Vol. 15, No. 1, March 2026, pp. 97~108
ISSN: 2089-4864, DOI: 10.11591/ijres.v15.i1.pp97-108

An edge AIoT system for non-invasive biological indicators estimation and continuous health monitoring using PPG and ECG signals

Hung K. Nguyen, Manh V. Pham
Key Laboratory for Smart Integrated Systems, University of Engineering and Technology, VietNam National University, Hanoi, VietNam

Article Info

Article history:

Received Sep 5, 2025
Revised Dec 3, 2025
Accepted Jan 11, 2026

Keywords:

Artificial intelligence of things
Biological indicators
Deep learning
Edge computing
Health monitoring

ABSTRACT

This paper presents the design and implementation of an artificial intelligence of things (AIoT)-based system that integrates deep learning and edge computing for real-time non-invasive health monitoring, focusing on the estimation of mean arterial pressure (MAP) alongside vital parameters such as heart rate (HR), blood oxygen saturation (SpO_2), and body temperature. Photoplethysmography (PPG) and electrocardiography (ECG) signals are acquired using low-power MAX30102 and AD8232 sensors, preprocessed with lightweight digital filters, and processed through a 1D convolutional neural network (CNN) deployed on a SEEED Studio XIAO ESP32S3 microcontroller. The model trained using the cuff-less blood pressure estimation dataset, achieved a mean absolute error (MAE) of 2.51 mmHg on the embedded microcontroller and 2.93 mmHg when validated against a standard blood pressure monitor. Experimental results demonstrate high accuracy, achieving a MAE below 5 mmHg, thereby meeting the AAMI and British Hypertension Society (BHS) Grade A standards for blood pressure measurement. The system achieves real-time inference with an average latency of 16 ms and efficient memory utilization, ensuring suitability for wearable and embedded devices. Physiological data are transmitted via Wi-Fi to a Firebase cloud platform and visualized through a cross-platform mobile application. The proposed system demonstrates strong potential for remote healthcare applications, particularly in continuous monitoring and early health risk detection.

Corresponding Author:

Hung K. Nguyen
Key Laboratory for Smart Integrated Systems, University of Engineering and Technology
VietNam National University
144 Xuan Thuy St., Cau Giay, Hanoi, Vietnam
Email: kiemhung@vnu.edu.vn

1. INTRODUCTION

Hypertension is one of the most prevalent chronic non-communicable diseases globally, affecting approximately 1.5 billion people. Its incidence continues to rise, particularly in low-income countries, while even in developed nations, hypertension remains underdiagnosed and poorly managed [1]. Many people who suffer from hypertension are unaware of their condition, leading to severe complications such as heart failure, stroke, and kidney failure. Continuous blood pressure monitoring is therefore essential for early detection and prevention of hypertension-related risks.

As cardiovascular diseases increasingly affect younger populations, continuous health monitoring systems have become critical for the timely identification of abnormal conditions [2]. The rapid advancement of artificial intelligence of things (AIoT) technologies has enabled real-time data acquisition, analysis, and health reporting through smart healthcare systems [3].

Recent studies have investigated blood pressure estimation using photoplethysmography (PPG) and electrocardiogram (ECG) signals. To overcome the limitations of manual feature extraction in pulse transit time (PTT)-based methods, this study employs lightweight convolutional neural networks (CNNs) for mean arterial pressure (MAP) estimation, leveraging their strong capability for automatic feature extraction and accurate prediction.

This paper presents an AIoT system that applies deep learning and edge computing to estimate MAP values and other biological indicators, such as heart rate (HR) and blood oxygen saturation (SpO_2). The system collects PPG signals from the MAX30102 sensor and ECG signals from the AD8232 module. A deep learning model is deployed on the SEEED Studio XIAO ESP32S3 microcontroller by using Edge Impulse [4] to estimate blood pressure values. The model is optimized for high performance on the resource-constrained microcontroller. The system uses the HTTP protocol and Wi-Fi wireless connection to transmit data to a server. A mobile application can retrieve real-time values and predict hypertension risk based on predefined thresholds.

The remainder of this paper is organized as follows. Section 2 provides an overview of related work. Section 3 describes the implementation method in detail. Section 4 discusses the experiments and evaluates the implementation's performance. Finally, section 5 provides a discussion and section 6 concludes the paper.

2. RELATED WORKS

2.1. Disease prediction based on biological indicators

Key biological indicators such as MAP, HR, SpO_2, and body temperature are vital for health assessment. MAP should remain above 65 mmHg to prevent hypoperfusion and organ failure, serving as a reliable marker of hypertension [5]–[7]. SpO_2 reflects blood oxygenation and helps detect respiratory or circulatory issues, while HR indicates cardiovascular regulation, with abnormal deviations from the normal 60–100 BPM range often linked to arrhythmia or hypertension. Body temperature deviations from the 36.1–37.2 °C range 36.1–37.2 °C suggest infection or fever. Continuous, non-invasive monitoring of these indicators with AI integration enables early diagnosis and remote healthcare.

2.2. Blood pressure measurement method

MAP [8] is a key physiological indicator of average blood pressure during a cardiac cycle and is crucial for diagnosing conditions such as circulatory shock, organ failure, and hypotension [5]. However, direct MAP measurement requires invasive equipment, limiting its use in home or remote healthcare. Recent studies have focused on non-invasive MAP estimation using PPG and ECG. PPG reflects changes in capillary blood volume, while ECG records cardiac electrical activity [6]. The time interval between the ECG's R-peak and a PPG feature, known as PTT, inversely correlates with blood pressure [9]. Higher BP shortens PTT due to vascular stiffening, whereas lower BP lengthens it. Additional features such as HR, PPG amplitude, and waveform slope provide insights into vascular resistance and elasticity. Traditional machine learning models [10]–[13] require manual feature extraction, limiting real-time applicability [14]. Deep learning, particularly CNNs, automatically extracts nonlinear features and jointly processes PPG–ECG inputs, enabling accurate, continuous, and non-invasive MAP estimation for wearable healthcare systems [14].

2.3. Heart rate measurement method

HR, expressed in beats per minute (BPM), is a fundamental physiological parameter that can be estimated using either PPG or ECG. In this study, HR estimation is derived from ECG signals because they accurately reflect cardiac electrical activity. The approach in Magrupov *et al.* [15] first detects R peaks after preprocessing the ECG with a band-pass filter to remove noise. The R–R interval, defined as the time between two consecutive R peaks, is then calculated, with unrealistic intervals excluded. HR is subsequently computed as (1):

$$HR = \frac{60}{R\text{–}R\ interval} \tag{1}$$

To enable efficient deployment on microcontrollers, peak detection is optimized using adaptive thresholding rather than computationally intensive algorithms [4], ensuring robustness and real-time performance.

2.3. SpO2 measurement method

SpO_2 is measured spectroscopically using a PPG sensor that emits red (RED) and infrared (IR) light. According to Tamura [16], both the AC and DC components of the PPG signal influence SpO_2 accuracy. As proposed by Stojanovic and Karadaglic [17], SpO_2 is estimated from the ratio (R) between RED and IR components as in (2), representing the proportion of oxygenated hemoglobin (HbO_2). HbO_2 absorbs more infrared light, while deoxygenated hemoglobin (Hb) absorbs more red light. A smaller R indicates higher HbO_2 concentration and increased SpO_2, whereas a larger R implies greater red-light absorption, lower oxygen binding, and reduced SpO_2 levels.

$$R = \frac{Red}{IR} = \frac{AC(red)/DC(red)}{AC(IR)/DC(IR)} \tag{2}$$

SpO2 will be calculated from this R ratio using to (3).

$$SpO_2 = 110 - 25R \tag{3}$$

3. IMPLEMENTATION METHOD

3.1. System overview

3.1.1. Hardware architecture

The proposed AIoT system is designed to continuously and non-invasively monitor key biological indicators, including MAP, HR, SpO_2, and body temperature. The system acquires PPG, ECG, and temperature signals as inputs. The system's structural components are illustrated in Figure 1. The PPG signal is captured using the MAX30102, an integrated module that measures HR and SpO_2, while the ECG signal is obtained from the AD8232 sensor. Body temperature is measured by the MLX90614 infrared sensor, enabling non-contact readings. All collected signals are processed by the SEEED Studio XIAO ESP32S3 microcontroller—a compact, Wi-Fi–enabled device optimized for embedded and wearable applications. The ESP32S3 handles signal preprocessing, executes a quantized deep learning model for MAP estimation, and computes HR, SpO_2, and body temperature through lightweight algorithms. The inference results are formatted in JSON and transmitted via Wi-Fi to the Firebase realtime database. A companion mobile application retrieves and visualizes real-time data, issues alerts when parameters exceed predefined thresholds, and supports continuous remote health monitoring. An integrated OLED display also provides local on-device readouts, enhancing system usability and reliability.

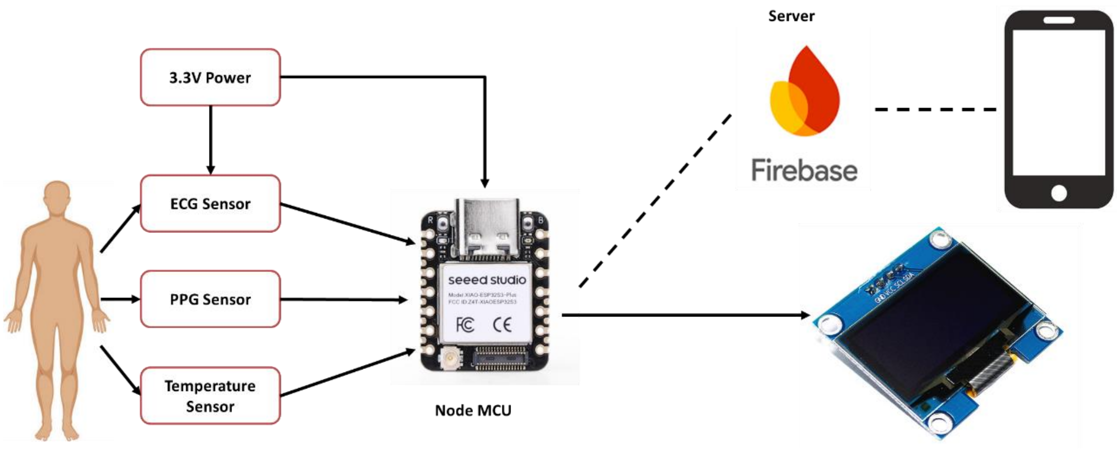


Figure 1. Overall system architecture

The proposed system is portable and operates autonomously, powered by four AAA batteries regulated through an LM256 buck converter to provide stable 5 V and 3.3 V for the ESP32-S3 and sensors. Wi-Fi is used for communication due to its high transmission rate (up to 150 Mbps) and ability to connect directly to cloud services like Firebase without intermediate gateways, reducing system complexity and cost. The ESP32-S3, with built-in Wi-Fi and support for HTTP, HTTPS, and MQTT, provides sufficient computational resources to execute TinyML models for real-time physiological signal processing. A cross-platform React Native mobile application integrated with Firebase enables seamless data synchronization, real-time visualization, and health alerts, supporting continuous and remote health monitoring.

3.1.2. Software overview

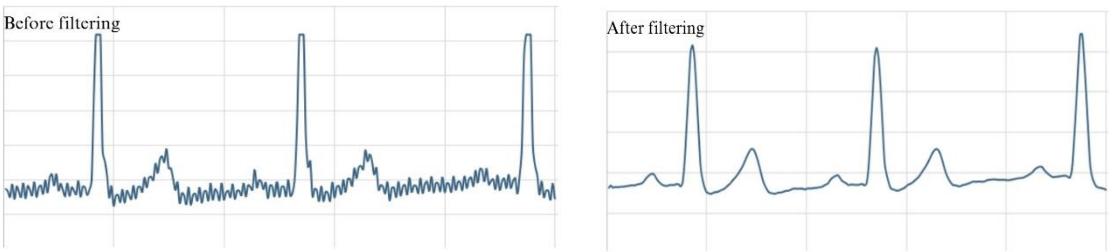
The proposed algorithm, shown in Figure 2, enables real-time estimation of MAP, HR, and SpO_2 using PPG and ECG signals acquired from dedicated sensors. The system software integrates signal acquisition, preprocessing via a moving-average filter, and inference using a lightweight CNN running on an embedded microcontroller. The estimated MAP, HR, SpO_2, and body temperature values are continuously compared against predefined safety thresholds. If the values remain within acceptable limits, the results are uploaded to a cloud database and displayed on a mobile application for user monitoring. In the event of abnormal indicator readings, the system triggers an immediate alert and delivers push notifications to the user's device, thereby supporting timely intervention and continuous health supervision.

```
Start
  Initialize system and sensors    // Setup microcontroller, configure MAX30102,
AD8232, Wi-Fi, etc.
  Loop:
     Read temperature, PPG and ECG signals      // Acquire raw signals from sensors
     Preprocess signals      // Apply moving average filter, normalization
     Estimate MAP using CNN model
     Estimate HR, SpO2      // Run inference on the microcontroller
     If (Indicators are within safe threshold) then
           Upload data to Firebase
           Display Indicators and signals on mobile application
        Else
           Trigger alert
           Send notification to mobile application
           Display alert status on mobile application
    End If
    Wait for next measurement cycle      // Optional: delay to control sampling rate
  End Loop
End
```

Figure 2. Pseudo-code describing software function

3.2. Preprocessing data from sensors

The MAX30102 and AD8232 sensors were configured to sample physiological signals at 125 Hz, ensuring compatibility with the reference dataset. The sampled ECG signal is shown in Figure 3. To enhance the signal quality (Figure 3(a)), preprocessing was performed on the microcontroller using a moving average filter, following the methodology of Kachuee [3]. Unlike low-pass or band-pass filters, which require complex coefficients and greater hardware resources, the moving average filter provides a lightweight and computationally efficient solution suitable for embedded systems such as the ESP32-S3. This approach effectively reduces random noise and motion artifacts while preserving the essential waveform morphology, thereby producing cleaner signals (Figure 3(b)) that are well-suited for subsequent deep learning–based analysis.



(a) (b)

Figure 3. ECG signal; (a) before filtering and (b) after filtering

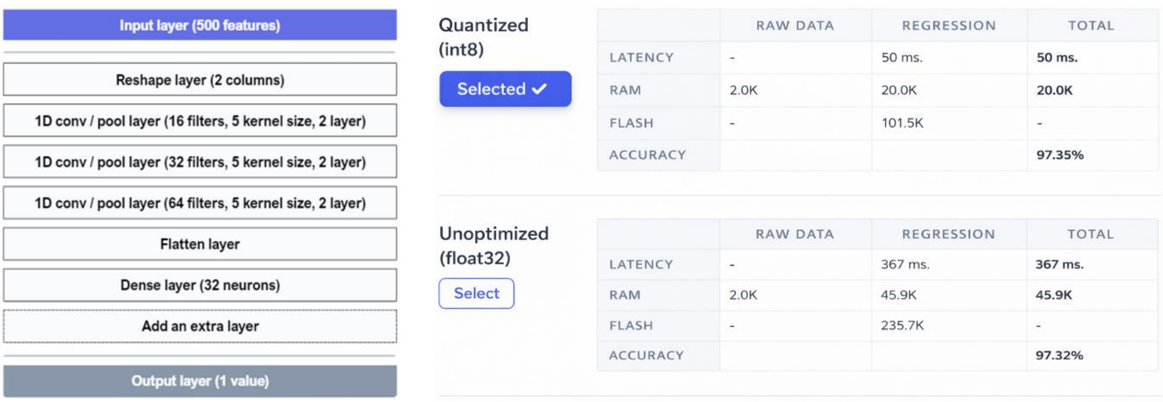
3.3. Blood pressure estimation using deep learning
3.3.1. The proposed convolutional neural network model

Deep learning enables automatic extraction of features relevant to blood pressure estimation and analysis of physiological signals [18]. CNNs, known for their ability to learn complex temporal and spatial patterns, can estimate MAP directly from ECG and PPG signals, making them ideal for real-time, wearable health-monitoring applications [14], [19]. Many studies employ hybrid CNN models, such as CNN–SVR or CNN–LSTM [14], [20]; however, these are unsuitable for deployment on resource-constrained microcontrollers. Long short-term memory (LSTM)requires extensive memory and sequential computation, resulting in slower inference and higher power consumption, while support vector regression (SVR) depends on nonlinear kernel functions and numerous support vectors, exceeding microcontroller unit (MCU) capacity. Therefore, lightweight architectures such as one-dimensional convolutional neural networks (1D-CNNs) or simple fully connected networks are preferred, as they offer high efficiency, fast inference, and full compatibility with TFLite Micro for embedded AI applications.

To evaluate the feasibility of deep learning for non-invasive blood pressure estimation, this study developed a 1D-CNN to estimate MAP from raw PPG and ECG signals on the Kaggle platform. The model leverages CNNs' hierarchical feature extraction capability through three Conv1D blocks of increasing complexity. The first block uses two convolutional layers (16 filters, kernel size 5) with MaxPooling1D to detect fundamental signal characteristics such as slopes and waveform contours. The second block doubles the number of filters per layer to 32 to extract complex temporal features, including waveform variability and peak intervals. The third block increases the filters per layer to 64 to capture nonlinear dependencies between biosignals and MAP. After feature extraction, a Flatten layer converts the feature maps into a one-dimensional vector, which is then processed by a fully connected dense layer with 32 neurons and ReLU activation. A final dense output layer provides the predicted MAP value. This architecture achieves a strong balance between computational efficiency and accuracy, making it suitable for wearable health-monitoring applications.

3.3.2. Deploying the model on the microcontroller unit

The configuration of CNN model is shown in Figure 4. The CNN model, after validation on the Kaggle platform, was imported into Edge Impulse, as shown in Figure 4(a) for training and quantization, and then deployed onto the Seeed Studio Xiao ESP32S3 microcontroller. Before deployment on the microcontroller, the trained model is quantized to reduce memory usage and improve inference speed. Quantization converts model weights and operations from high-precision (float32) to lower-precision (int8), significantly reducing resource consumption while maintaining comparable accuracy. As shown in Figure 4(b), when evaluated with a 10 mmHg signal threshold, the quantized model exhibits minimal accuracy loss compared to the original version, demonstrating its suitability for resource-constrained hardware. After training on the Edge Impulse platform, the quantized model is packaged as an Arduino library and deployed on the MCU. Once uploaded, the MCU processes physiological input signals such as PPG and ECG in real time, executing on-device inference to estimate blood pressure directly at the edge. This approach eliminates the need for continuous cloud-based computation, enabling low-latency, energy-efficient, and fully autonomous health monitoring.



(a) (b)

Figure 4. Configuration of the CNN model on; (a) edge Impulse and (b) comparison of models with and without quantization

3.3.3. Database

The proposed CNN model was trained on the cuff-less blood pressure estimation dataset [21], derived from the MIMIC II database [3], which contains raw ECG, PPG, and invasive arterial blood pressure (ABP) recordings collected in clinical settings. To ensure data reliability, extensive preprocessing was applied, including moving-average smoothing, removal of unreadable blood pressure segments, elimination of physiologically implausible HR, exclusion of discontinuous intervals, and filtering out PPG signals with abrupt waveform distortions. After preprocessing, the dataset provides three synchronized signal channels sampled at 125 Hz, totalling approximately 500 MB or 45 hours of recordings. Signal normalization was performed to standardize input ranges: PPG amplitudes were scaled to 0–4 within each 2-second window, while ECG values were normalized to –0.5–1.5. MAP was computed as the average blood pressure per cycle and used as the ground-truth label. This workflow ensures consistent, high-quality input for training the deep learning model, thereby ensuring optimal performance.

3.4. Implementation of heart rate estimation

In this work, HR is estimated from ECG signals, which provide a more reliable representation of cardiac electrical activity than PPG [15]. The algorithm in [22] consists of four main stages. First, the ECG signal is preprocessed using a noise-reduction filter. Second, R-peak detection is performed, enabling computation of R–R intervals, defined as the time differences between consecutive R peaks. Physiologically unrealistic intervals are removed through an outlier-rejection procedure. Third, HR is calculated as the inverse of the mean R–R interval and converted to beats per minute, see (1). To ensure efficient deployment on microcontrollers, R-peak detection is implemented using a simple amplitude-thresholding method, with the threshold set to 80% of the maximum value within each 2-second ECG window. This lightweight approach enables real-time operation, achieving an inference time of approximately 1–2 ms in the implementation, thus meeting the requirements for continuous monitoring in wearable and embedded systems.

3.5. Implementation of SpO_2 estimation

In this study, SpO_2 is estimated using signals acquired from the MAX30102 [23] sensor, which simultaneously records red and infrared light intensities. The calculation is based on separating the PPG signal into direct current (DC) and alternating current (AC) components [4]. The DC component represents the steady background light transmitted through tissue, obtained by averaging the signal over a fixed period (e.g., 2 seconds).

$$DC = \frac{1}{n}\sum_{i=1}^{n-1} PPG[n-i] \tag{4}$$

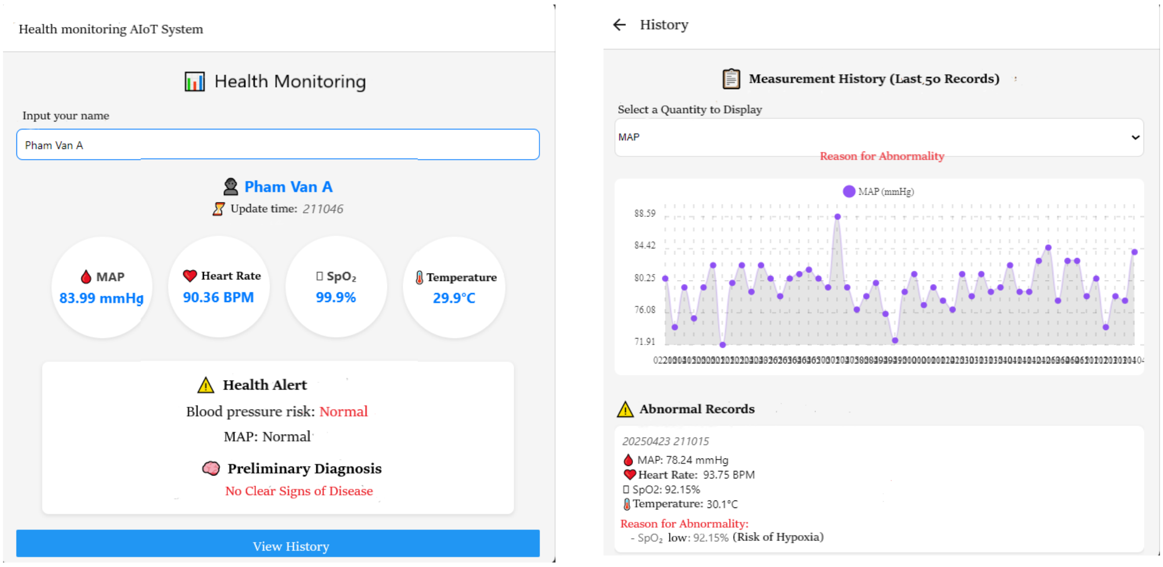
In contrast, the AC component reflects pulsatile variations associated with cardiac activity, computed as the difference between the maximum and minimum values of the PPG waveform.

$$AC = \max(PPG) - \min(PPG) \tag{5}$$

By combining the AC and DC components of both red and infrared signals, the SpO_2 value is derived using established equations, see (2) and (3), as described in sub-section 0. This algorithm is implemented as embedded functions on the ESP32-S3 microcontroller, enabling real-time estimation with an inference time of 1–2 ms. The low computational overhead ensures efficient operation for continuous health monitoring applications.

3.6. Disease prediction on mobile application

The mobile application provides a seamless platform for continuous monitoring of physiological parameters, including MAP, HR, SpO_2, and body temperature. Data transmitted from the measurement device to Firebase is retrieved through the REST API, enabling real-time synchronization and instantaneous updates. Beyond visualization, the application incorporates threshold-based analytics to detect abnormal conditions; for example, MAP values above 110 mmHg (stage 1 hypertension) or SpO_2 below 90% automatically trigger alert notifications. The interfaces, see Figure 5, support real-time data display (Figure 5(a)) and access to historical records (Figure 5(b)), allowing users to analyze long-term trends and review logged abnormal events. A key advantage of the system is its uninterrupted data acquisition: physiological measurements continue to be uploaded to Firebase even when the app is closed, the screen is locked, or network connectivity is interrupted. Once reconnected, the application automatically synchronizes historical data, ensuring reliable health tracking and timely detection of potential risks.



(a) (b)

Figure 5. Main interface of the mobile application; (a) real-time data and (b) history interface

4. EXPERIMENTS AND EVALUATIONS

4.1. Evaluation of the CNN model for MAP estimation

The CNN model for MAP estimation is evaluated using the following criteria: mean absolute error (MAE), mean square error (MSE), and the coefficient of determination (R^2).

4.1.1. Model training results in Kaggle

The model was trained on preprocessed input data split into 2-second samples, with 80% allocated to training and 20% to testing. Training was conducted over 100 epochs with a batch size of 32, using MSE as the loss function. As illustrated in Figure 6, both training and validation losses decreased sharply during the first 20 epochs, indicating efficient optimization and rapid learning of fundamental signal features. From epoch 20 onward, the losses continued to decline gradually, ultimately converging toward near-zero values. Between epochs 40 and 100, both curves remained stable with only minor fluctuations, and importantly, no evidence of divergence was observed. This stability confirms that the model avoids overfitting and maintains consistent performance on unseen data. Additionally, the minimal and consistent gap between training and validation loss highlights the strong generalization capability of the proposed architecture, demonstrating its effectiveness in learning complex relationships between physiological signals and MAP.

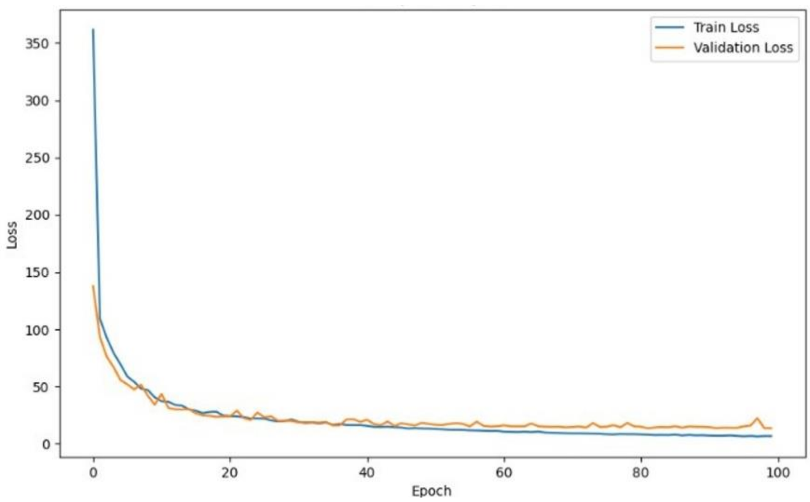


Figure 6. Losses of the CNN 1D model

Figure 7 shows a segment of data comparing the ground-truth MAP values computed from the dataset with the those estimated by the CNN model. In the plot, the blue line represents the actual MAP, while the red line indicates the predicted MAP. Overall, the predicted values closely follow the actual MAP trend, demonstrating that the model effectively learns and captures the patterns and amplitude of blood pressure variations. Moreover, a more comprehensive evaluation of the model is conducted using standard regression metrics, as shown in Table 1.

The proposed model achieved a MAE of 2.46 mmHg, substantially lower than traditional hand-crafted feature extraction methods and well within the AAMI [24] acceptance threshold of 5 mmHg for blood pressure measurement devices. The MSE of 13.55 mmHg² further indicates that prediction errors remain small on average. Additionally, the R^2 of 0.88 demonstrates that the model explains 88% of the variance in actual MAP values, reflecting strong predictive capability despite the inherent variability of biological signals.

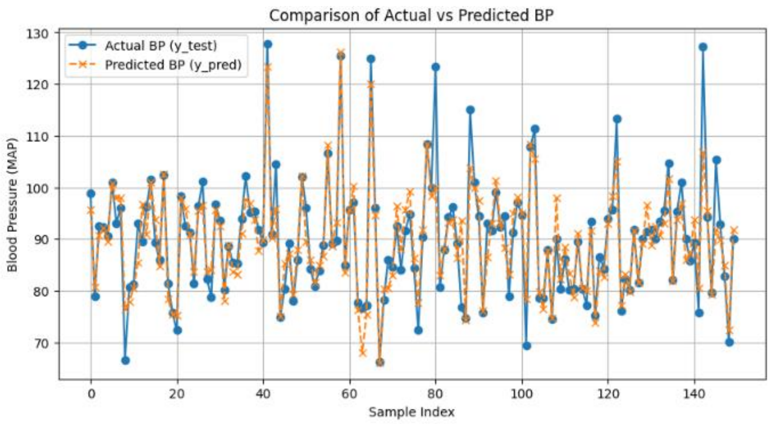


Figure 7. Comparison of actual and predicted MAP

Table 1. Performance metrics of the model on test dataset

Metric	Result
MAE	2.46 mmHg
MSE	13.55 mmHg^2
R^2	0.88

4.1.2. Model training results in edge impulse

The model was trained in Edge Impulse with 50 epochs, MSE as the loss function, a learning rate of 0.005, a batch size of 32, and 20% of the data reserved for validation. Training was performed on the Espressif ESP-EYE device, operating at 240 Hz, consistent with the SEEED STUDIO XIAO ESP32S3 platform. The trained model achieved a MSE of 17.06 mmHg², corresponding to a root mean squared error (RMSE) of 4.13 mmHg, and a MAE of 2.94 mmHg, which satisfies the AAMI standard (<5 mmHg). Furthermore, the explained variance score reached 0.88, indicating strong predictive capability. Performance was comparable to the Kaggle-trained model, confirming robustness on the cuff-less blood pressure estimation dataset. Resource analysis demonstrated the model's feasibility for embedded deployment, with an inference time of 50 ms per each 2-second data segment, RAM usage of 28.2 KB, well below the 320 KB available, and flash memory usage of 100.2 KB, significantly less than the 8 MB capacity. These results highlight the efficiency and practicality of deploying the CNN model for real-time, non-invasive blood pressure estimation on wearable devices.

4.1.3. Verification on XIAO ESP32S3 MCU with the testing dataset

The dataset used to test the model on the XIAO ESP32S3 MCU includes actual MAP values and input strings. The model performed inference with an average runtime of approximately 50 milliseconds per prediction. After performing inference on the dataset, the model's performance was evaluated using the MAE and MSE metrics, as shown in Table 2.

The evaluation shows the model's performance on the actual MCU closely matches its performance during training with Edge Impulse. The MAE of 2.511 mmHg demonstrates that the model maintains relatively low prediction error when running on hardware, meeting the AAMI standard for non-invasive blood pressure monitoring. The MSE is slightly lower than that in the training phase, suggesting that the test dataset used on the MCU contains fewer samples that are significantly deviant. The model's real-world performance was also evaluated using the British hypertension society (BHS) standard [25], as shown in Table 3.

Table 2. Model evaluation on XIAO ESP32S3 MCU

Metric	Result
MAE	2.51 mmHg
MSE	8.88 $mmHg^2$

Table 3. Performance evaluation compared with BHS standard

Parameter	mmHg		
	≤5	≤10	≤15
Result	82.3%	94.58%	99.02%
BHS-A	≥60%	≥85%	≥95%
BHS-B	≥50%	≥75%	≥90%
BHS-C	≥40%	≥65%	≥85%

These results indicate that the model deployed on the target hardware platform meets the grade A requirements of the BHS standard for blood pressure measurement. The trained model has been optimized effectively, maintaining stable performance with acceptable error margins. Given the system's objective of providing timely health warnings to users for early disease management, such error levels are considered acceptable and clinically relevant.

4.1.4. Compare with the standard BP device

Because conventional blood pressure monitors inherently exhibit measurement error, this evaluation aims to assess whether the proposed system can practically substitute cuff-based methods. Since MAP cannot be measured directly, it is estimated using the standard formula:

$$MAP = \frac{SBP}{3} - \frac{2 \times DBP}{3} \tag{6}$$

where SBP denotes systolic blood pressure and DBP denotes diastolic blood pressure.

Each blood pressure measurement with a standard device takes approximately 40 seconds. To provide a comprehensive assessment of the system, 20 model predictions were recorded over a 40-second interval, then averaged, and compared with the blood pressure monitor's results.

The system was tested on five individuals, including one with hypertension and four healthy individuals, with 20 measurements per subject at 3-minute intervals. For healthy individuals, predicted MAP values ranged from 85 to 93 mmHg, yielding an MAE of 2.93 mmHg and an MSE of 12.27 $mmHg^2$. Although slightly higher than training results, these errors remained well within the clinically acceptable margin of 5 mmHg. For the hypertensive subject, predicted MAP ranged from 106.45–117.34 mmHg, corresponding to Stage 1 hypertension, with an MAE of 3.56 mmHg and an MSE of 25.64 $mmHg^2$. While prediction errors were higher in this case, results remained clinically valid. Overall, the system provided faster, more comfortable, and non-invasive measurements compared to conventional cuff-based devices. With accuracy near the 5 mmHg threshold, the model demonstrates strong potential as a reliable alternative for continuous blood pressure monitoring.

4.2. Evaluation of HR and SpO_2 estimation

The estimated HR and SpO_2 values from the proposed AIoT system were validated against measurements obtained from the Pulse Oximeter LK87, a standard non-invasive device. Both systems were tested simultaneously on the same subjects, with results recorded after each measurement cycle. Performance evaluation using MAE and MSE, as summarized in Tables 4 and 5, demonstrates that the implemented algorithms provide HR and SpO_2 estimations closely aligned with the reference device.

Table 4. Evaluation of SpO_2

Metric	Result (%)
MAE	1.92
MSE	7.16

Table 5. Evaluation of HR

Metric	Result
MAE	2.77 BPM
MSE	12.15 BPM^2

4.3. Evaluation of response time

The proposed system delivers immediate and continuously updated health indicators by minimizing end-to-end response time from data acquisition to result display. Experimental evaluation shows an average inference time of 16 ms per prediction, meeting real-time monitoring requirements. With a sampling frequency of 125 Hz (≈8 ms per sample), batch processing combined with low-latency inference enables seamless data handling and timely alert generation. Reducing computational delay also reduces energy consumption, extending battery life in portable devices. These results demonstrate that the model provides both high accuracy and efficient runtime performance, supporting practical deployment on resource-constrained embedded platforms for continuous health monitoring.

5. DISCUSSION

This project applied a moving average filter to preprocess PPG and ECG signals, which were then fed into a pre-trained 1D-CNN model deployed on a microcontroller for non-invasive MAP estimation. The model achieved 97.32% accuracy within a 10 mmHg error margin on the training dataset and 94% accuracy compared to a standard blood pressure monitor. Inference averaged 16 ms per sample, demonstrating real-time processing capability. These results confirm the feasibility of implementing an embedded deep learning system for remote health monitoring, especially for home care, postoperative supervision, and chronic disease management. Compared with state-of-the-art models, as shown in Table 6, such as the CNN–SVR hybrid model [14], which achieves lower MAE values but is unsuitable for embedded deployment, the proposed 1D-CNN offers a favorable balance between accuracy and computational efficiency, outperforming traditional machine learning methods such as ANN and SVM reported by Kachuee e*t al.* [3]. Its compatibility with TFLite Micro further ensures efficient real-time inference on microcontrollers. Overall, the findings validate the system's practicality, offering a robust and scalable approach to non-invasive blood pressure monitoring and early health risk detection.

Table 6. Comparison of MAP estimation models

Model	MAE (mmHg)	MCU deployable
ANN [3]	>5	No
SVM [3]	>5	No
Hybrid CNN-SVR [14]	1.23 ± 2.45 (SBP), 3.08 ± 5.67 (DBP)	No
1D CNN (this work)	**2.46 (MAP)**	**Yes**

The dataset used in this study, derived from invasive clinical procedures, lacks diversity and remains limited to a small group of healthy individuals under simulated conditions. The absence of data from patients with cardiovascular, respiratory, or infectious diseases restricts real-world validation. To enhance generalizability, future datasets should include broader populations, various physiological states (rest, activity, exercise), and environmental conditions (lighting, temperature, time). Such diversity would improve model robustness and clinical relevance. Future work will focus on real-time health trend analysis and advanced AI models for disease risk prediction and early clinical decision support beyond threshold-based alerts. Various other models can be deployed to predict conditions such as tachycardia, fever, and influenza. According to ongoing research [9]-[17], the system will be further integrated with additional functions, enabling more comprehensive assessments of health conditions.

6. CONCLUSION

This paper presents an initial AIoT system for monitoring key health indicators, including MAP, HR, SpO_2, and body temperature, using deep learning and signal processing techniques. Edge computing principles were applied to deploy the trained model on embedded hardware, enabling real-time non-invasive measurements. Based on clinical research, specific threshold values were defined to generate timely health alerts. The system uses a moving-average filter to process PPG and ECG signals, which are then fed into a CNN-based deep learning model deployed directly on a microcontroller for MAP estimation. Tests using the training dataset achieved a prediction accuracy of 97.32% within a 10 mmHg margin and 94% compared to commercial reference devices. The model achieved a MAE of 2.51 mmHg on the embedded microcontroller and 2.93 mmHg when validated against a standard blood pressure monitor. Algorithms for HR and SpO_2 estimation run entirely on the MCU, with low MAE and MSE values: 2.77 BPM and 12.15 BPM^2 for HR, and 1.92% and 7.16% for SpO_2, respectively. The average inference time of 16 ms per sample demonstrates the system's capability for real-time health tracking. These promising results confirm the feasibility and potential of developing non-invasive health-monitoring and early-warning systems, providing a foundation for remote health supervision, especially for home-based elderly care.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Hung K. Nguyen	✓	✓				✓	✓			✓		✓	✓	✓
Manh V. Pham			✓	✓	✓			✓	✓		✓			

C : **C**onceptualization	I : **I**nvestigation	Vi : **Vi**sualization
M : **M**ethodology	R : **R**esources	Su : **Su**pervision
So : **So**ftware	D : **D**ata Curation	P : **P**roject administration
Va : **Va**lidation	O : Writing - **O**riginal Draft	Fu : **Fu**nding acquisition
Fo : **Fo**rmal analysis	E : Writing - Review & **E**diting	

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

INFORMED CONSENT

We have obtained informed consent from all individuals included in this study.

ETHICAL APPROVAL

The research related to human use has been complied with all the relevant national regulations and institutional policies in accordance with the tenets of the Helsinki Declaration and has been approved by the authors' institutional review board or equivalent committee.

DATA AVAILABILITY

Data availability is not applicable to this paper as no new data were created or analyzed in this study.

REFERENCES

[1] P. Verdecchia, C. Cavallini, dan F. Angeli, “Advances in the treatment strategies in hypertension: Present and future,” *Journal of Cardiovascular Development and Disease*, vol. 9, no. 3, p. 72, 2022, doi: 10.3390/jcdd9030072.

[2] E. O. Lopez, B. D. Ballard, and A. Jan, “Cardiovascular disease,” *Molecular and Cell Biology of Human Gene Therapeutics*, Dordrecht: Springer, 2025, doi: 10.1007/978-94-011-0547-7_15.

[3] M. Kachuee, M. M. Kiani, H. Mohammadzade and M. Shabany, "Cuff-less high-accuracy calibration-free blood pressure estimation using pulse transit time," *2015 IEEE International Symposium on Circuits and Systems (ISCAS)*, Lisbon, Portugal, 2015, pp. 1006-1009, doi: 10.1109/ISCAS.2015.7168806.

[4] R. A. Félix, A. Ochoa-Brust, W. Mata-López, R. Martínez-Peláez, L. J. Mena, dan L. L. Valdez-Velázquez, “Fast parabolic fitting: An R-peak detection algorithm for wearable ECG devices,” *Sensors*, vol. 23, no. 21, p. 8796, 2023, doi: 10.3390/s23218796.

[5] Z. Song *et al.*, “Mean arterial pressure trajectory with premature cardiovascular disease and all-cause mortality in young adults: The Kailuan prospective cohort study,” *Frontiers in Cardiovascular Medicine*, vol. 10, 2023, doi: 10.3389/fcvm.2023.1222995.

[6] D. Seok, S. Lee, M. Kim, J. Cho, dan C. Kim, “Motion artifact removal techniques for wearable EEG and PPG sensor systems,” *Frontiers in Electronics*, vol. 2, p. 685513, 2021, doi: 10.3389/felec.2021.685513.

[7] H. Kandil, A. Soliman, N. S. Alghamdi, J. R. Jennings, dan A. El-Baz, “Using mean arterial pressure in hypertension diagnosis versus using either systolic or diastolic blood pressure measurements,” *Biomedicines*, vol. 11, no. 3, p. 849, 2023, doi: 10.3390/biomedicines11030849.

[8] D. Daniel dan W. Daliah, “Physiology, mean arterial pressure,” *StatPearls*, Treasure Island, FL, USA: StatPearls Publishing, 2023, [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK538226/. (Accessed: Jun 30, 2025).

[9] H. Tang *et al.*, “Blood pressure estimation based on PPG and ECG signals using knowledge distillation,” *Cardiovascular Engineering and Technology*, vol. 15, pp. 39–51, 2024, doi: 10.1007/s13239-023-00695-x.

[10] R. Wang, W. Jia, Z. H. Mao, R. J. Sclabassi, dan M. Sun, "Cuff-free blood pressure estimation using pulse transit time and heart rate," in *Proceedings of the International Conference on Signal Processing*, Beijing, China, 2014, pp. 115–118, doi: 10.1109/ICOSP.2014.7014980.

[11] S. Ghosh, A. Banerjee, N. Ray, P. W. Wood, P. Boulanger, dan R. Padwal, "Continuous blood pressure prediction from pulse transit time using ECG and PPG signals," in *Proceedings of the IEEE Healthcare Innovation Point-of-Care Technologies Conference*, Cancun, Mexico, 2016, pp. 188–191, doi: 10.1109/HIC.2016.7797728.

[12] S. Haddad, A. Boukhayma, dan A. Caizzone, "Continuous PPG-based blood pressure monitoring using multi-linear regression," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 5, pp. 2096–2105, 2022, doi: 10.1109/JBHI.2021.3128229.

[13] T. Treebupachatsakul, A. Boosamalee, S. Shinnakerdchoke, S. Pechprasarn, dan N. Thongpance, "Cuff-less blood pressure prediction from ECG and PPG signals using Fourier transformation and amplitude randomization preprocessing for context aggregation network training," *Biosensors*, vol. 12, no. 3, p. 159, 2022, doi: 10.3390/bios12030159.

[14] S. Rastegar, H. Gholam Hosseini, dan A. Lowe, "Hybrid CNN–SVR blood pressure estimation model using ECG and PPG signals," *Sensors*, vol. 23, no. 3, p. 1259, 2023, doi: 10.3390/s23031259.

[15] T. Magrupov, M. Magrupova, S. Abdihalikov, dan D. Ripka, "ECG signal processing algorithms to determine heart rate," in *Proceedings of the IEEE International Conference on Electrical Engineering and Photonics*, St. Petersburg, Russia, 2020, pp. 132–135, doi: 10.1109/EExPolytech50912.2020.9243961.

[16] T. Tamura, "Current progress of photoplethysmography and SpO_2 for health monitoring," *Biomedical Engineering Letters*, vol. 9, no. 1, pp. 21–36, 2019, doi: 10.1007/s13534-019-00097-w.

[17] R. Stojanovic dan D. Karadaglic, "Design of an oximeter based on LED–LED configuration and FPGA technology," *Sensors*, vol. 13, no. 1, pp. 574–586, 2013, doi: 10.3390/s130100574.

[18] Y. LeCun, Y. Bengio, dan G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015, doi: 10.1038/nature14539.

[19] A. Tazarv dan M. Levorato, "A deep learning approach to predict blood pressure from PPG signals," in *Proceedings of the IEEE Engineering in Medicine and Biology Society Annual International Conference*, Mexico City, Mexico, 2021, pp. 5658–5662, doi: 10.1109/EMBC46164.2021.9629687.

[20] H. Mou dan J. Yu, "CNN–LSTM prediction method for blood pressure based on pulse wave," *Electronics*, vol. 10, no. 14, p. 1664, 2021, doi: 10.3390/electronics10141664.

[21] S. Hymel *et al.*, "Edge Impulse: An MLOps platform for tiny machine learning," *arXiv preprint*, arXiv:2212.03332, 2023, doi: 10.48550/arXiv.2212.03332.

[22] H. T. Mason *et al.*, "A complete pipeline for heart rate extraction from infant ECGs," *Signals*, vol. 5, no. 1, pp. 118–146, 2024, doi: 10.3390/signals5010007.

[23] S.-S. Oak dan P. Aroul, *How to Design Peripheral Oxygen Saturation (SpO_2) and Optical Heart Rate Monitoring Systems Using the AFE4403*, Application Report. Dallas, TX, USA: Texas Instruments, 2015.

[24] S. S. Mousavi, M. Firouzmand, M. Charmi, M. Hemmati, M. Moghadam, dan Y. Ghorbani, "Blood pressure estimation from appropriate and inappropriate PPG signals using a whole-based method," *Biomedical Signal Processing and Control*, vol. 47, pp. 196–206, 2019, doi: 10.1016/j.bspc.2018.08.022.

[25] R. Kundu, S. Biswas, dan M. Das, "Mean arterial pressure classification: A better tool for statistical interpretation of blood pressure related risk covariates," *Cardiology and Angiology: An International Journal*, vol. 6, no. 1, pp. 1–7, 2017, doi: 10.9734/CA/2017/30255.

BIOGRAPHIES OF AUTHORS

Hung K. Nguyen received the B.S. and M.S. degrees in electronic engineering from Vietnam National University, Hanoi, Vietnam, in 2003 and 2005, respectively. He received the Ph.D. degree in electronic engineering from Southeast University, Nanjing, China in 2013. He is currently a researcher and assistant professor at VNU Key Laboratory for Smart Integrated Systems, VNU University of Engineering and Technology. His research interests mainly include multimedia processing, reconfigurable computing, and VLSI and SoC designs. He can be contacted at email: kiemhung@vnu.edu.vn.

Manh V. Pham received the B.S. degree in electronic engineering from VNU University of Engineering and Technology in 2025. He is currently a researcher at VNU Key Laboratory for Smart Integrated Systems, VNU University of Engineering and Technology. His research interests mainly include embedded systems and IoT design and SoC design. He can be contacted at email: manhpv1688@gmail.com.