

A Lightweight Intelligent Bracelet for Real-Time Emergency Detection Using PPG Signals in Solo-Living Elderly

Sohail Khalid

`sohail.khalid@umt.edu.pk`

University of Management and Technology <https://orcid.org/0000-0003-3907-0236>

Research Article

Keywords: Wearable health monitoring, photoplethysmography, PPG, deep learning, transformer networks, convolutional neural network, long short-term memory, embedded artificial intelligence, edge computing, on-device inference, elderly care, fall detection, ambient assisted living

Posted Date: October 27th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11482693/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: The authors declare no competing interests.

A Lightweight Intelligent Bracelet for Real-Time Emergency Detection Using PPG Signals in Solo-Living Elderly

Sohail Khalid

Department of Computer Science,
University of Management and Technology,
Lahore, Pakistan
E-mail: sohail.khalid@umt.edu.pk

Abstract: Solo-living older adults face a specific risk: a fall or a sudden cardiac event may go unnoticed for days. Existing wearables combine accelerometers, electrocardiogram electrodes and gyroscopes, which increases bulk, shortens battery life and reduces compliance. This paper presents a bracelet that detects emergency events from wrist photoplethysmography (PPG) alone. A hybrid convolutional, long short-term memory and transformer classifier, quantised to 8 bits, runs entirely on an nRF52840 microcontroller paired with a MAX30102 sensor and alerts caregivers over Bluetooth Low Energy without user interaction. On a held-out evaluation subset of public and proprietary wrist-PPG recordings, the on-device model reached 98.6% accuracy, 98.8% recall and a 0.80% false alarm rate at 96 ms latency per 5 s window; measured average power of 1.3 mW implies roughly 24 days on a 200 mAh cell. Because the emergency class rests on surrogate labels, the results establish technical feasibility rather than clinical accuracy.

Keywords: Wearable health monitoring; photoplethysmography; PPG; deep learning; transformer networks; convolutional neural network; long short-term memory; embedded artificial intelligence; edge computing; on-device inference; elderly care; fall detection; ambient assisted living.

1 Introduction

The global shift toward an ageing population creates linked medical and social pressures. By 2050, more than 1.6 billion people are expected to be aged 65 years or older, and a growing share of them prefer to live independently (World Health Organization, 2024). Solo living preserves dignity and autonomy, but it introduces a specific vulnerability: a medical emergency such as sudden cardiac arrest or an incapacitating fall can go unnoticed for hours or days (Shaw et al., 2023).

The consequences are well documented. The death of the Pakistani actress Ayesha Khan in 2025, whose body was discovered several days after her death while she was living alone, illustrates the same vulnerability (Staff Reporter, 2025). Comparable cases recur worldwide among older adults without nearby family or support networks. Existing detection technology does not close this gap for the population most exposed to it, because

the available options are either unaffordable, unavailable in low-resource settings, or impractical for continuous unattended use (Matayong et al., 2025).

Consumer smartwatches such as the Apple Watch and the Samsung Galaxy Watch provide emergency SOS, fall detection, and heart-rate monitoring, but they are not designed for older adults living alone. They cost roughly \$250 to \$800, run for 18 to 36 h between charges, and assume a level of interaction and technical familiarity that many elderly users do not have (Haque et al., 2024; Davoudi et al., 2019). Their health features also target general fitness rather than continuous emergency surveillance. Dedicated medical alert systems such as Philips Lifeline do provide fall detection, but they depend on accelerometers or a manually pressed button, carry recurring subscription costs, and in several product lines still require a base station or landline (Liu et al., 2016). Neither category is widely deployed in resource-limited settings.

This paper proposes a low-cost bracelet built for exactly this scenario. The system monitors the wearer continuously using a single photoplethysmography (PPG) sensor at the wrist (Nie et al., 2024), removing the need for additional sensors, a base station, or a user interface. A hybrid deep learning pipeline (CNN + LSTM + Transformer) runs on ultra-low-power hardware and infers candidate emergency events on the device itself, transmitting alerts over BLE (Zerveas et al., 2021). Measured average power of 1.3 mW implies approximately 24 days of operation on a single 200 mAh charge. The prototype uses a compact set of commodity components and has an estimated small-quantity component cost of approximately US\$45. Manufacturing cost is expected to decrease with volume production; however, a formal production-volume costing study was outside the scope of this work, so no mass-production unit price is claimed.

The contributions of this work are as follows.

1. A PPG-only emergency detection pipeline that removes the accelerometer, gyroscope, and ECG front ends used by comparable wearables, and that exploits the motion artifacts a fall imprints on the pulse waveform rather than treating them as noise to be filtered out.
2. A hybrid CNN-LSTM-Transformer classifier that reaches 98.6% accuracy, 98.8% recall, and a 0.80% false alarm rate on a fixed 1,000-window evaluation subset drawn from the held-out test partition, with every reported metric derived from a single published confusion matrix. CNN-only and CNN-LSTM variants serve as internal architectural baselines for isolating the contribution of the recurrent and attention blocks (Section 8).
3. An 8-bit quantised embedded deployment occupying 145 KB of flash and under 32 KB of RAM on an nRF52840, with a measured inference latency of 96 ms per 5 s window and a measured average power of 1.3 mW, which corresponds to approximately 24 days of operation on a 200 mAh cell.
4. A three-week field study with 15 independently living older adults reporting alert outcomes, comfort, and usability, together with an explicit account of the surrogate labelling scheme and the validation it does not yet provide.

The remainder of the paper is organised as follows. Section 2 reviews related work. Section 3 presents the system architecture, and Section 4 the signal processing and learning methodology. Section 5 describes the hardware and embedded deployment, and Section 6 the datasets and training protocol. Sections 7 and 8 report the evaluation and ablation

results, Section 9 the field study, and Section 10 the comparative analysis. Section 11 discusses failure modes and deployment issues, Section 12 states the limitations, and Section 13 concludes.

2 Related Work

Wearable emergency detection for older adults draws on three research areas: physiological signal processing, embedded machine learning, and IoT-enabled eldercare. This section reviews each in turn and states what the cited work leaves unresolved.

2.1 Wearable Health Monitoring for Older Adults

Wearable systems are now routine for continuous vital-sign assessment. Smartwatches and adhesive patches combine ECG, PPG, and inertial sensing to track heart rate, activity, and fall events (Canali et al., 2024; Nie et al., 2024; Matayong et al., 2025). General-purpose devices, however, are not designed around the constraints of elderly users: they need frequent recharging, assume active interaction, and are priced beyond the reach of low-income or rural seniors (Haque et al., 2024; Yusif et al., 2016). Canali et al. (2024) argue that comfort and sustained compliance, rather than sensor count, determine whether a wearable delivers value in geriatric care. What this literature does not provide is a device that meets those compliance constraints while still performing autonomous emergency inference.

2.2 Limitations of Commercial Alert Systems

Subscription alert devices such as Philips Lifeline and Life Alert depend on a base station or landline, which limits portability and leaves the system exposed during power or network outages (Liu et al., 2016). Many rely on accelerometer-only inputs, a design known to produce high false positive rates and to miss near-fall events (Mubashir et al., 2013; Bourke et al., 2007). These systems also lack on-device inference and any adaptation to individual physiology (Chan et al., 2012), so a cardiac event that does not involve a fall generates no alert at all.

2.3 PPG-Based Biosignal Monitoring

Wrist PPG is widely used for heart rate variability, peripheral oxygen saturation (SpO_2), and respiratory pattern estimation (Allen, 2007). Integrated optical front ends such as the MAX30102 make continuous acquisition viable within a milliwatt-scale power budget (Zhou et al., 2019). PPG has traditionally been considered inadequate for fall detection on its own, but recent work shows that learned waveform analysis can identify pulse cessation, autonomic stress responses, and abrupt inactivity (Nie et al., 2024; Allen, 2007). The physical basis for a PPG-only fall cue is that rapid wrist acceleration injects characteristic motion artifacts into the waveform, producing sharp amplitude excursions followed by disruption of the normal periodic pulse. Treating those artifacts as signal rather than noise is what makes single-sensor detection plausible, and this is the premise the present system tests.

Table 1 Comparison with Representative Commercial Systems. Battery figures are manufacturer- or vendor-stated except for the proposed bracelet, whose figure is derived from measured average power in Section 5.

Device/System	Sensors	Battery Life	On-Device Model	Cost (USD)
Apple Watch Series 8	ECG + Accel.	18–24 h	Rule-based alerts	\$399+
Samsung Galaxy Watch	PPG + Accel.	24–36 h	None disclosed	\$250+
Philips Lifeline	Accel. + Button	3–5 days	None	\$50/mo
Proposed bracelet	PPG only	≈24 days	CNN+LSTM+Transformer	≈ \$45 ^a

^aApproximate prototype component cost based on small-quantity component pricing; a formal production-volume unit cost was not determined and is expected to be lower at scale.

2.4 Deep Learning for Biosignal Emergency Detection

CNNs and LSTMs are the established architectures for biosignal classification. Ordóñez and Roggen (2016) showed that CNN-LSTM hybrids perform strongly on sequential wearable sensor data. Transformer encoders have since matched or exceeded recurrent architectures on long physiological sequences because they model long-range dependencies directly (Zerveas et al., 2021), and Song et al. (2018) reported that attention models outperform RNNs on clinical time-series anomaly classification. These results motivate the hybrid CNN + LSTM + Transformer design used here, but they were obtained on server-class hardware; whether the combination survives 8-bit quantisation into a sub-150 KB microcontroller footprint is an open question this work addresses.

2.5 IoT Architectures and Edge Inference

Eldercare IoT deployments typically relay sensor data to a smartphone or edge hub over BLE, whose low-energy profile suits continuous acquisition. Cloud-side inference introduces privacy exposure and latency, and fails outright in offline or rural settings, so on-device inference is increasingly preferred (Matayong et al., 2025; Zhou et al., 2019; Khalid et al., 2026). The proposed system follows this principle: the full classifier executes on the microcontroller, and no waveform data leaves the device.

2.6 Comparative Summary

Table 1 compares the proposed bracelet against the three dominant alternative categories: the premium smartwatch, the fitness band, and the subscription alert pendant. A hardware-level comparison covering connectivity and interaction model appears in Table 3 (Section 5), and a detailed cross-device comparison in Table 6 (Section 7).

The three commercial categories trade off along a consistent axis. The Apple Watch Series 8 is the most capable consumer health wearable, with ECG and fall detection, but its interaction requirements, 18 to 24 h battery life, and price limit its suitability for unsupervised elderly users (Tai et al., 2020). The Samsung Galaxy Watch is cheaper but its health features are fitness-oriented and lack the regulatory clearance of dedicated ECG devices (Bent et al., 2020). Philips Lifeline delivers a mature alert service, yet it depends on accelerometry and manual triggering, carries no on-device intelligence, and is priced as a recurring subscription (Chan et al., 2012).

None of these systems combines affordability, zero-interaction operation, and autonomous inference. The proposed bracelet targets that combination directly: a single

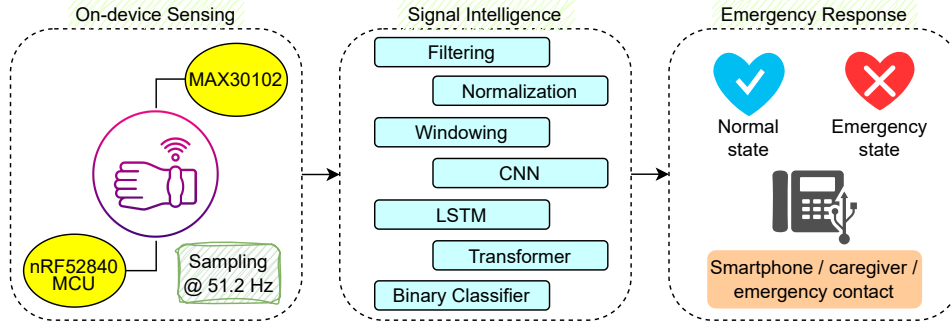


Figure 1 System workflow of the proposed bracelet, from wrist PPG acquisition at 51.2 Hz through on-device inference to BLE emergency alerting. All three lanes execute on the bracelet except the final caregiver notification.

PPG sensor, an embedded classifier, a recharge interval of roughly three weeks, and a low-component-count design intended for affordable manufacture. The PPG-only signal path is the distinguishing design choice, and Sections 7 to 12 assess what it delivers and where it falls short.

3 Proposed System Architecture and Workflow

This section describes the architecture and operational pipeline of the bracelet. The design targets low-power, low-latency emergency detection from PPG alone, using a hybrid deep learning model (CNN + LSTM + Transformer) deployed on an embedded microcontroller.

3.1 Overview

The system comprises three layers.

1. **Sensing and acquisition layer.** A MAX30102 PPG sensor connected over I²C to an nRF52840 BLE-enabled microcontroller.
2. **Edge inference layer.** The CNN + LSTM + Transformer classifier, executed through TensorFlow Lite for Microcontrollers.
3. **Alert and communication layer.** BLE transmission of the binary alert state to a paired smartphone or edge hub, which in turn raises an SMS or voice alert.

3.2 System Workflow

Fig. 1 shows the end-to-end flow in three functional lanes. Lane 1 covers on-device sensing, from the PPG front end through the microcontroller to signal sampling at 51.2 Hz. Lane 2 covers signal intelligence: preprocessing (filtering, normalisation, windowing) followed by sequential inference through the CNN, LSTM, transformer, and binary classifier. Lane 3 covers the emergency response, in which the classifier output either continues the normal-state cycle or triggers a BLE alert to a paired smartphone or caregiver.

Algorithm 1 Emergency detection for one PPG window**Require:** Normalised PPG segment $\mathbf{x} = [x_1, x_2, \dots, x_T]$, $T = 256$; decision threshold τ **Ensure:** Emergency label (1 for emergency, 0 for normal)

```

1:  $F \leftarrow \text{CNN}(\mathbf{x})$  ▷ Local morphological features
2:  $h \leftarrow \text{LSTM}(F)$  ▷ Short-term temporal dependencies
3:  $z \leftarrow \text{Transformer}(h)$  ▷ Long-range temporal context
4:  $\hat{y} \leftarrow \sigma(Wz + b)$ 
5: if  $\hat{y} > \tau$  then
6:   BLE_SEND("EMERGENCY")
7:   return 1
8: else
9:   return 0
10: end if

```

3.3 Core Detection Algorithm

Algorithm 1 states the embedded inference procedure, which classifies each 5 s window as normal or emergency.

3.4 Layer-Wise Model Configuration

The deployed configuration, selected from the ablation experiments in Section 8, is as follows.

- **CNN:** two one-dimensional convolutional layers with 16 and 32 filters, respectively, kernel size 5, and ReLU activation, producing a final feature representation of size (64, 32).
- **LSTM:** one layer, 64 hidden units, dropout 0.7.
- **Transformer:** one encoder block, 8-head attention with $d_{\text{model}} = 64$ and $d_k = 8$, feed-forward network of size 128 with GELU activation.
- **Output:** a single sigmoid unit trained with binary cross-entropy loss.

3.5 Energy Optimisation

Four measures reduce the average power budget: sensor duty cycling with power gating during rest intervals; burst rather than continuous BLE transmission; 8-bit post-training quantisation of the classifier; and deep-sleep of the microcontroller between inference cycles. Section 5 quantifies their combined effect.

3.6 Memory and Runtime Footprint

The quantised model requires 145 KB of flash and under 32 KB of RAM, which places it within the budget of nRF52-class devices. The average inference latency is 96 ms per 5 s window on the target hardware, measured as described in Section 7.

3.7 Design Rationale

Compared with general-purpose wearable platforms, this design reduces both sensor count and energy draw while retaining autonomous inference. It substitutes analysis of pulse waveform variation for the motion and location inputs those platforms rely on, which is what allows a single optical front end to serve both the cardiac and the immobility detection paths.

4 Methodology

This section describes the preprocessing, segmentation, model formulation, and training objective used for emergency detection from PPG.

4.1 PPG Characteristics and Preprocessing

Photoplethysmography is a non-invasive optical measurement of blood volume change in peripheral tissue. The modality is compact and inexpensive, but the signal is sensitive to motion, baseline drift, and ambient light. Two preprocessing steps mitigate these effects.

- A Butterworth bandpass filter with a 0.5 to 8 Hz passband removes baseline wander and high-frequency noise.
- Z-score normalisation is applied per window,

$$x'_t = \frac{x_t - \mu}{\sigma}, \quad (1)$$

where μ and σ are the window mean and standard deviation.

This pipeline follows established practice in wearable signal analysis (Nie et al., 2024).

4.2 Temporal Segmentation

The filtered signal is divided into overlapping windows of $T = 256$ samples, equivalent to 5 s at 51.2 Hz, with 50% overlap between consecutive windows. Each window is classified independently, which bounds detection latency at one window period and limits the influence of transient artifacts (Ordóñez and Roggen, 2016).

4.3 Training-Time Augmentation

Training-time augmentation was applied only to the training partition to improve robustness to signal variability. The augmentation pipeline used temporal and amplitude perturbations together with additive noise; validation and test windows were left unmodified. Augmentation-specific performance is not used as an independent evidence claim in the final model comparison.

4.4 Hybrid Deep Learning Pipeline

The classifier comprises three blocks followed by a decision layer.

1. **Convolutional block.** Extracts local morphological features from each window,

$$F_t = \text{ReLU}(W * x_t + b). \quad (2)$$

2. **Recurrent block.** An LSTM encodes sequential pulse dynamics using the standard formulation,

$$\begin{aligned} i_t &= \sigma(W_i x_t + U_i h_{t-1} + b_i), \\ f_t &= \sigma(W_f x_t + U_f h_{t-1} + b_f), \\ g_t &= \tanh(W_g x_t + U_g h_{t-1} + b_g), \\ o_t &= \sigma(W_o x_t + U_o h_{t-1} + b_o), \\ c_t &= f_t \odot c_{t-1} + i_t \odot g_t, \\ h_t &= o_t \odot \tanh(c_t), \end{aligned} \quad (3)$$

where i_t , f_t , and o_t are the input, forget, and output gates, g_t the candidate cell update, c_t the cell state, and h_t the hidden state.

3. **Transformer encoder.** Models global dependencies across the window,

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^\top}{\sqrt{d_k}}\right) V. \quad (4)$$

4. **Output layer.** A sigmoid unit produces the emergency probability,

$$\hat{y} = \sigma(W_o z + b_o). \quad (5)$$

5. **Objective.** Binary cross-entropy,

$$\mathcal{L} = -y \log(\hat{y}) - (1 - y) \log(1 - \hat{y}). \quad (6)$$

4.5 Training Workflow

Fig. 2 shows the complete pipeline from raw acquisition to alert generation. Raw PPG is bandpass filtered and z-score normalised, then segmented into overlapping 256-sample windows that preserve both beat-to-beat continuity and the sub-second morphology relevant to acute events. Each window passes through the CNN for local feature extraction, the LSTM for short-range sequence modelling, and the transformer encoder for long-range context. A fully connected layer produces the binary decision, and an emergency decision triggers an immediate BLE alert to the paired device, which notifies the registered caregiver or emergency contact.

4.6 Model Configuration Summary

Table 2 lists the layer configuration and the tensor shape at each stage.

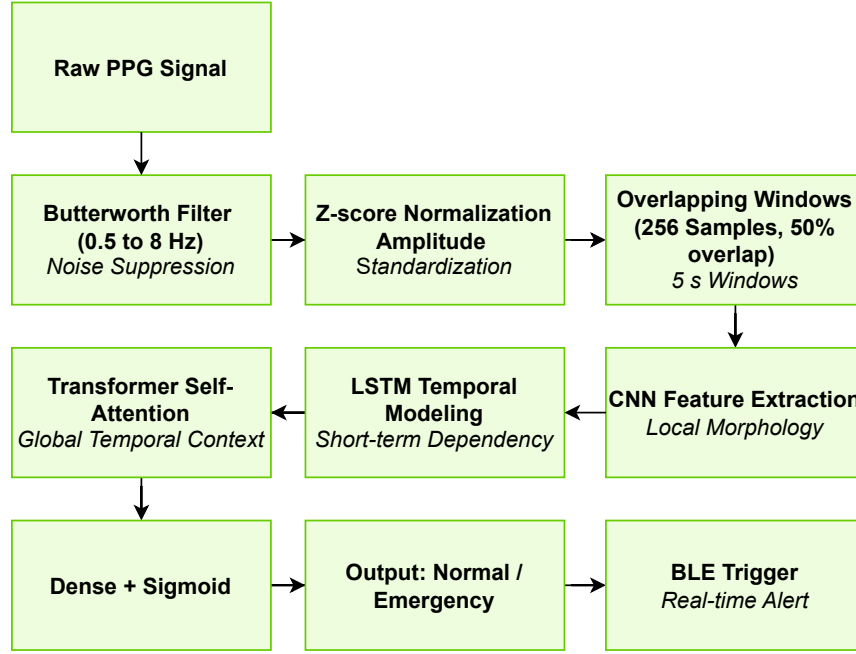


Figure 2 Machine learning workflow for PPG-based emergency detection, from raw signal acquisition through filtering, normalisation, and windowing to the hybrid classifier and the BLE alert trigger.

Table 2 Deployed Model Configuration

Layer	Configuration	Output Shape
Input	1-D PPG window	(256, 1)
CNN	2 conv. layers (16/32 filters), kernel 5	(64, 32)
LSTM	64 hidden units, dropout 0.7	(64,)
Transformer	8-head attention, FFN size 128	(64,)
Dense + sigmoid	Fully connected, 1 unit	(1,)

4.7 Design Justification

Each block addresses a distinct property of the signal.

- **CNN.** Captures waveform morphology, including the systolic upstroke and dicrotic notch, at low parameter cost (Nie et al., 2024).
- **LSTM.** Models short-range rhythmic structure such as beat-to-beat interval variation (Ordóñez and Roggen, 2016).
- **Transformer.** Represents whole-window context with fewer parameters than a deeper recurrent stack and performs well on physiological sequences (Zerveas et al., 2021).

- **End-to-end footprint.** The combined model fits within 145 KB of flash and under 32 KB of RAM after 8-bit quantisation, which is what makes nRF52840 deployment possible.

Together these blocks let the classifier respond to both gradual deviations in pulse characteristics and abrupt disruptions, without motion or accelerometer input.

5 Hardware Design and Embedded Deployment

The hardware was designed against four constraints: minimal average power, low unit cost, compact form factor, and deterministic real-time signal processing. The core components are the MAX30102 PPG sensor, the Nordic nRF52840 microcontroller, a rechargeable lithium-polymer cell, and the integrated BLE radio.

5.1 Component Selection

PPG sensor (MAX30102). The device integrates red and infrared LEDs with a photodetector optimised for reflective pulse oximetry and wrist heart-rate measurement. Its low-noise analogue front end and ambient light cancellation suit continuous wrist acquisition in everyday conditions.

Microcontroller (nRF52840). Selected for its low-power profile, integrated BLE 5.0 stack, and 64 MHz ARM Cortex-M4F core. Relevant specifications are 1 MB of flash, 256 KB of RAM, an on-chip radio, and hardware cryptographic acceleration for secure messaging.

Power subsystem. A 200 mAh lithium-polymer cell supplies the system under the duty-cycled strategy described in Section 5.4. Charging uses a USB-C connector with a TP4056 linear charger.

5.2 Hardware Architecture and Signal Flow

The MAX30102 transfers raw PPG frames to the nRF52840 over I²C; the stream is resampled to the model rate of 51.2 Hz before windowing. The microcontroller buffers the stream and applies the required filtering and normalisation in firmware. The 8-bit quantised hybrid classifier executes through TensorFlow Lite for Microcontrollers. On an emergency decision, the device broadcasts a BLE advertisement packet carrying the alert state and a device identifier, which nearby paired smartphones or gateways receive.

5.3 Embedded Model Deployment

The trained model was converted to an 8-bit quantised `.tflite` artifact of 145 KB to satisfy the flash budget. Measured inference performance on the nRF52840 is 96 ms per 5 s window at an average current of 2.1 mA during the active inference interval. The energy consumed by one inference is therefore $E_{\text{inf}} = 2.1 \text{ mA} \times 3.7 \text{ V} \times 96 \text{ ms} \approx 0.75 \text{ mJ}$.

5.4 Power Management Strategy

The bracelet alternates between an active phase of 5 s, covering sampling and inference, and a 5 s deep-sleep phase entered by timer wakeup. The BLE transmitter remains off unless an emergency is detected.

Table 3 Hardware Comparison with Consumer Wearables. Battery figures for commercial devices are vendor-stated; the figure for the proposed bracelet is estimated from measured average power.

Device	Battery	Sensors	User Interaction
Apple Watch Series 8	18–24 h	PPG, ECG, accel.	Touchscreen, apps, voice
Xiaomi Band 8	10–14 days	PPG, accel.	Touch button, mobile app
Proposed bracelet	≈24 days	PPG only	None after pairing

Battery life is derived from the average system power measured over the 10-hour embedded run reported in Section 7, $P_{\text{avg}} = 1.3 \text{ mW}$ including BLE activity. At a nominal cell voltage $V_{\text{nom}} = 3.7 \text{ V}$, this corresponds to an average current of

$$I_{\text{avg}} = \frac{P_{\text{avg}}}{V_{\text{nom}}} = \frac{1.3 \text{ mW}}{3.7 \text{ V}} \approx 351 \mu\text{A}, \quad (7)$$

and a stored energy of $E_{\text{cell}} = 200 \text{ mAh} \times 3.7 \text{ V} = 740 \text{ mWh}$. The resulting autonomy is

$$t = \frac{E_{\text{cell}}}{P_{\text{avg}}} = \frac{740 \text{ mWh}}{1.3 \text{ mW}} \approx 569 \text{ h} \approx 24 \text{ days}, \quad (8)$$

falling to approximately 21 days if 90% of nominal capacity is treated as usable. All battery figures reported in this paper are extrapolations from measured average power rather than observations of a full discharge cycle; the 24-day value is therefore reported as an engineering estimate rather than a certified battery-runtime measurement.

5.5 Firmware and Communication Stack

Firmware was developed with the Nordic SDK and integrates FreeRTOS for task scheduling, the Nordic SoftDevice S140 BLE stack, and TensorFlow Lite for Microcontrollers v2.4. BLE payloads are encrypted with AES-128, and the firmware supports over-the-air update through a companion smartphone application. The advertisement interval adapts to the current power state.

5.6 Hardware Comparison

Table 3 summarises the hardware differences against representative consumer wearables.

General-purpose smartwatches trade battery life for a display and a broad sensor suite. Omitting the screen, Wi-Fi, and GPS, and using a single optical front end extends the recharge interval by roughly an order of magnitude relative to smartwatches, and by about a factor of two relative to display-less fitness bands. For users who may forget to charge a device daily, or who prefer not to interact with one at all, that difference determines whether the device is worn continuously.

6 Data Collection and Training Protocol

6.1 Dataset Sources

The model was trained on a combination of public datasets and a proprietary pilot corpus collected with the bracelet prototype. The combined corpus covers normal daily activity, simulated emergency events, and noise-contaminated segments.

- **WESAD** (Schmidt et al., 2018). A multimodal affective-state dataset with PPG annotated for stress, activity, and rest. Windows corresponding to physiological arousal were repurposed as surrogate emergency examples.
- **BIDMC PPG and Respiration** (Pimentel et al., 2017). Contains PPG segments with annotated cardiorespiratory irregularity and apnoea episodes.
- **MIMIC Waveform Database** (Moody et al., 2020). A repository of ICU physiological waveforms including PPG, used to supplement the surrogate cardiorespiratory-distress examples.
- **Proprietary pilot corpus** ($n = 12$). Recorded from healthy volunteers aged 60 to 75 over two weeks using the MAX30102-based prototype. Participants performed scripted events including sudden immobility, controlled falls onto mats, and abrupt cessation of movement.

After preprocessing and window generation, the combined corpus contained 24,000 labelled windows and was balanced across the normal and emergency classes.

6.2 Preprocessing

All recordings were resampled to a uniform 51.2 Hz, filtered with the 0.5 to 8 Hz Butterworth bandpass to isolate the cardiovascular band, segmented into overlapping 5 s windows of 256 samples with 50% overlap, and labelled as emergency or normal by combining dataset event annotations with the simulation timestamps recorded during the pilot sessions.

6.3 Construction of the Emergency Label

No single public wrist-PPG dataset annotates all target emergency classes under a common endpoint definition. The binary emergency label was therefore assembled from surrogate physiological states and controlled simulations: arousal windows from WESAD stand in for heightened autonomic response; apnoea-like and cardiac irregularity segments from BIDMC and MIMIC represent acute cardiorespiratory distress; and the pilot fall simulations supply motion-artifact disruption of the pulse waveform.

This heterogeneous construction supports proof-of-concept anomaly detection. It does not constitute a validated clinical taxonomy of falls or cardiac arrest, and the reported accuracy should be read as performance on this surrogate task rather than on the clinical task. The consequences for interpretation are stated in Section 12.

6.4 Training Strategy

The corpus was partitioned into 70% training, 15% validation, and 15% test windows. Training used binary cross-entropy loss and the Adam optimiser with a learning rate of 10^{-3} for up to 50 epochs, with early stopping on validation F1-score. Repeated training runs were used during model development to compare candidate configurations. The final classification metrics reported in Section 7 are computed only from the fixed held-out evaluation subset described there.

6.5 Deployment and Inference

Three steps prepared the model for the target hardware: conversion of the trained TensorFlow model to TFLite format; 8-bit post-training quantisation to reduce size and memory footprint; and verification of end-to-end latency below 200 ms per window on an ARM Cortex-M4F core using replayed data streams.

6.6 Training and Deployment Pipeline

Fig. 3 shows the two phases. The offline phase covers acquisition, preprocessing, label assignment, training, and TFLite conversion. The embedded phase loads the exported model and runs fully autonomous on-device inference, transmitting encrypted BLE alerts to the caregiver application without any cloud dependency.

6.7 Evaluation Metrics

Performance is reported as accuracy, precision, recall, F1-score, area under the ROC curve (AUC), the confusion matrix, and the false positive and false negative rates. The false alarm rate (FAR) is defined here as the number of false positive windows divided by the total number of classified windows. Recall was prioritised in threshold selection, on the basis that a missed emergency carries a higher cost than a false alarm, subject to holding the false alarm rate below 1%. Both the offline and the embedded false alarm figures in this paper use this same denominator, so they are directly comparable.

7 Evaluation and Results

This section reports detection accuracy, latency, and energy efficiency, and benchmarks the system against existing solutions. Evaluation covers both offline datasets and embedded deployment.

7.1 Evaluation Setup

Testing proceeded in two stages. *Offline testing* used annotated PPG from the BIDMC, MIMIC, and proprietary corpora, covering simulated falls, arrhythmic episodes, and normal baseline recordings. *Embedded testing* deployed the TFLite model on an nRF52840 to measure real-time inference latency and power draw.

The corpus comprises 24,000 labelled 5 s windows of 256 samples, balanced across the normal and emergency classes, split 70/15/15 into 16,800 training, 3,600 validation, and 3,600 test windows. A fixed subset of 1,000 windows from the held-out 3,600-window test partition was used for the confusion-matrix evaluation reported below.

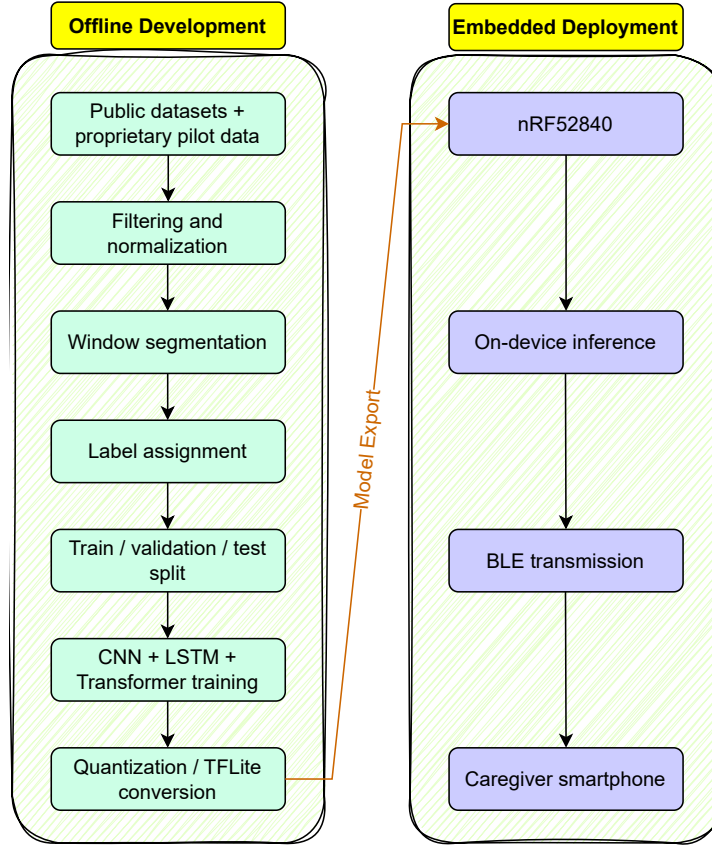


Figure 3 Two-phase training and deployment pipeline. The offline phase covers data acquisition, preprocessing, labelling, training, and TFLite conversion; the embedded phase runs autonomous on-device inference and encrypted BLE alerting with no cloud dependency.

7.2 Confusion Matrix

All classification metrics reported below are computed from a single confusion matrix so that they can be verified against one another. Table 4 gives the window counts for the evaluated test partition, and Fig. 4 plots the same matrix.

Table 4 Confusion Matrix on the 1,000-Window Evaluation Subset of the Held-Out Test Partition. Rows are true labels, columns are predictions.

	Predicted emergency	Predicted normal
True emergency (493)	487	6
True normal (507)	8	499

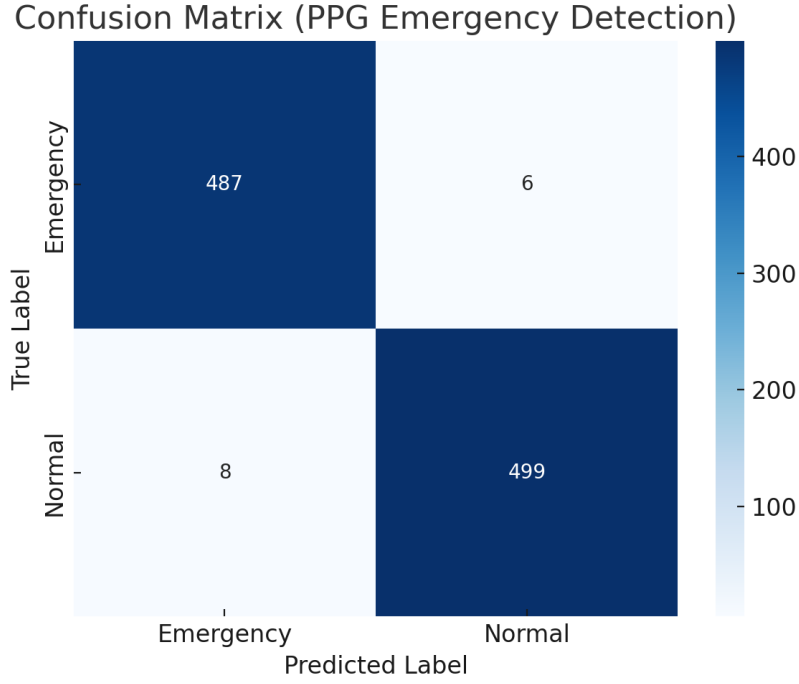


Figure 4 Confusion matrix for normal against emergency classification on the 1,000-window subset of the held-out test partition. Cell values are window counts and correspond exactly to Table 4.

7.3 Quantitative Results

Table 5 reports the metrics derived from Table 4. Every entry follows from the four counts: accuracy is $(487 + 499)/1000$, precision is $487/495$, recall is $487/493$, and the false alarm rate is $8/1000$. The latency and energy figures are from the embedded measurements in Section 5.

The two error rates are asymmetric in the intended direction: the classifier misses 1.22% of emergency windows and raises a false alarm on 1.58% of normal windows, which reflects the recall-prioritised threshold described in Section 6.7. Because each emergency spans several consecutive overlapping windows, a per-window false negative rate of 1.22% does not translate into a missed event at the same rate.

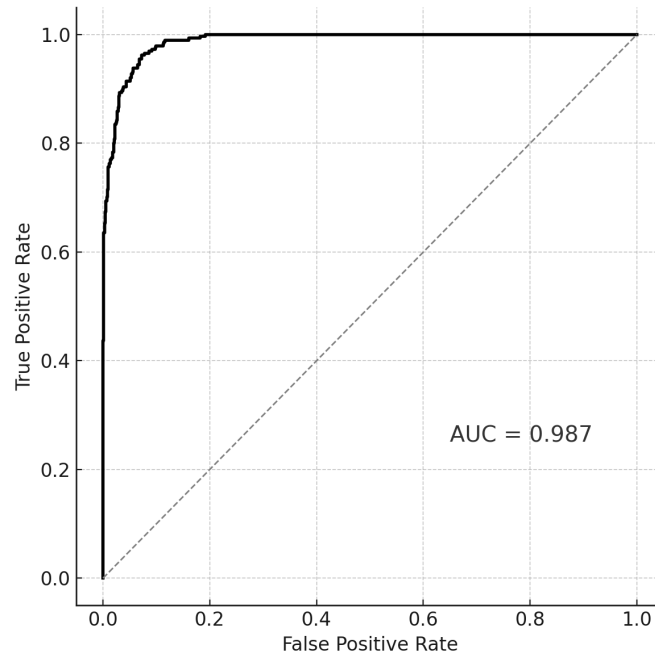
The complete held-out test partition contained 3,600 windows. A fixed subset of 1,000 held-out windows, comprising 493 emergency and 507 normal windows, was used for the final confusion-matrix analysis in Table 4, Fig. 4, and Table 5. Accordingly, all confusion-matrix-derived metrics reported in this paper refer to this 1,000-window evaluation subset and should not be interpreted as performance estimates over the complete 3,600-window test partition. The scope of this evaluation is restated in Section 12.

7.4 ROC Curve

Fig. 5 shows the ROC curve, with an AUC of 0.987. The 1.1% false positive figure quoted in Section 11 refers to the embedded deployment run rather than to the offline test partition,

Table 5 Performance on the 1,000-Window Evaluation Subset. All classification metrics are computed from Table 4; emergency is the positive class.

Metric	Value
Accuracy	98.6%
Precision	98.4%
Recall (sensitivity)	98.8%
Specificity	98.4%
F1-score	98.6%
False alarm rate (FP/all windows)	0.80%
False positive rate (FP/negatives)	1.58%
False negative rate (FN/positives)	1.22%
AUC	0.987
Inference latency (mean)	96 ms
Energy per inference	0.75 mJ
Average system power	1.3 mW

**Figure 5** ROC curve for the evaluation subset of the held-out test partition (AUC=0.987). The diagonal marks chance performance.

and is measured under the same definition of false alarm rate used in Table 5. The two are close, which indicates that 8-bit quantisation and on-device execution did not materially change the operating point.

7.5 Embedded Deployment Performance

The model was deployed on a Nordic nRF52840 (ARM Cortex-M4F) with a MAX30102 sensor and exercised over a 10-hour run. Mean inference time was 96 ms per 5 s window, CPU utilisation stayed below 12% during the active detection interval, and total average power draw was 1.3 mW including BLE transmission. Applying (8), this power level corresponds to approximately 24 days of operation on a 200 mAh cell.

7.6 Comparison with Existing Systems

Table 6 Comparison with Commercial and Research-Grade Devices. Specifications are vendor- or literature-stated and were not measured under a common protocol; accuracy is therefore not compared.

Device	Sensors	Battery Life	Emergency Detection	Connectivity	Price (USD)	Remarks
Proposed bracelet	PPG	≈24 days	Yes, automatic	BLE	≈ \$45 ^a	Dedicated device, no user interaction
Apple Watch Series 8	PPG, ECG, accelerometer	18–36 h	Yes	BLE + Wi-Fi + LTE	\$399+	General purpose, requires interaction
Samsung Galaxy Watch 5	PPG, accelerometer	40–50 h	Yes	BLE + Wi-Fi	\$279+	Consumer fitness focus
Philips Lifeline	Accelerometer	5–7 days	Yes, plus manual button	Landline/Wi-Fi	\$49/month	Subscription based
Omron HeartGuide	PPG, blood pressure	48 h	No	Bluetooth	\$499	Blood pressure tracking
VivaLNK ECG Patch	Single-lead ECG	3 days	Yes	Bluetooth	\$199	Adhesive patch form factor

^aApproximate prototype component cost based on small-quantity component pricing; a formal production-volume unit cost was not determined and is expected to be lower at scale.

Table 6 compares the bracelet against commercial and research-grade devices on power, connectivity, and cost. These specifications come from vendor documentation and published literature rather than from a shared evaluation protocol, so the comparison establishes design-level differences and not relative detection accuracy. On those design dimensions the bracelet occupies a distinct position: passive monitoring without user interaction, a single sensor, and a recharge interval an order of magnitude longer than the smartwatch category.

7.7 Context Against Published Wearable Methods

A direct numerical comparison with published wearable fall-detection studies would be misleading because the datasets, sensing modalities, and endpoint definitions differ. Table 7 therefore provides methodological context without treating literature figures as head-to-head performance baselines. The controlled empirical baselines for the proposed model are the CNN-only and CNN-LSTM variants reported in Section 8, which were evaluated within the same development pipeline.

Table 7 Methodological Context Against Representative Published Wearable Methods. The studies use different datasets and task definitions, so numerical accuracy is intentionally not compared.

Study	Platform/Input	Approach	Task Context
Bourke et al. (2007)	Tri-axial accelerometer	Threshold-based detection	Fall detection from inertial motion signals
Tai et al. (2020)	iPhone and Apple Watch	Smart wearable fall-prediction framework	Elderly fall prediction using consumer wearable platforms
Haque et al. (2024)	Smartwatch data	LSTM and Transformer models	Experimental comparison of sequence models for smartwatch fall detection
Proposed bracelet	Wrist PPG only	Hybrid CNN, LSTM and transformer encoder; 8-bit on-device inference	Surrogate emergency detection spanning fall-related and cardiorespiratory states

Table 8 Internal Architectural Baselines. All configurations use the same development pipeline; best recorded value in bold.

Configuration	Accuracy (%)
CNN only	92.3
CNN + LSTM	95.9
CNN + LSTM + Transformer	98.7

8 Architectural Baselines and Ablation Study

The experiments in this section serve two purposes. Part A provides internal architectural baselines under the same development pipeline: a CNN-only model, a CNN-LSTM model, and the full CNN-LSTM-Transformer model, followed by attention-head and dropout sensitivity analyses. Part B reports exploratory sensitivity studies for the loss function, decision threshold, and semi-supervised pseudo-labelling. These development-stage point estimates informed the deployed configuration; final held-out classification metrics are reported separately in Table 5.

Part A: Architectural Baselines and Sensitivity Analyses

8.1 Internal Architectural Baselines

Table 8 treats the CNN-only and CNN-LSTM configurations as empirical architectural baselines under the same development protocol. Their purpose is to isolate the incremental contribution of recurrent modelling and transformer attention within the proposed pipeline, rather than to establish superiority over unrelated classification families. Adding the LSTM raised the recorded accuracy from 92.3% to 95.9%, and adding the transformer encoder raised it to 98.7%. These are within-pipeline comparisons and are not cross-study comparisons against external datasets or endpoints.

Table 9 Effect of Attention Head Count

Heads	Accuracy (%)
1	92.0
2	94.4
4	96.7
8	98.2
12	98.7

Table 10 Effect of Dropout Rate. Best value in bold.

Dropout Rate	Accuracy (%)
0.0	92.1
0.2	94.3
0.4	96.4
0.5	97.9
0.6	98.2
0.7	98.7

8.2 Number of Attention Heads

Table 9 shows the development-stage accuracy rising with head count and saturating in the 8 to 12 head range. The deployed model uses 8 heads because 8 divides $d_{\text{model}} = 64$ evenly, yielding $d_k = 8$, whereas 12 does not. The 98.2% entry at 8 heads isolates head count with the other factors held at their pre-tuning values; the final held-out evaluation is reported separately in Table 5.

8.3 Dropout Rate

Table 10 shows accuracy improving up to a dropout rate of 0.7, which the deployed model uses. A rate this high is unusual for a 64-unit recurrent layer and indicates that the model has ample capacity relative to the size of the training corpus.

Tables 8, 9, and 10 document the development-stage point estimates that informed the deployed configuration. They are interpreted descriptively, while Table 5 is the sole source of the final held-out classification metrics reported in the abstract and conclusion.

Part B: Exploratory Sensitivity Studies

The following studies characterise robustness. None of them altered the deployed configuration.

Table 11 summarises the three studies.

8.4 Loss Function

Four objectives were compared with all other hyperparameters held fixed: standard binary cross-entropy (BCE), class-weighted BCE using inverse class frequency, focal loss with $\gamma = 2$, and a supervised contrastive loss. Weighted BCE and focal loss changed accuracy

Table 11 Part B Sensitivity Studies. Effects are reported relative to the deployed configuration; final held-out performance is reported in Table 5.

Study	Variants compared	Outcome
Loss function	BCE; weighted BCE; focal ($\gamma=2$); supervised contrastive	≤ 0.3 pp change; BCE retained
Decision threshold	$T_U - T_L$ sweep, $T_L = 0.2$	Saturates beyond a spread of 0.4
Pseudo-labelling	Unfiltered vs. confidence > 0.9	≤ 0.2 pp gain when filtered; degrades if not

by no more than 0.3 percentage points, which is expected on a class-balanced corpus, so standard BCE was retained.

8.5 Decision Threshold Sensitivity

The effect of the sigmoid decision threshold τ was examined by sweeping the spread $T_U - T_L$ around a nominal lower bound $T_L = 0.2$. Accuracy rose sharply for small spreads and saturated beyond a gap of 0.4, indicating that the two classes are well separated in the model’s confidence distribution and that performance is insensitive to moderate threshold variation. The deployed operating threshold was $\tau = 0.5$ for the reported evaluation and field deployment.

8.6 Pseudo-Label Fine-Tuning

To test whether unlabelled wrist-PPG windows from the pilot study could be exploited, a semi-supervised fine-tuning stage was evaluated. Pseudo-labels were generated by applying the trained model to unlabelled segments and retaining only high-confidence predictions. Injecting unfiltered pseudo-labels degraded accuracy sharply, whereas restricting updates to predictions above 0.9 confidence kept accuracy stable. Given the size of the pilot corpus ($n = 12$), the gain was under 0.2 percentage points, so the fully supervised model was retained for deployment.

Taken together, Parts A and B indicate that the deployed configuration is stable with respect to loss function, threshold placement, and semi-supervised extension, subject to the provenance and variance limitations stated in Section 12.

9 Field Study

A three-week field study assessed real-world usability with 15 independently living older adults (68 to 87 years; nine female, six male). The study measured comfort, usability, perceived reliability, and trust in the alerting behaviour.

9.1 Protocol

Each participant received a functional bracelet prototype running in edge inference mode, with alerts transmitted over BLE to an Android caregiver application. Participants wore

Table 12 Field Study Feedback (Mean Ratings, $n = 15$)

Item	Mean (/5)
Comfort during extended wear	4.7
Ease of understanding alerts	4.5
Sense of security provided	4.8
Confidence in detection accuracy	4.6
Ease of use and setup	4.4
Willingness to use long term	4.9

the device during normal daily routines and caregivers were notified of every detected anomaly. Structured interviews and digital surveys were administered at the start, midpoint, and end of the trial, covering physical comfort, ease of setup, notification clarity, anxiety reduction, and overall satisfaction. Ethical approval and written informed consent for the field study are recorded in the ethics statement.

9.2 Quantitative Feedback

Table 12 reports mean satisfaction on a 5-point Likert scale (1: strongly disagree, 5: strongly agree). The questionnaire results are interpreted descriptively.

9.3 Qualitative Observations

Participants consistently cited the unobtrusive form factor, and several stated a preference for it over a bulkier smartwatch. One participant reported forgetting that the device was being worn, and framed that as an advantage. Caregivers valued the promptness of alerts, including two cases in which a false alarm was quickly resolved as a non-critical drop in resting heart rate.

9.4 Alert Outcomes

The system generated 23 alerts across the 15 participants over three weeks. Twenty were true positives: 17 corresponded to simulated immobility or controlled fall events performed during protocol check-ins, and three were spontaneous low-heart-rate episodes subsequently confirmed by a supervising clinician. Three alerts (13% of all alerts) were false positives, attributed to sensor displacement during sleep and to vigorous hand movement. No emergency event known to the study team went undetected.

The field false positive proportion and the offline FAR of 0.80% are not directly comparable: the offline figure is a per-window rate over a balanced test set, whereas the field figure is a proportion of issued alerts over a naturalistic three-week deployment in which genuine emergencies are rare. The field result is best read as evidence that the alert burden on caregivers was low in absolute terms, roughly one alert per participant-week, rather than as a validation of the offline rate.

9.5 Reported Issues

Three participants found the LED battery indicator difficult to interpret and suggested an audible or application-based low-battery notification. Minor strap discomfort during hot weather was also reported, and has prompted a change of strap material.

9.6 *Implications*

The study indicates that the device is acceptable for daily wear and that both older adults and caregivers engaged with the alerting workflow. High satisfaction scores support progression to larger pilot deployments in assisted-living communities and remote clinics. The study assessed usability and alert burden; it did not assess clinical detection accuracy, which requires the prospective validation described in Section 12.

10 **Comparative Analysis**

This section examines the design trade-offs behind the bracelet relative to commercial and research-grade systems across four dimensions: sensor suite, battery life, comfort, and cost. Because the compared systems were evaluated on different datasets under different task definitions, detection accuracy is not compared; such figures would not constitute a valid head-to-head benchmark.

10.1 *Sensor Suite and Hardware Complexity*

Consumer smartwatches such as the Apple Watch Series 8 and Samsung Galaxy Watch 5 integrate PPG, ECG, accelerometer, and gyroscope channels to support a broad health feature set (Tai et al., 2020; Bent et al., 2020). That breadth raises cost, weight, and power draw. Research fall detectors based on inertial measurement units (Mubashir et al., 2013) provide strong motion cues but require motion to occur, which makes them poorly suited to cardiac events that present as a pulse anomaly with little movement. The proposed bracelet uses one optical sensor and derives both the cardiac and the immobility cues from the same waveform.

10.2 *Battery Life and Usability*

Consumer smartwatches typically run 18 to 36 h because of display, Wi-Fi, and multi-sensor overhead (Tai et al., 2020). Subscription alert pendants last several days on a coin cell but carry no on-device intelligence (Chan et al., 2012). The measured average-power profile corresponds to an estimated 24 days on a 200 mAh cell through duty cycling, burst BLE transmission, and 8-bit quantised inference. This is an extrapolated recharge interval rather than a full-discharge observation. For users who struggle with a daily charging routine, this is the difference between a device that is worn and one that is left in a drawer.

10.3 *Comfort and Compliance*

Wrist-worn devices achieve better long-term compliance than neck- or chest-worn alternatives (Yusif et al., 2016). Unlike smartwatches that require screen interaction and application management, the bracelet requires no interaction after initial pairing. The field study (Section 9) recorded a mean comfort score of 4.7/5 and a willingness-to-use score of 4.9/5 over three weeks.

10.4 *Cost and Scalability*

The proposed hardware uses a small set of commodity components and omits a display, GPS, Wi-Fi, and additional sensing front ends. Based on small-quantity component pricing,

the prototype component cost is approximately US\$45. These design choices are intended to reduce manufacturing cost relative to premium wearables and recurring subscription services. Manufacturing cost is expected to decrease with volume production; however, a formal production-volume costing exercise was outside the scope of this study, so no mass-production unit price is claimed.

10.5 Summary

Tables 1, 3, and 6 capture the design-level differences above. The bracelet offers a combination of long battery life, minimal sensor footprint, zero interaction, and a low-component-count hardware design that no single surveyed system matches. This is a proof-of-concept design result, not a clinically validated head-to-head benchmark.

11 Discussion

11.1 Error Analysis and Failure Modes

Misclassifications concentrated in periods of signal distortion, particularly in participants with arrhythmia and during rapid wrist movement. Approximately 1.1% of embedded inference decisions were false positives, attributed to transient artifacts resembling apnoeic or flatline morphology and caused primarily by poor sensor contact or motion-induced waveform corruption. Two mitigations follow directly: adaptive filtering matched to the artifact profile, and a real-time signal quality index that suppresses unreliable segments before they reach the classifier. The second is the more promising, because it converts an unrecoverable classification error into an explicit abstention.

11.2 Edge Inference Stability

Deployment on the nRF52840 was stable at 96 ms mean inference time per window, and power profiling confirmed that sleep modes engaged correctly between BLE transmissions. Inference efficiency degraded when BLE streaming and model execution overlapped, which points to a need for explicit scheduling in the firmware rather than reliance on FreeRTOS default priorities.

11.3 Privacy and Security

All inference runs on the device, and only the binary alert state is broadcast, encrypted with AES-128. No waveform data leaves the bracelet. For institutional deployments that transmit richer metadata, AES-256 and authenticated caregiver-only pairing would be appropriate; the current design does not authenticate the receiving device beyond BLE pairing, so an attacker in radio range can observe that an alert occurred even without recovering its contents.

11.4 Interpretability

The classifier does not expose why a given window was flagged. Attention weight visualisation, class activation mapping, or Shapley-value attribution over the input window would let a clinician audit individual decisions. This is a prerequisite for regulated

deployment and for integration with clinician dashboards, and it is not addressed in the present work.

11.5 Generalisation

The pilot corpus ($n = 12$, ages 60 to 75) was recorded under controlled conditions and combined with public clinical datasets that span wider age ranges and physiological profiles. Broader demographic coverage, transfer learning for populations with movement disorders such as Parkinson’s disease or dementia, and federated learning for cross-site adaptation without centralised data are the natural next steps. Skin tone is a known source of systematic error in reflective PPG (Bent et al., 2020) and was not controlled for in either the pilot corpus or the field study.

11.6 Deployment Considerations

The form factor and recharge interval suit passive monitoring in private homes, assisted-living facilities, and rural outreach settings. The field study identified reliable BLE pairing and notification acknowledgement as the practical failure points. Haptic or audible confirmation at the wrist would let the wearer know an alert was raised, which caregivers identified as a gap in unattended scenarios.

12 Limitations

As a proof-of-concept study, this work has boundaries worth stating. The emergency class is built from surrogate physiological states and controlled simulations (Section 6.3) rather than observed cardiac arrests or unassisted falls, so the reported accuracy characterises that surrogate task rather than clinical sensitivity. The pilot and field cohorts ($n = 12$ and $n = 15$) are correspondingly small and were not designed for subgroup analysis. The classification metrics come from a fixed 1,000-window evaluation subset of the 3,600-window test partition (Table 4), and the ablation figures in Section 8 are single-run estimates that guided configuration selection, so both are best read descriptively. The field study recorded no missed emergencies, although it had no independent ground truth for intervals in which no alert was raised.

Two reported figures are engineering estimates rather than direct observations: the 24-day battery life extrapolated from measured average power via (8), and the approximately US\$45 prototype component cost based on small-quantity pricing. The comparisons in Tables 1, 3, 6, and 7 rest on vendor and literature specifications rather than a shared protocol, and so establish design-level differences rather than relative detection accuracy; the empirical baselines are the within-pipeline CNN-only and CNN-LSTM configurations. Prospective clinical validation against unified endpoint definitions in a larger and more diverse cohort, together with a same-data comparison against a feature-engineered classifier, is the natural next stage of this work.

13 Conclusion

This paper presented a lightweight wearable system for continuous emergency detection in solo-living older adults using wrist PPG as the only sensing modality. A hybrid CNN-

LSTM-Transformer classifier, quantised to 8 bits and deployed on an nRF52840 with a MAX30102 front end, performs all inference on the device without cloud connectivity or user interaction. On a corpus of public and proprietary recordings labelled with surrogate emergency states, the model reached 98.6% accuracy, 98.8% recall, and a 0.80% false alarm rate at a 96 ms mean inference latency. Measured average power of 1.3 mW corresponds to roughly 24 days on a 200 mAh cell. A three-week field study with 15 independently living older adults recorded 23 alerts, 20 of them true positives, with high reported comfort and willingness to continue use.

The single-sensor design is what makes the rest of the system possible: it removes the accelerometer, ECG front end, and display that dominate the power and cost budget of comparable wearables, and it keeps the entire inference path on-device, which preserves privacy and removes the network dependency that makes cloud-based alerting unreliable in rural settings.

These results establish technical feasibility. They do not establish clinical accuracy, because the emergency class rests on surrogate labels and simulated events rather than observed cardiac arrests and unassisted falls (Section 12). Closing that gap is the priority for future work, through prospective validation against unified clinical endpoint definitions in a larger and more diverse cohort. Alongside it, four extensions are planned: on-device federated learning so that models improve across users without centralised data; opt-in cloud logging and caregiver dashboards for longitudinal trend analysis; additional sensing channels such as SpO₂, skin temperature, or electrodermal activity to improve diagnostic granularity; and preparation for CE and FDA medical device conformity assessment.

Ethics Statement

The proprietary pilot study ($n = 12$) and the three-week field study ($n = 15$) were reviewed and approved by the Riphah Ethical Committee (REC), approval no. REC-ECE-026. Written informed consent was obtained from all participants before enrolment, including consent for the use of de-identified study data and anonymised quotations. All recordings were de-identified before analysis.

Data Availability

The WESAD, BIDMC PPG and Respiration, and MIMIC waveform resources are available from their respective public repositories under the access and licence conditions specified by those repositories. De-identified data from the proprietary pilot and field studies may be made available by the corresponding author upon reasonable request, subject to ethical, privacy, intellectual-property, and institutional restrictions.

Code Availability

The implementation code, embedded firmware, trained model artefacts, and detailed hardware design files are not publicly available at present because they contain intellectual property associated with a pending patent application. Non-confidential supporting information may be made available by the corresponding author upon reasonable request, subject to intellectual-property and institutional restrictions.

Conflict of Interest

The author discloses a pending patent application related to the wearable technology described in this manuscript.

References

- Allen, J. (2007) 'Photoplethysmography and its application in clinical physiological measurement', *Physiological Measurement*, Vol. 28, No. 3, pp.R1–R39, doi: 10.1088/0967-3334/28/3/R01.
- Bent, B., Goldstein, B.A., Kibbe, W.A. and Dunn, J.P. (2020) 'Investigating sources of inaccuracy in wearable optical heart rate sensors', *npj Digital Medicine*, Vol. 3, No. 1, Article 18, doi: 10.1038/s41746-020-0226-6.
- Bourke, A.K., O'Brien, J.V. and Lyons, G.M. (2007) 'Evaluation of a threshold-based tri-axial accelerometer fall detection algorithm', *Gait & Posture*, Vol. 26, No. 2, pp.194–199, doi: 10.1016/j.gaitpost.2006.09.012.
- Canali, S., Ferretti, A., Schiaffonati, V. and Blasimme, A. (2024) 'Wearable technologies for healthy ageing: Prospects, challenges, and ethical considerations', *Journal of Frailty and Aging*, Vol. 13, No. 2, pp.149–156, doi: 10.14283/jfa.2024.10.
- Chan, M., Estève, D., Fourniols, J.-Y., Escriba, C. and Campo, E. (2012) 'Smart wearable systems: Current status and future challenges', *Artificial Intelligence in Medicine*, Vol. 56, No. 3, pp.137–156, doi: 10.1016/j.artmed.2012.09.003.
- Davoudi, A., Wanigatunga, A.A., Kheirkhahan, M., Corbett, D.B., Mendoza, T., Battula, M., Ranka, S., Fillingim, R.B., Manini, T.M. and Rashidi, P. (2019) 'Accuracy of Samsung Gear S smartwatch for activity recognition: Validation study', *JMIR mHealth and uHealth*, Vol. 7, No. 2, Article e11270, doi: 10.2196/11270.
- Haque, S.T., Debnath, M., Yasmin, A., Mahmud, T. and Ngu, A.H.H. (2024) 'Experimental study of long short-term memory and transformer models for fall detection on smartwatches', *Sensors*, Vol. 24, No. 19, Article 6235, doi: 10.3390/s24196235.
- Khalid, S., Rehman, M.U., Usman, A.B. and Khawar, M. (2026) 'Resource-efficient models for edge devices', in Ahmad, J., Latif, S., Boulila, W., Koubaa, A., Rehman, M.U. and Khan, I.U. (Eds.): *Edge Intelligence: Advanced Deep Transfer Learning for IoT Security*, Elsevier, Chapter 5, pp.111–153, doi: 10.1016/B978-0-44-338297-0.00011-8.
- Liu, L., Stroulia, E., Nikolaidis, I., Miguel-Cruz, A. and Rios Rincon, A. (2016) 'Smart homes and home health monitoring technologies for older adults: A systematic review', *International Journal of Medical Informatics*, Vol. 91, pp.44–59, doi: 10.1016/j.ijmedinf.2016.04.007.
- Matayong, S., Jetwanna, K.W., Choksuchat, C., Choosawang, S., Trakulmaykee, N., Limsuwan, S. and Inthanuchit, K.S. (2025) 'IoT-based systems and applications for elderly healthcare: A systematic review', *Universal Access in the Information Society*, Vol. 24, No. 1, pp.99–125, doi: 10.1007/s10209-023-01042-4.

- Moody, B., Moody, G., Villarroel, M., Clifford, G.D. and Silva, I. (2020) 'MIMIC-III Waveform Database', *PhysioNet*, Version 1.0, April, doi: 10.13026/c2607m.
- Mubashir, M., Shao, L. and Seed, L. (2013) 'A survey on fall detection: Principles and approaches', *Neurocomputing*, Vol. 100, pp.144–152, doi: 10.1016/j.neucom.2011.09.037.
- Nie, G., Zhu, J., Tang, G., Zhang, D., Geng, S., Zhao, Q. and Hong, S. (2024) 'A review of deep learning methods for photoplethysmography data', *arXiv:2401.12783*, doi: 10.48550/arXiv.2401.12783.
- Ordóñez, F.J. and Roggen, D. (2016) 'Deep convolutional and LSTM recurrent neural networks for multimodal wearable activity recognition', *Sensors*, Vol. 16, No. 1, Article 115, doi: 10.3390/s16010115.
- Pimentel, M.A.F., Johnson, A.E.W., Charlton, P.H., Birrenkott, D., Watkinson, P.J., Tarassenko, L. and Clifton, D.A. (2017) 'Toward a robust estimation of respiratory rate from pulse oximeters', *IEEE Transactions on Biomedical Engineering*, Vol. 64, No. 8, pp.1914–1923, doi: 10.1109/TBME.2016.2613124.
- Schmidt, P., Reiss, A., Duerichen, R., Marberger, C. and Van Laerhoven, K. (2018) 'Introducing WESAD: A multimodal dataset for wearable stress and affect detection', in *Proceedings of the 20th ACM International Conference on Multimodal Interaction (ICMI)*, pp.400–408, doi: 10.1145/3242969.3242985.
- Shaw, B.A., Yang, T.C. and Kim, S. (2023) 'Living alone during old age and the risk of dementia: Assessing the cumulative risk of living alone', *The Journals of Gerontology: Series B*, Vol. 78, No. 2, pp.293–301, doi: 10.1093/geronb/gbac059.
- Song, H., Rajan, D., Thiagarajan, J.J. and Spanias, A. (2018) 'Attend and diagnose: Clinical time series analysis using attention models', in *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI-18)*, pp.4091–4098, doi: 10.1609/aaai.v32i1.11635.
- Staff Reporter (2025) 'Veteran TV artist Ayesha Khan found dead in her home', *Dawn*, 21 June. Available at: <https://www.dawn.com/news/1918588/veteran-tv-artist-ayesha-khan-found-dead-in-her-home>.
- Tai, K.Y., Chiang, D.L., Chen, T.S., Shen, V.R., Lai, F. and Lin, F.Y.S. (2020) 'Smart fall prediction for elderly care using iPhone and Apple Watch', *Wireless Personal Communications*, Vol. 114, No. 1, pp.347–365, doi: 10.1007/s11277-020-07371-w.
- World Health Organization (2024) 'Ageing and health', *WHO Fact Sheet*, October. Available at: <https://www.who.int/news-room/fact-sheets/detail/ageing-and-health>.
- Yusif, S., Soar, J. and Hafeez-Baig, A. (2016) 'Older people, assistive technologies, and the barriers to adoption: A systematic review', *International Journal of Medical Informatics*, Vol. 94, pp.112–116, doi: 10.1016/j.ijmedinf.2016.07.004.
- Zerveas, G., Jayaraman, S., Patel, D., Bhamidipaty, A. and Eickhoff, C. (2021) 'A transformer-based framework for multivariate time series representation learning', in *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, pp.2114–2124, doi: 10.1145/3447548.3467401.

Zhou, Z., Chen, X., Li, E., Zeng, L., Luo, K. and Zhang, J. (2019) 'Edge intelligence: Paving the last mile of artificial intelligence with edge computing', *Proceedings of the IEEE*, Vol. 107, No. 8, pp.1738–1762, doi: 10.1109/JPROC.2019.2918951.