HUMAN SENSING USING LOW RESOLUTION THERMOPILE SENSORS

Shengjun (Eric) Ma

January 2026

Department of Electrical and Electronic Engineering

School of Engineering, Computer, and Mathematical Science

A THESIS SUBMITTED TO AUCKLAND UNIVERSITY OF TECHNOLOGY

IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Abstract

Interest in human sensing has grown substantially in recent years. This growth is driven by applications across diverse domains, including healthcare, residential aged care, security and surveillance, entertainment, and intelligent living environments. A fundamental requirement for many human sensing solutions is the use of device free methods that safeguard user privacy. Although various sensing modalities have been explored, each has its own advantages and limitations. Thermal sensing, which operates by detecting and measuring Infrared (IR) radiation emitted by the human body, is emerging as a promising option for device free human sensing.

This thesis presents a device-free and privacy-preserving framework for human sensing applications using the IR signal captured by low resolution thermopile sensors. Machine learning approaches are applied to perform two distinct tasks: localization and fall detection. We develop a scalable, networked thermopile sensing platform to perform these tasks. To the best of our knowledge, no open-source or commercially available hardware platform has been specifically designed for human sensing using low-resolution thermopile sensors. We address this gap by releasing all PCB schematics, firmware, and supporting code. This enables replication and further extension by the wider research community.

A 2D CNN–LSTM regression model was developed for localization and trained using data collected from eight participants. The model achieved median localization errors between 0.2 m and 0.3 m. The evaluation was performed on participants whose

data were not used for training, ensuring robust and generalizable performance. We systematically examine multiple ceiling- and wall-mounted sensor configurations, addressing limitations of prior studies that considered only a small number of sensors in fixed layouts. The results show a clear dependence of localization accuracy on both sensor number and placement. Experimental benchmarking against previous approaches indicates that the proposed algorithm provides improved localization accuracy. To assess scalability, four ceiling mounted sensors were deployed to cover a substantially larger area than reported in existing studies. The system operated reliably in this setting, demonstrating the practical viability of thermopile based localization.

We developed a fall detection system that employs machine learning models to automatically identify falls from the IR data captured by the thermopile sensors. Three machine learning classifiers, namely SVM, MLP and RF, were trained on data collected from five participants enacting fall scenarios and performing other activities such as standing and sitting. The accuracy of the fall detection algorithms was evaluated using data from an additional five participants who were not included in the training set. The SVM classifier demonstrated the most consistent performance within the 16 different thermopile sensor configurations investigated. The accuracy varies between 92.5% and 99.2% and depends on the number of sensors and their respective locations. A combination of ceiling- and wall-mounted sensors with overlapping fields of view offered the best balance between coverage and sensor count. However, our experimental results also show that accurate fall detection is achievable with a single sensor, provided the fall occurs within the sensor's clear field of view.

Keywords: Device-Free Localization, Fall Detection, Human Activity Recognition, Human Sensing, Indoor Localization, IR Sensing, Passive Localization, Thermal Sensor, Thermopile.

Contents

List of Tables

List of Figures

Attestation of Authorship

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person (except where explicitly defined in the acknowledgements), nor used artificial intelligence tools or generative artificial intelligence tools (unless it is clearly stated, and referenced, along with the purpose of use), nor material which to a substantial extent has been submitted for the award of any other degree or diploma of a university or other institution of higher learning.

Signature of candidate

Acknowledgements

This thesis represents the culmination of a unique academic journey that has been enriched by the environments of two distinguished institutions. I am grateful for the rare privilege of being trained across different academic landscapes, an experience that has broadened my perspective and deepened my research.

My deepest gratitude goes to my primary supervisor, Prof. Fakhrul Alam. His mentorship has been the bridge between these two chapters of my life. I am particularly indebted to him for his pivotal leadership during my transition from Massey University to Auckland University of Technology (AUT). Prof. Alam's commitment ensured that this move was not a disruption, but an opportunity for growth. Beyond his scholarly expertise, I am thankful for his empathy and proactive care for my well-being. His support gave me the stability to thrive in a new environment and turn challenges into achievements.

I am equally grateful to my co-supervisor, Dr. Daniel Konings, who has been a constant source of guidance from the very beginning. His presence across both institutions provided a vital sense of continuity. Dr. Konings' technical insights were instrumental to my work, but it was his kindness and concern for my personal journey that I value most. I appreciate his unwavering dedication to supporting me.

I would also like to express my sincere appreciation to Prof. Edmund Lai. His steadfast encouragement and invaluable advice served as a stabilizing force. His deep understanding of the field and his approachable nature provided a reassuring backdrop to my studies, offering wisdom and perspective when it was most needed.

Finally, I wish to acknowledge the distinct contributions of both universities to my development. I appreciate the foundational training and research initiation I received at Massey University, which set the stage for this work. Equally, I am profoundly grateful to AUT for providing the resources, infrastructure, and supportive community where this research ultimately flourished and came to fruition.

Co-Authored Works

All co-authors in the following table have approved these chapters for inclusion in Shengjun (Eric) Ma's doctoral thesis.

STUDENT AND SUPERVISOR APPROVALS

By signing you are confirming that the co-author contributions stated in the table(s) below are accurate.

Student Name	Shengjun Ma	**Signature**	[redacted]	**Date**	03/12/2025
Supervisor Name	Fakhrul Alam	**Signature**		**Date**	03/12/2025
Supervisor Name	Daniel Konings	**Signature**		**Date**	03/12/2025
Supervisor Name	Edmund M.-K. Lai	**Signature**		**Date**	03/12/2025

Chapter Number:	3
Manuscript Title:	Device-Free Localization Using Low-Resolution Thermopiles
Publication Status:	Submitted for Publication
Reference if published:	

AUTHOR SURNAME: (order as per manuscript)	**CONTRIBUTION** (May copy from the guidelines above)
Ma	80% (Conception and design of the project and output Acquisition of research data Analysis and interpretation of research data Drafting significant parts of the research Revising output)
Konings	8% (Interpretation of research data Revising output)
Lai	4% (Revising output)
Alam	8% (Conception and design of the project and output Revising output)

Chapter Number:	4
Manuscript Title:	Fall Detection Using Multiple Low-Resolution Thermopile Sensors and Machine Learning
Publication Status:	Published
Reference if published:	S. Ma, D. Konings, E. M. . -K. Lai and F. Alam, "Fall Detection Using Multiple Low-Resolution Thermopile Sensors and Machine Learning," in *IEEE Sensors Journal*, doi: 10.1109/JSEN.2025.3629734.

AUTHOR SURNAME: (order as per manuscript)	**CONTRIBUTION** (May copy from the guidelines above)
Ma	80% (Conception and design of the project and output Acquisition of research data Analysis and interpretation of research data Drafting significant parts of the research Revising output)
Konings	8% (Interpretation of research data Revising output)
Lai	4% (Revising output)
Alam	8% (Conception and design of the project and output Revising output)

Ethics Approval

The study was conducted at Auckland University of Technology, Auckland, from November 2024 to November 2025. The study was approved through Ethics Application 24/285 Thermopile Sensor Based Human Sensing on 8 November 2024 by the Auckland University of Technology Ethics Committee (AUTEC).

List of Abbreviations

Abbreviation	Definition
1D CNN	One-Dimensional Convolutional Neural Network
2D CNN	Two-Dimensional Convolutional Neural Network
ADL	Activities of Daily Living
AI	Artificial Intelligence
AoA	Angle of Arrival
AUT	Auckland University of Technology
AUTEC	Auckland University of Technology Ethics Committee
Bi-LSTM	Bidirectional Long Short-Term Memory
BN	Batch Normalization
CDF	Cumulative Distribution Function
CGAN	Conditional Generative Adversarial Network
CMOS	Complementary Metal-Oxide-Semiconductor
CNN	Convolutional Neural Network
CSI	Channel State Information
DCT	Discrete Cosine Transform
DCNN	Deep Convolutional Neural Network
DFL	Device-Free Localization
DL	Deep Learning
DNN	Deep Neural Network
ELM	Extreme Learning Machine

Abbreviation	Definition
ESD	Electrostatic Discharge
EVK	Evaluation Kit
FC	Fully Connected
FLIR	Forward Looking Infrared
FMCW	Frequency-Modulated Continuous-Wave
FNR	False Negative Rate
FoV	Field of View
FPR	False Positive Rate
GPS	Global Positioning System
GRU	Gated Recurrent Unit
GUI	Graphical User Interface
HAR	Human Activity Recognition
HSE	High-Speed External
HVAC	Heating, Ventilation, and Air Conditioning
I^2C	Inter-Integrated Circuit
IEEE	Institute of Electrical and Electronics Engineers
IoT	Internet of Things
IR	Infrared
ISM	Industrial, Scientific, and Medical
JTAG	Joint Test Action Group
k-NN	k-Nearest Neighbors
LDO	Low-Dropout Regulator
LLM	Large Language Model
LOS	Line of Sight
LSE	Low-Speed External
LSTM	Long Short-Term Memory
MCU	Microcontroller Unit
ML	Machine Learning

Abbreviation	Definition
MLP	Multilayer Perceptron
mmWave	Millimeter Wave
PC	Personal Computer
PCB	Printed Circuit Board
PCLK	Peripheral Clock
PIR	Passive Infrared
ReLU	Rectified Linear Unit
RF	Radio Frequency / Random Forest
RGB	Red, Green, Blue
RGB-D	Red, Green, Blue Depth
RMSE	Root Mean Squared Error
ROI	Region of Interest
RSSI	Received Signal Strength Indication
RTC	Real-Time Clock
SGD	Stochastic Gradient Descent
SR	Super-Resolution
SRCNN	Super-Resolution Convolutional Neural Network
SVM	Support Vector Machine
SVR	Support Vector Regression
TinyML	Tiny Machine Learning
TPE	Tree-structured Parzen Estimator
UART	Universal Asynchronous Receiver-Transmitter
USART	Universal Synchronous/Asynchronous Receiver-Transmitter
USB	Universal Serial Bus
Wi-Fi	Wireless Fidelity
WKNN	Weighted k-Nearest Neighbors

Chapter 1

Introduction

1.1 Human Sensing

Human sensing technology enables the observation and analysis of human movements within areas of activity such as homes and workplaces [1]. It captures and interprets human presence, activities, and behavioral patterns in real time through signals acquired from sensors. Insights from human sensing can support the design of smart environments and has wide potential in fields such as health, aged care, and engineering. Interest in this area has grown rapidly, leading to significant advances in recent years.

The rapid advancement of the Internet of Things (IoT) [2], Artificial Intelligence (AI) [3], and sensor technologies has accelerated the development of smart home applications [4]. This progress is especially significant in the context of ageing global population [5], which has intensified research into aged-care technologies. Human sensing, which combines sensor data with AI, is therefore emerging as a critical component of smart home environments and is poised for widespread adoption.

1.2 Human Sensing Taxonomy

Human sensing in indoor environments has expanded rapidly in recent years, with many applications now being deployed in real-world settings [6]. These applications cover a wide range of tasks. They include analyzing environmental occupancy to detect the presence of individuals [7]. They also include counting the number of people in a given space, localizing targets within the environment [8], and identifying specific individuals [9]. Other tasks involve implementing contactless gesture recognition [10], performing Human Activity Recognition (HAR) [8] and biometric measurement (e.g., gait, vital sign) [11, 12]. Figure 1.1 presents a visual summary of these major categories, followed by a brief discussion on each.

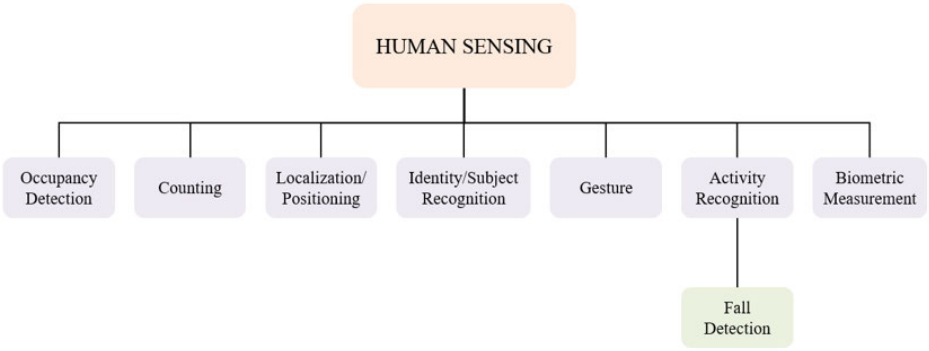


Figure 1.1: Classification of human sensing applications in indoor environments

- **Occupancy:** Detecting the presence or absence of individuals enables energy-efficient building operation through intelligent control of lighting, and HVAC systems. It can be used in security surveillance also.

- **Counting:** Having an accurate count of the people present in each region of a building can help with evacuation in an emergency, ensuring compliance with building safety codes. It can also be applied in smart scheduling of shared spaces (e.g., meeting rooms or classrooms) and retail analytics.

- **Localization/positioning:** Determining the location of individuals supports location-based services, (e.g., indoor navigation in hospitals). It can also help monitor vulnerable individuals (e.g., older adults), to ensure safety within defined spaces.

- **Identity recognition:** Determining the identity is concerned with distinguishing and tracking specific individuals within an environment. It is essential for multi-target localization, contact tracing, intrusion detection, and secure authentication in smart homes or workplaces.

- **Gesture recognition:** Contactless gesture recognition supports human–machine interaction, smart appliance control, and touch-free interfaces (reducing contamination risk). It also has applications in gaming and automotive interfaces.

- **Activity recognition:** It allows unobtrusive monitoring of Activities of Daily Life (ADL). A change in ADL can be a precursor to a cognitive or physical decline. It is also valuable in sports performance analysis. Gesture recognition is sometimes considered to be part of activity recognition. A common activity recognition application involves automated fall detection.

- **Biometric measurement:** Unlike identity recognition, which focuses on distinguishing individuals, biometric measurement emphasizes monitoring the body itself. The application includes contactless monitoring of vital signs such as heart rate, respiration, or gait stability. This enables personalized healthcare, early detection of medical risks, and continuous well-being assessment in smart and aged-care environments.

Figure 1.2 presents a taxonomy of human sensing modes, which are broadly divided into three major approaches. Table 1.1 outlines the key differences between the three major sensing modalities.

Vision-based sensing relies on cameras and/or video devices. While these techniques can achieve high accuracy, they require substantial computational resources and are limited to a specific Region of Interest (ROI). Their performance is further affected by environmental factors such as occlusion and lighting conditions. Most critically, they are often perceived as privacy-invasive, posing significant challenges for adoption in residential or care facilities.

Wearable sensors enable continuous monitoring directly from the body and therefore are not constrained to an ROI. While they may not be as privacy invasive as a vision-based system, they can be still physically obtrusive and require active user cooperation. They can only function if the user is wearing the device and ensuring it remains charged. On the positive side, they have relatively low computational need and are not impacted by environmental changes.

Ambient sensor-based device-free technique makes use of signals and physical phenomena in the surrounding environment, such as visible light, infrared or thermal signatures, capacitive or electric fields, and radio frequency signals. Device free sensing offers the advantage of minimal privacy intrusion. It does not require active user cooperation, which makes it less burdensome for the user. However, it is constrained to an ROI. Its performance can also be significantly influenced by environmental factors, such as interference.

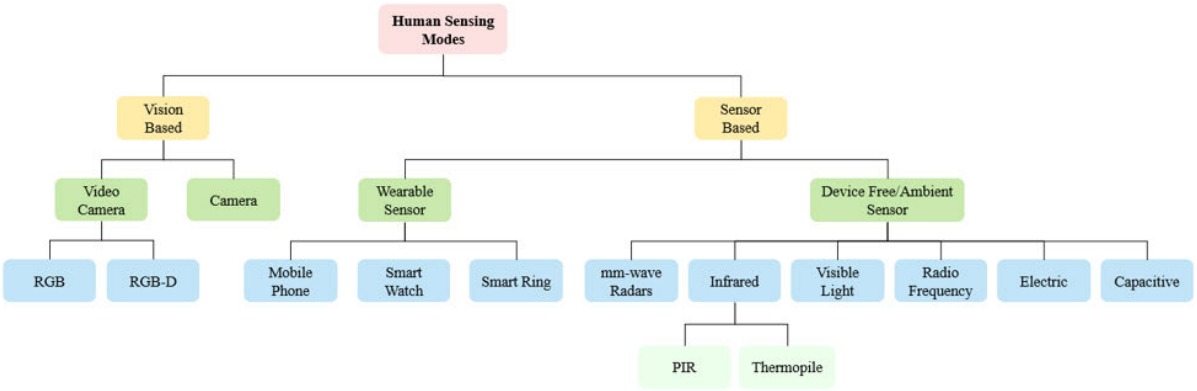


Figure 1.2: Taxonomy of human sensing based on sensing modality

Table 1.1: Brief high-level comparison of various modalities of human sensing

Modality	**Privacy**	**ROI**	**User Coop.**	**Env. Impact**	**Comp.**
Vision	Yes	Yes	No	Yes	High
Wearable	Obtrusive	No	Yes	No	Low
Device free	No	Yes	No	Yes	Moderate

Note: User Coop. denotes the need for user cooperation; Env. Impact refers to sensitivity to environmental conditions; Comp. indicates computational demand.

1.3 Motivation

The discussion presented here highlights the breadth of sensing modalities available for human sensing and how each modality offers a distinct approach of acquiring and processing data. It also illustrates the wide range of human sensing applications. Our research seeks to address some of the research challenges in designing human sensing systems that are accurate, unobtrusive, and robust in real-world environments. Low-resolution thermopile sensors present a promising, privacy-preserving alternative for human sensing application. The research work examines two critical human sensing applications: localization and automated fall detection. Prior research in this domain has

several limitations. Many studies rely on a limited number of sensors. This restricts the monitored area and also ignore the impact of different layouts. In addition, algorithms are frequently evaluated on the same participants used for training and calibration. These leave important questions about real-world generalizability unanswered. To address these critical gaps, this thesis systematically evaluates the true potential of a multiple thermopile-based system.

1.4 Research Contributions

The contribution of the work undertaken by this research can be summarized as follows:

1.4.1 Bespoke Thermopile Sensor Network

To date, no open-source or commercially available hardware platform has been specifically designed for human sensing using privacy preserving, low-resolution thermopile sensors. To address these gaps, we present the development and implementation of a scalable, thermopile sensor network. The full hardware design, including PCB schematics and supporting code, is made publicly available to ensure reproducibility and to enable adoption and extension by the research community. The developed platform is utilized to perform two major human sensing tasks: localization and fall detection.

1.4.2 Localization

Existing work on thermopile sensor-based localization systems has several limitations. Most prior studies employ a small number of sensors, which restricts the monitored area and limits scalability. The impact of sensor arrangement and mounting position on localization accuracy is not well understood. In addition, many reported systems are evaluated in small areas with few participants, leaving questions about real-world

applicability unanswered. Our work addresses these gaps by systematically evaluating many ceiling- and wall-mounted sensor configurations. We collected data from 8 participants and assessed performance and evaluated generalizability to unseen users. To demonstrate scalability, we deployed four ceiling-mounted sensors to cover a substantially larger area than previously reported in the literature. Our experimental results show median localization errors between 0.2 m–0.27 m, confirming the viability of multi-thermopile networks for practical indoor localization. An article on this work and the thermopile sensor network design is under review at the IEEE Sensors Journal (Q1).

1.4.3 Fall Detection Using Multiple-Thermopile Sensors

Existing work on thermopile sensor-based automated fall detection systems has two major limitations. The fall detection algorithms are often trained and tested on the same participants. This makes it difficult to comment on their generalizability. Our work evaluates the proposed algorithms on unseen individuals whose data was not used in training. This approach provides a stronger validation for real-world applicability. Prior studies use only a small number of thermopile sensors. The required number of sensors and the impact of spatial placement are not well understood. In contrast, our work systematically investigates a large number of ceiling- and wall-mounted sensor configurations. Our experimental results indicate that combining one ceiling-mounted and one wall-mounted sensor achieves the optimal balance between accuracy, cost, and implementation challenges. An article based on this work has been published in IEEE Sensors Journal (Q1).

1.5 Organization of the Thesis

The thesis is organized into five chapters. Chapter 1 has outlined the background, motivation, along with the contribution of the research undertaken. Chapter 2 reviews

existing work, with emphasis on thermopile-based human sensing, and summarizes key studies relevant to this field. Chapter 3 details the design and implementation of the thermopile-based sensor network and presents a machine-learning-based localization framework built on this platform. Experimental results demonstrate its effectiveness, with performance compared against reported studies. The proposed approach is shown to be scalable for large setup. Chapter 4 investigates automated fall detection using the same sensing infrastructure, evaluates performance against prior work, and offers guidance on optimal sensor placement. The proposed approach is shown to be robust and generalizable. Chapter 5 concludes the thesis and outlines directions for future research.

It should be noted that Chapters 3 and 4 are reformatted journal articles and therefore also include in situ literature reviews specific to their respective problem domains of localization and fall detection. Together with the review presented in Chapter 2, these embedded reviews offer a complete and up-to-date overview of thermopile-based human sensing research. Chapters 3 and 4 also include concluding discussions and future research directions related to the respective studies. These should be considered alongside, and in addition to, the overarching conclusions presented in Chapter 5.

Chapter 2

Literature Review

This chapter begins with a brief overview of vision- and wearable sensor-based approaches to human sensing. Vision-based systems raise significant privacy concerns, especially in residential settings. Wearable devices, on the other hand, depend on user compliance which limits their reliability. These and other limitations have motivated research into ambient sensor-based device-free techniques. This chapter introduces device-free human sensing and reviews the increasing role of infrared (IR) technologies in this area. Particular attention is given to thermopile sensors, which form the core sensing modality of this PhD research. As outlined in Chapter 1, this work focuses on indoor localization and automated fall detection. Accordingly, this chapter presents a detailed review of prior thermopile-based human sensing studies relevant to these two applications.

2.1 Computer Vision for Human Sensing

Human sensing based on computer vision [13, 14] has been studied for several decades. This has enabled many applications, including video surveillance [15], public safety [16], and gesture recognition [17]. Vision-based techniques utilize both RGB and RGB Depth

(RGB-D) cameras. Standard RGB images, composed of red, green, and blue channels in the visible spectrum, are captured by conventional cameras employing Complementary Metal-Oxide-Semiconductor (CMOS) sensors. Due to their lower cost, RGB cameras are more widely adopted for human sensing applications than their RGB-D counterparts. However, vision-based frameworks that use RGB-D data generally achieve higher accuracy, an advantage stemming directly from the additional spatial information provided by the depth channel.

Camera-based human sensing can achieve high accuracy because visual data provides a rich source of information. Computer vision and machine learning algorithms can analyze these detailed visual cues effectively. However, such methods face practical limitations. They are often hindered by line-of-sight occlusion, constrained viewpoints, variable lighting conditions, and insufficient sensor resolution in real-world deployments [18]. More critically, they raise privacy concerns. This perception is a major barrier to end-user adoption, particularly in residential and aged-care environments.

2.2 Wearable Sensor for Human Sensing

Wearable sensors are another common method for human sensing [19]. This trend is driven by advances in IoT and mobile computing. Such sensors are integrated into everyday devices like smartphones, smartwatches, and fitness trackers. They analyze signal changes linked to user activity, turning human sensing into an additional value-added feature.

A primary application area is fall detection [20]. Another key application is localization, where the near-ubiquitous smartphone can be used to track a person's position [21]. Furthermore, wearable technology has been widely used for activity and gesture recognition [22]. Wrist-worn devices, equipped with inertial sensors, accelerometers, and tilt switches, are used to classify a range of daily activities, including sleep/wake cycles

and specific hand gestures. Additional applications include person identification (e.g., security access cards).

Wearable technologies face several limitations. Users must consistently wear the devices for them to function. This requirement introduces issues of compliance and intrusiveness, as the user must be physically attached to or carry a device. Such barriers are particularly significant for older adults, a key demographic for many applications, who can resist carrying or wearing dedicated sensors [23]. The possibility of the device being forgotten or misplaced introduces additional practical risks that degrade system reliability. Battery life is another challenge. Although these devices are designed for low power consumption, continuous monitoring still depletes the battery, requiring frequent recharging and reducing long-term usability [24].

2.3 Device-Free Human Sensing

Device free human sensing technologies, typically using ambient sensors, have emerged as alternatives to vision- and wearable-based systems. They address privacy concerns while potentially providing robust performance and are independent of environmental lighting conditions. In addition, they require no active participation from monitored individuals. They operate passively in most configurations, requiring no wearable devices, or behavioral alterations from users. Individuals can proceed with normal activities while the sensing systems unobtrusively extract information to deduce occupancy, location, activity, gestures, and biometric parameters. This combination of privacy preservation, lighting-condition independence, and non-intrusive operation positions ambient sensing modalities as viable and desirable technologies for human sensing applications.

2.4 Device Free Sensing Modalities

2.4.1 Radio Frequency (RF) Sensing

Radio frequency–based sensing is well suited to widespread deployment due to the ubiquity of existing wireless infrastructure (e.g., Wi-Fi, Bluetooth). By analyzing Channel State Information (CSI) and Received Signal Strength Indication (RSSI), these systems capture the multipath propagation variations caused by the presence of the human body. Applications have matured from simple occupancy detection to more complex tasks like vital sign monitoring [25]. RSSI is a relatively coarse metric and has been used for occupancy estimation [26] and localization [27]. CSI is a finer grain metric and can recognize activities [28], detect fall events [29] and perform contactless vital signs monitoring [30].

The primary strength of this approach is the potential of repurposing existing wireless infrastructure. However, signals are highly sensitive to multipath propagation changes caused by everyday occurrences, such as rearranged furniture or the addition of absorptive materials [31]. Moreover, model portability remains a critical unresolved challenge. A model trained in one specific environment may not generalize to another without extensive retraining. Reliability is further compromised by interference from competing wireless networks [32] and long-term environmental drift [33].

To mitigate the interference issues and resolution limits associated with standard RF signals, higher frequency modalities have been investigated as an alternative approach.

2.4.2 mmWave Radar Systems

Millimeter-wave (mmWave) radars used in human sensing applications typically transmit Frequency-Modulated Continuous-Wave (FMCW) signals. The signal reflected back to the receiver by the target is then analyzed to perform tasks like localization [34],

activity recognition [35], fall detection [36], gesture recognition [37], and vital signs like heart and respiration rate monitoring [38].

mmWave based systems are unaffected by illumination, preserve privacy, and can offer high precision. However, cost and complexity are potential barriers to widespread adoption. Its limited Field of View (FoV) often necessitates multi-sensor arrays to cover a single room, exponentially increasing deployment costs. Furthermore, the high data rate necessitates specialized edge computing hardware, adding to the system complexity. Similar to Wi-Fi, mmWave performance is sensitive to sensor placement and environmental changes.

Because RF and mmWave radar systems share an inherent susceptibility to multipath propagation and wireless interference, alternative modalities have been explored to operate independently of electromagnetic wave reflection.

2.4.3 Capacitive Sensing

Capacitive sensing, sometimes also termed as electric field sensing, exploits the inherent electrical properties of the human body [39]. By instrumenting electrodes that detect body-generated or body-distorted electric fields, these systems infer human sensing tasks. Existing research includes localization [40] and identity recognition [9] using floor embedded capacitive sensing and basic hand gesture recognition using an array of capacitive proximity sensors. Ceiling- and wall-mounted capacitive sensors have been used for localization by Tariq et al. [41] and for occupancy detection and localization by Tang and Mandal [42].

Key strengths of this modality are that it does not depend on lighting conditions and is not affected by RF interference. Furthermore, the approach offers complete privacy preservation since it only captures electrical distortion information. However, sensing is limited to short distances making room-scale or building-scale monitoring challenging.

Environmental electrical noise can also substantially degrade system performance.

2.4.4 Vibration-Based Sensing

Structural vibration sensing detects human presence via footstep-induced vibrations captured by seismic sensors or geophones. This modality treats the floor surface as a large-scale sensing medium. Pan et al. performed person identification by extracting unique gait patterns from structural vibration signals [43]. Mirshekari et al. extended this concept for localization [44, 45]. The distinct advantage of this modality is its environmental robustness. Such systems are inherently immune to visual occlusion, lighting changes, or RF interference. A critical limitation is the inability to detect stationary subjects. Additionally, distinguishing human footsteps from background building noise (e.g., HVAC systems, door slams, or external traffic) requires sophisticated filtering and signal processing, often leading to high false alarm rates [39].

2.4.5 Visible Light-Based Sensing

Visible light-based human sensing detects optical signal disturbances caused by human presence and movement through shadowing or reflection. Recent research works have developed applications for localization, gesture, and activity recognition. Localization studies include "Watchers on the Wall," that utilizes wall-embedded light sensors to track targets by analyzing the shadowing effects cast on walls, achieving robust device-free tracking [46]. Similar sensing method has been utilized by the FieldLight [47] system for localization. Liao et al. demonstrated that visible light sensing can enable hand gesture recognition [48]. Yang et al. recently introduced an approach driven by generative language models (LLMs), for activity recognition [49].

However, visible light sensing imposes stringent line-of-sight (LOS) requirements

that fundamentally limit its practical applicability. Unlike RF signals that penetrate obstacles, light is easily blocked by furniture or non-target individuals, creating significant coverage dead zones. The existing studies were conducted under controlled lighting conditions, and their performance in real-world conditions is uncertain.

2.4.6 IR Sensing

The reliance on external illumination sources fundamentally restricts the applicability of visible light sensing in unlit environments. This limitation can be addressed by operating in the thermal spectrum. All objects emit infrared (IR) radiation, and the intensity of this emission increases with temperature. In the context of human sensing, two types of IR sensors are of primary interest: Passive Infrared (PIR) sensors and thermopiles. Algorithms interpret the IR signals from these sensors to enable a range of human-sensing applications.

2.4.7 PIR Sensors

PIR sensors based on the pyroelectric technique [50] are widely used for human sensing in applications such as security alarms and automatic lighting. Recent integrations with machine learning have enabled robust occupancy monitoring in smart buildings. They can distinguish between different occupancy levels rather than only detecting simple presence [51]. Although PIR sensors could theoretically be repurposed from existing security systems, those installations typically provide only binary motion triggers. Advanced human sensing applications require access to the analogue output of the PIR element [52]. This necessitates purpose-built PIR nodes. Furthermore, substantial modifications are required to detect stationary subjects [53] as PIR sensors only respond if there is movement across their FoV.

Several studies have investigated activity detection [54] and positioning [55] by

using the analogue output of bespoke PIR sensors. Further examples include the research conducted by Ngamakeur et al. who demonstrated that a Deep CNN-LSTM network could effectively learn spatial features, enabling precise indoor location [52]. Yuan et al. showed that fusing PIR data with other sensing modalities in an interpretable framework allows for granular human identification and activity recognition [56].

However, the use of bespoke PIR sensors for such advanced tasks diminishes the simplicity and low-cost advantages typically associated with PIR technology. Although the utilization of analogue signals improves performance, these methods increase computational overhead. In addition, PIR-based human sensing systems often require a high sensor density, with many sensors needed per unit area.

The literature review demonstrates that a wide range of sensing modalities are available for human sensing. However, each modality has inherent limitations that restrict its suitability to specific human sensing tasks. Table 2.1 summarizes the appropriateness, limitations, and unsuitability of different sensing modalities across the range of human sensing tasks.

Table 2.1: Comparison of sensing modalities and their suitability for different human sensing tasks.

Sensing Modality	**Occupancy**	**Localization**	**Identity**	**Activity**	**Gesture**	**Biometric**
RF	✓	✓	×	✓	~ (*CSI Only*)	~ (Respiration)
mmWave Radar	✓	✓	✓ (*Gait*)	✓	✓	✓ (*Heart/Resp*)
Capacitive	✓	✓	✓ (*Gait*)	~ (Basic)	✓	×
Vibration	~ (Motion only)	✓	~ (Footstep)	~ (Walking)	×	×
Visible Light	✓	✓	✓	~ (Walking)	~	×
PIR	~ (Motion only)	✓ (Array needed)	×	✓	×	×
Thermopile	✓	✓	~ (*not with low res.*)	✓	~ (*not with low res.*)	~ (*Temperature*)

Note: ✓ denotes suitable; ~ denotes limited suitability; × denotes not suitable. Parenthetical remarks indicate capabilities reported in prior studies.

2.5 Thermopile-based Human Sensing

Thermopile sensors are IR or thermal cameras formed by connecting multiple thermocouples in series. High-resolution thermal cameras provide detailed information but are expensive and can raise serious privacy concerns, much like regular cameras. In contrast, low-resolution thermopiles are inexpensive and inherently privacy-preserving, as they capture only coarse thermal patterns rather than identifiable features. Although less information-rich, they remain effective for human target capture, positioning, and activity recognition. This privacy advantage has made low-resolution thermopiles particularly attractive for human sensing applications [39]. Among these, the AMG8833 Grid-Eye with an 8×8 resolution is one of the most widely used [57].

2.5.1 Device-free Localization Using Thermopile Sensors

Device-free localization (DFL) or passive localization is the method of localizing or positioning a subject where the target does not actively participate in the localization process by carrying a device. Various RF [58] and non-RF [39] DFL techniques have been proposed. In this PhD work, we are focusing on infrared-based DFL using low resolution thermopile sensors.

Kuki et al. [59] perform subject tracking for a target area of 2.56 square meters, which achieved an accuracy (defined by mean error) between 0.15m and 0.35m. In the experiment, the experimental subjects passed through the execution of 11 medium movement patterns and four static patterns (for activity recognition).

Shetty et al. [60] proposed using thermal imaged captured by a ceiling mounted 8×8 Grid-Eye thermopile. Their proposed algorithm consisted of bicubic interpolation to improve the resolution of the image before blob tracking. While their experiments show the capability of localization and tracking of both stationary and moving subjects, unfortunately, they did not report any data on accuracy.

Tariq et al. [61] used a ceiling-mounted Omron D6T-44L thermopile sensor with a resolution of 4×4 pixels and achieved a 0.096 m Root Mean Squared Error. Their localization method was based on a supervised Machine Learning approach, using a Random Forest regressor trained on thermal pattern features to estimate subject position. Unfortunately, the coverage is limited to the FoV of a single thermopile of 2.48 m × 2.57 m, and only one participant was involved in the evaluation. For ground truth acquisition, they employed a Marvelmind ultrasonic system [62], which was also used to measure ground truth in our study.

Chen et al. [63] implemented a dual-function sensing system utilizing two wall-mounted thermopile sensors with a resolution of 16×4 pixels. Operating within a monitored area of 7 sqm, their system achieved an average localization error of 0.134 m through an Angle-of-Arrival (AoA) based algorithm. Furthermore, the authors extended the system's capability to fall detection using a k-Nearest Neighbours (KNN) classifier. Based on a dataset comprising 80 fall events and 80 regular movements, they reported a detection accuracy of 93%. However, it should be noted that the MLX90620 sensor model utilized in this study has since been discontinued by the manufacturer, limiting the direct reproducibility of this specific hardware configuration.

Chen et al. [64] used a table mounted Grid-Eye 8×8 thermopile in their proposed localization system. They used a Convolutional Neural Network to extract features from interpolated 32×32 images, which was used to train a Support Vector Machine. A localization RMSE of 0.19 m was achieved for a coverage area of 2.9 sqm. They also introduced an application that can infer the orientation of a person while seated on a chair.

Qu et al. [65] addressed the resolution limitations of wall-mounted Grid-Eye sensors by upscaling the raw thermal data into 71×71 pixel images. Their processing pipeline utilized Gaussian filtering for denoising and adaptive thresholding to separate targets from the background. By identifying the highest-temperature pixels and applying a

Kalman filter for trajectory smoothing, they reported an average localization error of 0.07 m. Notably, this study extended beyond single-occupant monitoring to track three simultaneous targets. While the authors acknowledged that variations in body size and clothing intensity alter thermal signatures, they observed that these signals remained relatively stable during detection intervals. Although the overall system accuracy faced challenges in complex scenarios, the methodology demonstrates the feasibility of multi-target tracking using low-resolution arrays.

Faulkner et al. used wall-mounted [66] and ceiling-mounted [40] Grid-Eye sensors to localize human subjects. For their wall-mounted work, MLP-based angular and radius coordinate models for the target with respect to thermopile sensors. These models were then used to localize individuals while within the coverage of a single thermopile or the joint common coverage of two or three thermopiles. Their maximum coverage area was 12 sqm and achieved median errors were between 0.1 and 0.35 meters. For their ceiling mounted work, they determined that Convolutional and Recurrent Neural Networks are the most accurate Machine Learning algorithms while achieving a median accuracy of as low as 0.032 m. However, their system's coverage is limited to the FoV of a single thermopile (8.3 sqm). For both of their works, Faulkner et al. utilized HTC Vive, a virtual reality system, for accurate ground truth recording.

A critical limitation observed across the reviewed literature is the restricted coverage area, which significantly hampers the ability to predict how these localization systems would perform in realistic, large-scale environments. Most studies validate their algorithms within small, controlled zones, such as the 2.56 m^2 area used by Kuki et al. or the single thermopile field-of-view (FoV) of approximately 6–8 m^2 seen in the works of Tariq et al. and Faulkner et al. Even multi-sensor setups, like the wall-mounted configuration by Chen et al., only cover about 7 m^2, effectively representing a small room or a specific zone rather than a full apartment.

Most of the reviewed DFL techniques have been validated only within small coverage areas, typically restricted to the FoV of a single sensor or a small array. This limitation makes it difficult to predict system performance in realistic scenarios, which involve larger spaces and potential environmental interferences. The scalability of these methods to fully cover a standard living room or apartment using low-resolution sensors remains an area requiring further investigation.

2.5.2 Fall Detection

This section provides an in-depth review of fall-detection research utilizing low-resolution thermopile sensors. Prior to this, we present a short summary of related work employing higher-resolution thermal imaging to establish the broader context.

Fall Detection with High-Resolution Thermopiles

Yu et al. [67] developed a fall detection system utilizing an 80×64 pixel thermopile imaging array. Their approach involved a back-projection algorithm to analyse thermal data. The system was evaluated using a dataset comprising 100 fall events and 100 non-fall events. The reported performance metrics were high, achieving a sensitivity of 98% and a specificity of 100%. However, the study's experimental design was limited to a single participant for both training and testing. This makes it difficult to assess the system's robustness and generalizability to a broader population.

Zhang et al. [68] proposed a system capable of identifying falls alongside static activities like standing and sitting, as well as dynamic activities such as walking and jumping. Their setup utilized two MLX90640 thermopile sensors (24×32 pixels) positioned to capture side and top views of the participants. The system achieved an accuracy of 97.5% for static activities and 94.6% for dynamic activities, including falls. As with the work of Yu et al., both training and testing were performed using data from

a single participant, which restricts the validation of the system's generalizability to different users.

Rezaei et al. [69] developed a fall detection system using two 24×32 pixel MLX90640 sensors mounted on a room's side wall and ceiling. They employed Deep Learning techniques for activity classification, location detection, and fall detection. Their best-performing algorithm achieved a binary classification accuracy of 97.9% and an F1-score of 94.5% for fall detection. While the study involved data from 10 participants, the algorithms were not evaluated on individuals excluded from the training set. Therefore, the system's generalizability to unseen subjects is unknown.

Kawashima et al. [70] proposed a fall detection and action recognition system using a single 16×16 pixel thermopile array mounted on the ceiling. Their method combined a Convolutional Neural Network (CNN) with Long Short-Term Memory (LSTM) layers to extract both spatial and temporal features from the thermal data. The system achieved an overall accuracy of 91.07%, with a specific fall detection accuracy of 93.06%. Notably, they validated their approach using a leave-one-person-out cross-validation protocol with a dataset of nine participants, ensuring that the test data came from an individual not seen during training.

Fall Detection with Low-Resolution Thermopiles

Mashiyama et al. [71] proposed an activity recognition and fall detection method using a ceiling mounted Grid-Eye sensor. Their system classified five activity types (No event, Stop, Walking, Fall, Sitting) using an SVM classifier trained on data from six volunteers, the training set consists of 50 falling, 30 sitting and 45 walking events for each subject. Their test set had 30 falling, 20 sitting, and 27 walking events per participant. The reported recognition accuracies were 100%, 94.8%, 99.99%, 100% and 78.6% for these activities, respectively. However, it is unclear how the system would perform on unseen subjects whose data were not used for training. In addition, no benchmarking was

conducted against other classifiers or against other published algorithms. This makes it difficult to place their results in context. The authors did examine the effect of varying sensor height, floor material and room temperature on recognition performance.

Adolf et al. [72] developed a system for body posture recognition and fall detection using a single Grid-Eye sensor. The system utilized the Inception-v3 Deep Learning model to classify postures such as standing, sitting, and lying down. It achieved a sensitivity of 85% for detecting lying postures (indicative of falls) and a specificity of 93% for the same class. The study included four participants. However, there was no testing on unseen individuals. Therefore, the system's ability to generalize to new users remains difficult to ascertain.

Fan et al. [73] proposed a fall detection system using a single Grid-Eye sensor and compared several Deep Learning methods, including LSTM networks and GRU. Their experiments covered falls in both perpendicular and parallel orientations relative to the sensor. They found that parallel falls were easier to classify, with their best model achieving an F1 score of 100%. Although training and testing data were collected separately, the study relied on a small sample size of only two participants, raising concerns about the system's robustness in real-world applications.

Tao et al. [74] introduced a fall and activity detection system using a single ceiling-mounted Grid-Eye sensor. They collected a dataset named "Infra-ADL2018" from eight participants performing seven distinct activities across 24 sessions. Using a leave-one-subject-out cross-validation approach, their system achieved an overall accuracy of 87.5% and a sensitivity of 100% for fall detection. The authors concluded that a single-sensor setup has limitations and suggested that a multi-sensor system providing multiple fields of view would improve robustness.

Chen et al. [63] implemented a fall detection application using two wall-mounted 16×4 infrared sensors placed at different locations. By integrating data from both sensors, the system aimed to capture three-dimensional information. Fall actions were

classified using a k-NN model. Tested with 80 fall actions and 80 normal actions, the system achieved a sensitivity of 95.25%, a specificity of 90.75%, and an overall accuracy of 93%. However, the system was not evaluated on unseen individuals.

Ogawa and Naito [75] proposed a fall detection system using a single ceiling-mounted thermopile with a resolution of 16×4. The system analyzed temperature distribution over a two-second window and employed a voting classifier to identify falls, achieving an accuracy of 97.75%. The study utilized a dataset of 1600 samples collected from 10 individuals. Like several other studies in this domain, it did not explicitly evaluate performance on unseen subjects.

2.5.3 Human Activity Recognition

This section reviews thermopile-based human activity recognition (HAR) studies. Although some of these works incorporate fall detection as part of broader activity classification frameworks, fall detection is not their primary focus. Nevertheless, these studies employ the same sensing modality and adopt similar evaluation practices. These studies provide useful contextual insights and help position the proposed approach within the broader research landscape.

Krishnan et al. [76] proposed an activity detection system utilizing a ceiling-mounted 24×32 infrared array sensor. They collected data at varying resolutions and applied Super-Resolution (SR) and denoising techniques to enhance image quality, alongside a Conditional Generative Adversarial Network (CGAN) for data augmentation. A hybrid model combining a CNN and LSTM was used to classify activities. Their experiments demonstrated that the super-resolution and denoising pre-processing significantly improved classification accuracy, particularly for lower-resolution inputs.

Muthukumar et al. [77] developed a hybrid Deep Learning model to detect human activities using wide-angle 24×32 sensors positioned on both the wall and ceiling. They

addressed the challenges of wide-angle lens distortion by employing a CNN to extract spatial features from individual frames, followed by an LSTM to analyze temporal sequences over a short window of less than one second. Their approach achieved high accuracy for activities such as falling and lying, effectively mitigating the issues caused by peripheral distortion in wide-angle sensors.

Yin et al. [78] develop an approach using LSTM models on data captured from a single Grid-Eye sensor. They achieve 96.3%-100% accuracy for activities including sitting, walking, standing, and jumping. The LSTM-based classifier demonstrates superior performance compared to several other classifiers (e.g., CNN, KNN, SVM, and ELM). However, it is difficult to generalize the performance since the training and testing were done on only one participant.

Gochoo et al. [79] developed a sophisticated yoga training system utilizing three strategically positioned Grid-Eye sensors—one ceiling-mounted and two wall-mounted at 0.9m height. Their DCNN approach achieved remarkable accuracy rates exceeding 98% for most of the 26 yoga poses analyzed. The system incorporated three convolutional layers with 32, 128, and 256 kernels, respectively, each using 3×3 kernel sizes with step size 1. Data collection involved 18 volunteers performing each pose twice for 20 seconds, generating 187,200 total samples across all participants.

Kuki et al. [59] approached activity recognition from a trajectory analysis perspective, developing a system capable of detecting and classifying 15 different movement patterns using ceiling-mounted AMG8833 sensors. Their fuzzy logic-based approach achieved accuracies ranging from 41.91% to 100% depending on the activity type, with stationary activities like sitting achieving perfect classification (100%) while dynamic activities like fast walking achieved lower accuracy (41.91%). The system effectively recorded human movement trajectories by tracking the center position of detected human body areas, with sitting, squatting, and step-by-step movements achieving the highest recognition rates.

Relatively high-resolution thermal thermopiles (120×160) have been utilized by Shih et al. [80] and Naik et al. [81] for activity recognition studies. Such studies were not considered due to their potential implications for privacy.

2.6 Additional Contextually Related Work

In addition to thermopile-based localization, fall detection and activity recognition studies, the literature includes several thermopile-based works that address related human sensing problems. Although these studies target different applications, they employ similar sensing principles and evaluation methodologies. These studies are reviewed briefly in this section to provide complementary context.

Basu et al. [82] developed a method to count the number of people and the direction of movement within the FoV of an 8×8 Grid-Eye sensor. They used SVM classification on connected component-based features and local peak counts. They counted the number of people with 80% accuracy. Subsequent cross-correlated analysis infers the direction of motion of objects in a set of scenes as left-to-right, right-to-left, top-to-bottom, and bottom-to-top. Ng et al. [83] used five Grid-Eye sensors distributed along the room's walls to detect a person in an area of 4.6m × 2.7m. However, they did not conduct any positioning or activity recognition tasks. Bharad et al. [84] implemented an occupancy detection method using Grid-Eye sensors. They used the sensor to monitor the temperature in an aisle or doorway. When a person passed by, the resulting temperature change beyond a set threshold was used to detect their presence. Jeong et al. [85] used the Grid-Eye sensors to achieve differentiation of human subjects from those in all target areas by analyzing target temperature, step distance, and size of thermal zones. Kallur et al. [86] proposed a method to pre-process the raw sensor data using interpolation and Gaussian filters, which can then potentially be used to locate human targets. Shubha and Shastrimath [87] proposed a real-time occupancy detection system

using an 8×8 Grid-Eye sensor interfaced with a Raspberry Pi. They employed a signal processing pipeline involving bicubic interpolation, Gaussian filtering, and adaptive thresholding to enhance the low-resolution thermal data and suppress background noise. Their experimental results demonstrated the system's capability to effectively detect multiple static and moving human targets while minimizing computation time.

Naser et al. [88] proposed and investigated an adaptive human segmentation and occupancy estimation framework. They performed data acquisition using MLX90640, which has a higher resolution (24×32) than the Grid-Eye. They employed a deep encoder-decoder convolutional neural network to segment people from the output of the thermopile sensor semantically. This framework allows the distance between the target and the sensor to be determined by estimating the area occupied by the target at different sensor positions. In a related study, Naser et al. [89] further developed the method for human distance estimation using the same thermal sensor array. However, they did not explicitly position or localize. However, their results indicated that their proposed framework was able to detect occupancy and vacancy with 100% accuracy. Their subsequent work [90] demonstrates that 24×32 thermal images can be reconstructed to high-resolution images to some extent raising privacy concerns when using such sensors. *This finding reinforces our decision to employ a significantly lower-resolution (8×8) thermopile sensor in this research.*

2.7 Summary

This chapter has reviewed the major sensing modalities for human sensing, including vision-based, wearable, and device-free approaches. Among device-free techniques, thermopile sensors offer a good balance of privacy preservation, cost-effectiveness, and sensing capability. The literature review has identified key gaps in existing thermopile-based research, particularly with respect to restricted coverage areas and the limited

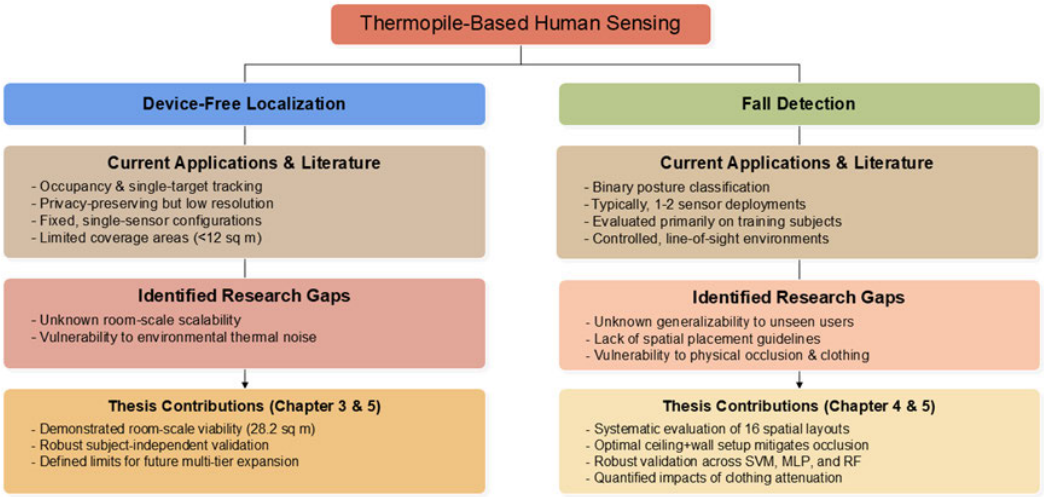


Figure 2.1: Thermopile-Based Human Sensing Gaps

evaluation on unseen subjects (see Figure 2.1). These identified gaps directly motivate the research presented in the following chapters, which aim to address these challenges through improved system design, deployment strategies, and evaluation methodologies.

Chapters 3 and 4 are reformatted journal articles and therefore include in situ literature reviews specific to their respective problem domains of localization and fall detection. As part of these reviews, summary tables of thermopile-based localization (Table 3.1) and thermopile-based fall detection studies (Tables 4.1 and 4.2) have been provided. Together with the review presented in this chapter, these embedded reviews offer a complete and up-to-date overview of thermopile-based human sensing research focusing on localization and fall detection.

Chapter 3

Device-Free Localization Using Low-Resolution Thermopiles

3.1 Abstract

Low-resolution thermopile sensors have emerged as a promising solution for privacy-preserving indoor localization. In this study, we investigate the localization performance of custom-designed thermopile sensors. We systematically evaluate multiple ceiling- and wall-mounted configurations, addressing the limitations of prior works that examined only a small number of sensors and fixed layout. A 2D CNN–LSTM regression model was trained and evaluated with data collected from 8 participants achieving median localization errors between 0.2 and 0.3 m. Our findings indicate a clear dependence of localization accuracy on the number and arrangement of sensors. Experimental benchmarking against previous approaches indicates that the proposed algorithm provides improved localization accuracy. To assess scalability, we deployed four ceiling-mounted sensors to cover a substantially larger area than the state-of-the-art and demonstrated that our proposed approach performs effectively in such a setting. All hardware design information and code are made publicly available to support continued research.

3.2 Introduction

Human sensing ranges from monitoring human presence, position, and movement, to analyzing human posture and identity [91]. It facilitates location based services [92] and ambient assisted living [93]. These capabilities enable a range of applications in smart environments, contributing to more interactive, safe, and supportive living conditions. The rapid development of computing and communication networks alongside sensing technologies are driving the advancement of human sensing technology. Indoor localization [39] is one of the key human sensing applications. Since GPS does not function reliably inside buildings, researchers have been investigating various alternative methods.

Indoor localization techniques can be broadly categorized into three types: camera or vision-based, wearable sensor or device-based and device-free. Localization using computer vision [14] has been an active research topic for several decades. Vision-based techniques can be very accurate due to a large amount of sensory input in the form of visual data. However, they cannot function in low light conditions and have privacy concerns in many application scenarios. Device-based techniques [89] utilize wearable sensors (e.g., smart watches, bracelets etc.) or smartphones. However, they rely on user cooperation and therefore cannot function if the user does not wear or carry them. There are also other inherent issues like requiring regular battery changing [94], which renders wearable sensors less reliable and inconvenient.

Device-free systems aim to mitigate some of the shortcomings of the vision- and device-based ones. Such techniques are gaining popularity for their privacy preserving nature, the ability to leverage pre-existing sensing infrastructure [95] for localizing individuals [39]. While many sensing methods have been investigated, they each have their respective strengths and weaknesses.

Radio frequency (RF) or wireless based localization techniques are the most widely

explored device-free localization [96, 97] technique in literature. However, concerns persist regarding their robustness, as these systems typically require frequent recalibration [98]. Also, from an implementation perspective, an effective localization solution is likely to incorporate multiple technologies. Non-RF methods have received comparatively less attention in literature and would benefit from additional research to assess their potential and applicability. Researchers have proposed to utilize the footstep induced vibrations, captured by seismic sensors, for human localization [44]. Smart floors based on triboelectric [99] and capacitive [40] sensors have also been developed for this purpose. In addition, wall mounted capacitive sensing has also been explored for localization [41]. Other sensing modalities for passive localization include visible light sensing based systems [47]. Thermal sensing represents another emerging modality, where passive detection of human Infrared (IR) emissions enables device-free localization using Passive Infrared (PIR) sensors [55] and Thermopiles [66].

PIR sensors leveraging pyroelectric detection of IR changes [50] have been widely used in security alarms and automatic lighting applications. Alongside localization [55] they have been utilized for human sensing applications like activity detection [54]. PIR sensors may appear suitable for repurposing from existing security systems. However, effective human-sensing applications typically require a much higher sensor density than these systems provide. Moreover, significant modifications to the sensor are necessary to make them capable of detecting a stationary subject [53].

Thermopiles, thermocouple based IR sensors [100], are commonly used as thermal camera. They can detect both stationary and moving subjects. High resolution thermal cameras are comparatively costly and can also be perceived as privacy invasive like a regular camera (see the picture of an individual captured by a high-resolution thermopile in Figure 3.1). In comparison, low-resolution thermopiles are inexpensive, not privacy invasive (see the picture of a subject captured by a low-resolution thermopile

in Figure 3.1). While they are not as information rich as their high resolution counterparts, they have become popular for localization applications [39]. By selecting lower resolution thermopiles, not only the privacy issue [66] can be eliminated but also can allow the implementation of applications with a relatively lighter workload of data processing. Moreover, the lighter computational requirement increases scalability, allowing deployment of multiple sensor arrays as an integrated sensing network.

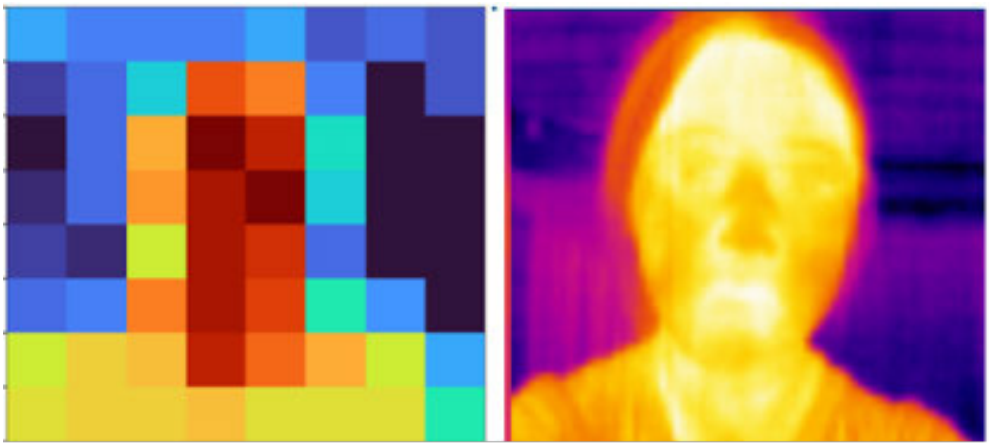


Figure 3.1: Image of a person taken by a Grid-Eye having an 8×8-pixel resolution (left) shows the inherent privacy-preserving aspect of low-resolution thermopiles. For comparison, a capture of the same subject from the same position with an 80×60-pixel resolution FLIR Radiometric Lepton 2.5 is provided. As can be observed, the higher resolution thermal camera can be perceived as not privacy-preserving.

3.3 Literature Review

Table 3.1 lists the previous studies on localization using thermopile sensors. In the localization literature, the terms accuracy and localization error are often used interchangeably. Both refer to statistical measures based on the Euclidean distance between the actual target position and the position estimated by the localization system.

Table 3.1: Summary of Localization Systems Using Low-resolution Thermopile Sensors.

Author	Sensor Model	Resolution	Number of Sensors	Coverage Area (m^2)	Accuracy (m)	Participants	Unseen?
Kuki et al. [101]	Omron D6T-44L	4×4	1	2.6	Mean: 0.10	4	No
Tariq et al. [61]	Omron D6T-44L	4×4	1	6.4	RMSE: 0.096	1	No
Chen et al. [63]	MLX90620	16×4	2	7.0	RMSE: 0.134	1	No
Chen et al. [102]	Grid-Eye	8×8	1	2.9	RMSE: 0.19	1	No
Faulkner et al. [66]	Grid-Eye	8×8	1–3	12.0	Median: 0.11–0.35	5	Yes
Faulkner et al. [103]	Grid-Eye	8×8	1	8.3	Median: 0.036–0.732	4	Yes
Proposed (Lab)	**Grid-Eye**	**8×8**	**4**	**5.8**	**RMSE: 0.29, Mean: 0.24, Median: 0.20**	**8**	**Yes**
Proposed (Lab)	**Grid-Eye**	**8×8**	**2**	**5.8**	**RMSE: 0.30, Mean: 0.25, Median: 0.21**	**8**	**Yes**
Proposed (Foyer)	**Grid-Eye**	**8×8**	**4**	**28.2**	**RMSE: 0.46, Mean: 0.35, Median: 0.27**	**1**	**No**

Note: Accuracy metrics are reported using the original definitions from each study (mean, RMSE, or median). All values are expressed in meters.

Kuki et al. [101] covered an area of 2.56 square meters with a single ceiling mounted 4×4 thermopile sensor (Omron D6T-44L). They achieved localization accuracy between 0.15 m and 0.35 m using a Support Vector Machine (SVM)-based algorithm. Tariq et al. [61] used a ceiling mounted 4×4 resolution (Omron D6T-44L) thermopile to achieve a 0.096 m Root Mean Squared Error (RMSE). They trained multiple neural network-based algorithms and found the One-Dimensional Convolutional Neural Network (1D CNN) to be the most accurate. Shetty et al. [60] proposed using low-resolution thermopile images captured by the Grid-Eye sensor for target tracking of foreground areas. Their experiments show localization and tracking of both stationary and moving subjects. However, they did not report any quantitative accuracy results. Chen et al. [102] used a table mounted Grid-Eye sensor to achieve a localization RMSE of 0.19 m. They extract features using a convolutional neural network to train a support vector machine-based algorithm. Faulkner et al. [103] used a single ceiling-mounted Grid-Eye sensor to localize human subjects accurately achieving accuracies between 0.03–0.7 m. They determined that convolutional and recurrent neural networks are the most accurate algorithms. It is important to note that all these studies used a single sensor installed at a fixed location. The coverage is also confined to between 2–9 m^2.

Chen et al. [63] implemented a localization application by using two 16×4 resolution thermopile sensors and mounting them on the wall at different locations. Their localization algorithm worked by estimating the Angle of Arrival (AOA) of each sensor's foreground temperature, and achieved a mean error of 0.134 m. Faulkner et al. [66] employed up to three wall-mounted Grid-Eye sensors to achieve median localization error between 0.13 m and 0.20 m. Multilayer Perceptron (MLP) models were trained to estimate the angular and radial coordinates of the target with respect to each thermopile sensor. The use of multiple sensors improved accuracy by capturing data from different viewpoints. However, the coverage area of these studies, defined by the overlapping FoVs of are relatively small, approximately between 7–12 m^2. Therefore, it is difficult

to ascertain the scalability of the systems.

The AMG88xx Grid-Eye family [104] is the most adopted thermopile sensor for privacy-preserving human sensing applications. Other than localization (see Table 3.1 for examples), it has been used for applications like movement tracking [82], determining the direction of facing of a subject [102], fall detection [71, 105], activity recognition [78], identifying yoga poses [79]. Example of other low-resolution thermopile includes MLX90620 (16×4 pixels resolution) which was used for localization and fall detection [63]. However, this device has been discontinued, and the suggested replacement, the MLX90640, offers a much higher 32×24 resolution. While this enables richer data capture, the sensor can potentially be intrusive to privacy. Alternatives such as the TPA81 (8×1 pixels) were used on people counting at doorways [106] and D6T-44L (4×4 pixels) were used for localization studies [101], [61]. However, their coarse resolution limits their suitability for sophisticated human sensing tasks.

The Grid-Eye has a field of view (FoV) of 60° for both horizontal and vertical directions [104]. When mounted on a ceiling at the typical residential height of 2.5 m, this translates to a sensing area of only 2.88 m $\times$ 2.88 m [103]. Expanding coverage therefore requires deploying multiple sensors in a network. However, prior studies have not demonstrated robust deployments involving more than three AMG8833 units. Moreover, in these works, additional thermopiles were used primarily to overlap FoVs for improved feature extraction, rather than to systematically expand coverage. At the hardware level, most reported implementations rely on general-purpose platforms rather than customized designs tailored for human sensing applications. For instance, Kowalski et al. [107] used a Raspberry Pi Zero with a multiplexer to control three Grid-Eye sensors. The study by Lin and Zhao [108] uses a single Grid-Eye sensor connected directly to a Raspberry Pi 4B. Other implementations include a Grid-Eye module directly hardwired to an STM32F103-based microcontroller board [39] or three breadboarded wireless nodes with WiPy2.0 modules [79]. Evaluation kits such as

the Atmel ATmega324P [105] as well as Arduino Uno EVK [102] have been used to interface with Grid-Eye Sensors. However, these platforms include superfluous interfaces, require manual I^2C wiring, and lack robustness for scalable deployment. Faulkner et al. [66], [103] utilized an STM32F103 microcontroller and a Grid-Eye sensor to develop a bespoke, network-capable sensing module. However, the design specifications required to reproduce this hardware are not publicly available.

To date, no open-source or commercially available hardware platform has been specifically designed for human sensing using Grid-Eye or other low-resolution thermopile sensors. We address this gap by developing a scalable, networked thermopile sensing platform with dedicated communication and data transfer between nodes and a central station. All PCB schematics, firmware, and supporting code are publicly released to ensure reproducibility and to enable adoption and extension by the research community.

Data is collected from eight participants, surpassing the participant counts reported in earlier thermopile localization studies. The accuracy of the localization algorithm is tested on participants whose data was not used for training ensuring that the reported performance is robust and generalizable. Unlike many prior studies that relied on only a few sensors in a single fixed layout, we evaluate multiple ceiling- and wall-mounted configurations. This provides us with insight about deployment in real-world environment.

We further demonstrate that the system can operate reliably over a substantially larger area than previously reported, highlighting the scalability and practical viability of thermopile-based localization. Together, these contributions establish a scalable thermopile-based sensing network that advances localization beyond the capabilities of prior single- or small-scale sensor deployments.

3.4 Sensor Development

3.4.1 Hardware

The STM32F415RGT6 micro-controller was selected as the MCU unit for this design. It has 1MB flash memory with a moderate mapped number of pins (64 pins), a processing speed of 168MHz, and a sufficient set of peripherals. The High-Speed External clock (HSE) was utilized with a 16 MHz external crystal. In addition, the Low-Speed External (LSE) clock was configured as a Real-Time Clock (RTC), enabling the device to operate in stop and standby modes with a 32.768 kHz crystal. Power is supplied through a typical 5 V USB bus. Since both the MCU and AMG8833 sensors require 3.3 V, a 5 V to 3.3 V regulator was used. Specifically, the AZ1084C low-dropout (LDO) linear regulator was selected due to its high current rating (5 A) and low dropout voltage (1.35 V). The regulator's larger TO-263-2 package facilitates efficient heat dissipation, ensuring stable operation under load. Table 3.2 lists the power consumption of active components.

Table 3.2: Current Estimation for All Essential Components of the Design [109, 110]

Unit	**Max. Current Consumption**	**Average Current Consumption**
MCU (STM32F415RGT6)	240mA	<120mA
AMG8833 Sensor	10mA	4.5mA
ESP32 Wi-Fi Module*	790mA	160–240mA

Note: *For future extension enabling wireless data communication.

3.4.2 Circuit Design

The MCU power supply is decoupled with a bank of 1 μF, 100 nF, and 4.7 μF capacitors connected in parallel [109]. The sensor circuit was designed following the manufacturer's reference schematic [110]. The corresponding MCU pin mapping is illustrated in Figure 3.2. In addition, the design incorporated reset switches for device ID configuration, power line polarity protection, and Electrostatic Discharge (ESD) safeguards.

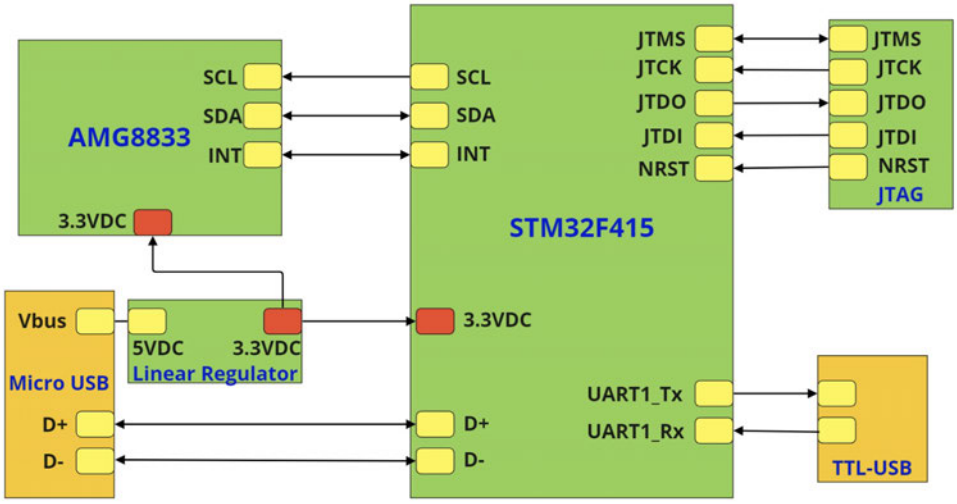


Figure 3.2: High-level hardware interface architecture showing the interconnections between the AMG8833 (Grid-Eye) thermopile sensor, STM32F415 microcontroller, power regulation, USB communication, and JTAG debugging interfaces.

3.4.3 PCB Layout

A four-layer PCB with a total thickness of 1.6 mm was designed using Altium Designer. The top layer carried low-speed signal traces, followed by a ground plane and a power plane, while the bottom layer was dedicated to high-speed signals. The crystal oscillator was placed in close proximity to the MCU pins to minimize signal delay and noise coupling [111]. Except for through-hole components, 0.8 mm vias were employed

for interlayer routing to simplify fabrication. Following fabrication and component assembly, the completed sensor module (Figure 3.3) was ready for firmware and software development.

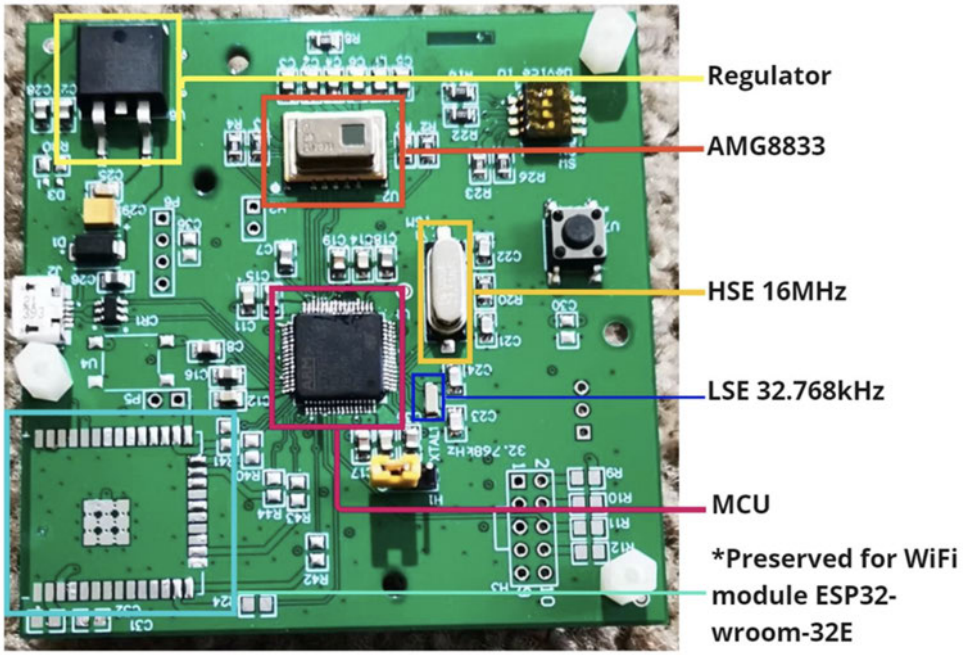


Figure 3.3: Annotated photograph of the custom-designed sensing board, showing the integrated AMG8833 (Grid-Eye) thermopile sensor, STM32 microcontroller unit, onboard voltage regulation, 16 MHz HSE and 32.768 kHz LSE crystal oscillators, and reserved footprint for future ESP32-WROOM-32E Wi-Fi module integration.

3.4.4 Firmware and Software

Figure 3.4 illustrates the embedded firmware data acquisition and communication flow for the sensor module. It consists of three major stages.

Initialization

The system initialization phase configures the core peripherals required for sensing, timing, and communication. System clock configuration was carried out using the high-level graphical interface by selecting external oscillators and re-scaling the parameters

to achieve 168 MHz for the system clock, 48 MHz for the USB clock and 42 MHz and 84 MHz for the peripheral clocks PCLK1. The RTC is also externally sourced using a 32.768 kHz crystal. The I²C interface was configured in master mode with 7-bit addressing at 400 kbps for communication with the AMG8833 sensor. The Universal Synchronous and Asynchronous Receiver-Transmitter (USART) was initialized for standard 115.2 kbps serial transmission, while the USB interface was set to virtual COM mode, enabling high-speed PC communication through a simplified serial protocol.

Main Function Loop

After initialization, Timer 3 is started in interrupt mode to ensure precise sampling intervals. Within an infinite loop, the system obtains the current timestamp from the RTC, reads 128 bytes of raw pixel data from the AMG8833 sensor via the I²C interface, then reads 2 bytes representing the ambient temperature from the AMG8833 (which can be used for temperature compensation), and introduces a short delay (10 ms) to maintain stable timing and processing.

Interrupt Routine

It handles data transmission, sending the acquired thermal and temperature data via USB or UART, ensuring non-blocking and timely communication. The sampling time interval for AMG8833 is 100 ms, but the PC oversamples the board at 25 Hz (40 ms) to minimize latency between new samples arriving at the GUI.

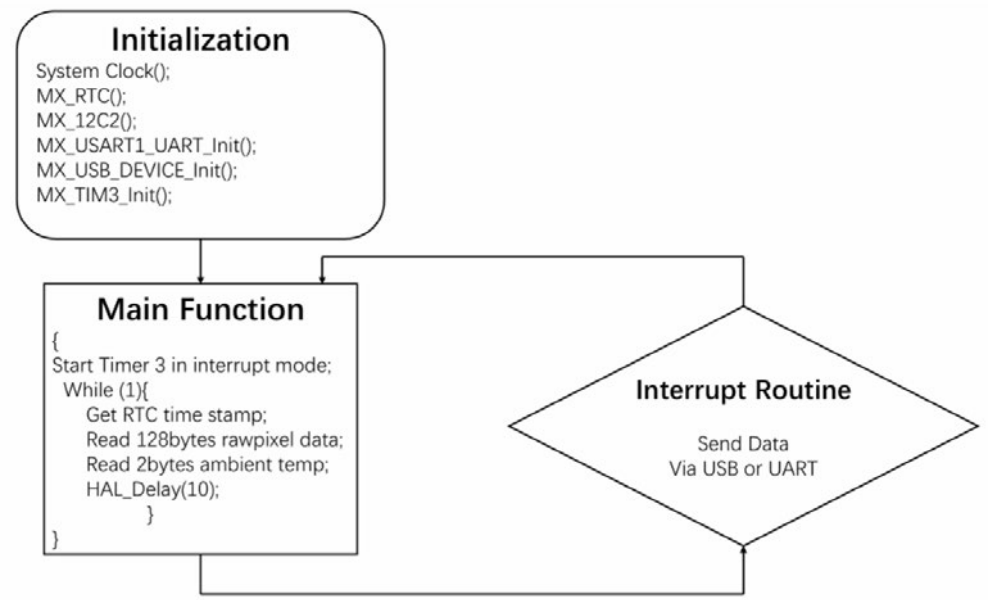


Figure 3.4: Flow diagram of the onboard data acquisition and communication process comprising three main stages

A Python GUI was created to acquire data from the sensor boards at 25 Hz over serial, log it to a csv file as shown in Figure 3.6, and display it graphically at 10 Hz, as shown in Figure 3.5.

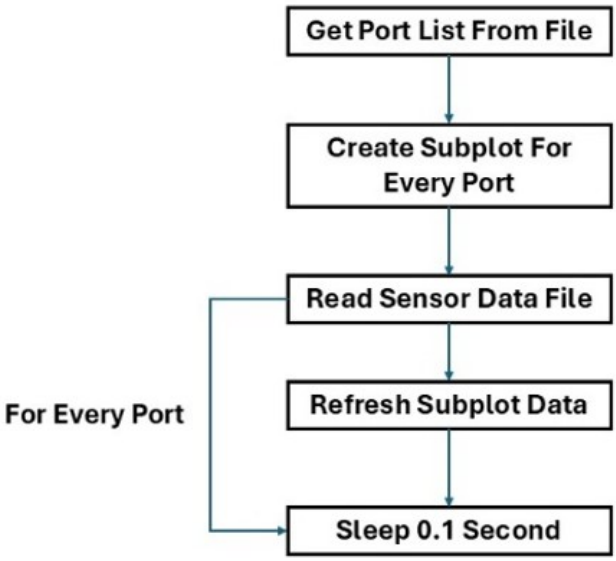


Figure 3.5: Flow diagram of the GUI display process. The program iterates through each sensor port, reads the latest data file, and refreshes the corresponding subplot at 10 Hz.

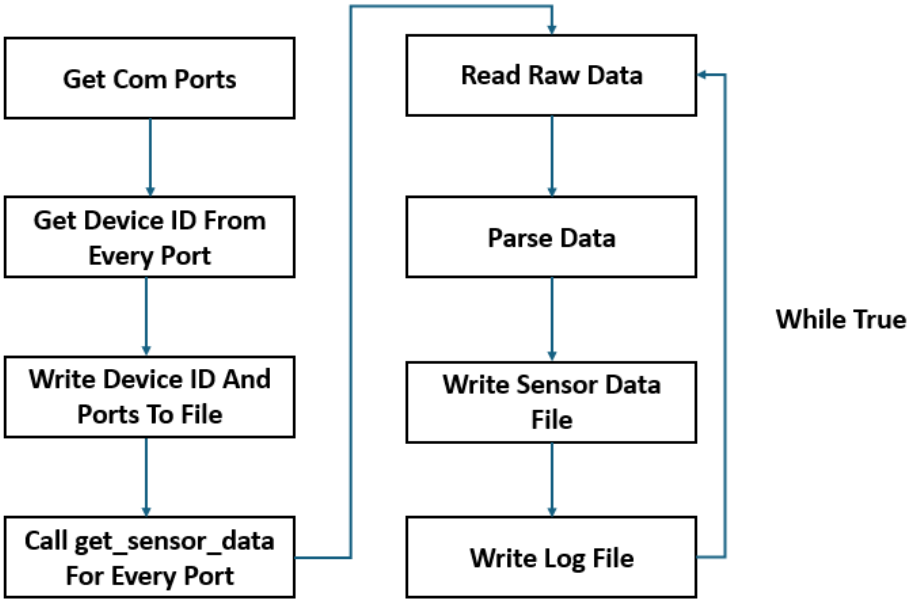


Figure 3.6: Program flow for data collection. It is used to create several independent sessions to collect data from multiple devices, write the data into separated log files, and also update the latest data frame from all the sensor devices to data files for display purposes

After completing the application development, the next step involved configuring the network. Prior to deployment, each sensor was individually tested. As all sensor data is processed on a central PC, a star-topology network was implemented. A USB 3.0 seven-port hub with an external power adapter was used to interface multiple sensor boards with the PC. To extend the communication range and enable wider coverage, a 10 m active USB cable was employed between the PC and the hub [66]. Each sensor device was then connected to the hub using 5 m USB cables.

3.5 System Setup

3.5.1 Sensor Arrangements

Six thermopile sensors were deployed to capture localization data in a laboratory environment. The sensor layout and physical setup are shown in Figure 3.7. Two sensors (C_1 and C_2) were mounted on a truss at a height of 2.3 m, emulating ceiling-mounted units in a real deployment. The remaining four sensors (W_1 – W_4) were placed on stands at a height of 1.3 m, representing wall-mounted sensors positioned around the perimeter of a room. The number of thermopiles used in this setup is likely much higher than what would be practical in real-world deployments. This allows to assess how many units are needed for adequate performance and to identify preferred sensor placement. Results from the two best performing 4-sensor arrangements and eight best 2-sensor arrangements are presented. These combinations are:

- **Comb1, Comb2:** Four-sensor configurations comprising two ceiling-mounted sensors and one wall-mounted sensor on the left wall (W_4). For Comb1, the remaining sensor is mounted on the right wall (W_2), whereas for Comb2 it is mounted on the back wall (W_3).
- **Comb3:** Two overhead sensors (C_1 and C_2).

- **Comb4–Comb7:** Two-sensor configurations in which both sensors are wall-mounted. Specifically, Comb4 uses W_4 and W_2; Comb5 uses W_1 and W_3; Comb6 uses W_4 and W_3; and Comb7 uses W_2 and W_1.

- **Comb8–Comb10:** Comprising one overhead sensor and one wall-mounted sensor; respectively (C_1, W_3), (C_2, W_1) and (C_1, W_4).

The coverage area for these sensor arrangements is $2.4\,\text{m} \times 2.4\,\text{m}$, which, although modest, is consistent with prior studies. In addition, we conducted a large-scale deployment spanning $28.2\,\text{m}^2$, representing a substantial increase in coverage compared with earlier thermopile-based studies. This expanded coverage is enabled by the capabilities of the proposed bespoke sensor network. Four thermopile sensors were installed on the ceiling of an office foyer at a height of $2.45\,\text{m}$. To ensure coverage and eliminate blind spots, adjacent sensor footprints were intentionally overlapped. The resulting layout formed a uniform 2×2 grid, creating a square sensing array covering approximately $28.2\,\text{m}^2$ (see Figure 3.9). This configuration is hereafter referred to as the *foyer setup*. Given the room dimensions, a mobile target could move beyond the effective detection range of wall-mounted thermopiles. Operational and health-and-safety restrictions precluded the use of stands for mounting the sensors. Therefore, ceiling-mounted installation was adopted as the only feasible configuration.

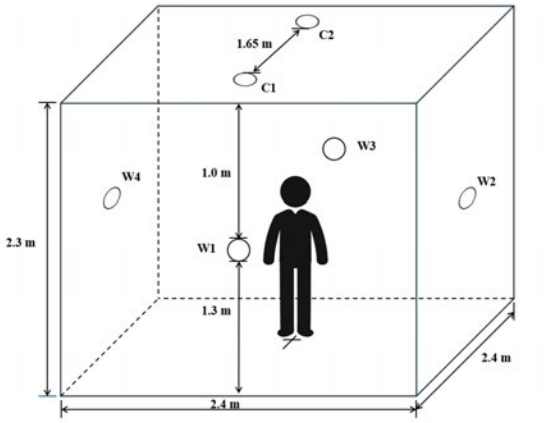


Figure 3.7: The sensor layout for the lab setup. Two ceiling-mounted sensors (C_1, C_2) are installed at 2.3 m height with 1.65 m spacing. Four wall-mounted sensors (W_1–W_4) are placed at 1.3 m height around the perimeter. The coverage area is $2.4\,\text{m} \times 2.4\,\text{m}$.

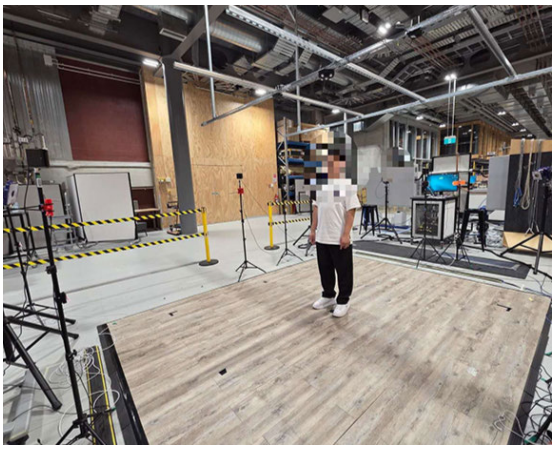


Figure 3.8: Photograph of the lab setup showing a participant standing within the sensing area. The thermopile sensors are mounted on tripod stands (wall-mounted) and on the overhead truss (ceiling-mounted). The participant's face is anonymized for privacy.

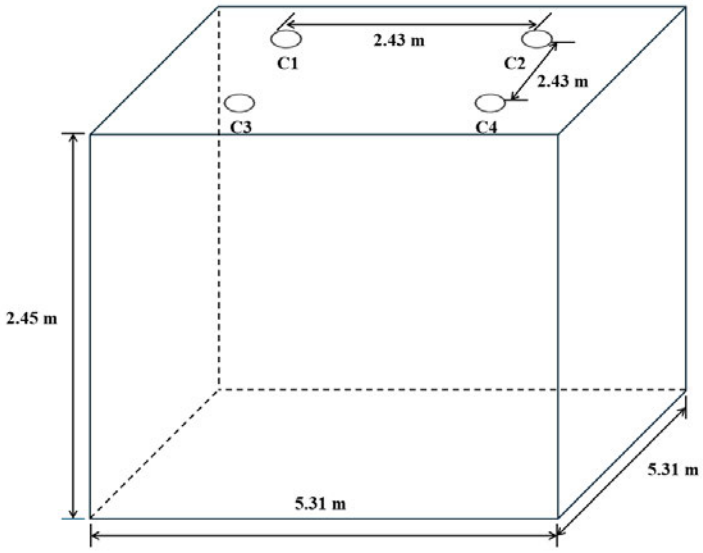


Figure 3.9: Layout of the 4 ceiling mounted sensors (C_1–C_4) for the foyer setup at a height of 2.45 m. The coverage area is approximately 28 m^2.

3.5.2 Data Collection

For the lab setup, data were collected from eight healthy participants with ages ranging from 22 to 38 years and heights varying from 1.65 m to 1.86 m. All participants provided informed consent. Figures 3.8 and 3.10 show a participant standing within the sensing area and the corresponding thermal data captured by the six thermopile sensors respectively. Each participant was instructed to walk two predefined trajectories: a spiral and a zigzag. Examples of these paths, as captured by the ground-truth system, are shown in Figures 3.11 and 3.12. To ensure sufficient variability for training and evaluation of the ML models, each participant completed five repetitions of both paths.

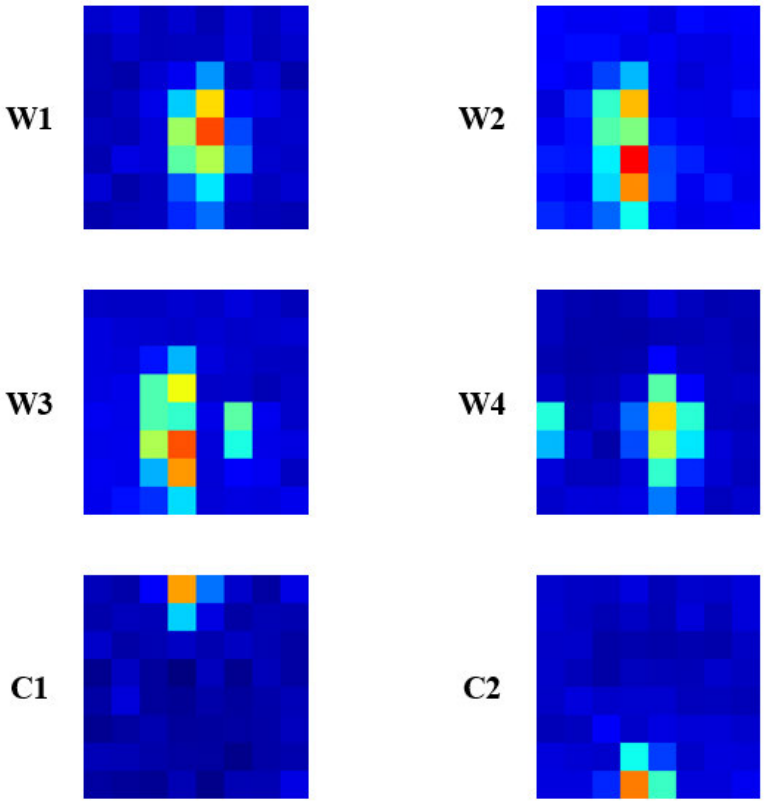


Figure 3.10: Thermal data captured by the six thermopile sensors (W_1–W_4: wall-mounted; C_1, C_2: ceiling-mounted) when a participant is standing within the sensing area (as shown in Figure 3.8).

For the foyer setup, a 1.65 m tall male volunteer walked five continuous circles centered under each of the four sensors. The actual trajectories as recorded by the ground truth system are shown in Figure 3.13.

To provide ground truth reference for evaluating the system's positioning accuracy, a high-precision ultrasonic sensor system, the Marvelmind Indoor Navigation System [112], was deployed. This system consisted of four ultrasonic beacons placed at the corners of the test area. The ground truth system sampled data at a frequency of 10 Hz to ensure accurate capture of the subject's movements. Strict temporal synchronization between the thermopile sensor array and the ultrasonic ground truth system is essential.

This was achieved through a bespoke Python-based control application that simultaneously initialized and triggered data acquisition across both systems. The recorded ground-truth data were then resampled to a common 10 Hz time base to align with the thermopile data. Linear interpolation, similar to that used for synchronizing the thermal data (please see Section 3.5.3), was applied. However, in this case, the interpolated quantities were the (x, y) positional coordinates rather than the pixel values.

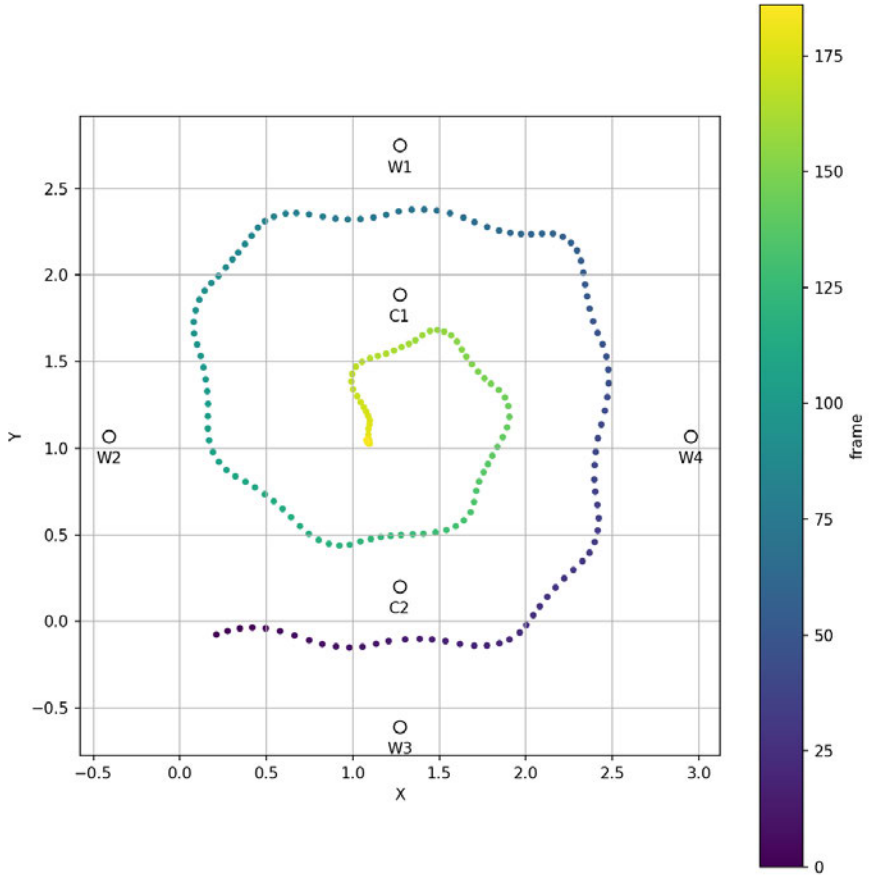


Figure 3.11: Example spiral trajectory of a participant in the laboratory setup, recorded by the ground-truth system. C_1, C_2 are ceiling-mounted sensors, and W_1–W_4 are wall-mounted sensors.

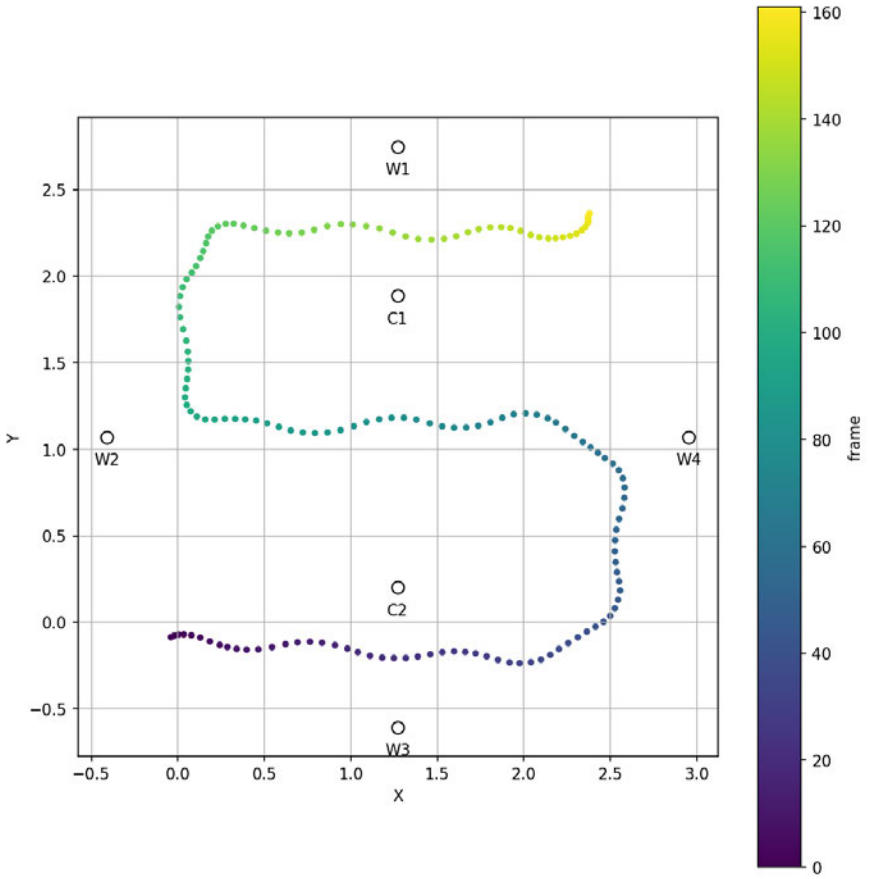


Figure 3.12: Example zigzag trajectory of a participant in the laboratory setup, recorded by the ground-truth system. C_1, C_2 are ceiling-mounted sensors, and W_1–W_4 are wall-mounted sensors.

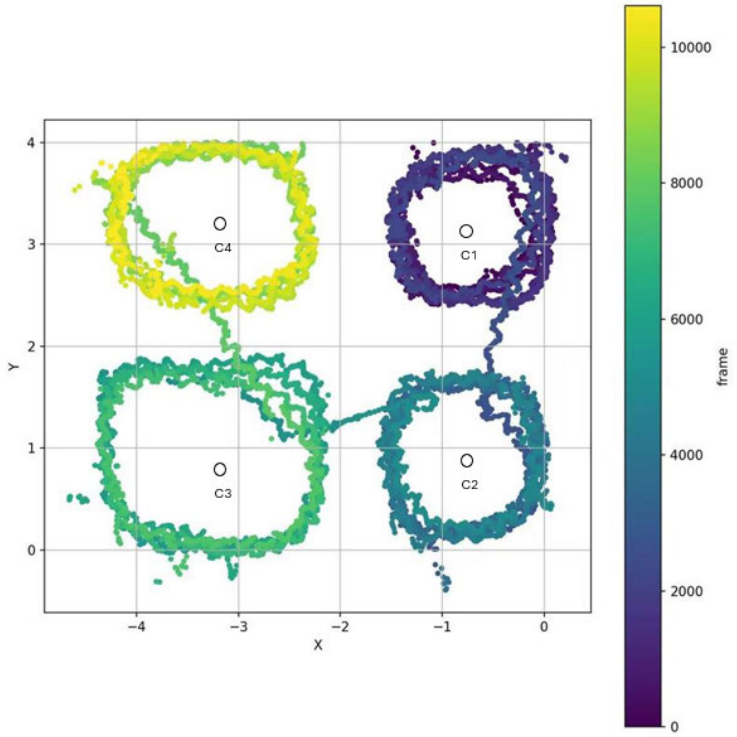


Figure 3.13: Circle trajectory for the foyer setup as captured by the ground-truth system with four ceiling-mounted sensors (C_1–C_4).

3.5.3 Data Processing

Raw data from each thermopile sensor are first background-corrected and then temporally synchronized to a common time base before being used for feature extraction and learning.

Background subtraction

Background temperature is subtracted from the raw sensor data to isolate action-specific thermal signatures. This is a standard technique in computer vision and is commonly used for thermopile-based localization [66].

For each thermopile sensor i and pixel (p, q) the background thermal map is estimated by averaging over M frames captured when the monitored area is empty so that

$$\bar{b}_i^{p,q} = \frac{1}{M} \sum_{m=1}^{M} b_i^{p,q}(m) \tag{3.1}$$

Here $b_i^{p,q}$ denotes the raw reading at pixel (p, q) during the background calibration period. $p \in \{1, 2, ..., 8\}$ and $q \in \{1, 2, ..., 8\}$ are the row and the column indices of the pixel. Before the trials commenced, background data were collected for 10 minutes from the unoccupied experimental area. At a sampling rate of 10 Hz, this yielded a total of 6,000 frames ($M = 6000$).

The background subtracted pixel value is then computed as

$$y_i^{p,q}(t_{i,k}) = z_i^{p,q}(t_{i,k}) - \bar{b}_i^{p,q} \tag{3.2}$$

Here, timestamp for the i^{th} thermopile is i: $\{t_{i,k}\}_{k=1}^{K_i}$ and $z_i^{p,q}$ is the raw temperature value captured at (p, q) pixel.

Synchronization

Due to minor jitter, the data streams from the thermopile sensors are not perfectly aligned in time. Temporal alignment is achieved by resampling the outputs of the thermopile sensors (each producing 8×8 frames at approximately 10 Hz) to a common 10 Hz time base. Because individual sensors can produce frames with slightly different timestamps, a global time domain $[T_{min}, T_{max}]$ is defined, and a uniformly spaced time vector, t_n, is generated so that $t_n = T_{min} + n\Delta t$ with $n = 0, 1, 2, ..., N$ and $\Delta t = 0.1s$. For each t_n within the original timestamp of $t_{i,k} \leq t_n \leq t_{i,k+1}$, the time aligned pixel is generated as

$$\hat{y}_i^{p,q}(t_n) = [1 - w_{i,k}^{p,q}(t_n)] y_i^{p,q}(t_{i,k}) + w_{i,k}^{p,q}(t_n) y_i^{p,q}(t_{i,k+1}) \tag{3.3}$$

Here $w_{i,k}^{p,q}(t_n)$ is the linear interpolation weight defined as

$$w_{i,k}^{p,q}(t_n) = (t_n - t_{i,k})/(t_{i,k+1} - t_{i,k}). \tag{3.4}$$

At the temporal boundaries of the data, bounded linear extrapolation is applied to maintain signal continuity, with values clipped to the observed range. The resultant synchronized dataset of 8×8 frames from all thermopiles is used for feature extraction and subsequent regression for localization.

3.5.4 Machine Learning

In this study, machine learning is employed to address the localization task through a regression framework. During the training phase, the regression model was trained on paired datasets consisting of thermopile sensor measurements and ground truth position coordinates. This enabled the model to learn the functional mapping between thermal signatures and actual spatial locations. Once trained, the model was used in the inference stage to estimate the target's position from new thermopile measurements that were not used during training. A 2D CNN–LSTM architecture was selected to capture both spatial and temporal features. Spatial features arise from location-dependent pixel distributions. Temporal features arise from changes in pixel values as the target moves. Similar architecture has also been proven to be effective in prior thermopile-based localization studies [103]. The architecture of the proposed 2D CNN–LSTM network is illustrated in Figure 3.14.

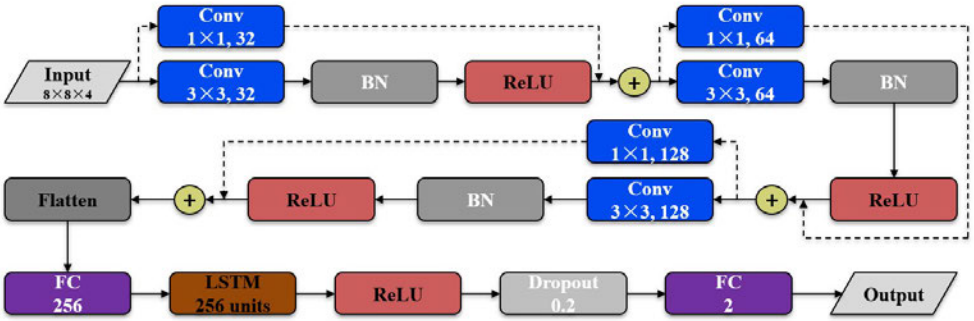


Figure 3.14: Architecture of the proposed 2D CNN–LSTM network. The network consists of three convolutional blocks with skip connections (showed using dotted lines) followed by a fully connected layer, an LSTM layer with 256 units, and a final output layer for position estimation. Each convolutional block includes batch normalization (BN) and ReLU activation.

The algorithms were trained on a computer with an AMD 9950x CPU and an NVIDIA RTX PRO 6000 Blackwell Max-Q GPU. A Bayesian optimizer was used to identify a network structure suitable for both datasets and related hyperparameter tuning. For the laboratory setup, data from eight participants were split into 4-2-2 training, validation, and test sets. This ensures that the algorithms are tested for *unseen users* whose data was not used for training. For the foyer setup, a 60%-20%-20% train/validation/test split was used. All data within each split were segmented into 1-s sequences (10 samples) for training.

The optimizer evaluated 120 network configurations spanning the hyperparameter ranges in Table 3.3. Each network was trained for up to 200 epochs, with early stopping applied when the validation error ceased improving for at least 30 epochs, returning the network with the lowest validation error. Networks were trained independently on the laboratory and foyer datasets. The objective function was the mean squared 2D Euclidean localization error, averaged across both environments and minimized by the optimizer.

Following optimization, the top five candidate architectures were evaluated using three-fold cross-validation on the full dataset. The best-performing structure is reported

in Table 3.3. This procedure provides a subject-independent performance estimate and assesses generalization to unseen individuals.

All optimization and structure selection were performed using four sensors in each environment to provide a structure that can work in both environments. Since the two environments differed in sensor density, the models were retrained in the laboratory setup using only two sensors to assess the impact of a sparser, and more realistic, deployment. All reported results use the hyperparameter values used in Table 3.3. Preliminary observations indicate that, for the two-sensor laboratory setup, reducing the LSTM size from 256 to 64 neurons does not noticeably affect performance. This suggests potential for future optimization of network size under sparser sensor configurations.

Two of the people in the lab setup recorded data both wearing jackets, and without jackets. The original hyperparameter optimization, and later 3 fold results included this data to fairly represent general performance under varying clothing conditions.

Table 3.3: Hyperparameter Search Space and Final Selected Values for the 2D CNN–LSTM Model

Algorithm	**Hyperparameter Range**	**Final Hyperparameter Values**
2D CNN–LSTM	CNN Section depth: [1, 4]	CNN Section depth: 3
	Filtersize: [2, 4]	Filtersize: 3
	Number of filters (1): $2^{[3,8]}$	Number of filters: 32
	Number of filters (2): $2^{[3,8]}$	Number of filters: 64
	Number of filters (3): $2^{[3,8]}$	Number of filters: 128
	Spatial Attention: [yes, no]	Spatial Attention: no
	Channel Attention: [yes, no]	Channel Attention: no
	Skip Connections: [yes, no]	Skip Connections: yes
	Pooling: [yes, no]	Pooling: no
	Pre-LSTM FC neurons: $2^{[3,10]}$	Pre-LSTM FC neurons: 256
	LSTM Section depth: [1, 2]	LSTM Section depth: 1
	LSTM neurons: $2^{[5,10]}$	LSTM neurons: 256
	Dropout: [0, 0.5]	Dropout: 0.2

3.6 Results and Discussion

The localization results for both setups are shown in Table 3.4. In this thesis, localization performance is evaluated using multiple statistical metrics, primarily the ***median error*** and ***Root Mean Squared Error (RMSE)***. While many existing studies only report one of these metrics (as summarized in Table 3.1), both are necessary for a comprehensive evaluation. The median error is utilized to represent the typical localization accuracy, as it provides robustness against statistical outliers. Conversely, RMSE is included due to its heightened sensitivity to large positional deviations, which may indicate underlying spatial coverage or Signal to Noise Ratio (SNR) issues. Consequently, two systems could achieve the same median error while exhibiting different RMSE values. When analyzed in conjunction with the 90th and 95th percentile errors ($\boldsymbol{P_{90}}$ and $\boldsymbol{P_{95}}$), these combined metrics offer a comprehensive assessment of the system's performance. For the lab setup, the sensor combinations are difficult to differentiate based on median errors alone. However, the RMSE, P_{90}, and P_{95} metrics indicate that the four-sensor configurations (Comb1 and Comb2) and two of the two-sensor configurations (Comb3 and Comb4) achieve the best performance. The remaining configurations exhibit a higher incidence of large localization errors, likely attributable to the inferior spatial coverage provided by those sensor arrangements.

Table 3.4: Localization Results for Foyer and Lab Setup

	Four-sensor setup			Two-sensor setup							
Metric	**Foyer**	**Comb1**	**Comb2**	**Comb3**	**Comb4**	**Comb5**	**Comb6**	**Comb7**	**Comb8**	**Comb9**	**Comb10**
RMSE (m)	0.46	0.29	0.31	0.34	0.30	0.37	0.35	0.35	0.43	0.34	0.37
Mean (m)	0.35	0.24	0.25	0.26	0.25	0.29	0.29	0.27	0.33	0.27	0.30
Median (m)	0.27	0.20	0.21	0.21	0.21	0.23	0.24	0.22	0.26	0.21	0.25
P_{90} (m)	0.71	0.42	0.45	0.48	0.48	0.55	0.54	0.50	0.64	0.49	0.58
P_{95} (m)	0.91	0.51	0.54	0.55	0.56	0.71	0.71	0.63	0.77	0.64	0.73

Note: All localization error statistics are reported in meters. Comb1 and Comb2 are four-sensor configurations; Comb3–Comb10 are two-sensor configurations.

The results presented in Tables 3.1 and 3.4 can be used to benchmark our proposed approach against earlier thermopile-based localization studies. While the reported accuracy of the proposed method for the lab setup appears slightly worse than values reported in some prior works, it is important to note that, with the exception of [66, 103], these studies do not report performance for unseen users. Consequently, their ability to generalize cannot be reliably assessed. It is likely that the performance of these systems would degrade when evaluated on users whose data were not included during training. Subsequent benchmarking results presented in Section 3.6.1 further show that the best-performing algorithm reported in [103] performs slightly worse than the proposed approach when evaluated under our experimental setups.

Considering the Foyer setup, we can observe a clear advancement in the scalability for localization low-resolution thermopile sensors. Previous works mostly relied on a single ceiling-mounted sensor (e.g., Kuki et al. [101], Tariq et al. [61]) or small multi-sensor configurations with overlapping FoV limited to less than 12 m^2 (Faulkner et al. [66]). In contrast, we employ four sensors serving more than twice the coverage of the next largest system. Despite its large coverage, the proposed system maintains high accuracy.

3.6.1 Experimental Benchmarking

In this section, we evaluate the approach proposed by Faulkner et al. [103] using our experimental setup.

Two sensor combinations

Table 3.5 benchmarks the proposed approach against the CNN–LSTM method reported by Faulkner et al. [103], which was the best-performing model in their study. For a fair comparison, the network was retrained using our dataset for the Comb3 and Comb4

sensor configurations, employing identical training, validation, and test splits as the proposed method.

For the ceiling-mounted Comb3 configuration, Faulkner et al.'s approach achieves a lower median error and a marginally lower mean error. However, the proposed method outperforms it across all remaining metrics. In particular, the substantially lower P_{90} and P_{95} errors indicate a comparatively reduced occurrence of large localization errors for our proposed approach. For the wall-mounted Comb4 configuration, the proposed approach consistently outperforms Faulkner et al.'s method across all evaluated metrics, again demonstrating markedly improved robustness by suppressing large errors.

Table 3.5: Benchmarking of two sensor arrangements in laboratory setup. All localization error statistics are reported in meters.

	Comb3		**Comb4**	
Metric	**Proposed**	**Faulkner et al. [103]**	**Proposed**	**Faulkner et al. [103]**
RMSE	**0.34**	0.37	**0.30**	0.39
Mean	0.26	**0.25**	**0.25**	0.31
Median	0.21	**0.17**	**0.21**	0.23
P_{90}	**0.48**	0.50	**0.48**	0.66
P_{95}	**0.55**	0.83	**0.56**	0.85

Note: For each metric and sensor arrangement, the best (lowest) error between the proposed method and Faulkner et al. [103] is highlighted in bold.

Four sensor combinations

Table 3.6 presents a corresponding benchmark for four-sensor configurations. As shown in the table, the proposed approach consistently achieves the best performance across all reported statistical metrics. Table 3.6 further shows that the proposed approach achieves significantly superior performance in the large-scale (foyer) setup.

It is worth noting that Faulkner et al.'s algorithm performed far worse in our two-

and four-sensor setups than in their original single-sensor system, where they reported a median error of 0.036 m. This highlights a key limitation of prior studies: their results may not reflect performance in larger, real-world spaces or in multi-sensor setups. Our experimental study addresses this gap through a multi-sensor system deployed also over a large coverage area.

We also implemented the neural network proposed by Tariq et al. [61]. Unfortunately, it performed poorly in our setting with median error well over a meter in both setups. This is likely because the network was developed for 4×4 thermal images captured by a single sensor. Therefore, it may have lacked the depth, capacity, and architectural components needed to extract discriminative features from the more complex inputs produced by our system. The models developed in [66] require a substantially larger amount of training data and therefore cannot be fairly trained or evaluated using our dataset. Unfortunately, other studies do not provide sufficient architectural details to enable reliable re-implementation and retraining for our experimental setup.

Table 3.6: Benchmarking of four sensor arrangements. All localization error statistics are reported in meters.

	Comb1		**Comb2**		**Foyer**	
Metric	**Proposed**	**Faulkner et al. [103]**	**Proposed**	**Faulkner et al. [103]**	**Proposed**	**Faulkner et al. [103]**
RMSE	0.29	0.38	0.31	0.40	0.46	0.56
Mean	0.24	0.27	0.25	0.31	0.35	0.41
Median	0.20	0.21	0.21	0.23	0.27	0.32
P_{90}	0.42	0.56	0.45	0.66	0.71	0.74
P_{95}	0.51	0.75	0.54	0.85	0.91	1.14

Note: Best (lowest) error is highlighted in bold.

3.6.2 Benchmarking Against Other Modalities

Localization performance reported across different sensing modalities varies widely in terms of accuracy and coverage (see Table 3.7). Performance spans from capacitive methods that achieve high accuracy within limited spatial areas to RF-based approaches that provide wider coverage at the expense of increased localization error. The thermopile-based system proposed in our work achieves a balanced compromise between these extremes. It maintains adequate spatial accuracy while supporting a practical sensing range suitable for real-world implementation.

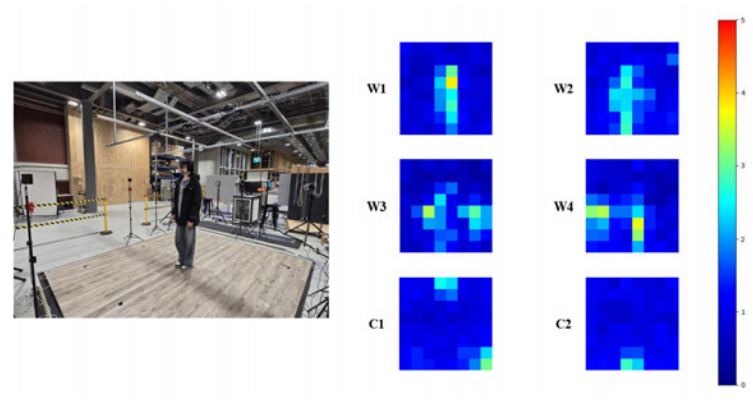
Table 3.7: Comparison of Localization Systems Across Common Sensing Modalities

Author	**Mode**	**ROI (sqm)**	**Accuracy (meter)**
WoW [46]	Visible Light	6.8	Median: 0.12
Fieldlight [47]	Visible Light	28–56	Median: 0.68–1.2
DeepPIRATES [113]	PIR	25–65	Mean 0.37–0.54
Ngamakeur et al. [52]	PIR	9	Mean: 0.2379
SpringLoc [27]	RF (RSSI)	25–44	Median: 0.6–1.2
Hi-loc [97]	RF (CSI)	120	Mean: 0.65–1.03
Mirshekari et al. [44]	Vibration	20	Median 0.38
Alajlouni et al. [114]	Vibration	26	80 percentile: 0.7
CapLoc [40]	Capacitive	2.9	Median: 0.026
Tariq et al. [41]	Capacitive	3.24	Mean: 0.307, RMSE: 0.251
Proposed	Thermopile	28.2	Median: 0.27, Mean: 0.35, RMSE: 0.46

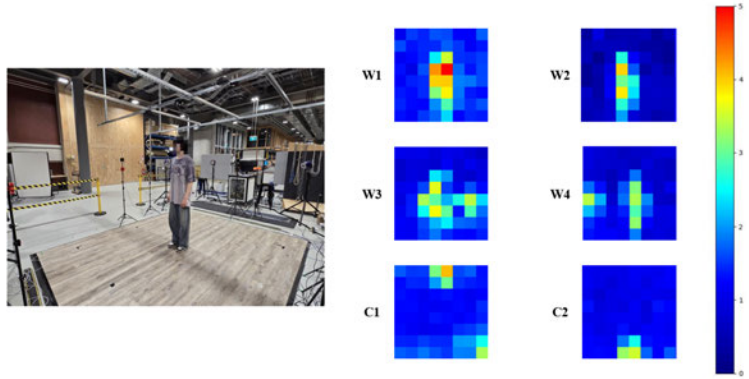
3.6.3 Impact of Clothing

Thermopile sensors operate by capturing the IR emission of the human body. Thick clothing can attenuate this emission, thereby degrading sensor measurements and localization accuracy. In climate-controlled environments, large variations in clothing are unlikely. However, this assumption does not hold universally. It is therefore important to investigate the impact of clothing on localization performance.

As noted in Section 3.5.4, two participants in the lab setup contributed data both while wearing jackets and without jackets. Figure 3.15 illustrates the difference in thermal signatures captured by the sensors when a participant is wearing a jacket versus without one. The attenuation of IR emission due to the jacket is clearly noticeable in terms of reduced thermal contrast.



(a) Participant wearing a jacket



(b) Participant without a jacket

Figure 3.15: Comparison of thermal signatures captured by six thermopile sensors (W_1–W_4: wall-mounted; C_1, C_2: ceiling-mounted) for the same participant with and without a jacket. The colorbar indicates temperature difference from background in °C. Wearing a jacket attenuates the IR emission, resulting in lower thermal contrast.

To investigate the impact of clothing on localization performance, the final network was retrained using data from the six participants without jackets as the training set, with no separate validation set. The two participants who had data available in both clothing conditions were withheld from training and used exclusively for testing. This evaluation was conducted using datasets collected under the Comb3 and Comb4 sensor configurations, as well as the best-performing four-sensor arrangement Comb1. The resulting performance differences for unseen participants who were either wearing a jacket or not are summarized in Table 3.8.

As shown, all three configurations exhibit only minor degradation in median and mean localization errors when the test subject is wearing a jacket. However, heavy clothing like jacket attenuates IR radiation from the human body. This reduces the thermal contrast captured by the thermopile sensors, particularly as the sensor–target distance increases. Therefore, such clothing has a pronounced effect on the tail of the error distribution, with the P_{90}–P_{95} metrics increasing substantially across all sensor configurations. This indicates that while typical localization accuracy is largely preserved, wearing a jacket significantly increases the likelihood of large localization errors.

Table 3.8: Impact of clothing on localization performance. All localization error statistics are reported in meters.

	Comb1		**Comb3**		**Comb4**	
Metric	**Jacket On**	**Jacket Off**	**Jacket On**	**Jacket Off**	**Jacket On**	**Jacket Off**
RMSE	0.33	0.26	0.42	0.29	0.35	0.27
Mean	0.27	0.21	0.33	0.21	0.29	0.22
Median	0.21	0.18	0.26	0.18	0.25	0.19
P_{90}	0.52	0.40	0.62	0.36	0.57	0.42
P_{95}	0.62	0.49	0.73	0.45	0.70	0.46

Note: Comb1 is a four-sensor configuration; Comb3 and Comb4 are two-sensor configurations.

3.7 Conclusion and Future Work

This paper presented the design, implementation, and evaluation of a bespoke thermopile-based indoor localization system using low-resolution AMG8833 Grid-Eye sensors. A functional sensing network comprising up to six sensors was successfully deployed, and localization was performed using a 2D CNN–LSTM regression model.

3.7.1 Summary of Findings

For the lab setup covering $2.4\,\text{m} \times 2.4\,\text{m}$, the best performance was achieved with the four-sensor configurations (Comb1 and Comb2), yielding median errors of 0.20 m and 0.21 m, respectively. Among the two-sensor configurations, the ceiling-mounted arrangement (Comb3) and the opposite wall-mounted arrangement (Comb4) achieved comparable performance with median errors of 0.21 m. The proposed method consistently outperformed algorithms proposed in earlier research, particularly in reducing large localization errors as indicated by improved P_{90}–P_{95} metrics.

The impact of clothing on localization performance was investigated. While typical localization accuracy (median and mean errors) remained largely preserved when participants wore jackets, heavy clothing significantly increased the likelihood of large localization errors due to attenuated IR radiation from the human body. This finding is important for practical deployments in environments where clothing variations are expected.

For the large-scale foyer setup, localization was performed with four ceiling-mounted sensors covering an area of approximately 28.2 m^2, which is significantly larger than any prior thermopile-based study. A median error of 0.27 m was achieved with the CNN–LSTM network. The benchmarking results demonstrated that algorithms proposed by others for single-sensor setups suffer significant degradation when applied to larger multi-sensor deployments, highlighting the importance of our multi-sensor

approach.

Compared to other sensing modalities, the proposed thermopile system achieves a balanced trade-off between localization accuracy and coverage area. The system offers advantages including privacy preservation, device-free operation, and robustness to lighting conditions.

3.7.2 Future Work

Several directions for future research have been identified. First, the current implementation addresses single-person localization only. Future studies can extend the proposed approach to handle multi-target scenarios, which requires addressing the additional challenges of thermal signature overlapping and multi-target data association.

Second, the system was evaluated in two environments. Future research could evaluate the performance across a broader range of scenarios with varying environmental conditions, furniture layouts, and participant demographics to validate the generalization capability more thoroughly.

Third, connectivity improvements are possible. The current USB-based setup has a maximum cable length of approximately 5 m [115]. Ethernet-based connectivity could extend this to 100 m [116], while wireless networking using the ESP32 Wi-Fi module (already provisioned on the PCB) could enable even more flexible deployments.

Finally, the developed thermopile system can be extended to other human sensing tasks. Our preliminary results show 100% occupancy detection accuracy. Future studies can explore occupancy estimation (counting multiple individuals) and activity recognition, areas where low-resolution thermopile-based research remains limited.

Chapter 4

Fall Detection Using Multiple Low-Resolution Thermopile Sensors and Machine Learning

4.1 Abstract

Automated fall detection enhances the safety of older adults and individuals with health conditions that make them prone to falling. One promising option is based on the detection and measurement of the Infrared (IR) radiation emitted by people. Thermopiles are IR sensors that have received significant interest from researchers in recent years for automated fall detection in a privacy-preserving manner. This study aims to develop and evaluate a fall detection system using the data captured by multiple ceiling and wall-mounted thermopile sensors with overlapping coverage. Unlike prior studies that use only a small number of thermopile sensors, this work systematically investigates 16 different ceiling- and wall-mounted sensor configurations. Three machine learning classifiers are trained with the data collected from 5 participants enacting fall poses and other activities like standing and sitting. The accuracy of the fall

detection algorithms is tested on additional 5 participants whose data was not used for training. The accuracy of algorithms varies between 92.50% and 99.20% and depends on the number of sensors and their respective locations. We utilize this information to assess how many sensors are required for adequate performance and also to provide insight into the desired locations of the sensors. Our experimental results indicate that combining one ceiling-mounted and one wall-mounted sensor achieves the optimal balance between accuracy, cost, and implementation challenges.

4.2 Introduction

Falls present a significant health risk for older adults and individuals with mobility limitations, chronic health conditions, impacting both their physical health and quality of life. They also place a considerable burden on healthcare systems and aged-care facilities. A significant percentage of falls are fatal or can become fatal within 12 months [117]. In addition to serious detriment to health, fear of falling and the loss of independence can lower the quality of life [118]. Automated fall risk prediction and proactive prevention remain an aspirational goal of fall research. Unfortunately, no robust solution currently exists. The risk of suffering a future fall is currently estimated through clinical assessment tests [119] in hospitals/laboratories incurring cost and inconvenience for the subject and the health services. Several studies (e.g., [120]) have explored wearable airbag technologies that deploy upon detecting pre-impact falls. This can lessen the impact and thereby reduce the severity of the injury. Commercially available products exist [121, 122]. Their usage remains confined to pilot programs [123] with ongoing research into their acceptability [124]. However, they have yet to see broad endorsement from the health professionals. Recent findings also highlight the need to consider possible adverse outcomes [125]. Ongoing research is also exploring the use of pre-impact detection to enable robotic exoskeletons to arrest

falls [126]. However, as with wearable airbags, their acceptability and efficacy require further investigation [127]. Time spent lying on the floor often determines the length of hospital stay, and subsequent permanent institutionalisation [128]. Consequently, the pragmatic and widely accepted approach continue to be timely fall detection followed by prompt assistance.

Development of automated fall detection systems has benefited greatly from the rapid advancements of the IoT, Machine Learning, and technologies for human activity sensing [91]. Figure 4.1 shows various automated fall detection approaches that have been proposed in the literature [129]. Camera based systems [130] does not require users to wear a device. However, they are perceived to be invasive of privacy. They also require adequate lighting, which is challenging as many falls occur at night [131]. On the other hand, wearable devices are often not worn at night [128, 129, 131], which limits applicability for users with nocturia. Some require client activation after a fall, which may not be possible if the victim loses consciousness, is in a confused state, or has cognitive impairments. Automated fall detection research needs to address these issues.

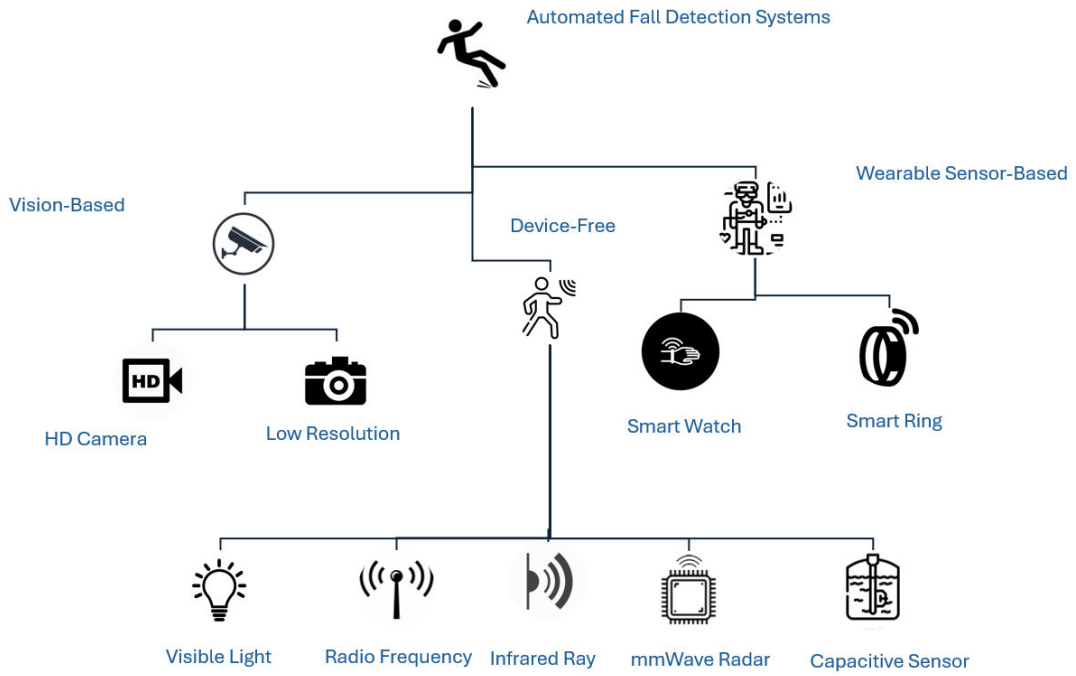


Figure 4.1: Sensing modalities for automated fall detection found in the literature.

A method proposed in [132] leveraged on the fact that human activities impact radio wave propagation. It makes use of mmWave signals at high frequencies (typically at 60-64GHz or 76-81GHz) and detect their reflections off objects by the receiver. Machine learning models are trained on the data extracted from the reflected signal to classify fall events and distinguish them from other activities [133]. The advantage of using mmWave is that it is unaffected by changes in illumination and it is privacy-preserving. However, they are sensitive to changes in the environment, such as rearranged furnitures [134]. Furthermore, the equipment tends to be relatively expensive. A less costly way is to use ISM band radios such as Wi-Fi, and Zigbee. The RSSI or CSI from Wi-Fi systems could be used to detect falls [135]. However, such signals are vulnerable to interference from other wireless networks and to changes in the surroundings. Alternative approaches make use of pressure [136], capacitance [40], electric field [137] and vibration [138] sensing on the floors. While these systems preserve privacy and

work independently of lighting conditions, their installation and maintenance costs are relatively high. They also require direct physical contact with the floor.

A promising approach is to sense the thermal IR radiation emitted by humans. An important aspect of IR sensing is its ability to function reliably regardless of lighting conditions. Thermopile sensors, often used for non-contact temperature measurements through IR sensing, have been proposed for activity recognition and fall detection [139].

Thermopiles are thermocouple-based IR sensors [140] that are widely used to construct thermal cameras. The number of sensing elements would determine the resolution of the thermal image. High-resolution thermal cameras are comparatively costly and may also be perceived as privacy invasive. However, low-resolution thermopiles, due to their privacy-preserving nature, have become popular for human sensing applications [141]. A lower resolution thermopile also has a smaller number of sensing channels (pixels), which reduces the computational burden.

Unlike prior studies that use only a small number of thermopile sensors, our work systematically investigates 16 different ceiling- and wall-mounted sensor configurations. This enables a deeper understanding of how sensor count and placement affect fall detection accuracy, providing practical guidance for robust system deployment.

Most existing studies (e.g., [132, 142–150]) train and test fall detection algorithms on the same participants, limiting generalizability. In contrast, our work trains three machine learning classifiers on data from five participants and evaluates performance on a separate group of five *unseen* individuals, achieving a detection accuracy of up to 99.2%. This approach provides a stronger validation for real-world applicability.

4.3 Fall Detection Using Thermopile Sensors

Thermopile sensors have been utilized in various human sensing applications, including occupancy detection [88], people counting [84], localization [61], activity recognition [139], and fall detection. These sensors can be classified by their resolution. This section provides a review on fall detection research using both high-resolution and low-resolution thermopiles. Tables I and II summarize these works respectively. Table II also highlights our work's robustness and efficacy in terms of number of sensors, number of participants, evaluation methodology with unseen participants and accuracy.

4.3.1 High Resolution Thermopiles

Several researchers make use of relatively high resolution thermopiles for activity recognition including fall detection. Yu et al. [142] introduced a fall detection system that employs an 80×64 pixel thermopile imaging array to detect falls. The system was tested with 100 fall events and 100 non-fall events, achieving a sensitivity of 98% and a specificity of 100%. A human activity recognition system capable of identifying fall events along with other static activities such as standing and sitting, and dynamic activities like walking and jumping has been proposed in [144]. This system uses two MLX90640 thermopile array sensors positioned for side and top views, each capturing 24×32 thermal data of human participants. The system achieved 97.5% accuracy for static activities and 94.6% for dynamic activities including falls. While both systems achieved high accuracies, the training and testing are limited to a single participant. Therefore it is difficult to ascertain their robustness and generalizability across diverse users.

Rezaei et al. [151] developed a fall detection system using two MLX90640 infrared thermal sensors, each with a 24×32-pixel resolution. Positioned on a room's side wall and ceiling, these sensors capture thermal data for activity classification, location

detection, and fall detection. Among the various Machine Learning algorithms trained, the best performing algorithm achieved 97.9% accuracy and a 94.5% F1-score in binary classification of fall. The experiment involved 10 participants. However, the algorithms were not tested with individuals who were not part of the training. Therefore, its generalizability is unknown.

Another fall detection and action recognition system using a 16×16-pixel thermopile array mounted on the ceiling was proposed by Kawashima et al. [145] This system utilizes a combination of CNN and LSTM layers to learn both spatial and temporal features. Experimental results showed an overall accuracy of 91.07%, with a 93.06% accuracy specifically for fall detection. They employed a leave-one-person-out cross-validation approach with a dataset involving nine participants. In this setup, the data from one participant was excluded during training and later used for evaluation, while the data from the remaining eight participants was used for training.

Table 4.1: Summary of Fall Detection Systems Using High-Resolution Thermopile Sensors

Author	**Sensor**	**Res.**	**No.**	**Algorithm**	**Performance**	**Part.**	**Uns.**
Kawashima et al. [145]	D6T-1616 OMRON	16×16	1	CNN+LSTM	Accuracy: 93.06%	6	No
Yu et al. [142]	HTPA80x64d	80×64	1	Back projection	Sensitivity: 98%, Specificity: 100%	1	No
Morawski et al. [143]	Not Specified	32×32	3	DL for bed exits	Accuracy: 96.51%	1	No
Zhang et al. [144]	MLX90640	24×32	2	DCNN + 3-axis	Accuracy: 94.6%	12	No
Rezaei et al. [132]	MLX90640	24×32	2	Binary classif. DL	Accuracy: 97.9%, F1-score: 94.5%	8	No
Naser et al. [88]	MLX90640	32×24	1	LSTM, Bi-LSTM	Accuracy: 99.7%	8	Yes
Riquelme et al. [152]	MLX90640, D6T-8L	32×24, 1×8	6	Bi-LSTM	Accuracy: 93%	6	Yes

Note: Res.=Resolution, No.=Number of Sensors, Part.=Participants, Uns.=Unseen.

4.3.2 Low Resolution Thermopiles

Several researchers have employed lower resolution thermopile sensors in their studies, with the AMG8833 Grid-EYE sensor (8×8 resolution) emerging as the most popular choice. Another popular low-resolution thermopile is the MLX90620 [153] (16×4 pixels resolution) which has now been discontinued.

Adolf et al. [146] developed a body posture recognition and fall detection system using a single Grid-EYE sensor. This system classifies body postures, including standing, sitting, and lying down based on low-resolution thermal images. Using the Inception-v3 Deep Learning model, the system achieved sensitivities of 85% for detecting lying postures, which correspond to falls, and a specificity of 93% for the same class. The experiments involved four participants with no testing on unseen individuals, which makes it difficult to ascertain the system's ability to generalize to diverse populations.

Another single Grid-EYE sensor based fall detection system is proposed by Fan et al. [148] This system processes the thermal data using several Deep Learning methods, including LSTM networks and Gated Recurrent Units (GRU). They conducted experiments in two different configurations, and testing falls in both perpendicular and parallel orientations relative to the sensor. They observed that parallel falls are easier to classify, with their best-performing model achieving an F1 score of 100%. The experiments involved two participants, with training and testing data collected separately. However, the small sample size limits the system's ability to generalize to diverse populations, raising concerns about its robustness in real-world applications.

Tao et al. [147] proposed a fall and activity detection system using a single ceiling mounted Grid-EYE sensor. Their dataset, named Infra-ADL2018, contains recordings from eight participants performing seven distinct activities across 24 sessions. Using leave-one-subject-out cross-validation, the system achieved an overall accuracy of

87.5% and 100% sensitivity for fall detection. They highlight the need for a multi-sensor system providing multiple FoVs to improve robustness in real-world applications.

The method proposed by Mashiyama et al. [150] is able to recognize various human activities, including fall, stopping, walking, sitting and no event. It consists of one ceiling-mounted Grid-EYE sensor. They used a dataset with six participants, with 50 fall events per individual in the training set and 30 in the test set. The results showed 100% accuracy in fall detection, demonstrating the effectiveness of the sensor in identifying falls alongside other activities, even with varying conditions such as sensor height and room temperature. However, sitting activities achieved the lowest recognition accuracy at 78.6%, with sitting motions sometimes misclassified as falls. Furthermore, the dataset's reliance on the same participants for training and testing raises concerns about the method's generalizability to unseen subjects.

Chen et al. [153] implemented a localization and fall detection application by using two low-resolution infrared sensors with a resolution of 16 x 4 and mounting them on the wall at different locations. By integrating the two sensors placed on the wall, the system attempts to capture three-dimensional image information. For fall detection, features from the falling action are extracted and classified using a k-NN model. Tested with 80 fall actions and 80 normal actions, this system achieved a sensitivity of 95.25%, a specificity of 90.75%, and an overall accuracy of 93%. Again, the system has not been tested on unseen individuals.

In [149], a fall detection system using a single ceiling mounted thermopile with a resolution of 16×4 has been proposed. By analyzing temperature distribution over a two-second window, the system can identify falls using machine learning algorithms, including a voting classifier, achieving an accuracy of 97.75%. The study involved a total of 1600 data samples, collected from 10 different individuals. But it has also not been explicitly evaluated on unseen subjects.

Table II highlights the comparative advantages and robustness of our approach.

Unlike prior studies, our work explores sensor placement strategies through multi-sensor deployment and demonstrates that a ceiling–wall configuration achieves superior performance. A more detailed discussion on benchmarking against prior studies can be found in the Results and Discussions section.

Table 4.2: Summary of Fall Detection Systems Using Low-Resolution Thermopile Sensors

Author	**Sensor**	**Res.**	**No.**	**Algorithm**	**Performance**	**Part.**	**Uns.**
Adolf et al. [146]	Grid-EYE	8×8	1	DNN	Sensitivity: 85%	4	No
Tao et al. [147]	Grid-EYE	8×8	1	DCT+SVM	Specificity: 99.21%	8	No
Fan et al. [148]	Grid-EYE	8×8	1	MLP, LSTM, GRU	Perpend.: 79%, Parallel: 99%	2	No
Mashiyama et al. [150]	Grid-EYE	8×8	1	SVM	Accuracy: 100%	6	No
Ogawa et al. [149]	MLX90640, D6T-8L-06	16×4	2	Voting classifier	Accuracy: 94.6%	12	No
Chen et al. [154]	MLX90620	16×4	2	AoA+k-NN	Sensitivity: 95.25%	5	Yes
Proposed	**Grid-EYE**	**8×8**	**6**	**MLP, SVM, RF**	**Accuracy: 98%**	**10**	**Yes**

Note: Res.=Resolution, No.=Number of Sensors, Part.=Participants, Uns.=Unseen.

4.4 System Setup

We investigated fall detection systems using various combinations of six thermopiles. In this section, we describe our experimental setup, data collection, and classifier training in detail.

4.4.1 Choice of Hardware

As highlighted in Section 4.3.2, Grid-EYE stands out as the preferred sensor for thermopile-based fall detection applications and it is also used in our system. A 4-layer PCB was developed for the sensor module. The STM32F415RGT6 [155] was selected as the microcontroller. Data from the sensing hardware are acquired at a frequency of 25Hz over serial link. Figure 4.2 shows one of the bespoke sensor modules.

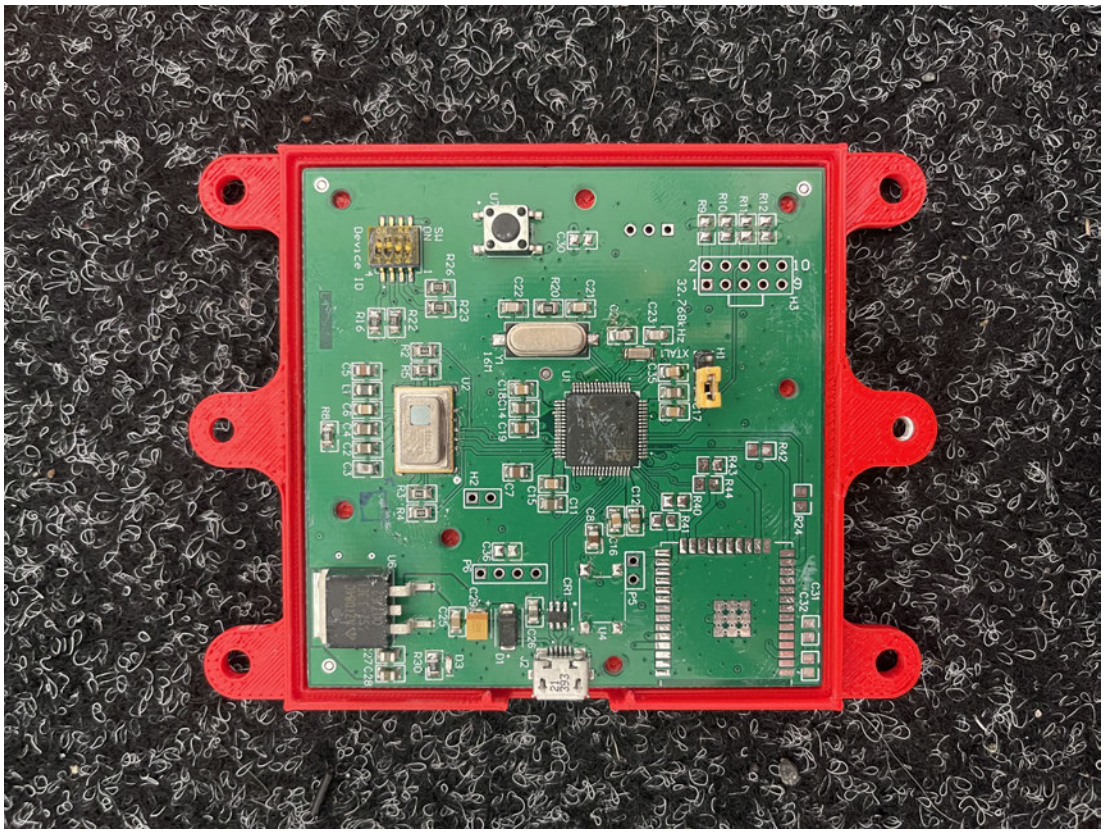


Figure 4.2: The bespoke sensor module with the Grid-EYE thermopile

4.4.2 Thermopile Arrangements

Six thermopiles, two mounted on the ceiling and four on the walls, are utilized in our experimental setup. Deploying sensors on the walls and ceiling ensures a multi-angled, spatially aware setup that captures detailed temperature data. This enhances the system's ability to detect falls accurately across the monitored environment. This larger number of sensors allows us to investigate and evaluate the number of sensors required for adequate performance and also provide insight into the desired locations of the sensors. However, it should be noted that the number of thermopiles used in this experimental setup is likely significantly higher than what would be practical in a real-world settings.

Each sensor collects data independently but synchronously, allowing real-time monitoring of temperature changes from different angles. The wall-mounted sensors provide lateral coverage, capturing temperature changes at mid-height. This lateral perspective complements the top-down view provided by the two ceiling-mounted sensors, which offer a broad overview of the entire floor area.

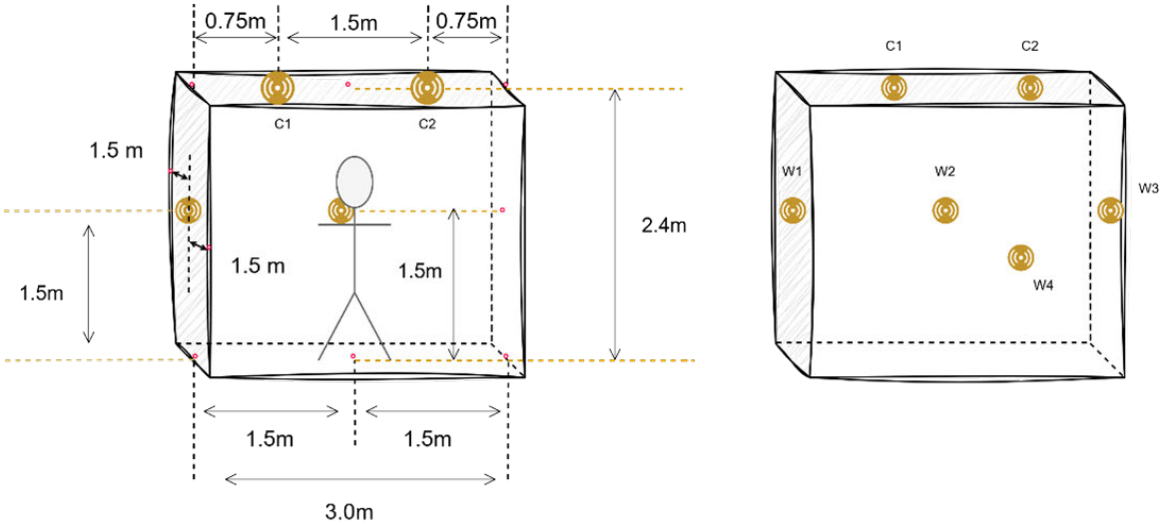


Figure 4.3: The experiment occurs in a controlled lab with a floor area of 3.0 x 3.0 meters. Six thermopile sensors are positioned around the room to capture IR data from all angles. Two sensors are mounted on the ceiling at a height of 2.4 meters and spaced 1.5 meters apart to capture a top view of the area. Four sensors are placed on each wall at a height of 1.5 meters and tilted 30° downwards to capture ground-level temperature changes. Each wall sensor is centred on the wall, 1.5 meters from the edge, ensuring the entire floor area coverage.

This study aims to gain insight into the effects of sensor arrangements on the accuracy of fall detection. Sixteen different sensor configurations are considered. They are labelled as "Comb1" through to "Comb16" and their locations are shown in Figure 4.4. In particular,

- **Comb1, Comb5, Comb6, Comb7, Comb8, and Comb9:** These configurations use wall-mounted sensors placed at varying positions on all four walls (W1, W2, W3, W4) without a ceiling sensor. This setup tests the model's performance in scenarios where only wall-based temperature data is available, focusing on lateral coverage and lower-angle detection.

- **Comb2, Comb3, and Comb4:** These combinations include one or more ceiling sensors, allowing sufficient top-down head coverage, at the cost of a smaller field of view (as the sensors are located 2.4m above the ground).

- **Comb12:** Includes all six ceiling and wall-mounted sensors in an evenly spaced arrangement, providing comprehensive coverage across the room. This combination is a reference setup for benchmarking the system's accuracy in detecting falls across different parts of the monitored area.

- **Comb10, Comb11, Comb13, Comb14, Comb15, Comb16:** These combinations use a mix of both ceiling and wall-mounted sensors to explore the influence of various placement options on fall detection accuracy.

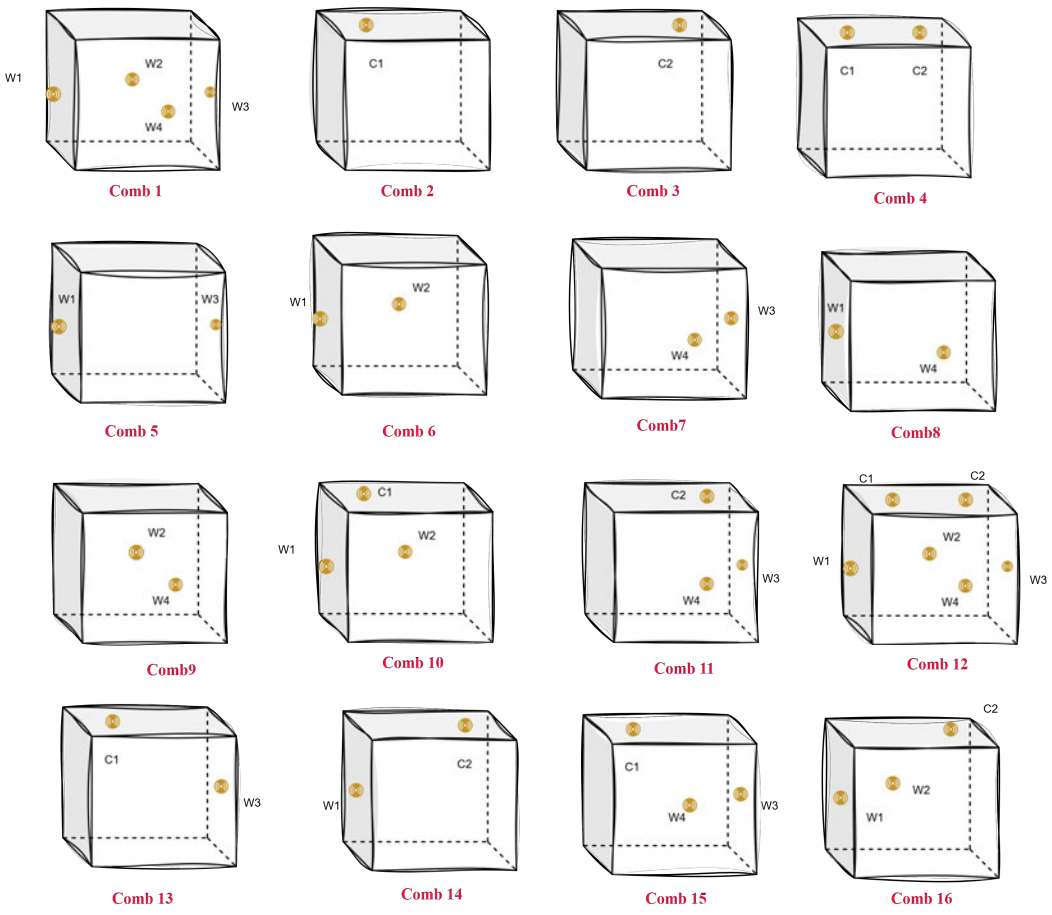


Figure 4.4: Sensor Layout Combinations and locations

4.4.3 Data Collection

Data was collected from ten participants with an equal number of males and females who all gave informed consent. Their ages ranged from 20 to 40 years, with heights from 1.55 to 1.90 meters, and weights from 45 to 80 kilograms. Due to health and safety concerns, it is not possible to ask older adults to simulate falls. In fact, our ethics approval explicitly requires us to recruit healthy volunteers. However, it is important to emphasize that we are not trying to detect pre-impact fall or the act of falling. Our goal is to detect the aftermath (post-impact fall), when a person is already on the ground. We do not expect major differences if the subject were an older adult in a real fall. It should be noted that falls are very hard to reproduce in realistic conditions. When trying to enact the act of fall, healthy volunteers usually fall onto mats, which is not realistic. They are also asked to fall at slower speeds than in real life. The study by Chu et al. [156] provides a clear example of this limitation. Participants wore short and long-sleeved shirts, pants, and shorts to introduce variability in heat distribution across different types of clothing. Additionally, for two participants, a cardboard sheet was placed under their feet to vary the ground temperature slightly. Data collected from five of the ten participants were used for training and the remainder for testing. Of the two participants with the cardboard sheet, data from one was included in the training set, while the other was allocated to the test set. This setup ensured the model was trained and evaluated under diverse environmental conditions, enhancing its generalization ability.

Raw temperature data from each sensor were sampled at an adjusted rate of 10 Hz, capturing all 64 pixel values per frame. Each sensor is connected with cables of the same length and material, minimising variations. A Python-based software enables real-time monitoring of temperature data. It also generates a continuous heatmap from the sensor data, helping to visualize temperature changes across the area.

Each participant performed the following tasks:

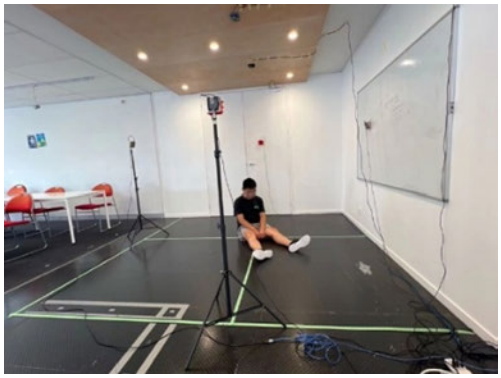
1. standing still for 30 seconds, repeated 3 times, resulting in 900 frames (10 x 30 x 3) captured by each thermopile per participant;

2. sitting on the floor for 30 seconds, repeated 3 times, produced another 900 frames;

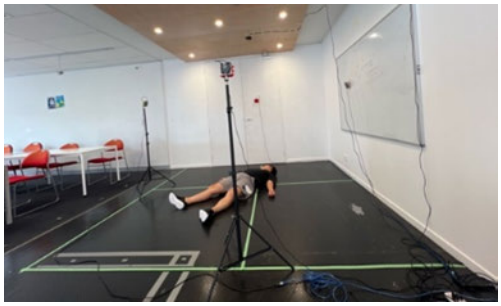
3. lying down on the floor for 30 seconds, repeated 24 times, generating 7200 frames (10 x 30 x 24) per thermopile per participant.

The lying-down activity involved variations such as lying on the back and the tummy, with different arm, leg, and body positions to simulate diverse fall scenarios. These repetitions ensured the model was exposed to various conditions, improving its ability to recognize different fall patterns. Room temperature data without the presence of participants is also collected to establish the background conditions.

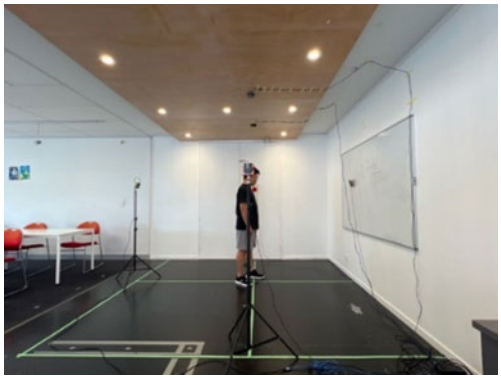
Figure 4.5 show one of the authors performing the three activities. The corresponding thermopile data are shown in Figure 4.6.



(a) Sitting Pose

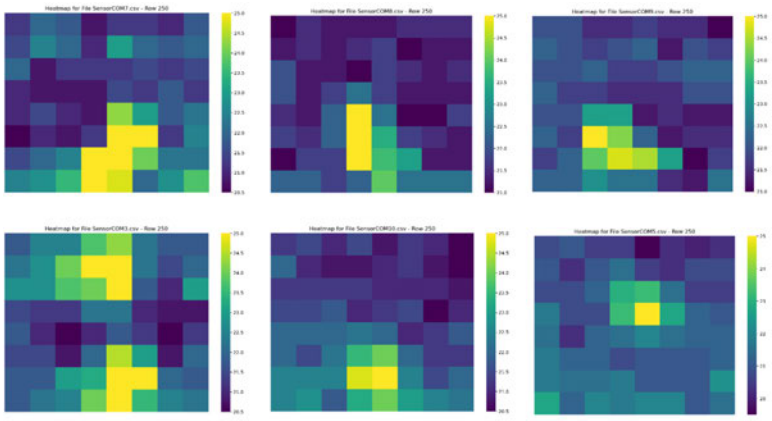


(b) Lying Down ("Fall") Pose

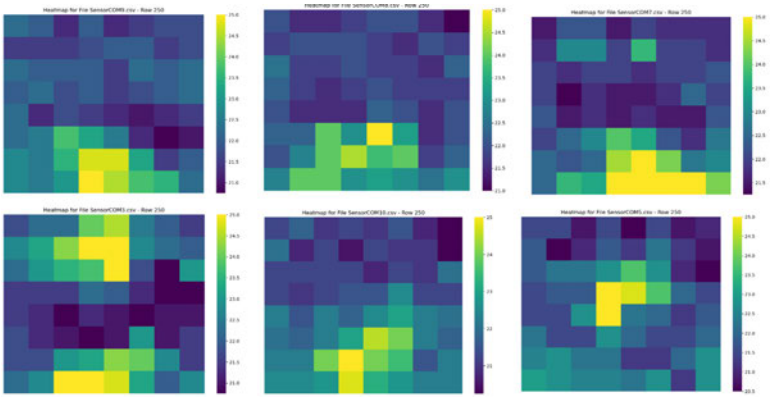


(c) Standing Pose

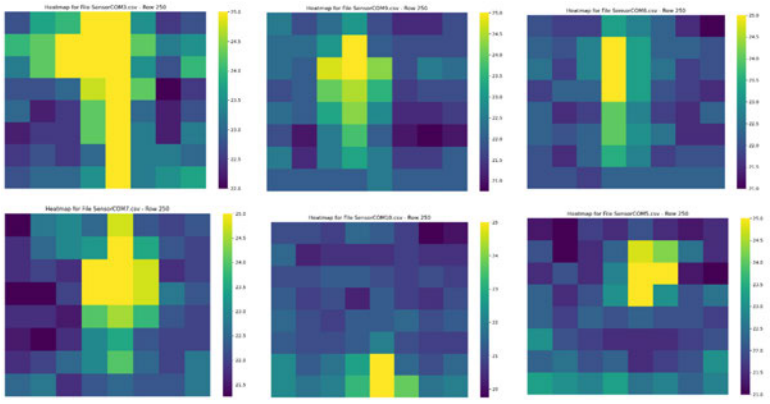
Figure 4.5: One of the authors reenacting the activities performed by the volunteers



(a) Sitting



(b) Lying Down



(c) Standing

Figure 4.6: Samples of data captured by the thermopile sensors. Top row: two ceiling-mounted sensors (C1, C2) and one wall-mounted sensor (W1). Bottom row: the remaining wall-mounted sensors (W2, W3, W4). Here, COMP7, COMP3, COMP9, and COMP8 correspond to W1, W2, W3, and W4 respectively, while COMP10 and COMP5 correspond to C1 and C2

4.4.4 Data Preprocessing

Minor timing discrepancies across sensor readings are resolved by linear interpolation. In this way, the data from each sensor are aligned in time. An equal number of samples were taken from both fall and non-fall poses, to maintain balanced training data. Data preprocessing involved three key steps: background subtraction [87], synchronization, and interpolation.

`Background Subtraction`: To isolate the human thermal signature, a static background template was subtracted from each raw thermal frame. The background value for each pixel, $\tilde{b}_i^{p,q}$, was computed as the mean of M = 72,000 samples collected over a 120-minute baseline period when the test area was unoccupied:

$$\tilde{b}_i^{p,q} = \frac{1}{M} \sum_{m=1}^{M} b_i^{p,q}(m) \tag{4.1}$$

This prolonged baseline ensures robustness against slow ambient thermal drift. The background-subtracted pixel value, where $z_i^{p,q}$ represents the raw data, is then:

$$y_i^{p,q}\left(t_{i,k}\right) = z_i^{p,q}\left(t_{i,k}\right) - \tilde{b}_i^{p,q} \tag{4.2}$$

`Synchronization and Interpolation`: The data streams from the six thermopiles that were synchronized to a common 10 Hz time base. For any target timestamp t_n falling between sensor timestamps $t_{i,k}$ and $t_{i,k+1}$ the pixel value was computed via linear interpolation:

$$\hat{y}_i^{p,q}(t_n) = \left[1 - w_{i,k}^{p,q}(t_n)\right] y_i^{p,q}(t_{i,k}) + w_{i,k}^{p,q}(t_n) y_i^{p,q}(t_{i,k+1}) \tag{4.3}$$

where the interpolation weight $w_{i,k}^{p,q}(t_n)$ is defined as $(t_n - t_{i,k})/(t_{i,k+1} - t_{i,k})$, normalizing the position of t_n within the interval. At the temporal boundaries of the data, bounded

linear extrapolation was applied to ensure a continuous signal, with values clipped to the observed data range. The resultant synchronized dataset of 8×8 frames from all thermopiles at 10 Hz was used for subsequent feature extraction and classification.

4.4.5 Machine Learning Approaches

Three different machine learning classifiers are trained to detect falls in this study. They are Support Vector Machine (SVM), Multi-Layer Perceptron (MLP), and Random Forest (RF). They perform binary classification to differentiate between normal and fall poses. SVM operates by finding the maximum margin boundary between different classes. MLP is a neural network-based approach that learns complex patterns in data through multiple layers of processing units. RF makes decisions by building multiple decision trees and taking the majority vote from these trees for classification. These classifiers were chosen due to their effectiveness in handling complex pattern recognition, making them suitable for accurately identifying fall incidents. Since the focus of this work is on identifying static poses, shallow structures that do not require temporal information are adequate.

The algorithms were trained on a computer with an AMD 9950x CPU and an NVIDIA RTX A100 GPU. An 80/20 training/validation split was utilized during training. The Optuna optimization framework, using a Tree-structured Parzen Estimator (TPE) Sampler was used for hyperparameter selection. The search ranges for each hyperparameter are given in Table 4.3. The final sets of hyperparameters are given in Table 4.4.

Table 4.3: Hyperparameter Search Ranges

Model	Hyperparameters	Search Range
MLP	number of layers	1, 2, 3, 4, 5
	hidden layer sizes	32, 64, 128, 256, 512
	activation	relu, tanh, logistic
	alpha	1e-5 – 1e-1
	learning rate init	1e-5 – 1e-1
	batch size	32, 64, 128, 256
	momentum	0.0 – 0.99
SVM	c	1e-5 – 1e-1
	kernel	linear, poly, rbf, sigmoid
	degree	2 – 8
	gamma	1e-5 – 1e-1
	coef0	-0.1 – 0.1
	shrinking	true, false
RF	number of estimators	100 – 1000
	max depth	2, 20
	min samples split	2, 10
	min samples leaf	1, 10
	max features	Sqrt, log2, none
	bootstrap	true, false

Table 4.4: Subset of Tuned Hyperparameters

Comb.	Model	Hyperparameters
Comb1	MLP	layers: 1; hidden: 30; activation: tanh; alpha: 0.0001; lr: 0.003; batch: 200; momentum: 0
	SVM	C: 2.3; Kernel: poly; Degree: 3; Gamma: 0.060; coef0: 0.55; shrinking: False
	RF	estimators: 650; depth: 17; split: 4; leaf: 5; features: sqrt; bootstrap: false
Comb2	MLP	layers: 1; hidden: 64, 128; activation: relu; solver: adam; alpha: 0.005; lr: 0.002; batch: 64
	SVM	C: 2.5; Kernel: linear; Degree: 3; Gamma: scale; coef0: 0.5; shrinking: True
	RF	estimators: 800; depth: 12; split: 5; leaf: 3; features: sqrt; bootstrap: true
Comb12	MLP	layers: 4; hidden: 256, 128, 64, 32; activation: tanh; solver: adam; alpha: 0.0001; lr: 0.001; batch: 32
	SVM	C: 0.03; Kernel: poly; Degree: 4; Gamma: 0.0007; coef0: 1.0; shrinking: True
	RF	estimators: 1000; depth: 12; split: 4; leaf: 1; features: sqrt; bootstrap: true
Comb13	MLP	layers: 2; hidden: 64, 32; activation: relu; solver: adam; alpha: 0.0007; lr: 0.0018; batch: 64
	SVM	C: 0.6; Kernel: poly; Degree: 3; Gamma: 0.003; coef0: 0.5; shrinking: True
	RF	estimators: 400; depth: 9; split: 5; leaf: 2; features: sqrt; bootstrap: true
Comb14	MLP	layers: 3; hidden: 128, 64, 32; activation: tanh; solver: sgd; alpha: 0.0003; lr: 0.0025; batch: 64
	SVM	C: 1.1; Kernel: rbf; Degree: 3; Gamma: 0.001; coef0: 0; shrinking: True
	RF	estimators: 450; depth: 10; split: 6; leaf: 2; features: none; bootstrap: false

Table 4.5: Two-Class Classifier Results

Comb.	**MLP**	**SVM**	**RF**	**Comb.**	**MLP**	**SVM**	**RF**
Comb1	95.5	98.3	93.5	Comb9	95.2	98.4	91.6
Comb2	91.0	92.8	81.8	Comb10	96.3	98.8	94.1
Comb3	90.9	92.5	81.0	Comb11	96.5	98.9	94.4
Comb4	91.3	92.8	84.8	Comb12	97.0	**99.2**	95.0
Comb5	92.5	98.2	94.4	Comb13	93.5	97.9	92.0
Comb6	95.8	94.7	85.0	Comb14	94.2	98.2	92.7
Comb7	95.9	98.1	92.4	Comb15	94.9	98.5	93.0
Comb8	92.7	97.8	93.2	Comb16	95.6	98.6	93.5

All values are accuracy (%). Best result (99.2%) achieved by Comb12-SVM.

Data for 5 participants were selected for training the models. Those for the remaining 5 participants are used for testing. This allows us to evaluate whether the models could generalize effectively to unseen subjects.

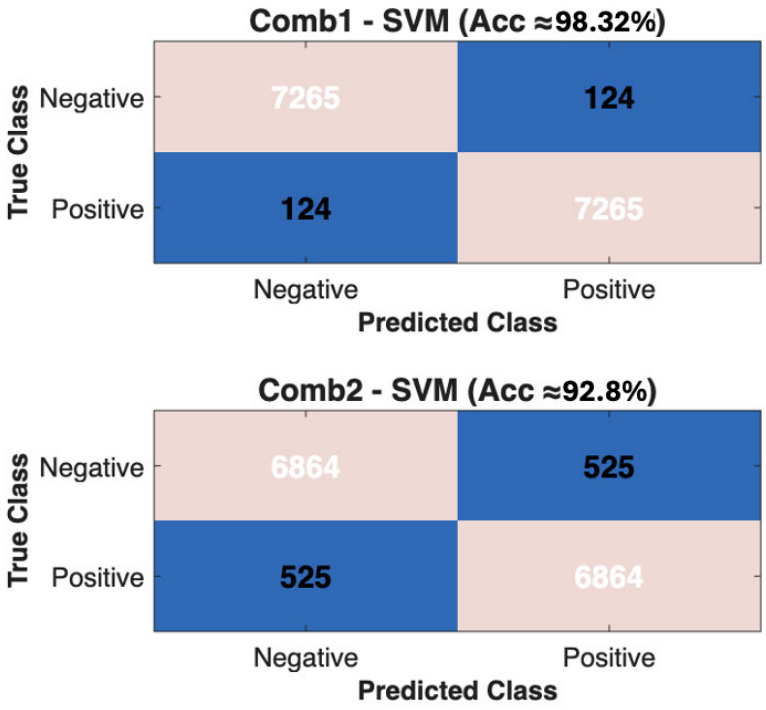


Figure 4.7: Confusion Matrix - Part 1

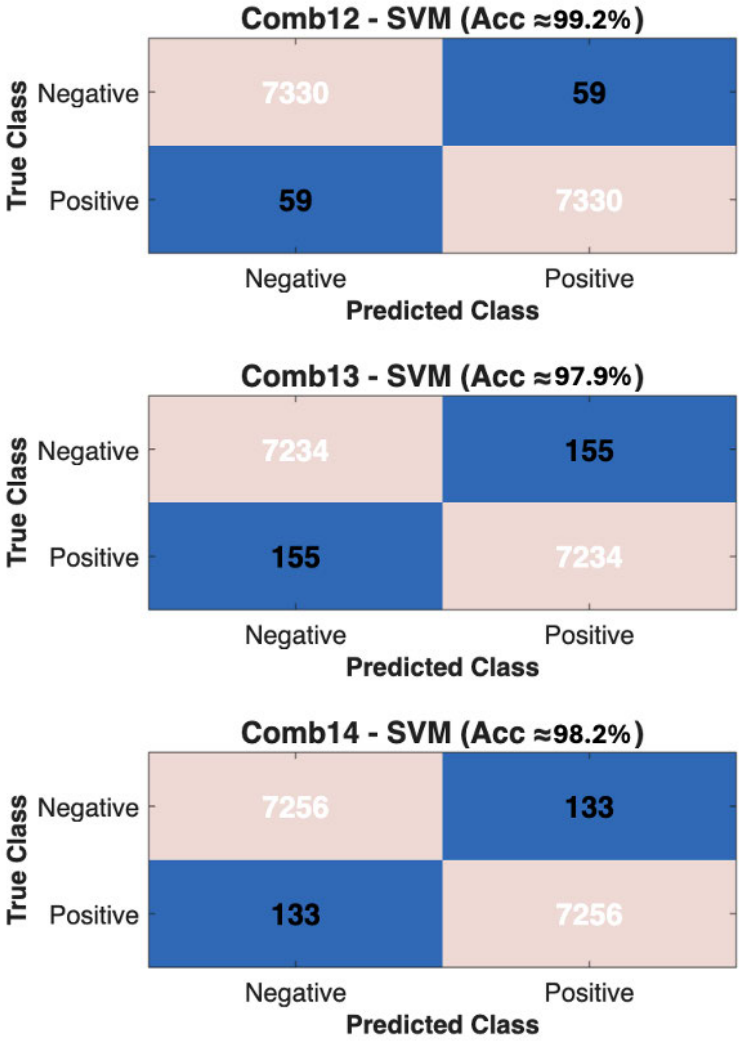


Figure 4.8: Confusion Matrix - Part 2

4.5 Results and Discussion

The classification results are shown in Table 4.5. Overall, having more sensors does offer improved classification performance. However, it should be noted that at least two classification models in each sensor configuration produced accuracies above 90%.

In general, the wall-mounted sensors outperformed the ceiling sensors, which we hypothesize is due to their increased field of view. Furthermore, there was only a 6.7% decrease in classification accuracy when going from the 6 sensor SVM model

in "Comb12" to the single sensor SVM model in "Comb3". This suggests that a single sensor may be adequate to provide acceptable classification of fall events within its field-of-view. But having more sensors would help mitigate the risk of complete sensor occlusion during fall events from furniture or other objects present within the environment.

In terms of the classifier models, the SVM demonstrated the most consistent performance among the 16 different thermopile sensor configurations. The accuracy of the best-performing algorithms ranged from 92.5% for a single ceiling-mounted sensor to 99.2% for all six sensors. It is evident from the results presented in Table 4.5 that accuracy improves noticeably when moving from a single sensor (Comb2 and Comb3) to a two-sensor configuration, provided that one of the sensors is wall-mounted (e.g., Comb5, Comb7, Comb8, Comb9, Comb13, and Comb14). In such cases accuracy increases by at least 5% compared to the single sensor configuration. By contrast, a two-sensor arrangement with both sensors ceiling-mounted (Comb4) offers no significant improvement. Moreover, increasing the number of sensors beyond two yields only diminishing returns. We therefore recommend deploying two sensors, ideally combining one ceiling-mounted and one wall-mounted unit (e.g., Comb13 or Comb14) with overlapping coverage. This configuration achieved approximately 98% accuracy in our experimental setup and offers the best balance between accuracy, cost, and implementation challenges. This arrangement also enhances robustness to occlusion by leveraging complementary perspectives.

Figures 4.7 and 4.8 show that prediction performance remains balanced between fall/non-fall events, irrespective of whether wall sensors (Comb1), a single ceiling sensor (Comb2), multiple ceiling sensors (Comb4), a mixture of wall and ceiling sensors (Comb13, Comb14), or all sensors (Comb12) are utilized. Comb 12 records the lowest misclassification, measured by False Positive Rate (FPR) and False Negative Rate (FNR), with FPR/FNR $\approx 0.8\%$. This is followed by Comb 10 at FPR/FNR $\approx 1.2\%$.

Comb 1 is slightly higher at FPR/FNR $\approx 1.7\%$, whereas Comb 2 and Comb 4 peak at FPR/FNR $\approx 7.1\%$. Our recommended combinations, Comb 13 and Comb 14, perform strongly, with misclassification rates of FPR/FNR $\approx 2.1\%$ and FPR/FNR $\approx 1.8\%$, respectively.

An analysis of these misclassification events provides further insight into the operational limitations of the system. False positives occur when non-fall activities are incorrectly classified as falls. In our experimental scenarios, these instances primarily involved the "sitting on the floor" activity. This misclassification occurs because the lowered thermal profile of a seated individual can closely resemble a lying posture, particularly from the limited perspective of certain wall-mounted sensors. Conversely, false negatives occur when actual fall events go undetected. As will be detailed in the subsequent clothing impact analysis, a substantial number of false negatives arise from infrared attenuation caused by heavy clothing (e.g., long sleeves and long pants), which significantly reduces the thermal contrast required by the classifiers. Furthermore, while the controlled laboratory setup ensured clear lines of sight, practical residential deployments would introduce partial or complete occlusion of the fallen subject by furniture. This can be a cause of false negative detections.

Two-sensor configurations, Comb 13 and Comb 14, achieve an accuracy of approximately 98%, exceeding that reported by Ogawa et al. [149]. While Mashiyama et al. [150] and Fan et al. [148] report higher accuracy, their studies involved fewer participants. More importantly, unlike our work, their results were not evaluated on unseen users. Therefore, the systems' ability to generalize to new individuals remains uncertain.

Adolf et al. [146] and Chen et al. [153] report sensitivity of their fall detection algorithms. Our two-sensor configuration achieves sensitivity (97.9%-98.2%) significantly higher than that of Adolf et al. It should be noted that their study involved only four participants and did not evaluate performance on unseen users. Chen et al. report a

sensitivity of 95.25%, which is slightly lower than ours. They also tested generalization to unseen users. However, their experiments included only five participants.

Tao et al. [147] reports slightly higher specificity than ours. However, their fall detection algorithm was not tested for unseen users, making it unclear how well their systems would adapt to entirely new individuals.

We evaluated the CNN-based algorithm (MICNN) developed by Zhang et al. [144] for our two-sensor configurations, Comb13 and Comb14. It should be noted that the thermopile sensor employed in their study had a higher resolution 24×32 compared to the low-resolution (8×8) thermopiles used in our work. To ensure compatibility with their structure, we had to independently resize the images captured by our ceiling- and wall-mounted thermopiles. Bicubic interpolation was used to resize them to 24×24, before using four columns of zero padding on either side of the images to match the 24×32 size of the original images.

The MICNN network was trained using data from the same five participants as our algorithm. We used an 80/20 train–validation split and tested the model on the remaining five participants. The test accuracies for Comb13 and Comb14 were 96.52% and 92.43%, respectively. These values are lower than what were achieved with our proposed SVM model (97.9% and 98.2% respectively). MICNN achieved validation accuracies above 99% for both combinations. This indicates that the network is likely overfitting. The complex structure of the model probably requires more training data to generalize effectively.

An interesting impact of clothing was observed during our analysis. For Comb13 and Comb14, classifiers were trained with data from three participants wearing shorts and half sleeved shirts. Testing was then performed on three additional participants who wore full sleeved shirts and long pants. The results shown in Table 4.6 indicate that the accuracy and the sensitivity of all the algorithms degrade quite significantly. This is mainly due to a higher number of false negatives, meaning a substantial number

of falls are not being detected. Training only on a "best-case" scenario (short sleeves) seems to have constrained the network's ability to generalize. However, when the training–testing configuration was reversed, the performance did not degrade in the same way. Participants wearing long-sleeved shirts and long pants produce coarser training images due to lower and less defined IR contrast across pixels. This forces the models to rely on to more robust global features (hypothesized as body proportions, centroid location) as the finer-grained features (hypothesized as exact limb position) are unavailable. This leads to more generalized models and ultimately better performance.

The participant locations were fixed during data collection with all poses enacted approximately at the center of the monitored area. In real-world settings, falls can occur at arbitrary positions. To address this, we performed a dedicated analysis to simulate a less constrained environment. Focusing on the ceiling-mounted thermopile arrays (which provide a top-down view), we augmented our original dataset by applying a series of 2D translations and rotations to the occupant postures. This process generated synthetic data representing individuals in varied locations and orientations, expanding the dataset eightfold while ensuring data integrity. Evaluation on this augmented dataset yielded encouraging results. The MLP and the SVM models maintained performance within 1% of the original fixed-location results (for Comb 2 and Comb 3). The random forest classifier showed a slight performance decrease of less than 3%, which we consider acceptable given the significantly expanded test conditions. This demonstrates that our approach, particularly with the MLP and SVM models, possesses a strong degree of robustness to positional variation and generalizes well to less constrained scenarios.

Table 4.7 compares the fall detection performance of our Comb13 and Comb14 sensor configurations with several state-of-the-art sensing modalities, including mmWave radar, WiFi CSI, and wearable sensors. Both Comb13 and Comb14 demonstrate competitive accuracy (above 97.9%), precision, recall, and specificity levels. This benchmark

contextualizes the performance of our approach for reliable fall detection.

Table 4.6: Impact of clothing on fall detection

Training	**Model**	**Acc. (%)**	**Prec. (%)**	**Sens. (%)**	**Spec. (%)**
Comb13 Configuration					
Short → Long	RF	89.36	99.18	78.45	99.40
	SVM	85.11	99.22	69.49	99.50
	MLP	88.04	92.21	81.97	93.62
Long → Short	RF	96.33	99.69	92.95	99.71
	SVM	98.75	99.71	97.79	99.71
	MLP	95.26	99.79	90.72	99.81
Comb14 Configuration					
Short → Long	RF	83.32	98.00	66.75	98.75
	SVM	86.21	97.92	72.78	98.58
	MLP	84.36	97.55	69.12	98.40
Long → Short	RF	92.49	98.36	86.46	98.55
	SVM	94.17	95.99	92.22	96.13
	MLP	92.02	94.44	89.35	94.71

Note: Short = Short sleeves, shorts; Long = Long sleeves, full pants. Arrow indicates training → testing clothing.

Table 4.7: Comparative Fall Detection Performance of Various Sensor Modalities

Author	**Sensor**	**Algorithm**	**Performance**	**Part.**	**Uns.**
Qian et al. [157]	One Wearable	Multilevel threshold	Accuracy: 94.88%	20	No
Wang et al. [158]	Two wearable cam	RF/SVM+MLP	Accuracy: 91.8%, Sensitivity: 93.6%, Specificity: 89.2%	5	Yes
Rezaei et al. [132]	mmWave radar	CNN	Accuracy: 92.3%, Recall: 89.1%, Precision: 80.1%	8	Yes
Yao et al. [159]	mmWave radar	CNN+SVM	Accuracy: 97.9%, Recall: 98.3%, Precision: 97.5%	21	Yes
Chu et al. [156]	Wi-Fi CSI	CNN	Accuracy: 97.9%, Precision: 97.3%, Recall: 98%	22	No
Jang et al. [160]	Wi-Fi CSI	GRU	Accuracy: 94.16%	2	Yes
Proposed	**Thermopile**	**SVM**	**Accuracy: 98.2%, Precision: 98.21%, Recall: 98.2%**	**10**	**Yes**

Note: Part.=Participants, Uns.=Unseen.

4.6 Conclusion and Future Works

We developed and evaluated a fall detection system that employs machine learning models to automatically identify falls from the IR data captured by multiple low-resolution thermopile sensors. Three machine learning classifiers, namely SVM, MLP, and RF, were trained on data collected from five participants enacting fall scenarios and performing other activities such as standing and sitting. The accuracy of the fall detection algorithms was evaluated using data from an additional five participants who were not included in the training set. This approach contrasts with most prior studies and provides a more realistic evaluation of generalization. The SVM classifier demonstrated the most consistent performance within the 16 different thermopile sensor configurations investigated. The accuracy of the best-performing algorithms ranged from 92.50% for a single ceiling mounted sensor to 99.20%, depending on both the number and placement of sensors. Sensor number and location were found to have a noticeable impact on detection accuracy. A two-sensor combination of ceiling- and wall-mounted sensors with overlapping fields of view offered the best balance between coverage and sensor count. Overall, the proposed system achieved better and more robust fall detection performance than those reported in comparable studies using low-resolution thermopile sensors. Moreover, our results show that the proposed algorithms generalize better to our dataset than the deeper CNN-based network reported in previous work. The performance achieved by our thermopile-based system is also comparable to that of fall detection approaches employing other sensing modalities such as wearable sensors, mmWave radar, and Wi-Fi CSI. Collectively, these findings highlight the strong potential of the proposed system based on low-resolution thermopile sensors as a reliable and non-intrusive solution for fall detection.

Future studies could examine whether the algorithms' performance improves with additional data, either by involving more participants or utilizing data augmentation

techniques. It will also be interesting to test the algorithms for a larger population and comprehensively explore the generalizability of the ML algorithms. When evaluating the models' accuracy on the five test participants, the trained models were left unchanged. If additional data collection is feasible after the initial training phase, transfer learning could potentially enhance performance. Therefore, future research could also investigate the efficacy of transfer learning in adapting the algorithms to new users.

There may be scope for investigating more sophisticated ML algorithms in the future. However, the accurate performance of reasonably simple ML structures suggests that the added computational complexity for marginal improvements in accuracy may not be justified. Moreover, our evaluation of Zhang et al.'s [144] CNN-based network indicates that deeper architectures likely require substantially larger training datasets to generalize effectively. The implemented models operated as binary classifiers distinguishing falls from non-fall events, where the non-fall group comprised two distinct activity types. However, the classifiers did not differentiate between these two activities. Future studies could focus on incorporating additional activities to improve the robustness of the fall detection algorithms.

Our experiments were conducted with volunteers wearing clothing appropriate for the ambient temperature. However, it should be noted that thermopile sensors rely on the users emitting IR, which is influenced by the clothes they are wearing. For instance, individuals wearing thick clothing may not be easily detected by the sensor unless they are in very close proximity. However, as suggested by Faulkner et al. [103] if heavy clothing is worn due to lower ambient temperature, the greater temperature difference between body and environment may reduce this effect. This natural mitigation warrants further investigation into how clothing influences the performance of thermopile-based fall detection systems. It should be noted that in climate-controlled environments, users are unlikely to wear very thick clothing, and significant variations in apparel between individuals are not expected. Furthermore, every sensing modality has inherent

limitations. Infrared sensors are affected by clothing. Vision-based systems are degraded in low light. RF-based systems, such as Wi-Fi CSI, are sensitive to in-band interference. mmWave radars struggle in cluttered environments. Thermopile sensors do not suffer from these particular issues. It should be noted that fall detection systems using ambient sensors (Thermopile, Wi-Fi CSI, mmWave Radar) or camera are Region of Interest (ROI) specific. They will not be able to detect falls outside the monitored room. In contrast, wearable sensor-based approaches techniques are not ROI specific. However, their effectiveness relies on users consistently carrying and maintaining charged devices, which is not always the case [161]. Under such circumstances, wearable-based fall detection may fail. Unfortunately, no fall detection system is entirely failsafe, and each has scenarios where performance breakdown can occur. A robust fall detection system could potentially be created through multimodal sensor fusion, which combines different modalities to exploit their strengths and offset their weaknesses. Such multimodal approaches remain largely unexplored in literature and can be explored in a future study.

A relatively simple background subtraction method was utilized in our work to reduce the environment induced differences. More advanced background subtraction methods proposed in the literature (e.g., outlined in [61], [162]) can be an avenue to pursue in a future extension of this work. However, in climate-controlled dwellings, environment induced differences may not be significant and thus may offer minimal benefit.

It is important to note that it is not possible to completely eliminate the effects of obstacles such as furniture even with having a ceiling mounted sensor in a two-sensor configuration. IR emission can get blocked, and if a person is fully behind such an obstacle, the thermopile sensors will neither detect their presence nor identify a fall. The proposed system therefore requires line of sight to function properly. Ceiling-mounted sensors are important to maintain line of sight within the sensor's FoV. Moreover, furniture produces no IR signals of its own and is thus transparent to thermopile sensors.

As a result, even if the room layout is mapped to aid the algorithm, any rearrangement of furniture will not be detected automatically and will require recalibration. Given these considerations, the impact of occlusion, sensor placement, and algorithm robustness to occlusion and furniture rearrangement warrants a separate, dedicated study in a future paper.

Ecological validation was examined by augmenting the dataset for two single-sensor ceiling-mounted configurations. Augmenting the dataset for wall-mounted sensors or multi-sensor configurations is considerably more challenging. Future work should therefore evaluate multi-sensor setups using experimental data collected from participants enacting activities at varied locations within the coverage area.

Chapter 5

Conclusion and Future Works

5.1 Conclusion

This research develops and validates a robust, privacy-preserving human sensing system that uses low-resolution thermopile sensors to capture infrared emissions from human occupants. The proposed system supports multiple tasks within a unified framework, enabling efficient and versatile operation.

The proposed approaches, presented results and discussions can be summarized in the following key points:

- A custom low-resolution (8×8 pixel) thermopile sensor node was developed and used to implement a scalable sensor network. Using this network, two key human-sensing tasks, localization and fall detection, were successfully demonstrated. All hardware design files, firmware, and software resources have been made publicly available to support further research.

- Thermopile-based localization can achieve high accuracy, with median localization errors in the range of 0.2 m–0.3 m in a modest-sized laboratory environment for unseen users. A median localization error of 0.27 m is achieved over a 28

m^2 area using four ceiling-mounted sensors, demonstrating the scalability of the proposed approach. The achieved accuracy and coverage jointly indicate that the proposed localization system outperforms existing thermopile-based systems and is comparable to state-of-the-art approaches employing other sensing modalities.

- The proposed thermopile-based automated fall detection demonstrates high accuracy and robustness. Using one wall-mounted and one ceiling-mounted sensor, the system covers approximately 9 m^2 and achieves a fall detection accuracy of 98% using an SVM based classifier. The optimal configuration was identified by evaluating 16 spatial sensor combinations. The algorithm also generalizes well to unseen users, accurately detecting falls from individuals whose data were not included in training. The proposed approach demonstrates superior performance across several key metrics when compared with existing thermopile-based studies. The proposed method also performs at a level comparable to systems built on other sensing modalities.

To synthesize the primary outcomes of this research, Table 5.1 provides a concise integrative summary linking the core configurations, performance metrics, and contributions across both the localization and fall detection applications developed on the bespoke thermopile network.

The study conducted on localization and the development of the thermopile network is currently under review at a Q1 Journal:

S. Ma, D. Konings, E M.-K. Lai, and F. Alam, “Device-Free Localization Using Low-Resolution Thermopiles”, *Under Review*, IEEE Sensors Journal (Q1, IF 4.5).

The investigation on fall detection conducted in this thesis has directly resulted in the publication of the following journal article in a Q1 journal:

S. Ma, D. Konings, E M.-K. Lai, and F. Alam, "Fall Detection Using Multiple Low-Resolution Thermopile Sensors and Machine Learning", November 2025, IEEE Sensors Journal (Q1, IF 4.5).

Table 5.1: Integrative summary of research contributions across localization and fall detection tasks.

Task & System Parameters	**Key Research Contributions**
Device-Free Localization *Algorithms:* 2D CNN-LSTM (Regression) *Deployment:* Foyer (4 nodes, 28.2 m^2) Lab (2-4 nodes, 5.8 m^2)	• Developed a scalable, networked thermopile sensing platform, releasing all PCB schematics, firmware, and supporting code to enable replication and further extension. • Deployed four ceiling-mounted sensors to cover a substantially larger area (28.2 m^2) than reported in existing studies to assess scalability. • Evaluated the models on participants whose data were not used for training, ensuring robust and generalizable performance with median localization errors between 0.2 m and 0.3 m.
Fall Detection *Algorithms:* SVM, MLP, and RF *Deployment:* 16 spatial configurations (9.0 m^2 coverage)	• Systematically investigated 16 different ceiling- and wall-mounted sensor configurations, demonstrating that a combination with overlapping fields of view offers the best balance between coverage and sensor count. • Evaluated the algorithms using data from unseen participants who were not included in the training set, providing a stronger validation for real-world applicability. • Achieved an accuracy of approximately 98% for the recommended two-sensor setup while having more sensors to help mitigate the risk of complete sensor occlusion.

5.2 Future Research Direction

In this section, we provide some suggestions for extending this research. These should be considered alongside, and in addition to, the future research directions identified in Chapters 3 and 4.

This thesis focused on two key human sensing tasks: localization and fall detection. Future work could extend the system to a broader range of applications. As noted in Chapter 4, the fall detection framework can be expanded to recognize additional activities. Our preliminary studies also indicate that the thermopile network can detect occupancy with near-perfect accuracy (~100%), providing a strong basis for more comprehensive investigation. Future research may also explore occupancy estimation and counting. Expanding the system to support multiple capabilities would substantially increase its practical value.

In Chapter 3, we proposed incorporating wireless data transfer and Ethernet connectivity into the developed sensor nodes to enhance network scalability. Integrating a wireless module should be relatively straightforward, as the PCB already includes the necessary provisions and power budget. In contrast, enabling Ethernet functionality would require migrating to a microcontroller with native Ethernet support (e.g., STM32F746VE [163]), while retaining the overall system architecture.

While the foyer deployment successfully demonstrated the system's operational viability within a 28.2 m^2 area, extending this architecture to whole-house deployments introduces specific hardware and algorithmic limitations. From a hardware perspective, expanding the physical footprint of the sensor network is constrained by maximum USB cable lengths and the reliability of data communication over extended distances. Furthermore, the scalability of the current localization algorithm is limited by sensor resolution dependencies. As discussed in Section 3.6.1, Faulkner et al.'s algorithm [103] developed for an 8×8 resolution, scaled to the 16×16 foyer deployment with

only slight performance degradation. However, the algorithm developed by Tariq et al. [61] for 4×4 resolution did not scale up. To mitigate these hardware and algorithmic constraints in larger residential environments, future deployments will require a multi-tier localization architecture. Under this approach, an entire house would be divided into discrete zones; a primary tier would identify the specific room or zone occupied by the user, while a secondary tier would execute precise localization strictly within that active zone.

Improved data preprocessing can enhance signal quality and, in turn, increase the accuracy of downstream Machine Learning tasks. Chapter 4 proposed utilizing more advanced background subtraction methods as a future research direction. Another potential improvement is the interpolation of low-resolution thermal images captured by the thermopile sensors. Several thermopile-based human sensing studies (e.g., Chen et al. [64]) have adopted this approach, drawing on techniques from classical computer vision.

Tiny Machine Learning (TinyML) [164] enables Machine Learning inference to run directly on resource-constrained edge devices such as microcontrollers. It is therefore well-suited for the human sensing applications studied in this thesis. Deploying TinyML models can substantially reduce runtime energy consumption and minimize reliance on continuous network connectivity. Future research could focus on developing TinyML-based human sensing applications that operate entirely on the thermopile nodes, eliminating the need to transmit raw data to cloud infrastructure. Such on-device processing would also enhance system security and reduce the risk of data breaches.

The number of participants in our experimental studies was greater than those used in prior thermopile-based research. Nevertheless, future work could involve a larger and more diverse participant pool to further strengthen the evaluation. Conducting experiments across multiple environments would further enhance robustness and improve the generalizability of the developed algorithms. However, expanding the system to diverse

and dynamic environments inevitably introduces multi-target scenarios. In these contexts, performing continuous localization exclusively with low-resolution thermopiles is fundamentally challenging. Because the sensors are inherently privacy-preserving, they do not capture sufficient anatomical or visual features to establish or maintain individual user identities. Consequently, it becomes exceedingly difficult to keep track of specific individuals when their trajectories intersect and thermal signatures temporarily merge (cross-overs). While resolving this data association problem could technically be achieved by fusing the thermopile data with a secondary, feature-rich sensing modality, introducing such technologies would directly compromise the privacy-preserving advantage that motivated the adoption of the system.

One recognized limitation of this study, which reflects a wider challenge within the discipline of device-free human sensing, is the absence of formal statistical significance testing between the various model configurations and sensor arrangements. This is primarily a result of the experimental constraints associated with recruiting a large participant cohort, which yielded insufficient data to perform robust significance tests. Furthermore, the current evaluation methodology did not implement cross-fold validation across all investigated sensor combinations. This precluded the reporting of per-fold variance or standard deviation for metrics such as the median error and RMSE. Given the high number of configuration combinations analyzed, a full cross-fold statistical summary was deemed impractical for the current scope. Consequently, comparative claims in this thesis are grounded in practical significance. Future research should focus on the acquisition of larger datasets and the implementation of multi-fold validation frameworks to enable rigorous statistical benchmarking.

The data captured by the thermopile sensors is affected by occlusions and by the clothing worn by users, as discussed in Chapters 3 and 4. In Chapter 3, a brief study is also presented to evaluate the impact of clothing on localization performance. Future

research should examine these factors in a systematic and comprehensive manner. Additionally, investigating the impact of thermal interference would be a valuable direction for further study. In practical deployments, this extends to broader environmental thermal interference, such as localized space heaters and direct sunlight. While static thermal noise from these elements can be cancelled using background subtraction algorithms, mitigating the effects of dynamic temperature changes remains a challenge. Future iterations of the system could incorporate prior spatial information about the monitored area, utilizing the known locations of fixed heat sources to further filter out environmental thermal interference.

Our review of the literature indicates that publicly available thermopile-based human sensing datasets are scarce. Future research should therefore prioritize collecting comprehensive datasets and making them accessible to the wider research community. Unfortunately, our ethics approval did not permit public release of the data collected in this study.

The low-resolution thermopiles used in this study are inherently privacy-preserving. However, it is essential to validate this perception through direct input from end users. For any sensing technology to be adopted, it must be viewed as useful, usable, and reliable across diverse user groups. Future work should therefore include studies that examine public perceptions and acceptance of the proposed system. Such feedback would help identify potential barriers to adoption and inform appropriate mitigation strategies.

References

[1] Y. Nakamura, "Human sensing," in *Field Informatics: Kyoto University Field Informatics Research Group*. Springer, 2012, pp. 39–53.

[2] A. A. Laghari, K. Wu, R. A. Laghari, M. Ali, and A. A. Khan, "A review and state of art of internet of things (iot)," *Archives of Computational Methods in Engineering*, vol. 29, no. 3, pp. 1395–1413, 2022.

[3] C. Zhang and Y. Lu, "Study on artificial intelligence: The state of the art and future prospects," *Journal of Industrial Information Integration*, vol. 23, p. 100224, 2021.

[4] M. Alaa, A. A. Zaidan, B. B. Zaidan, M. Talal, and M. L. M. Kiah, "A review of smart home applications based on internet of things," *Journal of Network and Computer Applications*, vol. 97, pp. 48–65, 2017.

[5] G. W. Leeson, "The growth, ageing and urbanisation of our world," *Journal of Population Ageing*, vol. 11, no. 2, pp. 107–115, 2018.

[6] B. Dong, V. Prakash, F. Feng, and Z. O'Neill, "A review of smart building sensing system for better indoor environment control," *Energy and Buildings*, vol. 199, pp. 29–46, 2019.

[7] L. Rueda, K. Agbossou, A. Cardenas, N. Henao, and S. Kelouwani, "A comprehensive review of approaches to building occupancy detection," *Building and Environment*, vol. 180, p. 106966, 2020.

[8] F. Zafari, A. Gkelias, and K. K. Leung, "A survey of indoor localization systems and technologies," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2568–2599, 2019.

[9] D. Konings, F. Alam, N. Faulkner, and C. de Jong, "Identity and gender recognition using a capacitive sensing floor and neural networks," *Sensors*, vol. 22, no. 19, p. 7206, 2022.

[10] L. Guo, Z. Lu, and L. Yao, "Human-machine interaction sensing technology based on hand gesture recognition: A review," *IEEE Transactions on Human-Machine Systems*, vol. 51, no. 4, pp. 300–309, 2021.

[11] J. Zhang *et al.*, "A survey of mmwave-based human sensing: Technology, platforms and applications," *IEEE Communications Surveys & Tutorials*, vol. 25, no. 4, pp. 2052–2087, 2023.

[12] X. Wang, C. Yang, and S. Mao, "On csi-based vital sign monitoring using commodity wifi," *ACM Transactions on Computing for Healthcare*, vol. 1, no. 3, pp. 1–27, 2020.

[13] H.-B. Zhang *et al.*, "A comprehensive survey of vision-based human action recognition methods," *Sensors*, vol. 19, no. 5, p. 1005, 2019.

[14] A. Morar and et al., "A comprehensive survey of indoor localization methods based on computer vision," *Sensors*, vol. 20, no. 9, p. 2641, 2020.

[15] G. Sreenu and S. Durai, "Intelligent video surveillance: a review through deep learning techniques for crowd analysis," *Journal of Big Data*, vol. 6, no. 1, pp. 1–27, 2019.

[16] B. W. Wirtz, J. C. Weyerer, and C. Geyer, "Artificial intelligence and the public sector—applications and challenges," *International Journal of Public Administration*, vol. 42, no. 7, pp. 596–615, 2019.

[17] M. Oudah, A. Al-Naji, and J. Chahl, "Hand gesture recognition based on computer vision: a review of techniques," *Journal of Imaging*, vol. 6, no. 8, p. 73, 2020.

[18] M. R. López, O. Sergiyenko, and V. Tyrsa, "Machine vision: approaches and limitations," *Computer Vision*, vol. 395, p. 42, 2008.

[19] Y. Wang, S. Cang, and H. Yu, "A survey on wearable sensor modality centred human activity recognition in health care," *Expert Systems with Applications*, vol. 137, pp. 167–190, 2019.

[20] Z. Jiang *et al.*, "Fall detection systems for internet of medical things based on wearable sensors: A review," *IEEE Internet of Things Journal*, vol. 11, no. 21, pp. 34 797–34 810, 2024.

[21] S. Subedi and J.-Y. Pyun, "A survey of smartphone-based indoor positioning system using rf-based wireless technologies," *Sensors*, vol. 20, no. 24, p. 7230, 2020.

[22] G. Li and H. Sato, "Handwritten signature authentication using smartwatch motion sensors," pp. 1589–1596, 2020.

[23] T.-H. Tsai, W.-Y. Lin, Y.-S. Chang, P.-C. Chang, and M.-Y. Lee, "Technology anxiety and resistance to change behavioral study of a wearable cardiac warming system using an extended tam for older adults," *PloS one*, vol. 15, no. 1, p. e0227270, 2020.

[24] J.-S. Lee and H.-H. Tseng, "Development of an enhanced threshold-based fall detection system using smartphones with built-in accelerometers," *IEEE Sensors Journal*, vol. 19, no. 18, pp. 8293–8302, 2019.

[25] Y. Ma, G. Zhou, and S. Wang, "Wifi sensing with channel state information: A survey," *ACM Computing Surveys*, vol. 52, no. 3, pp. 1–36, 2019.

[26] F. Demrozi, C. Turetta, F. Chiarani, P. H. Kindt, and G. Pravadelli, "Estimating indoor occupancy through low-cost ble devices," *IEEE Sensors Journal*, vol. 21, no. 15, pp. 17 053–17 063, 2021.

[27] D. Konings, F. Alam, F. Noble, and E. M. Lai, "Springloc: A device-free localization technique for indoor positioning and tracking using adaptive rssi spring relaxation," *IEEE Access*, vol. 7, pp. 56 960–56 973, 2019.

[28] J. Chen, Z. Wei, Y. Tong, H. Jiang, X. Miao, and C. Yin, "Skeleton-based human activity recognition with wifi csi using a hybrid approach combining convolutional neural network and long short term memory," *Multimedia Systems*, vol. 30, no. 6, p. 381, 2024.

[29] Y. Kim, W. Jeon, and D. Jeong, "Amfall: Wifi csi amplitude-based fall detection using denoised scalograms," *IEEE Internet of Things Journal*, 2025.

[30] J. C. Soto, I. Galdino, E. Caballero, V. Ferreira, D. Muchaluat-Saade, and C. Albuquerque, "A survey on vital signs monitoring based on wi-fi csi data," *Computer Communications*, vol. 195, pp. 99–110, 2022.

[31] H. Wang, D. Zhang, Y. Wang, J. Ma, Y. Wang, and S. Li, "Rt-fall: A real-time and contactless fall detection system with commodity wifi devices," *IEEE Transactions on Mobile Computing*, vol. 16, no. 2, pp. 511–526, 2016.

[32] D. Konings, N. Faulkner, F. Alam, F. Noble, and E. M. Lai, "The effects of interference on the rssi values of a zigbee based indoor localization system," pp. 1–5, 2017.

[33] K. Kayano, S. Ohta, T. Nishio, D. Yoda, K. Taniguchi, and T. Nabetani, "Experimental evaluation of long-term concept drift and its mitigation in wifi csi sensing," in *2025 IEEE 22nd Consumer Communications & Networking Conference (CCNC)*. IEEE, 2025, pp. 1–2.

[34] A. Shastri *et al.*, "A review of millimeter wave device-based localization and device-free sensing technologies and applications," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 3, pp. 1708–1749, 2022.

[35] C. Yu, Z. Xu, K. Yan, Y.-R. Chien, S.-H. Fang, and H.-C. Wu, "Noninvasive human activity recognition using millimeter-wave radar," *IEEE Systems Journal*, vol. 16, no. 2, pp. 3036–3047, 2022.

[36] A. Rezaei *et al.*, "Unobtrusive human fall detection system using mmwave radar and data driven methods," *IEEE Sensors Journal*, vol. 23, no. 7, pp. 7968–7976, 2023.

[37] G. Tang, T. Wu, and C. Li, "Dynamic gesture recognition based on fmcw millimeter wave radar: Review of methodologies and results," *Sensors*, vol. 23, no. 17, p. 7478, 2023.

[38] M. H. Shirazi, S. Yongchareon, A. Singh, and J. Ma, "A survey on machine learning approaches for vital sign monitoring using radar," *Measurement*, vol. 253, p. 117707, 2025.

[39] F. Alam, N. Faulkner, and B. Parr, "Device-free localization: A review of non-rf techniques for unobtrusive indoor positioning," *IEEE Internet of Things Journal*, vol. 8, no. 6, pp. 4228–4249, 2020.

[40] N. Faulkner, B. Parr, F. Alam, M. Legg, and S. Demidenko, "Caploc: Capacitive sensing floor for device-free localization and fall detection," *IEEE Access*, vol. 8, pp. 187 353–187 364, 2020.

[41] O. B. Tariq, M. T. Lazarescu, and L. Lavagno, "Neural networks for indoor human activity reconstructions," *IEEE Sensors Journal*, vol. 20, no. 22, pp. 13 571–13 584, 2020.

[42] X. Tang and S. Mandal, "Indoor occupancy awareness and localization using passive electric field sensing," *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 11, pp. 4535–4549, 2019.

[43] S. Pan *et al.*, "Footprintid: Indoor pedestrian identification through ambient structural vibration sensing," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 1, no. 3, pp. 1–31, 2017.

[44] M. Mirshekari, J. Fagert, S. Pan, P. Zhang, and H. Y. Noh, "Obstruction-invariant occupant localization using footstep-induced structural vibrations," *Mechanical Systems and Signal Processing*, vol. 153, p. 107499, 2021.

[45] ——, "Step-level occupant detection across different structures through footstep-induced floor vibration using model transfer," *Journal of Engineering Mechanics*, vol. 146, no. 3, p. 04019137, 2020.

[46] N. Faulkner, F. Alam, M. Legg, and S. Demidenko, "Watchers on the wall: Passive visible light-based positioning and tracking with embedded light-sensors on the wall," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 5, pp. 2522–2532, 2019.

[47] D. Konings, N. Faulkner, F. Alam, E. M.-K. Lai, and S. Demidenko, "Fieldlight: Device-free indoor human localization using passive visible light positioning and

artificial potential fields," *IEEE Sensors Journal*, vol. 20, no. 2, pp. 1054–1066, 2019.

[48] Z. Liao *et al.*, "Gesture recognition using visible light on mobile devices," *IEEE/ACM Transactions on Networking*, vol. 32, no. 4, pp. 2920–2935, 2024.

[49] Y. Yang, Z. Liu, Y. Chen, B. Yan, Y. Sun, and T. Feng, "Visible light human activity recognition driven by generative language model," *Information Fusion*, vol. 121, p. 103159, 2025.

[50] N. Neumann and V. Banta, "P12-comparison of pyroelectric and thermopile detectors," in *Proceedings IRS² 2013*, 2013, pp. 139–143.

[51] A. Shokrollahi, J. A. Persson, R. Malekian, A. Sarkheyli-Hägele, and F. Karlsson, "Passive infrared sensor-based occupancy monitoring in smart buildings: a review of methodologies and machine learning approaches," *Sensors*, vol. 24, no. 5, p. 1533, 2024.

[52] K. Ngamakeur, S. Yongchareon, J. Yu, and Q. Z. Sheng, "Deep cnn-lstm network for indoor location estimation using analog signals of passive infrared sensors," *IEEE Internet of Things Journal*, vol. 9, no. 22, pp. 22 582–22 594, 2022.

[53] L. Wu and Y. Wang, "Shuttered passive infrared sensor for occupancy detection: Exploring a low power electro-mechanical driving approach," in *Smart Materials, Adaptive Structures and Intelligent Systems*, vol. 51951. American Society of Mechanical Engineers, 2018, p. V002T05A010.

[54] X. Luo, Q. Guan, H. Tan, L. Gao, Z. Wang, and X. Luo, "Simultaneous indoor tracking and activity recognition using pyroelectric infrared sensors," *Sensors*, vol. 17, no. 8, p. 1738, 2017.

[55] B. Yang, Q. Wei, and L. Yuan, "Location ambiguity resolution and tracking method of human targets in wireless infrared sensor network," *Infrared Physics & Technology*, vol. 96, pp. 174–183, 2019.

[56] L. Yuan *et al.*, "Interpretable passive multi-modal sensor fusion for human identification and activity recognition," *Sensors*, vol. 22, no. 15, p. 5787, 2022.

[57] A. Torres, "Adafruit amg8833 8x8 thermal camera sensor," 2021. [Online]. Available: https://static.rapidonline.com/pdf/73-0174.pdf

[58] R. C. Shit, S. Sharma, D. Puthal, and A. Y. Zomaya, "Ubiquitous localization (ubiloc): A survey and taxonomy on device free localization for smart world," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3532–3564, 2019.

[59] M. Kuki, H. Nakajima, N. Tsuchiya, K. Kuramoto, S. Kobashi, and Y. Hata, "Mining multi human locations using thermopile array sensors," in *2013 IEEE 43rd International Symposium on Multiple-Valued Logic*. IEEE, 2013, pp. 59–64.

[60] A. D. Shetty, B. Shubha, and K. Suryanarayana, "Detection and tracking of a human using the infrared thermopile array sensor—"grid-eye"," in *2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT)*. IEEE, 2017, pp. 1490–1495.

[61] O. B. Tariq, M. T. Lazarescu, and L. Lavagno, "Neural networks for indoor person tracking with infrared sensors," *IEEE Sensors Letters*, vol. 5, no. 1, pp. 1–4, 2021.

[62] Marvelmind Robotics, "Indoor positioning system (with 2cm accuracy)," Tallinn, Estonia, 2025.

[63] W.-H. Chen and H.-P. Ma, "A fall detection system based on infrared array sensors with tracking capability for the elderly at home," pp. 428–434, 2015.

[64] Z. Chen, Y. Wang, and H. Liu, "Unobtrusive sensor-based occupancy facing direction detection and tracking using advanced machine learning algorithms," *IEEE Sensors Journal*, vol. 18, no. 15, pp. 6360–6368, 2018.

[65] D. Qu, B. Yang, and N. Gu, "Indoor multiple human targets localization and tracking using thermopile sensor," *Infrared Physics & Technology*, vol. 97, pp. 349–359, 2019.

[66] N. Faulkner, F. Alam, M. Legg, and S. Demidenko, "Device-free localization using privacy-preserving infrared signatures acquired from thermopiles and machine learning," *IEEE Access*, vol. 9, pp. 81 786–81 797, 2021.

[67] L. Yu, H. Chen, H. He, H. Nie, X. Zhai, and B. Xiong, "A fall detection system based on a thermopile imaging array and a back projection algorithm," in *2020 IEEE International Conference on Electro Information Technology (EIT)*. IEEE, 2020, pp. 60–65.

[68] T. Zhang, B. Yang, and N. Gu, "Indoor human activity perception based on infrared thermopile array sensor," in *2024 IEEE 19th Conference on Industrial Electronics and Applications (ICIEA)*. IEEE, 2024, pp. 1–6.

[69] A. Rezaei, M. C. Stevens, A. Argha, A. Mascheroni, A. Puiatti, and N. H. Lovell, "An unobtrusive human activity recognition system using low resolution thermal sensors, machine and deep learning," *IEEE Transactions on Biomedical Engineering*, vol. 70, no. 1, pp. 115–124, 2022.

[70] T. Kawashima *et al.*, "Action recognition from extremely low-resolution thermal image sequence," in *2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*. IEEE, 2017, pp. 1–6.

[71] S. Mashiyama, J. Hong, and T. Ohtsuki, "Activity recognition using low resolution infrared array sensor," in *2015 IEEE International Conference on Communications (ICC)*. IEEE, 2015, pp. 495–500.

[72] J. Adolf, M. Macas, L. Lhotska, and J. Dolezal, "Deep neural network based body posture recognitions and fall detection from low resolution infrared array sensor," in *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE, 2018, pp. 2394–2399.

[73] X. Fan, H. Zhang, C. Leung, and Z. Shen, "Robust unobtrusive fall detection using infrared array sensors," in *2017 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*. IEEE, 2017, pp. 194–199.

[74] L. Tao, T. Volonakis, B. Tan, Y. Jing, K. Chetty, and M. Smith, "Home activity monitoring using low resolution infrared sensor," 2018, arXiv preprint arXiv:1811.05416.

[75] Y. Ogawa and K. Naito, "Fall detection scheme based on temperature distribution with ir array sensor," pp. 1–5, 2020.

[76] A. M. Krishnan, M. Bouazizi, and T. Ohtsuki, "An infrared array sensor-based approach for activity detection, combining low-cost technology with advanced deep learning techniques," vol. 22, no. 10. MDPI, 2022, p. 3898.

[77] K. Muthukumar, M. Bouazizi, and T. Ohtsuki, "A novel hybrid deep learning model for activity detection using wide-angle low-resolution infrared array sensor," *IEEE Access*, vol. 9, pp. 82 563–82 576, 2021.

[78] C. Yin, J. Chen, X. Miao, H. Jiang, and D. Chen, "Device-free human activity recognition with low-resolution infrared array sensor using long short-term memory neural network," *Sensors*, vol. 21, no. 10, p. 3551, 2021.

[79] M. Gochoo and et al., "Novel iot-based privacy-preserving yoga posture recognition system using low-resolution infrared sensors and deep learning," *IEEE Internet of Things Journal*, vol. 6, no. 4, pp. 7192–7200, 2019.

[80] K. Naik, T. Pandit, N. Naik, and P. Shah, "Activity recognition in residential spaces with internet of things devices and thermal imaging," vol. 21, no. 3. MDPI, 2021, p. 988.

[81] ——, "Activity recognition in residential spaces with internet of things devices and thermal imaging," vol. 21, no. 3. MDPI, 2021, p. 988.

[82] C. Basu and A. Rowe, "Tracking motion and proxemics using thermal-sensor array," *arXiv preprint arXiv:1511.08166*, 2015.

[83] H. M. Ng, "Human localization and activity detection using thermopile sensors," in *Proceedings of the 12th international conference on Information processing in sensor networks*, 2013, pp. 337–338.

[84] P. Bharad, E. K. Lee, and D. Pompili, "Towards a reconfigurable cyber physical system," in *2014 IEEE 11th International Conference on Mobile Ad Hoc and Sensor Systems*. IEEE, 2014, pp. 531–532.

[85] Y. Jeong, K. Yoon, and K. Joung, "Probabilistic method to determine human subjects for low-resolution thermal imaging sensor," in *2014 IEEE Sensors Applications Symposium (SAS)*. IEEE, 2014, pp. 97–102.

[86] D. C. Kallur, "Human localization and activity recognition using distributed motion sensors," Ph.D. dissertation, Oklahoma State University, 2014.

[87] B. Shubha and V. V. D. Shastrimath, "Real-time occupancy detection system using low-resolution thermopile array sensor for indoor environment," *IEEE Access*, vol. 10, pp. 130 981–130 995, 2022.

[88] A. Naser, A. Lotfi, and J. Zhong, "Adaptive thermal sensor array placement for human segmentation and occupancy estimation," *IEEE Sensors Journal*, vol. 21, no. 2, pp. 1993–2002, 2020.

[89] R. S. Naser, M. C. Lam, F. Qamar, and B. Zaidan, "Smartphone-based indoor localization systems: A systematic literature review," *Electronics*, vol. 12, no. 8, p. 1814, 2023.

[90] A. Naser, A. Lotfi, and J. Zhong, "Calibration of low-resolution thermal imaging for human monitoring applications," *IEEE Sensors Letters*, vol. 6, no. 3, pp. 1–4, 2022.

[91] T. Teixeira, G. Dublon, and A. Savvides, "A survey of human-sensing: Methods for detecting presence, count, location, track, and identity," *ACM Computing Surveys (CSUR)*, vol. 5, no. 1, pp. 59–69, 2010.

[92] F. Furfari and et al., "Discovering location based services: A unified approach for heterogeneous indoor localization systems," *Internet of Things*, vol. 13, p. 100334, 2021.

[93] G. Cicirelli, R. Marani, A. Petitti, A. Milella, and T. D'orazio, "Ambient assisted living: a review of technologies, methodologies and future perspectives for healthy aging of population," *Sensors*, vol. 21, no. 10, p. 3549, 2021.

[94] S. Z. Homayounfar and T. L. Andrew, "Wearable sensors for monitoring human motion: a review on mechanisms, materials, and challenges," *SLAS TECHNOLOGY: Translating Life Sciences Innovation*, vol. 25, no. 1, pp. 9–24, 2020.

[95] A. Calatroni, D. Roggen, and G. Tröster, "A methodology to use unknown new sensors for activity recognition by leveraging sporadic interactions with primitive sensors and behavioral assumptions," ETH Zurich, Tech. Rep., 2010.

[96] V. Tilwari, S. Pack, M. Maduranga, and H. Lakmal, "An improved wi-fi rssi-based indoor localization approach using deep randomized neural network," *IEEE Transactions on Vehicular Technology*, 2024.

[97] Y. Ruan, L. Chen, X. Zhou, G. Guo, and R. Chen, "Hi-loc: Hybrid indoor localization via enhanced 5g nr csl," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–15, 2022.

[98] A. Bilge, E. Ergen, B. Soner, and S. Coleri, "Scalable wi-fi rss-based indoor localization via automatic vision-assisted calibration," *arXiv preprint arXiv:2509.22869*, 2025.

[99] Q. Shi and et al., "Deep learning enabled smart mats as a scalable floor monitoring system," *Nature communications*, vol. 11, no. 1, p. 4609, 2020.

[100] A. Aryal and B. Becerik-Gerber, "A comparative study of predicting individual thermal sensation and satisfaction using wrist-worn temperature sensor, thermal camera and ambient temperature sensor," *Building and Environment*, vol. 160, p. 106223, 2019.

[101] M. Kuki, H. Nakajima, N. Tsuchiya, and Y. Hata, "Human localization using thermopile sensors," in *2012 IEEE 42nd International Symposium on Multiple-Valued Logic*. IEEE, 2012, pp. 316–321.

[102] Z. Chen, Y. Wang, and H. Liu, "Unobtrusive sensor-based occupancy facing direction detection and tracking using advanced machine learning algorithms," *IEEE Sensors Journal*, vol. 18, no. 15, pp. 6360–6368, 2018.

[103] N. Faulkner, D. Konings, F. Alam, M. Legg, and S. Demidenko, "Machine learning techniques for device-free localization using low-resolution thermopiles," *IEEE Internet of Things Journal*, vol. 9, no. 19, pp. 18 681–18 694, 2022.

[104] Panasonic Industry, "Infrared array sensor grid-eye," https://industrial.panasonic.com/ww/products/pt/grid-eye, accessed.

[105] P. Hevesi, S. Wille, G. Pirkl, N. Wehn, and P. Lukowicz, "Monitoring household activities and user location with a cheap, unobtrusive thermal sensor array," in *Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing*, 2014, pp. 141–145.

[106] J. L. Honorato, I. Spiniak, and M. Torres-Torriti, "Human detection using thermopiles," in *2008 IEEE Latin American Robotic Symposium*. IEEE, October 2008, pp. 151–157.

[107] C. Kowalski, K. Blohm, S. Weiss, M. Pfingsthorn, P. Gliesche, and A. Hein, "Multi low-resolution infrared sensor setup for privacy-preserving unobtrusive indoor localization," in *Proceedings of the 5th International Conference on*

Information and Communication Technologies for Ageing Well and e-Health (ICT4AWE), 2019, pp. 183–188.

[108] Y. Lin and Q. Zhao, "Human occupancy monitoring and positioning with speed-responsive adaptive sliding window using an infrared thermal array sensor," *Sensors*, vol. 25, no. 1, p. 129, 2024.

[109] STMicroelectronics, "Stm32f415rg high-performance arm cortex-m4 mcu with 1 mbyte flash," Datasheet, 2020, docID022152 Rev 9.

[110] Panasonic, "Amg8833 infrared array sensor grid-eye," Datasheet, 2017, specifications subject to change without notice.

[111] STMicroelectronics, "Getting started with stm32 mcu hardware development," Application Note AN4661, 2021, rev 5.

[112] Marvelmind Robotics, "Marvelmind indoor navigation system operating manual," https://marvelmind.com/pics/marvelmind_navigation_system_manual.pdf, 2024, accessed.

[113] T. Yang, P. Guo, W. Liu, X. Liu, and T. Hao, "Deeppirates: A training-light pir-based localization method with high generalization ability," *IEEE Access*, vol. 9, pp. 86 054–86 061, 2021.

[114] S. Alajlouni and P. Tarazaga, "A new fast and calibration-free method for footstep impact localization in an instrumented floor," *Journal of Vibration and Control*, vol. 25, no. 10, pp. 1629–1638, May 2019.

[115] USB Implementers Forum, "Universal serial bus specification revision 2.0," USB-IF, 2000, section 7.1.19, Cable Length.

[116] IEEE Computer Society, "Ieee standard for ethernet," IEEE Std 802.3-2018, 2018.

[117] G. Forbes, S. Massie, and S. Craw, "Fall prediction using behavioural modelling from sensor data in smart homes," vol. 53, no. 2, 2020, pp. 1071–1091.

[118] R. Luque, E. Casilari, M.-J. Morón, and G. Redondo, "Comparison and characterization of android-based fall detection systems," *Sensors*, vol. 14, no. 10, pp. 18 543–18 574, 2014.

[119] M. Hemmatpour, R. Ferrero, B. Montrucchio, and M. Rebaudengo, "A review on fall prediction and prevention system for personal devices: evaluation and experimental results," *Advances in Human-Computer Interaction*, vol. 2019, 2019.

[120] S. Jung, Y. Lee, J. Kim, H. Song, S. Oh, J. Jang, and J. Hong, "Analysys of a wearable active airbag for falling injury protection of the elderly," in *Frontiers in Biomedical Devices*, vol. 87752. American Society of Mechanical Engineers, 2024, p. V001T04A012.

[121] Helite, "Hip'guard - wearable airbag for hip protection," https://helite.com/hipguard/, 2023, accessed: 2025.

[122] Prop, "Smart airbag pro for fall protection," https://propairbag.com/, 2023, accessed: 2025.

[123] B. Nemeth, M. van der Kaaij, R. Nelissen, J. K. van Wijnen, K. Drost, and G. J. Blauw, "Prevention of hip fractures in older adults residing in long-term care facilities with a hip airbag: a retrospective pilot study," *BMC Geriatrics*, vol. 22, no. 1, p. 547, 2022.

[124] C. Onyejekwe, H. Chong, C. Wilson, F. Kim, D. C. Mackey, K. M. Sibley, and S. N. Robinovitch, "Perceptions of wearable hip protectors among canadian community-dwelling old adults," *Canadian Journal on Aging*, pp. 1–7, 2025.

[125] E. Skelly, R. Argent, D. Williams, O. Sahota, and F. Dockery, "Case report of a hip-protecting, wearable airbag contributing to a serious adverse event in an older adult," *BMC Geriatrics*, vol. 25, no. 1, p. 611, 2025.

[126] M. Raitor, S. W. Ruggles, S. L. Delp, C. K. Liu, and S. H. Collins, "Lower-limb exoskeletons appeal to both clinicians and older adults, especially for fall prevention and joint pain reduction," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 32, pp. 1577–1585, 2024.

[127] R. Claeys, E. Embrechts, A. Bourazeri, R. Debeuf, M. Firouzi, M. Eggermont, and D. Beckwée, "Understanding the perspectives of older adults and physiotherapists on home-based lower-limb exoskeletons," *Wearable Technologies*, vol. 6, p. e31, 2025.

[128] K. D. S. Miguel and G. Lewin, "Brief report: Personal emergency alarms: What impact do they have on older people's lives?" *Australasian journal on ageing*, vol. 27, no. 2, pp. 103–105, 2008.

[129] P. Vallabh and R. Malekian, "Fall detection monitoring systems: a comprehensive review," *Journal of Ambient Intelligence and Humanized Computing*, vol. 9, no. 6, pp. 1809–1833, 2018.

[130] Z. Zhang, C. Conly, and V. Athitsos. (2015) A survey on vision-based fall detection. Doi:10.1145/2769493.2769540. presented at the Proceedings of the 8th ACM International Conference on PErvasive Technologies Related to Assistive Environments - PETRA '15.

[131] C. Magota *et al.*, "Seasonal ambient changes influence inpatient falls," *Age and ageing*, vol. 46, no. 3, pp. 513–517, 2017.

[132] A. Rezaei *et al.*, "Unobtrusive human fall detection system using mmwave radar and data driven methods," *IEEE Sensors Journal*, vol. 23, no. 7, pp. 7968–7976, 2023.

[133] A. Soumya, C. K. Mohan, and L. R. Cenkeramaddi, "Recent advances in mmwave-radar-based sensing, its applications, and machine learning techniques: A review," *Sensors*, vol. 23, no. 21, p. 8901, 2023.

[134] S. Hu, S. Cao, N. Toosizadeh, J. Barton, M. G. Hector, and M. J. Fain, "Radar-based fall detection: A survey," *IEEE Robotics & Automation Magazine*, 2024.

[135] Y. Wang, K. Wu, and L. M. Ni, "Wifall: Device-free fall detection by wireless networks," *IEEE Transactions on Mobile Computing*, vol. 16, no. 2, pp. 581–594, 2016.

[136] G. Feng, J. Mai, Z. Ban, X. Guo, and G. Wang, "Floor pressure imaging for fall detection with fiber-optic sensors," vol. 15, no. 2, 2016, pp. 40–47.

[137] S.-B. Jeon *et al.*, "Self-powered fall detection system using pressure sensing triboelectric nanogenerators," *Nano Energy*, vol. 41, pp. 139–147, 2017.

[138] J. Clemente, F. Li, M. Valero, and W. Song, "Smart seismic sensing for indoor fall detection, location, and notification," *IEEE journal of biomedical and health informatics*, vol. 24, no. 2, pp. 524–532, 2019.

[139] C. Yin, J. Chen, X. Miao, H. Jiang, and D. Chen, "Device-free human activity recognition with low-resolution infrared array sensor using long short-term memory neural network," *Sensors*, vol. 21, no. 10, p. 3551, 2021.

[140] A. Aryal and B. Becerik-Gerber, "A comparative study of predicting individual thermal sensation and satisfaction using wrist-worn temperature sensor, thermal camera and ambient temperature sensor," *Building and Environment*, vol. 160, p. 106223, 2019.

[141] F. Alam, N. Faulkner, and B. Parr, "Device-free localization: A review of non-rf techniques for unobtrusive indoor positioning," *IEEE Internet of Things Journal*, vol. 8, no. 6, pp. 4228–4249, 2020.

[142] L. Yu, H. Chen, H. He, H. Nie, X. Zhai, and B. Xiong, "A fall detection system based on a thermopile imaging array and a back projection algorithm," in *2020 IEEE International Conference on Electro Information Technology (EIT), 2020*, 2020, pp. 60–065.

[143] I. Morawski, W.-N. Lie, L. Aing, J.-C. Chiang, and K.-T. Chen, "Deep-learning technique for risk-based action prediction using extremely low-resolution thermopile sensor array," vol. 33, no. 6, 2022, pp. 2852–2863.

[144] T. Zhang, B. Yang, and N. Gu, "Indoor human activity perception based on infrared thermopile array sensor," in *2024 IEEE 19th Conference on Industrial Electronics and Applications (ICIEA), 2024*, 2024, pp. 1–6.

[145] T. Kawashima *et al.*, "Action recognition from extremely low-resolution thermal image sequence," in *2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), 2017*, 2017, pp. 1–6.

[146] J. Adolf, M. Macas, L. Lhotska, and J. Dolezal, "Deep neural network based body posture recognitions and fall detection from low resolution infrared array sensor," in *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2018*, 2018, pp. 2394–2399.

[147] L. Tao, T. Volonakis, B. Tan, Y. Jing, K. Chetty, and M. Smith, "Home activity monitoring using low resolution infrared sensor," 2018, arXiv preprint.

[148] X. Fan, H. Zhang, C. Leung, and Z. Shen, "Robust unobtrusive fall detection using infrared array sensors," in *2017 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI), 2017*, 2017, pp. 194–199.

[149] Y. Ogawa and K. Naito, "Fall detection scheme based on temperature distribution with ir array sensor," pp. 1–5, 2020.

[150] S. Mashiyama, J. Hong, and T. Ohtsuki, "Activity recognition using low resolution infrared array sensor," pp. 495–500, 2015.

[151] A. Rezaei, M. C. Stevens, A. Argha, A. Mascheroni, A. Puiatti, and N. H. Lovell, "An unobtrusive human activity recognition system using low resolution thermal sensors, machine and deep learning," *IEEE Transactions on Biomedical Engineering*, vol. 70, no. 1, pp. 115–124, 2022.

[152] F. Riquelme, C. Espinoza, T. Rodenas, J.-G. Minonzio, and C. Taramasco, "ehomeseniors dataset: an infrared thermal sensor dataset for automatic fall detection research," *Sensors*, vol. 19, no. 20, p. 4565, 2019.

[153] W.-H. Chen and H.-P. Ma, "A fall detection system based on infrared array sensors with tracking capability for the elderly at home," in *2015 17th International Conference on E-health Networking, Application & Services (HealthCom), 2015*, 2015, pp. 428–434.

[154] Z. Chen, Y. Wang, and H. Liu, "Unobtrusive sensor-based occupancy facing direction detection and tracking using advanced machine learning algorithms," *IEEE Sensors Journal*, vol. 18, no. 15, pp. 6360–6368, 2018.

[155] STMicroelectronics, "Stm32f415xx reference manual," *STMicroelectronics*, 2011. [Online]. Available: https://www.st.com/resource/en/datasheet/stm32f415xx.pdf

[156] Y. Chu, K. Cumanan, S. K. Sankarpandi, S. Smith, and O. A. Dobre, "Deep learning-based fall detection using wifi channel state information," *IEEE Access*, vol. 11, pp. 83 763–83 780, 2023.

[157] Z. Qian *et al.*, "Development of a real-time wearable fall detection system in the context of internet of things," *IEEE Internet of Things Journal*, vol. 9, no. 21, pp. 21 999–22 007, 2022.

[158] X. Wang, E. Talavera, D. Karastoyanova, and G. Azzopardi, "Fall detection with a nonintrusive and first-person vision approach," *IEEE Sensors Journal*, vol. 23, no. 22, pp. 28 304–28 317, 2023.

[159] Y. Yao *et al.*, "Fall detection system using millimeter-wave radar based on neural network and information fusion," *IEEE Internet of Things Journal*, vol. 9, no. 21, pp. 21 038–21 050, 2022.

[160] K. Jang, C. Sun, J. Zhou, Y. Seo, Y. Kim, and S. Choi, "A study on the lightweight and fast response gru techniques for indoor continuous motion recognition based on wi-fi csi," *IEEE Access*, 2025.

[161] V. R. Xefteris, A. Tsanousa, G. Meditskos, S. Vrochidis, and I. Kompatsiaris, "Performance, challenges, and limitations in multimodal fall detection systems: A review," *IEEE Sensors Journal*, vol. 21, no. 17, pp. 18 398–18 409, 2021.

[162] A. A. Trofimova, A. Masciadri, F. Veronese, and F. Salice, "Indoor human detection based on thermal array sensor data and adaptive background estimation," *Journal of Computer and Communications*, vol. 5, no. 4, pp. 16–28, 2017.

[163] STMicroelectronics, "STM32F746VE high-performance and DSP with FPU, Arm Cortex-M7 MCU with 512 kbytes of flash memory, 216 MHz CPU, art accelerator, L1 cache, TFT," https://www.st.com/en/microcontrollers-microprocessors/stm32f746ve.html, 2024, accessed: 2025.

[164] P. Warden and D. Situnayake, "Tinyml: Machine learning with tensorflow lite on arduino and ultra-low-power microcontrollers," 2019.