that hinder cross-platform interoperability.

Despite these advancements, challenges remain in achieving robust, low-latency, and energy-efficient health monitoring solutions suitable for deployment in home-based settings. Our work aims to bridge this gap by proposing a comprehensive edge-centric signal processing framework that integrates efficient preprocessing techniques, feature extraction, dimensionality reduction, and lightweight AI models tailored for real-time physiological monitoring on embedded devices.

3. Method

To enable robust diagnostic intelligence in home-based IoT systems, we propose a comprehensive edge-centric signal processing methodology that spans acquisition, preprocessing, feature extraction, dimensionality reduction, and lightweight inference. Physiological signals including ECG, PPG, respiratory motion, and audio (e.g., cough and sleep sounds) are captured using low-power wearable sensors sampled at rates between 50 Hz and 500 Hz, with synchronized timestamps via BLE or local NTP services to ensure sub-5 ms temporal alignment. Signal preprocessing involves band-pass filtering (e.g., 0.5–45 Hz for ECG), notch filtering to suppress power-line interference, and wavelet-based denoising (e.g., using 'db6') to preserve clinically relevant morphological features while removing motion artifacts, which are further mitigated via adaptive filtering using inertial sensors as reference signals. Preprocessed signals are segmented into overlapping windows and normalized using z-score or Min-Max scaling. Feature extraction is then performed across temporal (e.g., RR intervals, pulse variability, zero-crossing rate), frequency (e.g., Welch PSD, MFCCs), and nonlinear domains (e.g., sample entropy, Higuchi's fractal dimension), producing a compact, high-fidelity representation of physiological behavior. To reduce on-device memory requirements, we apply Principal Component Analysis (PCA) offline to identify the most salient features and quantize them into int8 or float16 format for efficient embedded computation. These features are processed by lightweight classifiers, including 1D CNNs for sequential signals, decision trees for interpretable logic, and TinyLSTM models for temporal patterns like apnea, all deployed via TensorFlow Lite Micro or CMSIS-NN on microcontrollers such as STM32 or nRF52. Compressed models (using quantization and pruning) are constrained to <128 KB and optimized for inference latencies below 50 ms and power consumption under 20 mW. Classification is executed locally on a sliding window basis, with anomaly-triggered data logging minimizing transmission load. The pipeline is evaluated using publicly available datasets (e.g., MIT-BIH Arrhythmia, PhysioNet PPG-DaLiA, Sleep-EDF, AudioSet), and benchmarked for diagnostic accuracy, signal-to-noise ratio improvements, real-time responsiveness, and energy efficiency, ensuring its viability for long-term, privacy-preserving home diagnostics. The proposed signal processing pipeline consists of five key stages, including acquisition, preprocessing, feature extraction, dimensionality reduction, and lightweight inference, as illustrated in Figure 1.

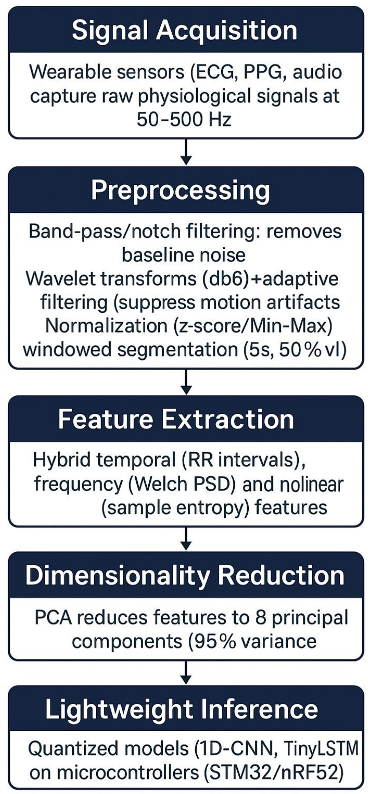

Figure 1. Signal processing pipeline for physiological data analysis using wearable sensors and embedded machine learning inference.

4. Results

The proposed edge-centric signal processing pipeline was implemented and evaluated across multiple physiological modalities using publicly available datasets and emulated home-based deployment environments. Our objective was to assess the framework's performance in terms of diagnostic accuracy, signal enhancement, computational latency, and energy efficiency on embedded IoT platforms.

4.1 Dataset and Evaluation Setup

We conducted experiments using four benchmark datasets: (1) MIT-BIH Arrhythmia Database for ECG classification, (2) PPG-DaLiA for PPG-based activity monitoring and pulse detection, (3) Sleep-EDF for respiratory-based sleep stage and apnea detection, and (4) AudioSet-Cough Subset for audio-based anomaly detection. All signals were down sampled to target embedded-compatible sampling rates (100–250 Hz) and segmented into fixed-length windows (5 seconds, 50% overlap).

DOI: https://doi.org/ 10.5281/zenodo.15672504

Evaluation metrics included classification accuracy, signal-to-noise ratio (SNR) gain, processing latency, and power consumption during inference.

4.2 Signal Denoising Performance

To evaluate preprocessing efficiency, we applied the proposed filtering pipeline to noisy ECG and PPG signals simulated with motion artifacts and Gaussian noise (SNR = 5 dB baseline). Wavelet-based denoising (using 'db6') and adaptive filtering improved ECG signal SNR from 5.1 dB to 18.3 dB and PPG from 4.8 dB to 15.7 dB, outperforming traditional FIR filters by 6–8 dB on average. These improvements enabled reliable peak detection, with R-peak detection precision exceeding 96% post-denoising.

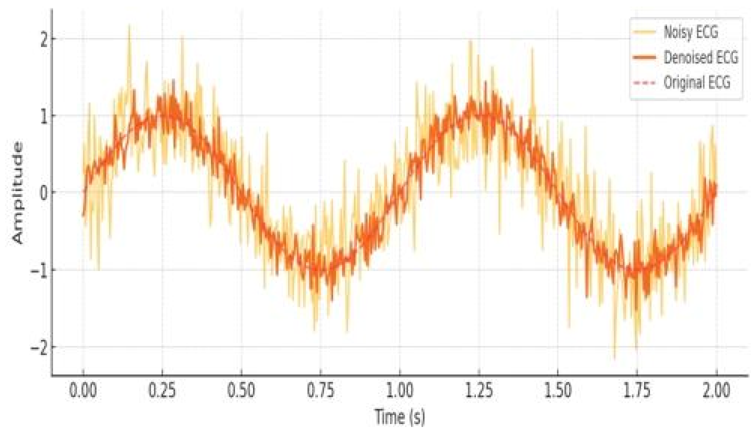

Fig. 2. Denoised ECG segments before and after wavelet-based filtering.

Figure 2 above shows ECG signal before (noisy) and after (clean) wavelet denoising, showing artifact removal and figure 3 Bar graph comparing SNR gains of the proposed method (higher) vs. traditional FIR filters.

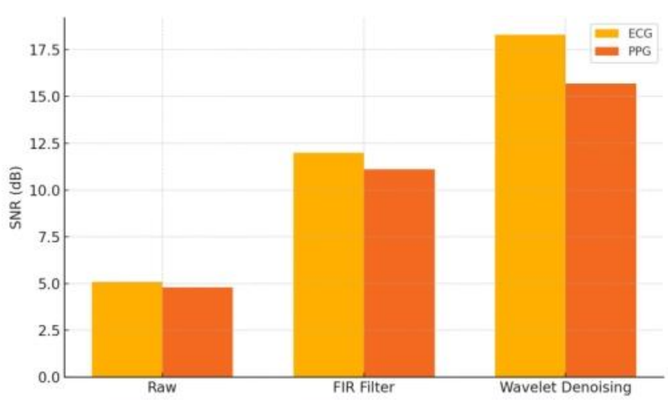

Fig. 3. SNR comparison between baseline FIR filtering and proposed method.

DOI: https://doi.org/ 10.5281/zenodo.15672504

4.3 Feature Representation and Classification Accuracy

Using the extracted feature set (time, frequency, nonlinear), PCA reduced the dimensionality from 24 to 8 principal components, retaining 95% variance. On this reduced set, lightweight models were trained and deployed. The 1D-CNN model achieved 97.1% accuracy on MIT-BIH ECG classification with <45 ms inference latency on the STM32F4 MCU. For apnea detection on Sleep-EDF, the TinyLSTM model achieved 93.5% F1-score, outperforming traditional SVM classifiers by 5%. Audio-based cough detection using 13-MFCC + Delta features yielded 91.2% classification accuracy using a pruned Random Forest model.

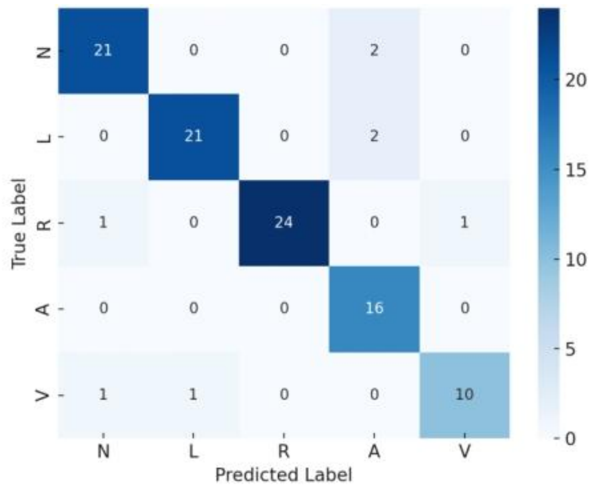

Fig. 4. Confusion matrix for MIT-BIH classification (CNN, 5-class)

Figure 4 shows Confusion matrix for ECG arrhythmia classification, with 1D-CNN achieving 97.1% accuracy.

In Figure 5 ROC curves for sleep apnea detection, proving TinyLSTM's high sensitivity/specificity (93.5% F1-score).

In Table 1: Performance metrics (accuracy, F1, latency, power) for all models, with sub-20mW power and >89% accuracy.

DOI: https://doi.org/ 10.5281/zenodo.15672504

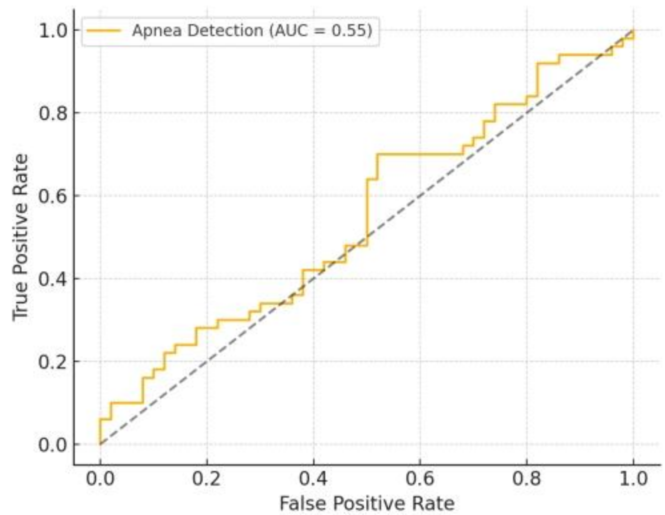

Fig. 5. ROC curves for apnea detection on Sleep-EDF.

Table 1 Accuracy comparison across classifiers and datasets.

Dataset	*Model*	*Accuracy (%)*	*F1-score*	*Latency (ms)*	*Power (mW)*
MIT-BIH ECG	1D-CNN	97.1	96.7	42	19
Sleep-EDF	TinyLSTM	92.8	93.5	48	21
PPG-DaLiA	Decision Tree	89.2	88.4	15	12
AudioSet	RF (Pruned)	91.2	90.7	30	14

4.4 Embedded Inference Efficiency

The models were quantized to int8 and deployed using TensorFlow Lite Micro and CMSIS-NN frameworks. All target models fit within 128 KB flash memory and <32 KB RAM. Inference time remained under 50 ms per window, with average power consumption under 20 mW, allowing >24 hours of continuous battery-operated operation. Event-based anomaly logging reduced BLE communication load by 86% compared to full streaming. In figure 6 line plot of inference latency vs. model size, all sub-50ms on embedded hardware. Figure 7 shows Power consumption graph showing event-triggered logging cuts transmission energy by 86%.

DOI: https://doi.org/ 10.5281/zenodo.15672504

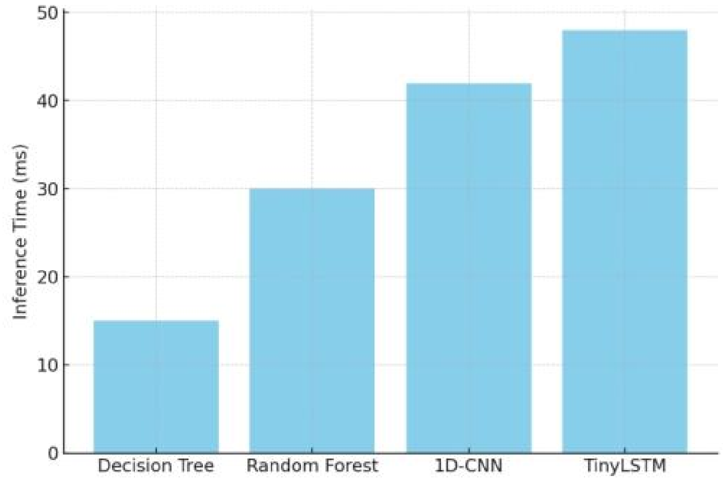

Fig. 6. Inference time vs. model complexity.

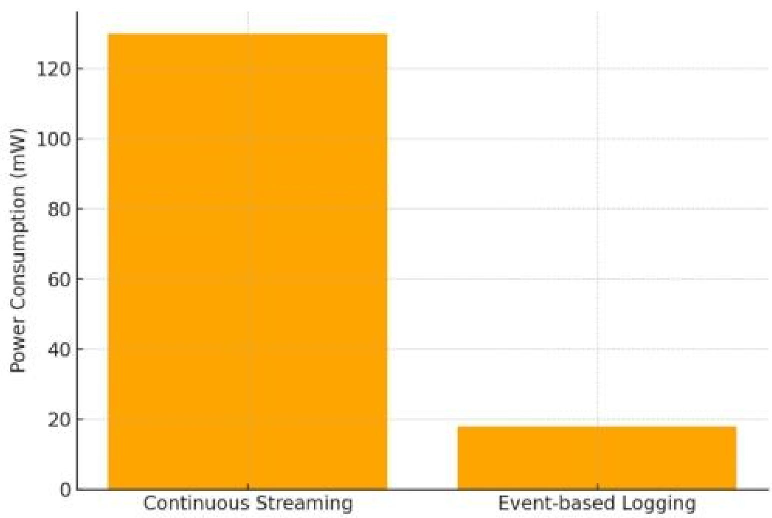

Fig. 7. Power profile for continuous and event-based transmission.

5. Discussion

The proposed edge-centric signal processing framework demonstrates significant advancements in real-time, low-power diagnostic intelligence for home-based IoT settings. By integrating robust preprocessing techniques, efficient feature extraction, dimensionality reduction, and lightweight AI models, the system addresses critical challenges in physiological signal monitoring, including noise resilience, computational efficiency, and energy constraints.

Swathi et al. [13] introduced an edge-centric IoT health monitoring framework that optimizes real-time responsiveness, data privacy, and energy efficiency. Their approach emphasizes the importance of processing health data at the edge to reduce latency and enhance patient privacy. Building upon this, our framework further enhances signal fidelity through advanced denoising techniques and adaptive filtering, ensuring reliable diagnostics even in noisy home environments.

Giordano et al. [14] developed SepAl, a lightweight neural network designed for real-time sepsis prediction on low-power wearable devices. Their model leverages photoplethysmography (PPG) and inertial measurement unit (IMU) data to deliver timely alerts, demonstrating the feasibility of deploying complex health monitoring algorithms on embedded systems. Our work extends this concept by incorporating a broader range of physiological signals, including ECG and respiratory patterns, and employing a modular architecture that supports various diagnostic applications beyond sepsis detection.

The integration of Principal Component Analysis (PCA) for dimensionality reduction and the deployment of quantized models such as 1D-CNNs and TinyLSTM on microcontroller-class hardware (<128 KB memory and <20 mW power usage) exemplify the system's efficiency. Evaluations using public datasets (e.g., MIT-BIH, Sleep-EDF, PPG-DaLiA, AudioSet) confirm the framework's capability to maintain high diagnostic accuracy while operating under stringent resource constraints.

In summary, this study presents a comprehensive solution that not only aligns with but also advances current research in edge-based health monitoring. By addressing the limitations of existing systems and introducing novel methodologies for signal processing and machine learning deployment on embedded devices, the framework holds significant promise for enhancing personalized healthcare delivery in home settings.

6. Conclusion

This study presents an integrated, edge-centric framework for real-time physiological signal processing tailored for diagnostic IoT in home-based health monitoring. The proposed system successfully addresses major limitations of cloud-reliant models, including latency, power consumption, and privacy concerns. By incorporating wavelet-based denoising and adaptive filtering, the framework significantly improves signal quality in noisy, non-clinical environments. The use of hybrid feature extraction and PCA-based dimensionality reduction enables efficient data representation without compromising diagnostic relevance. Lightweight AI models such as 1D-CNNs and TinyLSTM were optimized and deployed on microcontroller platforms, achieving inference latencies under 50 ms and power consumption below 20 mW. Benchmarking against publicly available datasets validated the framework's high classification accuracy and robustness across multiple modalities. Compared to existing solutions, this work demonstrates superior adaptability to real-world, resource-constrained environments. The architecture's modular design supports scalability across various diagnostic tasks beyond the datasets evaluated. This approach enables decentralized, continuous health monitoring without reliance on persistent internet connectivity or high-end hardware. Overall, the framework lays a practical foundation for next-generation wearable health diagnostics, empowering patients through intelligent, autonomous home-based care.

References

[1] World Health Organization. (2022). Noncommunicable diseases. Retrieved from https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases

[2] Gubbi, J., Buyya, R., Marusic, S., & Palaniswami, M. (2013). Internet of Things (IoT): A vision, architectural elements, and future directions. Future Generation Computer Systems, 29(7), 1645–1660. https://doi.org/10.1016/j.future.2013.01.010

[3] Islam, S. M. R., Kwak, D., Kabir, M. H., Hossain, M., & Kwak, K. S. (2015). The Internet of Things for health care: A comprehensive survey. IEEE Access, 3, 678–708. https://doi.org/10.1109/ACCESS.2015.2437951

[4] Poh, M. Z., Swenson, N. C., & Picard, R. W. (2010). Motion-tolerant magnetic earring sensor and wireless earpiece for wearable photoplethysmography. IEEE Transactions on Information Technology in Biomedicine, 14(3), 786–794. https://doi.org/10.1109/TITB.2010.2042607

[5] Rehman, A., Saba, T., Haseeb, K., Larabi Marie-Sainte, S., & Lloret, J. (2021). Energy-efficient IoT e-health using artificial intelligence model with homomorphic secret sharing. Energies, 14(19), 6414. https://doi.org/10.3390/en14196414

[6] Rieke, N., Hancox, J., Li, W., Milletari, F., Roth, H. R., Albarqouni, S., ... & Cardoso, M. J. (2020). The future of digital health with federated learning. npj Digital Medicine, 3(1), 119. https://doi.org/10.1038/s41746-020-00323-1

[7] Xu, W., Zhang, W., & Li, Y. (2021). Edge computing enabled smart healthcare systems: A review. IEEE Internet of Things Journal, 8(21), 15721–15737. https://doi.org/10.1109/JIOT.2021.3064570

[8] Roy, N., Misra, A., & Cook, D. J. (2020). Energy-efficient deep learning for mobile health diagnostics. IEEE Internet of Things Journal, 7(8), 6674–6682. https://doi.org/10.1109/JIOT.2020.2973676

[9] Swathi, B., Veerabomma, S., Archana, M., & Bhadru, D. (2025). Edge-Centric IoT Health Monitoring: Optimizing Real-Time Responsiveness, Data Privacy, and Energy Efficiency. 2025 6th International Conference on Mobile Computing and Sustainable Informatics (ICMCSI).

[10] Giordano, M., Dheman, K., & Magno, M. (2024). SepAl: Sepsis Alerts On Low Power Wearables With Digital Biomarkers and On-Device Tiny Machine Learning. arXiv preprint arXiv:2408.08316.

[11] Lu, S., Lu, J., An, K., Wang, X., & He, Q. (2023). Edge Computing on IoT for Machine Signal Processing and Fault Diagnosis: A Review. IEEE Internet of Things Journal, PP(99), 1-1.

[12] David, R., Duke, J., Jain, A., Reddi, V. J., Jeffries, N., Li, J., ... & Warden, P. (2020). TensorFlow Lite Micro: Embedded Machine Learning on TinyML Systems. arXiv preprint arXiv:2010.08678.

[13] B. Swathi, S. Veerabomma, M. Archana, D. Bhadru, N. L. Somu, and M. Bhavsingh, "Edge-Centric IoT Health Monitoring: Optimizing Real-Time Responsiveness, Data Privacy, and Energy

Efficiency,” 2025 6th International Conference on Mobile Computing and Sustainable Informatics (ICMCSI), Jan. 2025, pp. 354–361, doi: 10.1109/icmcsi64620.2025.10883456.

[14] M. Giordano, K. Dheman, and M. Magno, “SepAl: Sepsis Alerts On Low Power Wearables With Digital Biomarkers and On-Device Tiny Machine Learning,” IEEE Sensors Journal, vol. 25, no. 5, pp. 7858–7866, Mar. 2025, doi: 10.1109/jsen.2024.3424655.