International Journal of Engineering & Extended Technologies Research (IJEETR)

|ISSN: 2322-0163| www.ijeetr.com | A Bimonthly, Peer Reviewed, Scholarly Indexed Journal |

| Volume 8, Issue 2, March - April 2026 |

DOI:10.15662/IJEETR.2026.0802069

Tiny ML-Enabled Edge Computing for Autonomous Health Monitoring and Anomaly Detection

M. Vichithra[1], Mr. A. Syed Aslam[2], Mr. R. Shanjay[3], Mr. S. Vengatachalam[4], Mr. F. Siraj Mohammath[5]

Department of Computer Science and Engineering, R P Sarathy Institute of Technology, Salem, Tamil Nadu, India

Publication History: Received: 25.02.2026; Revised: 20.03.2026; Accepted: 25.03. 2026; Published: 28.03.2026.

ABSTRACT: This paper presents a decentralized, AI-powered wearable system for real-time physiological monitoring and anomaly detection. By leveraging TinyML (machine learning on resource-constrained edge devices), the proposed system eliminates latency and dependency on cloud-based healthcare solutions. The architecture integrates a multi-sensor array capturing both hemodynamic and kinematic data with an ultra-low-power microcontroller to perform on-device inference for cardiac anomalies and fall detection.

A key contribution is the integration of a hybrid solar-battery energy model, ensuring sustainable operation in resource-limited environments. Experimental results demonstrate that edge-based intelligence significantly reduces emergency response time through autonomous, multi-channel alert mechanisms. This work contributes to accessible and reliable healthcare by enabling high-accuracy, offline diagnostic support for geriatric and remote patient monitoring.

KEYWORDS: TinyML, Edge Computing, Anomaly Detection, Wearable Health Systems, Autonomous Alerting, Geriatric Care, Microcontroller Inference.

I. INTRODUCTION

The rapid growth of Internet of Things (IoT) devices and embedded systems has enabled continuous health monitoring beyond traditional clinical environments. However, most existing wearable systems rely heavily on cloud infrastructure, leading to increased latency, privacy concerns, and dependence on stable internet connectivity.

These limitations are particularly critical in rural areas and for elderly populations, where reliable network access cannot be guaranteed. TinyML, a specialized field of machine learning focused on deploying models on microcontrollers with constrained resources, offers a practical solution. By enabling on-device inference, TinyML reduces latency, enhances privacy, and ensures system reliability even in offline environments.

This paper proposes a fully autonomous wearable health monitoring system that integrates TinyML-based anomaly detection, fall detection, and real-time emergency alerting within a compact, energy-efficient device. The system operates entirely offline while delivering rapid, life-critical responses through multi-channel alert mechanisms including SMS, buzzer, and mobile app notifications.

II. RELATED WORK

Previous research in wearable health monitoring has largely relied on cloud-based computation. Rajput et al. [1] implemented ECG anomaly detection using LSTM models on cloud platforms, achieving high accuracy but with latency unsuitable for emergency scenarios. Alam et al. [2] introduced a fog computing approach to reduce latency; however, it still required continuous connectivity.

With advancements in microcontroller capabilities, edge-based approaches have emerged. Warden and Situnayake [3] demonstrated TinyML feasibility using TensorFlow Lite on microcontrollers, proving meaningful neural inference achievable within a 256 KB RAM budget. Kartsch et al. [4] applied convolutional neural networks for fall detection with low latency and high sensitivity, reporting 94.1% sensitivity with sub-10ms inference latency.

Despite these advancements, existing solutions do not integrate cardiac monitoring, fall detection, autonomous alerting, and sustainable power systems into a single offline-capable wearable device. This work addresses that critical gap.

III. PROBLEM STATEMENT

Current wearable health systems face several critical limitations that reduce their effectiveness for high-risk populations:

Late Detection: Cardiac anomalies often go undetected due to reliance on intermittent or cloud-based monitoring, resulting in preventable complications and fatalities.

Lack of Fall Detection: Many devices do not provide reliable or real-time fall detection, leaving elderly patients without emergency coverage during daily activities.

Delayed Alerts: Emergency notifications are often delayed due to processing and connectivity constraints, causing critical delays in medical intervention.

Cloud Dependency: Systems fail in environments with limited or no internet access, making them unreliable in rural areas, disaster zones, and remote locations.

IV. PROPOSED SYSTEM ARCHITECTURE

A. Hardware Components

The system is built on an ultra-low-power ARM Cortex-M4 microcontroller with 256 KB SRAM. This processor was selected for its hardware floating-point unit, which accelerates on-device neural network inference. The sensor suite includes: (1) MAX30102 Pulse Oximeter and Heart Rate Sensor for hemodynamic data acquisition at 100 samples/sec; (2) MPU-6050 Six-Axis Accelerometer and Gyroscope for fall detection and activity classification; (3) DS18B20 Digital Temperature Sensor for core body temperature monitoring; (4) SIM800L GSM Module for SMS alerts independent of Wi-Fi; and (5) a 100mW Solar Panel with 3.7V LiPo Battery for hybrid power supply.

B. System Workflow

The system operates through four sequential stages: (1) Data Collection from the multi-sensor array at a configurable sampling rate; (2) Preprocessing and normalization including mean subtraction, windowing, and feature extraction on the microcontroller; (3) TinyML Model Inference producing a classification label (Normal, Cardiac Anomaly, Fall Detected, or Hyperthermia); and (4) Automated Response via simultaneous activation of SMS alerts through the GSM module, an audible buzzer, and mobile application push notifications.

V. TINYML MODEL DESIGN

A. Model Architecture

A one-dimensional Convolutional Neural Network (1D-CNN) is used for time-series analysis of physiological signals. The architecture comprises three convolutional blocks (Conv1D, Batch Normalization, ReLU activation, MaxPooling) followed by two fully connected layers and a softmax output layer with four classes. The model is trained using TensorFlow 2.x and subsequently quantized to 8-bit integer precision using TensorFlow Lite post-training quantization. The quantized model occupies 48 KB of flash memory with an average inference latency of 6.2 ms on target hardware.

B. Dataset

The model is trained using three datasets: the MIT-BIH Arrhythmia Dataset for cardiac signal classification, the SisFall Dataset for fall detection from accelerometer sequences, and synthetically augmented temperature readings for hyperthermia detection. A total of 42,000 labeled samples (400 ms windows, 50% overlap) were used with a 70/15/15 split for training, validation, and testing respectively.

VI. EXPERIMENTAL RESULTS

A. Model Performance

Table I presents the classification performance of the quantized TinyML model on the held-out test set. The system demonstrates high sensitivity across all four target classes, with F1-scores exceeding 92% for the life-critical fall detection and cardiac anomaly classes.

Class	Precision	Recall	F1-Score
Normal	96.4%	97.1%	96.7%
Cardiac Anomaly	93.2%	94.8%	94.0%
Fall Detected	91.7%	93.5%	92.6%
Hyperthermia	95.1%	94.3%	94.7%

Table I: Classification Performance of the TinyML Model

B. System Performance Metrics

End-to-end emergency response latency, measured from anomaly onset to alert delivery, averaged 47 ms compared to over 2,300 ms for a comparable cloud-based baseline. Power consumption in active monitoring mode was measured at 38 mW, enabling approximately 52 hours of continuous operation on the battery alone. With solar charging under standard indoor illumination (500 lux), the system achieved indefinite self-sustaining operation.

VII. CONCLUSION

This paper presents a TinyML-based wearable health monitoring system capable of real-time cardiac anomaly detection, fall detection, and autonomous emergency alerting without reliance on cloud infrastructure. The 1D-CNN model, quantized to 8-bit integer arithmetic, achieves F1-scores exceeding 92% across all target classes while maintaining a mean inference latency of 6.2 ms and operating at 38 mW power consumption.

The proposed solution has strong potential in rural healthcare, elderly monitoring, home healthcare, and disaster response scenarios. Future work includes model personalization using federated learning for patient-specific adaptation, and further hardware miniaturization to improve wearability and user comfort.

REFERENCES

1. Rajput, S. et al. (2021). Cloud-based ECG anomaly detection using LSTM networks for remote patient monitoring. IEEE Journal of Biomedical and Health Informatics, 25(3), 812-821.
2. Alam, M. et al. (2017). Fog computing approach for latency reduction in wearable health monitoring systems. Proc. IEEE Fog World Congress, pp. 1-6.
3. Warden, P. and Situnayake, D. (2019). TinyML: Machine Learning with TensorFlow Lite on Arduino and Ultra-Low Power Microcontrollers. O'Reilly Media, Sebastopol, CA.
4. Kartsch, V. et al. (2018). CNN-based fall detection with low latency on a GAP8 edge processor. Proc. IEEE International Symposium on Circuits and Systems (ISCAS), pp. 1-5.
5. Howard, A. et al. (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861.
6. LeCun, Y., Bengio, Y. and Hinton, G. (2015). Deep learning. Nature, 521, pp. 436-444.
7. Landreani, B. et al. (2016). Accelerometric measurement of heart rate by smartphones: Reliability and validity. Journal of Medical Systems, 40(8), p. 184.
8. C.Nagarajan and M.Madheswaran - 'Stability Analysis of Series Parallel Resonant Converter with Fuzzy Logic Controller Using State Space Techniques'- Taylor &Francis, Electric Power Components and Systems, Vol.39 (8), pp.780-793, May 2011. DOI: 10.1080/15325008.2010.541746
9. C.Nagarajan and M.Madheswaran - 'Experimental verification and stability state space analysis of CLL-T Series Parallel Resonant Converter' - Journal of Electrical Engineering, Vol.63 (6), pp.365-372, Dec.2012. DOI: 10.2478/v10187-012-0054-2

10. C.Nagarajan and M.Madheswaran - 'Performance Analysis of LCL-T Resonant Converter with Fuzzy/PID Using State Space Analysis'- Springer, Electrical Engineering, Vol.93 (3), pp.167-178, September 2011. DOI 10.1007/s00202-011-0203-9
11. S.Tamilselvi, R.Prakash, C.Nagarajan,"Solar System Integrated Smart Grid Utilizing Hybrid Coot-Genetic Algorithm Optimized ANN Controller" Iranian Journal Of Science And Technology-Transactions Of Electrical Engineering, DOI10.1007/s40998-025-00917-z,2025
12. S.Tamilselvi, R.Prakash, C.Nagarajan," Adaptive sliding mode control of multilevel grid-connected inverters using reinforcement learning for enhanced LVRT performance" Electric Power Systems Research 253 (2026) 112428, doi.org/10.1016/j.epsr.2025.112428
13. S.Thirunavukkarasu, C. Nagarajan, 2024, "Performance Investigation on OCF and SCF study in BLDC machine using FTANN Controller," Journal of Electrical Engineering And Technology, Volume 20, pages 2675–2688, (2025), doi.org/10.1007/s42835-024-02126-w
14. C. Nagarajan, M.Madheswaran and D.Ramasubramanian- 'Development of DSP based Robust Control Method for General Resonant Converter Topologies using Transfer Function Model'- Acta Electrotechnica et Informatica Journal , Vol.13 (2), pp.18-31,April-June.2013, DOI: 10.2478/aeei-2013-0025.
15. C.Nagarajan and M.Madheswaran - 'DSP Based Fuzzy Controller for Series Parallel Resonant converter'- Springer, Frontiers of Electrical and Electronic Engineering, Vol. 7(4), pp. 438-446, Dec.12. DOI 10.1007/s11460-012-0212-0.
16. C.Nagarajan and M.Madheswaran - 'Experimental Study and steady state stability analysis of CLL-T Series Parallel Resonant Converter with Fuzzy controller using State Space Analysis'- Iranian Journal of Electrical & Electronic Engineering, Vol.8 (3), pp.259-267, September 2012.
17. C.Nagarajan and M.Madheswaran, "Analysis and Simulation of LCL Series Resonant Full Bridge Converter Using PWM Technique with Load Independent Operation" has been presented in ICTES'08, a IEEE / IET International Conference organized by M.G.R.University, Chennai.Vol.no.1, pp.190-195, Dec.2007
18. Suganthi Mullainathan, Ramesh Natarajan, "An SPSS and CNN modelling based quality assessment using ceramic materials and membrane filtration techniques", Revista Materia (Rio J.) Vol. 30, 2025, DOI: https://doi.org/10.1590/1517-7076-RMAT-2024-0721
19. M Suganthi, N Ramesh, "Treatment of water using natural zeolite as membrane filter", Journal of Environmental Protection and Ecology, Volume 23, Issue 2, pp: 520-530,2022
20. Abadi, M. et al. (2016). TensorFlow: A system for large-scale machine learning. Proc. 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI), pp. 265-283.