MINISTRY OF EDUCATION AND TRAINING

VIETNAM ACADEMY OF SIENCE AND TECHNOLOGY

GRADUATE UNIVERSITY OF SIENCE AND TECHNOLOGY

DAO TO HIEU

RESEARCH ON METHODS FOR HUMAN ACTIVITY RECOGNITION AND INDOOR POSITIONING TO SUPPORT SEARCH AND RESCUE IN EMERGENCY SITUATIONS

SUMMARY OF DISSERTATION ON
ELECTRICAL, ELECTRONICS AND TELECOMMUNICATION ENGINEERING

Major: **Control and Automation engineering**

Code: **9 52 02 16**

Hanoi - 2026

The dissertation is completed at: Graduate University of Science and Technology, Vietnam Academy of Science and Technology

Supervisors:

1. Supervisor 1: Prof. Tran Duc Tan, Phenikaa University
2. Supervisor 2: Dr. Tran Duc Nghia, Institute of Information Technology

Referee 1:……

Referee 2:……

Referee 3:……

The dissertation will be examined by Examination Board of Graduate University of Science and Technology, Vietnam Academy of Science and Technology at…….. (time, date, year…)

This dissertation can be found at:
1. Graduate University of Science and Technology Library
2. National Library of Vietnam

LIST OF THE PUBLICATIONS RELATED TO THE DISSERTATION

[CT1]. N. T. Thu, **D.-T. Dao*,** B. Q. Bao, D. nghia Tran, P. V. Thanh, and D.-T. Tran, 2022, "Real-time wearable-device based activity recognition using machine learning methods," International Journal of Computing and Digital Systems, vol. 12, no. 1, pp. 321–333, DOI: 10.12785/ijcds/120126 **[SCOPUS, Q3].**

[CT2]. **T.-H. Dao**, H.-Y. Hoang, V.-N. Hoang, D.-T. Tran, and D.-N. Tran*, 2022, Human activity recognition system for moderate performance microcontroller using accelerometer data and random forest algorithm, *EAI Endorsed Transactions on Industrial Networks and Intelligent Systems*, vol. 9, no. 40, pp. 1–18, DOI: 10.4108/eetinis.v9i4.2571 **[SCOPUS, Q3].**

[CT3]. **T.-H. Dao,** D.-N. Tran, V.-H. Bui, V. Son Nguyen, D. K. Hoa*, P. Van Thanh, and D.-T. Tran, 2025, RFAR: A real-time firefighter activity recognition system using wearable accelerometer, *IEEE Sensors Journal*, vol. 25, no. 17, pp. 33 674–33 691, 2025, DOI: 10.1109/JSEN.2025.3593466 **[SCIE Q1].**

[CT4]. **T.-H. Dao**, D.-N. Tran, Q.-T. Hoang, H.-D. Vu, D. T. Huy, and D.-T. Tran, 2023, Developing Real-time Automatic Step Detection On A Low-Cost, Performance-Constrained Microcontroller, *2023 IEEE Statistical Signal Processing Workshop*, pp. 150–154, ISBN: 978-1-6654-5245-8, DOI: 10.1109/SSP53291.2023.10207955. **[SCOPUS, Q4].**

[CT5]. V.-H. Bui, D.-N. Tran, **T.-H. Dao***, T.-A. Le, V.-A. Tran, T.-C. Bui, and D.-T. Tran, 2025, Development of an Indoor Position Monitoring System for Rescue Personnel during Emergencies, *2025 8th International Conference on Multimedia Analysis and Pattern Recognition*, pp. 196–201, ISSN: 979-8-3315-5466-8, DOI: 10.1109/MAPR67746.2025.1113393 **[SCOPUS, Q4].**

[CT6]. **T.-H. Dao,** D.-N. Tran, V.-H. Bui, H.-D. Vu, and D.-T. Tran, 2025, IPWSEP-Investigation of an indoor positioning system based on wearable sensors for emergency response, *Hanoi University of Industry Journal of Science and Technology*, vol. 61, pp. 1-11 [accepted].

INTRODUCTION

1. RESEARCH NECESSITY

The rapid pace of urbanization poses significant challenges to construction safety and fire prevention worldwide [1]. The gradual degradation over time of rescue and firefighting systems can lead to severe consequences [2] when incidents occur. Large-scale and complex fires have increased the risk to firefighters during search and rescue operations. They are frequently exposed to high temperatures [3], toxic smoke [4], as well as the risk of structural collapse [5]. In addition, they must operate in environments with limited visibility [6] and impaired auditory conditions [7]. In fire scenarios, firefighters themselves may become victims and require timely assistance.

challenge to command and coordination in rescue operations. Therefore, the development of a support system for firefighters operating indoors is essential. Accordingly, the doctoral candidate has chosen to pursue the research topic: *Research on Methods for Human Activity Recognition and Indoor Positioning to Support Search and Rescue in Emergency Situations*.

2. OBJECTIVES OF THE STUDY

- To develop an algorithm for a real-time human activity recognition system for firefighters. The proposed algorithm is designed to be deployed on wearable devices integrated with inertial sensors and resource-constrained microcontrollers, enabling energy efficiency, reduced response latency, and suitability for specific firefighter activities.
- To develop a location tracking model deployed on wearable devices for rescue personnel operating inside buildings and infrastructures, based on advanced machine learning and signal processing techniques.

3. METHODOLOGY

- The study focuses on selecting microcontrollers, Ultra Wideband (UWB) transceivers, and sensors such as inertial and barometric sensors to develop a low-cost wearable device.
- The study investigates machine learning and signal processing techniques to evaluate their advantages and limitations in processing and classifying biosignals as well as firefighter activity states.
- The study explores appropriate signal analysis and fusion techniques to estimate position based on motion-related parameters such as movement angles, altitude, and traveled distance.

- The proposed algorithms will be validated using both publicly available datasets and self-collected datasets, followed by further refinement to improve system accuracy.

4. SUBJECTS AND SCOPE

The study focuses on monitoring the activity states and indoor positions of firefighters. The proposed algorithms aim to operate within a self-contained system that does not rely on pre-existing infrastructure. Information regarding the estimated location and current status of firefighters is transmitted to the command center based on the building's pre-designed map data.

5. CONTRIBUTIONS

- To propose an algorithm that integrates an adaptive sliding window with an activity transition detection mechanism, aiming to optimize signal segmentation and enhance the quality of input features. The algorithm is applied to develop a real-time human activity recognition system, optimized for the specific operational environment of firefighters, with the objective of improving system accuracy and response latency in automated monitoring.
- To propose a low-complexity step detection algorithm based on the fusion of inertial sensor data and UWB measurements. The proposed algorithm aims to develop an efficient pedestrian motion estimation model, thereby improving the accuracy and stability of indoor positioning and tracking systems.

6. DISSERTATION STRUCTURE

▷ **INTRODUCTION**

▷ **CHAPTER 1: OVERVIEW**

▷ **CHAPTER 2: REAL-TIME HUMAN ACTIVITY RECOGNITION SYSTEM FOR FIREFIGHTERS**

▷ **CHAPTER 3: WEARABLE-BASED INDOOR POSITIONING METHOD**

▷ **CONCLUSIONS**

CHAPTER 1. OVERVIEW

1.1. Overview of human activity recognition

Human activity recognition (HAR) is currently approached mainly through two directions: computer vision [20] and wearable sensor-based signals [21]. Among these, wearable sensors have attracted significant attention due to their advantages,

such as compact size, low cost, and the ability to rapidly detect changes in motion dynamics [23–26].

Activity recognition based on accelerometer data is commonly applied in healthcare monitoring [17, 21, 24, 31–34]; however, it also shows strong potential for real-time activity tracking and hazard detection for firefighters [29]. Recent studies have applied deep learning (DL) techniques [30, 33, 40, 41] to improve recognition accuracy. Nevertheless, deep learning methods typically require substantial computational resources and high energy consumption, which reduces operational time and necessitates larger batteries, thereby limiting the design of compact wearable devices for firefighters. The adoption of lightweight machine learning (ML) models that can be deployed directly on microcontrollers presents a feasible solution, enabling resource optimization while maintaining a balance between accuracy, processing speed, and energy efficiency.

1.1.1. Challenges

Research on HAR has numerous applications, such as monitoring physical activity in sports [42] and supervising and preventing sedentary behavior in humans [43]. The placement of sensors (Fig. 1.1) significantly affects the performance of the recognition model, as well as the accuracy and the types of activities being monitored.

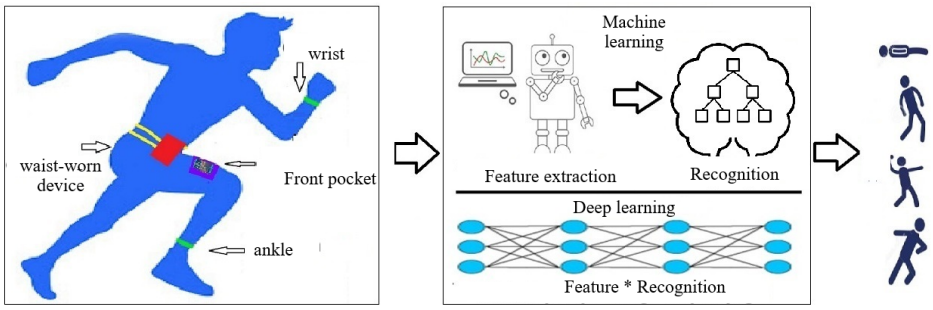

Figure 1.1: General framework of HAR and machine learning methods using wearable sensor data.

1.1.2. Wearable-based human activity recognition

To meet the requirements of continuous and non-intrusive monitoring, wearable devices are widely used due to their ability to collect motion data at low cost while preserving user privacy [21, 24]. The selection of machine learning algorithms that can be embedded on microcontrollers remains a significant challenge. For real-time HAR applications, particularly on low-power microcontroller platforms, traditional machine learning methods are a suitable choice due to their simplicity, low computa-

tional cost, and efficient deployability.

1.1.3. Human activity recognition for firefighters

Human activity recognition for firefighters (FAR) focuses on handling high-intensity activities with continuous and dynamic changes [29, 30]. Most studies employing wearable sensors for firefighters primarily focus on fall detection. Several studies on FAR [9, 11, 38, 52] aim to enhance safety and operational efficiency by monitoring physical activities, detecting anomalies, and providing real-time information in emergency situations.

1.1.4. Edge-based machine learning

1.1.4.1. Lightweight Machine Learning Methods

The deployment of HAR algorithms on wearable devices for firefighters poses significant challenges in terms of hardware resource constraints [45, 55–58]. Microcontroller units (MCUs), which are commonly used in sensor-integrated wearable devices, typically have limited memory capacity and low processing speed to ensure energy efficiency. Therefore, this dissertation prioritizes the use of lightweight machine learning methods, such as Decision Trees (DT) and Random Forests (RF), for classification and regression tasks.

1.1.4.2. Lightweight machine learning deployment techniques on microcontrollers

To embed trained models into embedded systems, intermediate tools are required to compile models into standard C/C++ code. However, the resource overhead becomes significant, as microcontrollers must allocate a considerable portion of memory to store interpreter libraries and operators. This leads to inefficiency for recognition tasks, making such approaches less suitable for real-time data processing on resource-constrained hardware, while also increasing energy consumption.

Traditional machine learning algorithms such as DT, RF, and XGBoost are adopted to better accommodate real-time data processing on microcontrollers. Libraries that transform the prediction function $f(x)$ from Python environments (Scikit-learn) into a pure C/C++ function $f_C(x)$, such as m2cgen, emlearn, and micromlgen, are commonly used.

1.2. Overview of indoor positioning systems in emergency situations

Accurate determination of the location and status of each firefighter during operations plays a crucial role for commanders in making timely and effective tactical

decisions [59]. For indoor positioning systems (IPS), there are three main approaches: (i) methods based on pre-established infrastructure for tracking and positioning; (ii) methods that do not require prior infrastructure; and (iii) hybrid approaches that integrate data from wearable sensors with communication technologies.

1.2.1. IPS based on pre-deployed infrastructure

Infrastructure-based indoor positioning methods [64–67] require the prior deployment of anchor nodes or fixed positioning devices within the operating environment.

1.2.2. Infrastructure-independent IPS

Research on infrastructure-free indoor positioning focuses on leveraging user-carried or body-mounted devices to collect sensor data and transmit positioning information to external processing systems [59, 71].

1.2.3. Hybrid indoor positioning systems

To overcome the limitations of both positioning approaches, hybrid indoor positioning systems (Hybrid Indoor Positioning System – HIPS) have emerged as a balanced and promising solution [80, 81].

1.3. Evaluation methods

The study employs evaluation methods such as accuracy (acc) and sensitivity (sen) [24], specificity (spe) [24], precision (pre) [82], F1-micro score ($F1_{mi}$) [24], K-fold cross-validation, mean absolute error (Mean Absolute Error - MAE), mean squared error (Mean Squared Error - MSE) [78], and standard deviation of error (SD) [83].

1.4. Research gap

- Existing HAR studies have not sufficiently addressed activity recognition at transition points between actions.
- Limitations of step estimation models in handling nonlinear and random walking patterns.
- Current research trends in applying UWB technology are primarily focused on non-contact approaches, with limited optimization for wearable device usage.

1.5. Research directions

The dissertation proposes algorithms to enhance the accuracy of step counting, high-precision step length estimation, detection of turning events and directions, as

well as vertical position estimation. The integration of these components forms the foundation of a high-accuracy inertial positioning system, contributing to improved monitoring efficiency and enhanced safety for rescue personnel in emergency environments.

CHAPTER 2. REAL-TIME HUMAN ACTIVITY RECOGNITION SYSTEM FOR FIREFIGHTERS

2.1. Real-time firefighter activity recognition system model

2.1.1. System architecture

The proposed system is capable of recognizing multiple types of activities, including walking (WA), standing (ST), lying (LY), sitting (SI), jogging (JO), stoop walking (WS), crawling (CR), slithering (SL), falling (FA), downstairs (DS), and upstairs (US). Fig. 2.1 illustrates the overall architecture and operational workflow of the proposed real-time firefighter activity recognition system (Real-time firefighter activity recognition system - RFAR).

2.1.2. Dataset

2.1.2.1. Private dataset

The data were sampled at a frequency of 100 Hz, with a total duration of 371 minutes and a size of over 100 MB. The dataset is publicly available on the IEEE DataPort repository[1].

2.1.2.2. Public datasets

This study utilizes three publicly available datasets — [1]MobiFall [84], [2]SFDLA [85], and [3]UniMiB-SHAR [86], which contain similar activities, for comparison in order to ensure diversity and objectivity in the study.

2.1.3. Data processing

2.1.3.1. Sensor noise and processing

The raw signal $\overline{s}_{raw}[n] = (\overline{a}_x[n], \overline{a}_y[n], \overline{a}_z[n])$ at sample time n is described as follows:

$$s_{raw}[n] = \mathbf{M}^* \cdot s_{true}[n] + \mathbf{a}_{offset} + \eta[n], \tag{2.1}$$

[1]https://dx.doi.org/10.21227/9w3p-v638

[1]https://www.kaggle.com/datasets/kmknation/mobifall-dataset-v20

[2]https://kilyos.ee.bilkent.edu.tr/~billur/

[3]https://www.kaggle.com/datasets/alirezacman/unimib-shar

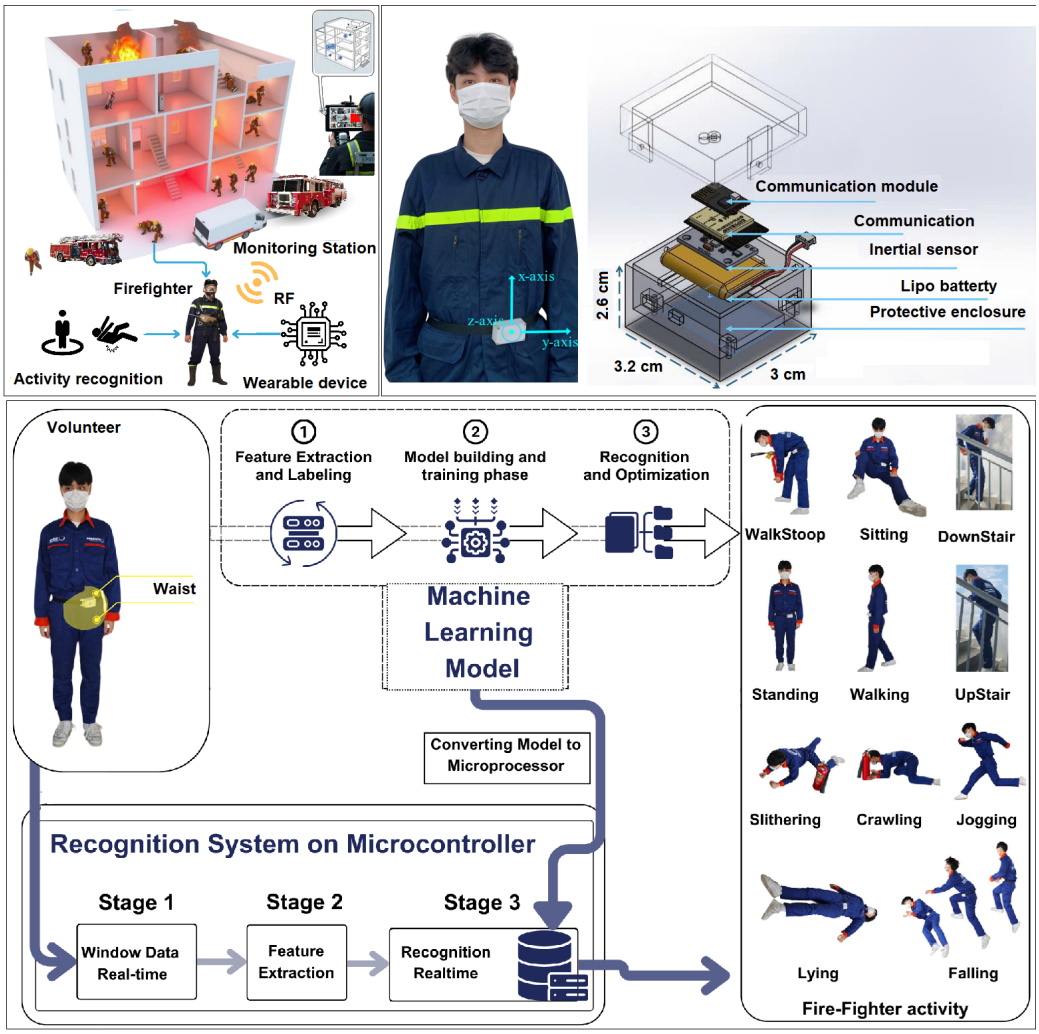

Figure 2.1: The RFAR system is designed to enable commanders to monitor the activities of firefighters in real time.

where $s_{true}[n]$ is the actual kinematic acceleration vector generated by the user; M* is the scale factor matrix; a_{offset} is the bias vector; and η represents random noise arising from mechanical vibrations and electronic noise in the integrated circuitry.

Each activity is assumed to be represented by the signal $\overline{s}[n] = (\overline{a}_x[n], \overline{a}_y[n], \overline{a}_z[n])$, where $\overline{N}$ denotes the length of the signal sequence $\overline{s}[n]$ and $n \in [1, \overline{N}]$. This signal sequence $\overline{s}[n]$ is recorded by an accelerometer mounted at the waist of the subject with a sampling frequency of 100 Hz during activity execution. After downsampling to 50 Hz, the resulting signal sequence $s[n]$ is segmented into time windows of 3 seconds, corresponding to 450 samples across the three acceleration axes x, y, and z. All segmented activity data are then divided into training and testing sets with a ratio of 60/40.

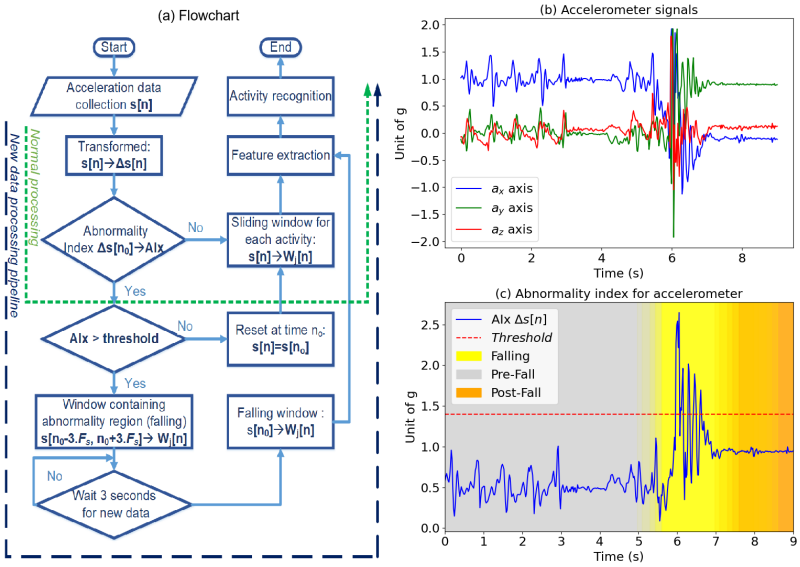

Figure 2.2: Flowchart of the anomaly detection and data processing pipeline (a), three-axis acceleration data describing a subject moving and then falling (b), and anomaly detection-based data processing (c).

2.1.3.2. Proposed data processing pipeline for fall detection

The falling activity is segmented separately based on the detection of abnormal values when the signal exceeds a predefined threshold, thereby generating a signal region containing this activity. Fig. 2.2 illustrates the proposed new pipeline for anomaly detection in data processing. The modified processing flow is highlighted in green, representing the corresponding new processing procedure.

2.1.4. Feature extraction

To optimize computational efficiency, this study considers only 12 types of time-domain features with low complexity for acceleration data. The feature set includes: energy (ε), mean absolute deviation (mad), signal magnitude area (sma), mean value (μ), standard deviation (σ), simple square integral (ssi), average amplitude change (aac), maximum value (max), interquartile range (iqr), median absolute deviation ($mead$), range (γ), and complexity (hc). Ineffective features are removed, and 12 suitable feature types are retained. The results of this optimization process identify the 30 most significant feature components across each signal axis.

2.1.5. Labelling

The data labeling process involves segmenting accelerometer data into windows and assigning labels corresponding to specific firefighter activities. The dataset consists of pairs of windows and their corresponding labels, denoted as $(W_1[n], L_1), (W_2[n], L_2), \ldots, ($

2.1.6. Random forest model deployment on microcontrollers

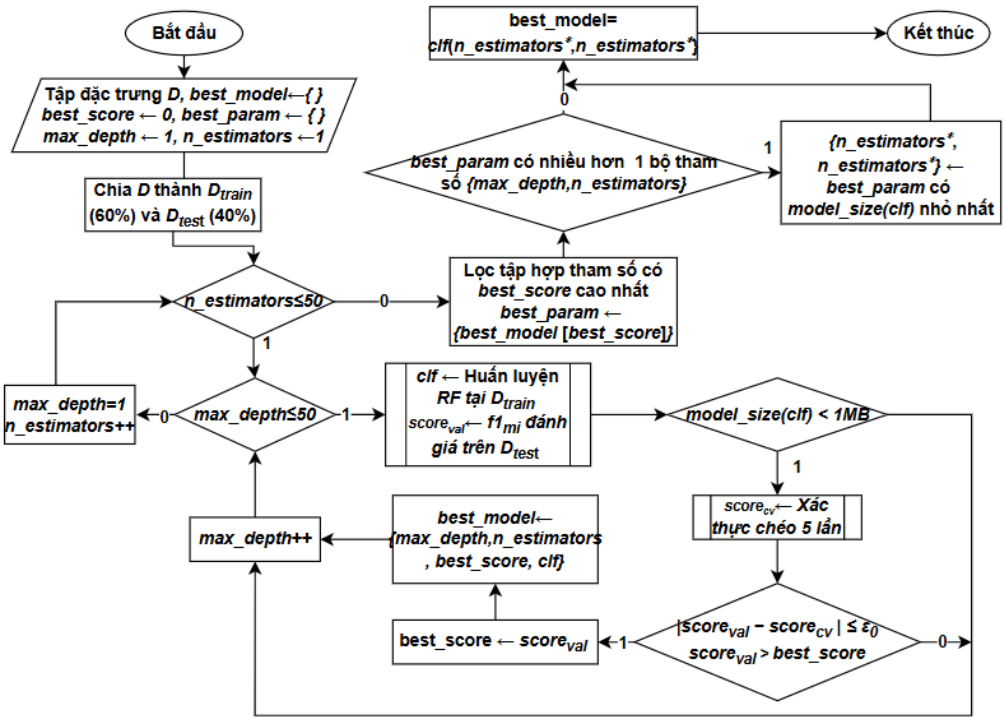

Figure 2.3: Hyperparameter optimization for machine learning models

Algorithm 2.3 describes the hyperparameter tuning[1] process for the RF model, based on the F1-score micro ($F1_{mi}$) metric and constrained such that the model size does not exceed 1 MB.

2.2. Experiments and results

2.2.1. Experimental results on private dataset

Under memory constraints, the RF model demonstrates a good balance between high accuracy, model size, and processing speed. With n_estimators $= 17$ and max_depth $= 14$, the model occupies only 0.921709 MB while achieving $acc = 99.3\%$ and above 96% across all other metrics. In addition, RF exhibits fast processing speed, with a training time of 0.093589 s on 4171 data windows (training) and an inference time of 0.010951 s on 2774 windows (testing), corresponding to approximately 3.95 μs per prediction. RF maintains more balanced performance compared to other models, while XGB and Light gradient boosting machine (LBM) also achieve stable and comparable results

[1] `https://github.com/HieuSSALAB/FireFighter` (Accessed: 12:00 10/10/2025)

Table 2.1: Hyperparameter configurations of recognition models on private dataset

Model	Hyperparameters	acc	sen	spe	pre	$F1_{mi}$	$F1_{ma}$	Size	Traning	Testing
RF	n_estimators $= 17$, max_depth $= 14$	99.3	97.0	99.6	97.1	96.1	97.0	0.921709	0.093589	0.010951
XGB	n_estimators $= 37$, max_depth $= 14$	99.2	96.4	99.5	96.6	95.5	96.5	0.938040	0.083459	0.047790
DT	max_depth $= 16$	98.5	93.3	99.1	93.6	91.7	93.5	0.056939	0.026039	0.009996
GBDT	n_estimators $= 35$, max_depth $= 4$	98.8	94.5	99.3	94.9	93.4	94.7	0.967428	21.003737	0.014934
SVM	$C\_ = 29$	98.9	94.8	99.4	94.6	93.7	94.7	0.247342	21.748170	0.019596
LBM	n_estimators $= 46$, max_depth $= 22$	99.1	96.1	99.5	96.4	95.2	96.2	0.966480	0.226004	0.007965

[1] $C\_$ is the regularization parameter for SVM.
[2] Units: % for acc, sen, spe, pre, $F1_{mi}$, $F1_{ma}$; training/testing time: seconds; model size:

2.2.2. Recognition results on public datasets

2.2.2.1. Evaluation on the SFDLA dataset

For the SFDLA (Simulated Falls and Daily Living Activities) dataset, data collected from female participants were excluded. In addition, sensor data from locations such as the head, right wrist, right thigh, and chest were also removed, as these positions are not aligned with the design objective of a waist-mounted wearable recognition model. After preprocessing, the signals from the waist-mounted sensor were segmented into 3-second windows. The overall performance of the proposed model on the SFDLA dataset is quite satisfactory, with $acc = 99.4\%$, $sen = 95.7\%$, $spe = 99.6\%$, $pre = 95.9\%$, and $F1_{mi} = 96.6\%$.

2.2.2.2. Evaluation on MobiFall

Most activities are correctly recognized, and misclassification cases are limited. However, distinguishing stair-related activities remains challenging. The RF model achieves $acc = 99.2\%$, $sen = 92.2\%$, $spe = 99.5\%$, $pre = 94.2\%$, and $F1_{mi} = 96.6\%$.

2.2.2.3. Evaluation on UniMiB-SHAR

A major challenge of the UniMiB-SHAR (University of Milano Bicocca Smartphone-based Human Activity Recognition) dataset is that the smartphone is placed in the front pocket, and its orientation varies significantly across data collection sessions. Therefore, the activities are preprocessed and mapped to a common orientation when similar actions begin. With sen and pre only reaching 71 – 72% and $F1_{mi}$ of 78.2%,

the model faces difficulties in recognizing different types of falls. However, the proposed model is still capable of accurately distinguishing between ADLs and fall-related activities, without misclassifying fall actions as ADLs.

2.3. Device deployment and experimental evaluation

2.3.1. Sensor calibration

To ensure the accuracy of the inertial sensor-based system, the calibration process is conducted in two stages: sensor error (noise) calibration and initial orientation calibration of the wearable device.

2.3.2. Experiments with volunteers

During rescue training exercises, the volunteers performed 11 activities under various conditions, repeating each activity multiple times. A demonstration video[1] was created to present the real-time recognition capability of the proposed model after deployment on the microcontroller. Some activities may exhibit slight display delays due to signal transmission and reception errors.

The proposed model achieves $acc = 99.3\%$ and $F1_{mi} = 96.1\%$ on the private dataset, while real-time experiments yield $acc = 99.4\%$ and $F1_{mi} = 96.2\%$. Minor differences are also observed in sen (95.9% and 96.4%) and pre (96.4% and 96.6%). The negligible differences between the experiments indicate that the recognition performance of the proposed model remains stable when deployed in real-world conditions.

2.3.3. Inference time, latency and computational complexity

During operation, the accelerometer samples tri-axial data at a frequency of 50 Hz, and the system collects 450 samples (150 samples per axis) within each 3 second window. After data acquisition is completed, the feature extraction process is performed with approximately 84507 lines of instructions, taking about 8858 μs. The proposed model, with 140838 executed instructions, completes the inference process within approximately 733 – 738 μs. This model achieves a recognition latency of about 9.561–9.596 ms on the ESP32 microcontroller, outperforming related studies [24, 56, 57, 59, 95] on TinyML applications for microcontrollers.

2.3.4. Evaluation of critical misclassification cases

Misclassification between normal activities (e.g., WA and JO) has a negligible impact on the estimation of activity intensity. Experimental results show that FA samples

[1] `https://youtu.be/S4hAzgEo-NA` (Accessed: 14:00 10/10/2025)

are correctly recognized. Meanwhile, some samples of the SL activity are misclassified as LY. These minor misclassifications involving SL are considered acceptable within the safety limits of the real-time monitoring system.

2.3.5. Performance of the proposed model

The proposed model demonstrates good and stable recognition performance on the private dataset, with both acc and $F1_{mi}$ exceeding 96%. For the public datasets, MobiFall and SFDLA achieve comparable results ($F1_{mi} = 96.6\%$). However, UniMiB yields significantly lower performance, with $F1_{mi} = 78.2\%$. The primary reason is that the UniMiB dataset is collected with varying smartphone orientations, increasing the complexity of the recognition task.

Table 2.2: Comparison with existing fall detection solutions for firefighters.

Study	sen **(%)**	spe **(%)**	acc **(%)**	GM **(%)**
Proposed model	100	100	100	100
Dinh et al. [22]	100	98.88	99.49	99.40
Chai et al. [29]	92.25	94.59	94.10	93.40
Thanh et al. [9]	95.89	100	97.96	95.24

If daily living activities are considered as non-fall cases, the proposed method achieves 100% recognition of the FA activity across the metrics sen, spe, acc, and GM ($GM = \sqrt{spe \times sen}$).

2.3.6. Sliding window and sampling frequency

Although a 6 second window achieves an $F1_{mi}$ score above 99%, it makes it difficult to accurately recognize firefighter activities. In addition, a sampling frequency of $F_s = 50\ Hz$ is considered appropriate, as it reduces energy consumption compared to 100 Hz while still maintaining reliable recognition accuracy.

2.3.7. Feature optimization

Feature selection based on their importance has helped reduce the number of features [89,90], computational complexity [21], and model size, while the proposed model still maintains high performance [24]. In the context of prolonged rescue operations, increasing the number of features leads to longer computation time, negatively affecting recognition performance in real-time data processing, while the improvement in accuracy is negligible and may even decrease.

2.4. Conclusion of chapter 2

The contents of Chapter 2 have been published in three works [CT1], [CT2], and [CT3]. In particular, [CT1] proposes a simple HAR model deployment approach on low-power microcontrollers for basic activities. Subsequently, [CT2] investigates and develops an optimized signal processing pipeline for microcontrollers with moderate performance. By integrating the proposed algorithms, [CT3] extends and completes the recognition system by constructing a dedicated dataset consisting of 11 firefighter-specific activities.

CHAPTER 3. WEARABLE-BASED INDOOR POSITIONING METHOD

3.1. Overview of the implementation stages of the proposed IPS system

3.1.1. System architecture

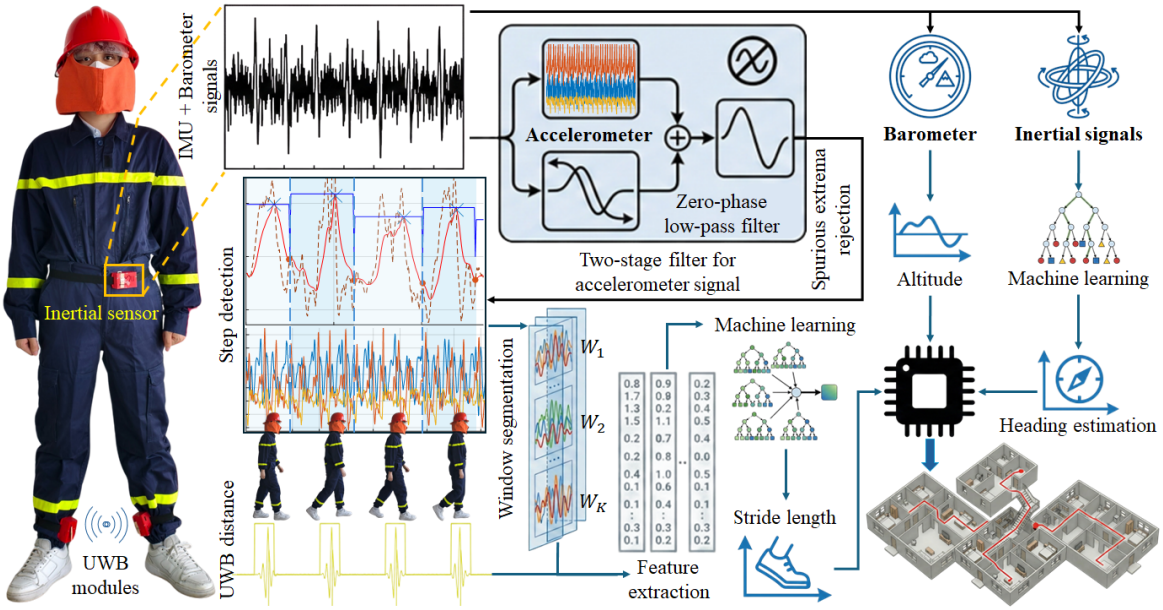

Figure 3.1: Overview of the implementation stages of the proposed IPS system

This study develops a position tracking system based on wearable devices and an unmanned aerial vehicle (Unmanned Aerial Vehicle - UAV). The system includes wearable devices integrated with inertial measurement units comprising 9 axes (accelerometer ADXL345, gyroscope ITG3200, magnetometer HMC5883L), a barometric sensor MS-5837, and a DW3000 module utilizing UWB technology. The wearable devices are compactly designed with dimensions of $4.8 \times 4.5 \times 4$ cm^3, and are placed at the waist and ankle to collect body motion data under various activity states, as well as information on altitude, traveled distance, and movement direction. The system is

powered by a 3.7 V - 2000 mAh lithium battery.

3.1.2. Processing and implementation steps

Fig. 3.1 illustrates the system deployment process. This study aims to develop a real-time indoor positioning system based on pedestrian position estimation techniques. The system focuses on three main tasks: (i) accurate step length estimation, (ii) movement direction classification, and (iii) floor-level position estimation based on building design data.

3.1.3. UWB calibration

3.1.3.1. Multipath processing

- To determine this parameter, two devices (Anchor and Tag) are placed at a standard reference distance of $1\ m$.
- The system applies a moving average filter with a window size of 10 samples to smooth the data and eliminate outliers caused by instantaneous multipath effects.

3.1.3.2. Time synchronization and clock drift evaluation

Time misalignment between devices affects the data fusion process. Therefore, the system employs a handshake protocol via BLE during initialization. The central device transmits a command packet containing a timestamp (t_0). Upon receiving the command, the peripheral devices synchronize their internal timers. Continuous operation tests over 30 minutes show that the time drift between modules remains within an acceptable range ($< 1\ ms$), which is smaller than the sampling period ($20\ ms$ corresponding to a frequency of 50 Hz), thereby ensuring data integrity during fusion.

3.1.4. Method for estimating traveled distance

3.1.4.1. Dataset

A stride length dataset[1] was collected from volunteers at Phenikaa University, including 10 male students with physical characteristics similar to firefighters. Each participant wore three identical wearable devices communicating via Bluetooth Low Energy.

[1] `https://doi.org/10.21227/emvz-3t63` (Accessed: 12:00 12/12/2025).

3.1.4.2. Movement distance computation

The signal preprocessing procedure is performed to transform the raw acceleration data in order to reduce noise (Fig. 3.2) and improve the accuracy of motion phase detection. The study employs a two-stage filtering process: i/ the first stage includes raw signal transformation and noise reduction using a zero-phase low-pass filter (Zero phase low pass filter - ZLF); ii/ the second stage focuses on removing false peaks and detecting signal segments corresponding to each stride.

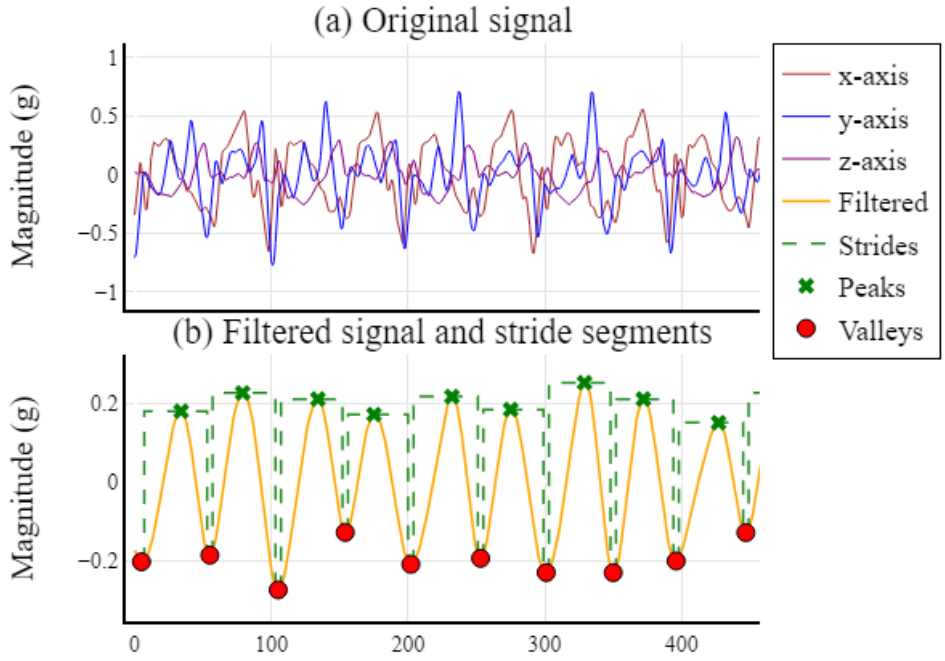

Figure 3.2: Signal segmentation process during stride execution: (a) raw three-axis acceleration data; (b) filtered signal and extracted stride segments.

Stage 1 - zero-phase low-pass filter

The tri-axial acceleration signals $a_x[i], a_y[i], a_z[i]$ with length N are transformed into a magnitude sequence ($a[i]$) by computing the signal magnitude according to equation 3.1, expressed in units of g ($g = 9.8\ m/s^2$).

$$a[i] = \sqrt{a_x[i]^2 + a_y[i]^2 + a_z[i]^2} \tag{3.1}$$

The study applies a ZLF consisting of two stages, where $k \in 1, 2$ denotes the two filtering stages, and $a[n]$ represents the filtered output corresponding to the input signal $a[i]$. The general transfer function is formulated as in equation 3.2:

$$H_k(z) = \frac{b_{0k} + b_{1k}z^{-1} + b_{2k}z^{-2}}{1 + a_{1k}z^{-1} + a_{2k}z^{-2}}, \tag{3.2}$$

At stage 1, each forward-filtered value $y_{fwd}[i]$ is computed from the input value $a[i]$

through two steps:

- Step 1: $out_1[i] = b_{01}a[i] + b_{11}a[i-1] + b_{21}a[i-2] - a_{11}out_1[i-1] - a_{21}out_1[i-2]$.
- Step 2: $y_{fwd}[i] = b_{02}out_1[i] + b_{12}out_1[i-1] + b_{22}out_1[i-2] - a_{12}y_{fwd}[i-1] - a_{22}y_{fwd}[i-2]$.

Stage 2 eliminates phase delay. The signal sequence $y_{fwd}[i]$ is reversed to obtain $y_{rev}[i]$, and then filtered once again:

- First reversal: $y_{rev}[i] = \tilde{y}_{fwd}[i] = y_{fwd}[N-1-i]$.
- Backward filtering: The signal $y_{rev}[i]$ is further processed through two steps similar to the forward stage to produce the sequence $y_{back}[i]$.
- Second reversal: $a[n] = \tilde{y}_{back}[i] = y_{back}[N-1-i]$.

Stage 2 - removal of false extrema and identification of stride signal regions

False peaks can significantly affect the accuracy of step detection results in this study. In this approach, true peaks are identified using a pair of sliding windows with equal size on both sides of each data point. As the walking pattern becomes slower and less stable, an adaptive sliding window mechanism is employed instead of a fixed one. The window size is automatically adjusted based on the step frequency measured in previous cycles.

In each step cycle, when the two feet are close to each other (pre-step or double-support phase), the minimum distance is d_{min}; when the step is completed and the stance foot and swing foot are maximally separated, the maximum distance is d_{max}. These two states form the geometry of a triangle, and by applying the Pythagorean theorem, we obtain:

$$d = \sqrt{d_{\text{max}}^2 - d_{\text{min}}^2} \tag{3.3}$$

3.1.5. Movement direction estimation algorithm

By exploiting the relationship between the sign and amplitude variation of the gyroscope signal and the rotation direction, the DT model can rapidly determine the instantaneous movement direction without performing angular integration. The trained model with the highest accuracy is described and exported as a header file, and is represented as shown in Fig. 3.3.

The confusion matrix of movement angles shows that the model achieves high and stable accuracy across most orientation classes. In particular, the straight movement angle (0°), which has the largest number of samples (348 samples), is predicted correctly without confusion with other angles. Major movement angles such as 45°,

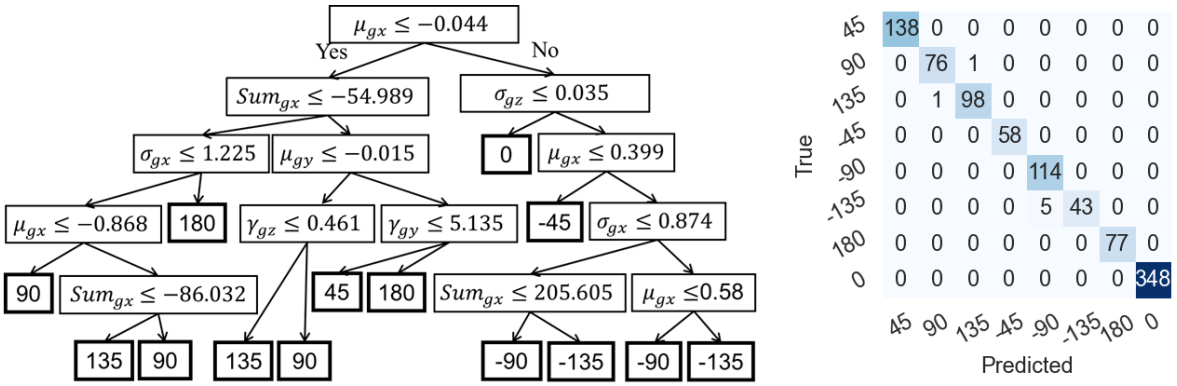

Figure 3.3: Decision tree flowchart with five levels capable of recognizing rotation angles with high accuracy.

$-45°$, $-90°$, and $180°$ also demonstrate strong recognition performance with low misclassification rates.

Minor confusion occurs only in a few cases between angles with similar inertial signal patterns, such as between $90°$ and $135°$ (1 misclassified sample), or between $-90°$ and $-135°$ (5 misclassified samples). Overall, these results demonstrate that using machine learning models (such as the DT decision tree described) to classify instantaneous movement direction, instead of traditional angular integration, is a highly effective solution. It eliminates cumulative drift and ensures high reliability for firefighter positioning systems.

3.1.6. Altitude estimation and floor determination algorithm

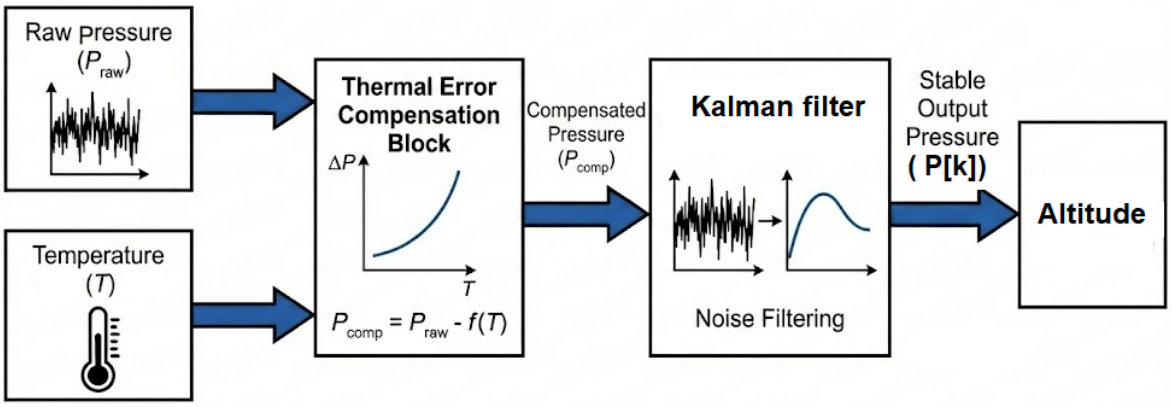

Figure 3.4: Implementation diagram of temperature drift compensation algorithm for pressure

At each sampling instant t_k, the algorithm is executed following the procedure illustrated in Fig. 3.4. The raw pressure data from the sensor ($P_{raw}[k]$) and the temperature around the pressure sensor ($T_{meas}[k]$) are collected. To reduce errors, the temperature is normalized as the deviation relative to the reference temperature T_{ref}

averaged at the measurement time:

$$\Delta T[k] = T_{meas}[k] - T_{ref} \tag{3.4}$$

Next, the study applies Horner's scheme to derive the formulation for computing the error compensation term $\delta P[k]$:

$$\delta P[k] = \Delta T[k] \times (k_1 + k_2 \times \Delta T[k]) \tag{3.5}$$

The compensated pressure $P[k]$ is continuously updated using a Kalman filter:

$$P[k] = P_{raw}[k] - \delta P[k] \tag{3.6}$$

The relative altitude (H_k) of the firefighter is computed based on atmospheric pressure[1] using equation 3.7 [9] after temperature-based pressure compensation.

$$H_k = 44330 \times \left(1 - \left(\frac{P[k]}{P_0}\right)^{\frac{1}{5.255}}\right) \tag{3.7}$$

The computed altitude (H_k^*) is used after calibration by subtracting the bias parameter ($Bias_k$):

$$H_k^* = H_k - Bias_k \tag{3.8}$$

Information for floor-level determination includes:

- Height map (h_{F_k}).
- Processed altitude (H_k^*).
- Predicted activity (S_{RFAR}).

If $S_{RFAR} \in \{\text{US}, \text{DS}\}$, the system activates the transition mechanism (ΔF_k) and determines the number of steps C_{seq}:

$$F_k \leftarrow F_k + \Delta F_k \quad \text{if} S_{RFAR} \in \{\text{US}, \text{DS}\}$$

$$C_{seq} = \begin{cases} C_{seq} + 1 & \text{if} S_{RFAR}^{(k)} = S_{RFAR}^{(k-1)} \in \{\text{US}, \text{DS}\} \\ 0 \end{cases} \tag{3.9}$$

[1] https://www.sciencedirect.com/topics/engineering/pressure-altitude accessed 15/1/2026.

$$\Delta F_k = \begin{cases} +1 & \text{if}(S_{RFAR} = \text{US}) \wedge (C_{seq} \geq \overline{\gamma}_{min}) \wedge (H_k^* - h_{F_{k-1}} \geq \delta_{th}) \\ -1 & \text{if}(S_{RFAR} = \text{DS}) \wedge (C_{seq} \geq \overline{\gamma}_{min}) \wedge (H_k^* - h_{F_{k-1}} \leq -\delta_{th}) \\ 0 & \text{khác.} \end{cases} \tag{3.10}$$

In the case where $S_{RFAR} =$ others, the system activates the floor-locking mechanism ($F_k = F_{k-1}$) and compensates for pressure drift errors.

3.1.7. Quaternion-based indoor positioning system modeling

At time step k, the system uses the input data $\mathbf{u}_k = [d_k, \alpha_k, H_k^*]^T$ to compute the new state $\mathbf{x}_k^* = [\mathbf{p}_k^T, \mathbf{q}_k^T]^T = [x_k, y_k, z_k, q_{0,k}, q_{1,k}, q_{2,k}, q_{3,k}]^T$. The movement direction angle α_k is determined using the DT model (f_{DT}), with the output being one of eight discrete angles: $\alpha_k = f_{DT}(feat_k) \in \{0, \pm 45^\circ, \pm 90^\circ, \pm 135^\circ, 180^\circ\}$. The discrete angle α_k is represented as a quaternion $\mathbf{q}_k$ to describe the movement orientation in the OXY plane. The system state is updated as follows:

$$\begin{cases} x_k = x_{k-1} + d_k \cos(\alpha_k) \\ y_k = y_{k-1} + d_k \sin(\alpha_k) \\ z_k = H_k^* \\ q_{0,k} = \cos\left(\dfrac{\alpha_k}{2}\right) \\ q_{1,k} = 0 \\ q_{2,k} = 0 \\ q_{3,k} = \sin\left(\dfrac{\alpha_k}{2}\right) \end{cases} \tag{3.11}$$

At each time step k, the floor number F_k is determined by comparing H_k^* with the height intervals $\Psi(H_k^*)$ according to the building floor database:

$$F_k = \Psi(H_k^*) = i, \quad H_{floor}^{(i)} \leq H_k^* < H_{ceiling}^{(i)}$$

3.1.8. Digital map for building information visualization

To enable monitoring of firefighter positions inside buildings, a map is required to visualize the structure and layout of the building, allowing the command center to conveniently monitor and guide firefighters. All buildings are designed prior to construction and include detailed information such as total height, floor heights, locations of staircases and elevators, etc. These data provide valuable support for locating victims inside the building.

3.2. Results and evaluation

3.2.1. Experimental scenario

The experimental environment is set up in a typical corridor with limited dimensions (width of 2 m), surrounded by reinforced concrete walls commonly found in evacuation routes. To comprehensively evaluate the performance of the proposed algorithm in floor-level positioning under simulated fire conditions, experiments are conducted in the A4 building, which has 7 floors with a minimum height of 3 m per floor. The evaluation includes:

- Verification of the stability of the step distance estimation algorithm.
- Verification of the movement direction and trajectory estimation algorithm.
- 3D positioning experiments and altitude noise filtering.

3.2.2. Position tracking results and visualization on digital map

3.2.2.1. Evaluation of trajectory deviation

The digitized floor plan (Fig. 3.5) is used to define the experimental route with actual dimensions of 104 m in length and 22.52 m in width. The test volunteers perform straight-line motion, allowing the system to collect stride data from the stationary state ($d_{\min}$) to the maximum stride length ($d_{\max}$).

The experimental results are summarized in Table 3.1, where the performance of the proposed IPS is analyzed in detail across five trajectories with lengths varying from 104 m to 171 m. The experiments are conducted with specific turning maneuvers and translational movements to simulate search and rescue operations, including rotation angles of $0°$, $\pm 45°$, $\pm 90°$, $\pm 135°$, and reverse turns of $180°$. The actual distances of the trajectories range from 104 m to 171 m, while the corresponding estimated values exhibit only minor deviations, with differences ranging from 0.46 m to 1.22 m. The average absolute error is only 0.818 m, indicating that the distance estimation algorithm introduces negligible deviation. Notably, longer trajectories such as the blue (162 m) and red (171 m) paths show slightly higher absolute errors, which is consistent with the cumulative error characteristics of inertial positioning systems as the traveled distance increases.

3.2.2.2. Evaluation of floor determination algorithm in buildings

When the ambient temperature increases from $30°$C to $70°$C at the same floor (either floor 1 or floor 7), the pressure sensor records variations in atmospheric pressure

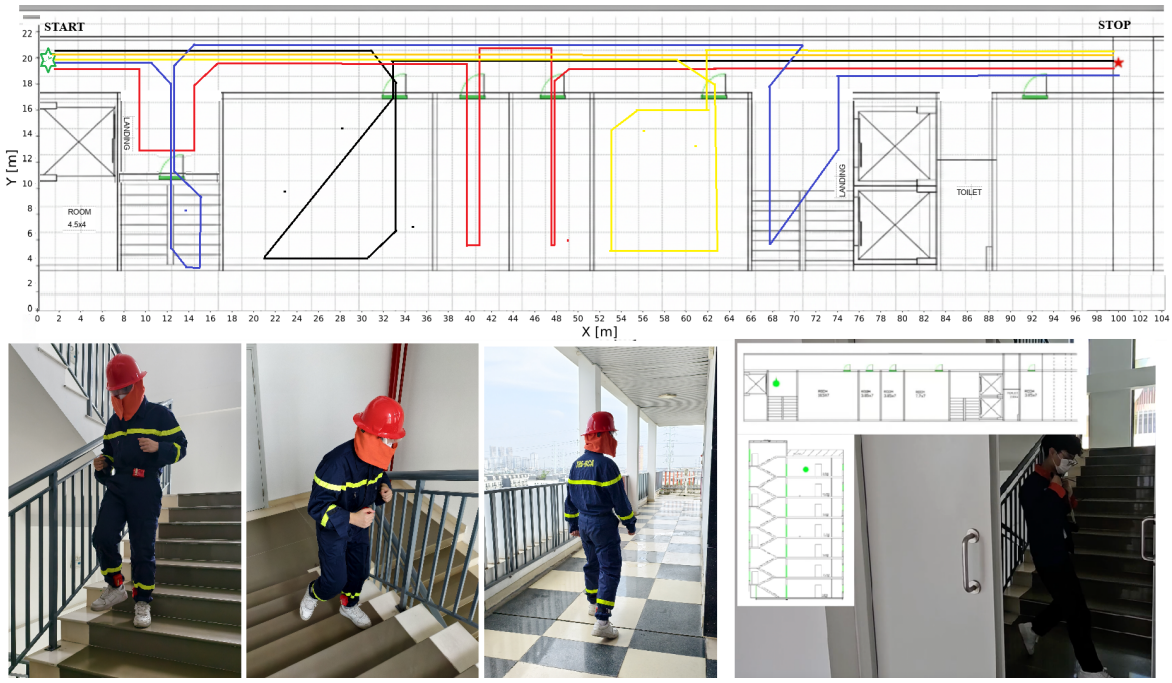

Figure 3.5: Comprehensive experiments on real-time position estimation and trajectory description when volunteers use wearable sensors in the A4 building at Phenikaa University.

Table 3.1: Distance estimation and cumulative position error analysis

Method/trajectory	**Orange**	**Black**	**Yellow**	**Blue**	**Red**	**Average**
Distance (m)	104	143	146	162	171	—
Estimation (m)	103.54	142.37	145.2	160.78	170.02	—
Deviation (m)	0.46	0.63	0.8	1.22	0.98	**0.818**
MAE (cm)	0.523	0.521	0.506	0.645	0.676	**0.574**
MSE (cm^2)	0.753	0.678	0.776	0.977	1.011	**0.839**
SD%	0.583	0.539	0.774	0.881	0.631	**0.682**
MAE%	0.442	0.441	0.548	0.761	0.573	**0.553**

due to local thermal fluctuations. However, after applying compensation and calibration algorithms, the estimated altitude (H_k^*) remains stable. At floor 1, the error of H_k^* fluctuates below 0.4 m. Meanwhile, at floor 7, H_k^* reaches 18.6 m with a deviation of 0.6 m compared to the actual value. When performing upward and downward movements under normal temperature conditions, the compensation factor $Bias_k$ remains small and stable (0.1–0.2 m), allowing the corrected altitude H_k^* to closely follow the actual floor height. The error of H_k^* consistently remains below 0.2 m, which is

significantly smaller than the minimum floor height of 3 m. In the combined movement scenario with thermal disturbances, sudden temperature increases simulating fire conditions at floor 2 (65°C) and floor 7 (70°C) cause the initial altitude H_{raw} to deviate significantly due to temperature effects on the sensor. However, the pressure compensation algorithm δP combined with the Kalman filter brings H_k back close to the true value, effectively eliminating thermal noise. When a measured altitude of 16 m is detected, deviating from the reference map value of 15 m, the system automatically adjusts $Bias_k$ from 0.4 to 0.7. This timely adjustment enables subsequent measurements at floor 7 and during descent to closely match the actual altitude, thereby ensuring high accuracy in floor positioning (F_k).

3.2.3. Evaluation of energy consumption and operating time

The current consumption measurement (I_{state}) is conducted using a dedicated current measurement device connected in series with the power supply of the wearable device. The system can operate continuously for more than 14 hours in search mode and approximately 9.5 hours under continuous high-intensity data transmission conditions. Compared to the typical duration of a firefighting shift or rescue mission (6–8 hours), the 2000 mAh battery capacity fully satisfies the system's energy requirements.

3.2.4. Evaluation of environmental condition impacts

The device is designed to be resistant to sweat and humidity; however, moisture inside protective clothing may still lead to condensation within the device. Therefore, the device is oriented toward an IP67-rated design for water and dust resistance, including sealing gaskets, minimizing external exposure points, and applying protective coatings on the circuit board.

3.2.5. Evaluation of latency and real-time responsiveness

The latency caused by computation and data transmission $T_{feat} + T_{infer} + T_{trans} \approx 11$ ms accounts for only about 2% to 3% of the total duration of a single step cycle. The system can complete processing and transmit results almost instantaneously upon the completion of each step.

3.2.6. Comparison with related studies

When power supply is lost or communication nodes are damaged by high temperatures, positioning systems [62, 67, 68] become completely inoperative. These models require substantial computational resources, leading to increased energy con-

sumption and latency when deployed on wearable embedded devices. The proposed system achieves continuous operation exceeding 9 hours with low processing latency (below 11 ms). This demonstrates a well-balanced trade-off between positioning accuracy, energy efficiency, and practical deployability.

3.3. Chapter conclusion

Chapter 3 presents a system supporting firefighters during operations. The system enables indoor positioning in emergency situations through a hybrid approach combining PDR and UWB. The contents of this chapter have been published in [CT4, CT5, CT6]. In particular, [CT4] presents a low-complexity step detection algorithm suitable for real-time processing. [CT5] describes a proposed solution for 2D positioning. [CT6] proposes a building-level position tracking model based on building maps and the integration of IMU, UWB, and barometric sensors.

CONCLUSIONS

General conclusion. The dissertation addresses the critical problem of safety monitoring for firefighters operating in harsh environments where satellite-based positioning systems are unavailable. This study makes two main contributions:

- It proposes an algorithm integrating an adaptive sliding window with an activity transition detection mechanism to optimize signal segmentation and enhance input feature quality. The algorithm is applied to develop a real-time human activity recognition system, optimized for the specific operational conditions of firefighters, with the goal of improving accuracy and response speed of the automated monitoring system.
- It proposes a low-complexity step detection algorithm based on the fusion of inertial sensor data and ultra-wideband measurements. This algorithm enables the development of an efficient pedestrian motion estimation model, improving the accuracy and stability of indoor positioning and tracking systems.

Limitations. The proposed system still has several limitations:

- The current system has a relatively large size due to manual assembly, which may cause inconvenience during field deployment.
- The dataset is collected from male volunteers with physical characteristics similar to professional firefighters.
- Pressure variations caused by extremely high temperatures (chimney effect) may introduce larger errors that are not fully modeled in the current system.

Recommendations and future work. To enhance applicability and further improve the system, the following research directions are proposed:

- Optimize hardware design, including integrated circuit optimization, to reduce device size and weight.
- Expand and diversify datasets across different user groups to improve model generalization and accuracy.
- Investigate the integration of mesh network communication protocols to support simultaneous tracking of multiple firefighters, enabling cooperative data transmission when signal degradation occurs.
- Develop a comprehensive rescue ecosystem through integrated software and hardware solutions, enabling seamless connectivity between wearable devices and the field command center.