This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2026.3727948

Device-Free Localization Using Low-Resolution Thermopiles

Shengjun Ma *(Graduate Member, IEEE)*, Daniel Konings *(Senior Member, IEEE)*, Edmund M.-K. Lai *(Life Senior Member, IEEE)* and Fakhrul Alam *(Senior Member, IEEE)*

***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 0.17–0.22 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 perform effectively in such a setting. All hardware design information and code are made publicly available to support continued research.**

***Index Terms*—Device-Free Localization (DFL), Human Sensing, Indoor Positioning System (IPS), Infrared Sensing, Passive Localization, Thermopile Sensor**

I. Introduction

HUMAN sensing ranges from monitoring human presence, position, and movement, to analyzing human posture and identity [1]. It facilitates location based services [2] and ambient assisted living [3], enabling applications in smart environments that support more interactive, safe, and assisted living. Indoor localization [4] is one such key application; since GPS is unreliable indoors, a range of alternative methods have been investigated.

Indoor localization techniques can be broadly categorized as vision-based, device-based, and device-free. Vision-based localization [5] can be highly accurate owing to the richness of visual data, but it fails in low light and raises privacy concerns. Device-based techniques [6] use wearables (e.g., smart watches, bracelets) or smartphones, but they rely on user cooperation and require regular battery charging [7], which limits their reliability and convenience.

Device-free systems mitigate these shortcomings and are gaining popularity for their privacy-preserving nature and ability to leverage pre-existing sensing infrastructure [8] for localization [4]. Many sensing modalities have been investigated, each with distinct strengths and weaknesses.

Radio frequency (RF) or wireless based techniques are the most widely explored device-free localization [9], [10] approach, but concerns persist regarding their robustness, as these systems typically require frequent recalibration [11]. Non-RF methods have received comparatively less attention and merit further investigation. Researchers have proposed using footstep-induced vibrations captured by seismic sensors for human localization [12]. Smart floors based on triboelectric [13] and capacitive [14] sensors have also been developed for this purpose. In addition, wall mounted capacitive sensing has also been explored for localization [15]. Other sensing modalities for passive localization include visible light sensing based systems [16]. Thermal sensing represents another emerging modality, where passive detection of human Infrared (IR) emissions enables device-free localization using Passive Infrared (PIR) sensors [17] and Thermopiles [18].

PIR sensors, which use pyroelectric detection of IR changes [19], are widely used in security alarms and automatic lighting, and have also been applied to localization [17] and activity detection [20]. However, effective human sensing typically requires a much higher sensor density than security installations provide, and detecting a stationary subject requires significant sensor modification [21].

Thermopiles, thermocouple-based IR sensors [22], act as thermal cameras and detect both stationary and moving subjects. High-resolution thermal cameras are costly and, like regular cameras, can be perceived as privacy-invasive, whereas low-resolution thermopiles are inexpensive and privacy-preserving (Figure 1). Although less information-rich, they have become popular for localization [4]. Their

(Corresponding author: Fakhrul Alam.)

Shengjun Ma is with the School of Business Technologies, New Zealand Skills & Education College (NZSE), Auckland 1010, New Zealand (e-mail: eric@nzse.ac.nz).

Daniel Konings is with the Department of Mechanical Engineering, School of Engineering, Computer and Mathematical Sciences, Auckland University of Technology, Auckland 1010, New Zealand and also with AUT Institute of Biomedical Technologies, Auckland University of Technology, Auckland 1010, New Zealand (e-mail: daniel.konings@aut.ac.nz).

Edmund M-K Lai is with the Department of Data Science and AI, Auckland University of Technology, Auckland 1010, New Zealand (e-mail: edmund.lai@aut.ac.nz).

Fakhrul Alam is with the Department of Electrical & Electronic Engineering, School of Engineering, Computer and Mathematical Sciences, Auckland University of Technology, Auckland 1010, New Zealand (e-mail: fakhrul.alam@aut.ac.nz).

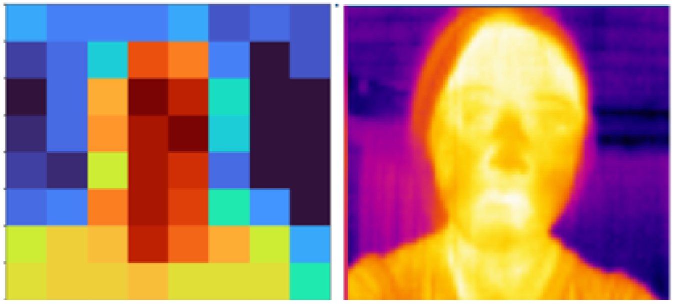


Fig. 1. Image of a person taken by a Grid-Eye having an 8 x 8 pixel resolution (left) shows the inherent privacy-preserving aspect of low-resolution thermopiles. For comparison, a capture of the same subject from the same position with an 80x60 pixels resolution FLIR Radiometric Lepton 2.5 is provided. As can be observed, the higher resolution thermal camera can be perceived as not privacy-preserving.

low resolution mitigates privacy concerns [18] and their light processing load improves scalability, enabling deployment of multiple sensor arrays as an integrated sensing network.

II. Literature Review

Table I 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.

Kuki et al. [23] 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. [24] 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. [28] used Grid-Eye images for foreground target tracking of both stationary and moving subjects but reported no quantitative accuracy. Chen et al. [26] used a table-mounted Grid-Eye sensor to achieve a localization RMSE of 0.19 m, extracting CNN features to train a support vector machine. Faulkner et al. [27] used a single ceiling-mounted Grid-Eye sensor to achieve accuracies between 0.03–0.7 m, finding convolutional and recurrent networks most accurate. All these studies used a single fixed sensor, with coverage confined to 2–9 m^2.

Chen et al. [25] used two wall-mounted 16×4 thermopile sensors and estimated the Angle of Arrival (AOA) of each sensor's foreground temperature, achieving a mean error of 0.134 m. Faulkner et al. [18] employed up to three wall-mounted Grid-Eye sensors to achieve median errors between 0.13 m and 0.20 m, training Multilayer Perceptron (MLP) models to estimate the target's angular and radial coordinates relative to each sensor. Although multiple viewpoints improved accuracy, the coverage of these studies, defined by overlapping FoVs, remained small (approximately 7–12 m^2), making system scalability difficult to ascertain.

The AMG88xx Grid-Eye family [29] is the most adopted thermopile sensor for privacy-preserving human sensing applications. Other than localization (see Table I for examples), it has been used for applications like movement tracking [30], determining the direction of facing of a subject [26], fall detection [31], [32], activity recognition [33], identifying yoga poses [34]. Example of other low-resolution thermopile includes MLX90620 (16×4 pixels resolution) which was used localization and fall detection [25]. 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 [35] and D6T-44L (4×4 pixels) were used for localization studies [23], [24]. 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 [29]. 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 [27]. 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 implementations rely on general-purpose platforms rather than designs tailored for human sensing: a Raspberry Pi Zero with a multiplexer driving three Grid-Eye sensors [36], a single Grid-Eye on a Raspberry Pi 4B [37], a Grid-Eye hardwired to an STM32F103 board [4], breadboarded WiPy2.0 wireless nodes [34], and evaluation kits such as the Atmel ATmega324P [31] and Arduino Uno [26]. These platforms include superfluous interfaces, require manual I^2C wiring, and lack robustness for scalable deployment. Faulkner et al. [18], [27] 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.

A. Contribution

To date, thermopile-based human localization studies have largely been limited to small-scale proof-of-concept demonstrations using custom hardware, a limited number of participants, and fixed sensor configurations. This work advances the field along four directions.

First, we develop a dedicated open-source thermopile sensing platform for human localization. To the best of our knowledge, no publicly available hardware platform currently exists for networked human sensing using Grid-Eye or similar low-resolution thermopile sensors. By releasing the PCB schematics, firmware, and supporting software lowers the barrier to reproducible experimentation. Second, the study moves beyond the single-layout evaluations common in prior work by systematically investigating multiple ceiling- and wall-mounted configurations, providing guidance on deployment

TABLE I
SUMMARY OF LOCALIZATION SYSTEMS USING LOW-RESOLUTION THERMOPILE SENSORS.

Author	Sensor Model	Resolution	Number of Sensors	Coverage Area (m^2)	Accuracy (m)	Number of Participants	Unseen User?
Kuki et al. [23]	Omron D6T-44L	4×4	1	2.6	Accuracy: 0.15–0.35	4	No
Tariq et al. [24]	Omron D6T-44L	4×4	1	6.4	RMSE: 0.096	1	No
Chen et al. [25]	MLX90620	16×4	2	7.0	Mean: 0.134	1	No
Chen et al. [26]	Grid-Eye	8×8	1	2.9	RMSE: 0.19	1	No
Faulkner et al. [18]	Grid-Eye	8×8	1–3	12.0	Median: 0.13–0.20	5	Yes
Faulkner et al. [27]	Grid-Eye	8×8	1	8.3	Median: 0.036–0.732	4	Yes
Proposed (Lab)	Grid-Eye	8×8	4	5.8	RMSE: 0.369, Mean: 0.242, Median: 0.17	8	Yes
Proposed (Lab)	Grid-Eye	8×8	2	5.8	RMSE: 0.388, Mean: 0.253, Median: 0.171	8	Yes
Proposed (Foyer)	Grid-Eye	8×8	4	28.2	RMSE: 0.436, Mean: 0.346, Median: 0.302	1	No

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

trade-offs. Third, the system is evaluated using participant-independent testing on data from eight participants, exceeding the scale of many earlier studies and providing stronger evidence of generalization to unseen users. Finally, we demonstrate reliable operation over a substantially larger sensing area than previously reported, shifting the focus from incremental accuracy gains in small laboratory setups toward the scalability and real-world deployment potential of distributed thermopile sensing networks.

III. SENSOR DEVELOPMENT

A. Hardware

The STM32F415RGT6 microcontroller 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 standard 5 V USB bus, while the MCU and AMG8833 sensors operate at 3.3 V.A low-dropout (LDO) linear regulator was therefore used for voltage conversion. Although switch-mode regulators offer higher efficiency, linear regulation was preferred because of its lower circuit complexity and reduced switching noise and electromagnetic interference (EMI). Given the modest current consumption of the sensing node (Table II), the resulting power dissipation remained acceptable while providing sufficient margin for future extensions such as wireless communication modules. An AZ1084C regulator was selected due to its adequate current handling capability and stable operation under the expected load conditions.

TABLE II
CURRENT ESTIMATION FOR ALL ESSENTIAL COMPONENTS OF THE DESIGN [38], [39]

Unit	Max. Current Consumption	Average Current Consumption
MCU	240mA	Estimate $<$ 120mA
AMG8833	10mA	4.5mA
ESP32 Wi-Fi Module*	790mA	160–240mA

Note: *For future extension enabling wireless data communication.

B. Circuit Design

The MCU power supply is decoupled with a bank of 1 μF, 100 nF, and 4.7 μF capacitors connected in parallel [38]. The sensor circuit was designed following the manufacturer's reference schematic [39]. The corresponding MCU pin mapping is illustrated in Figure 2. In addition, the design incorporated reset switches for device ID configuration, power line polarity protection, and Electrostatic Discharge (ESD) safeguards.

C. 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

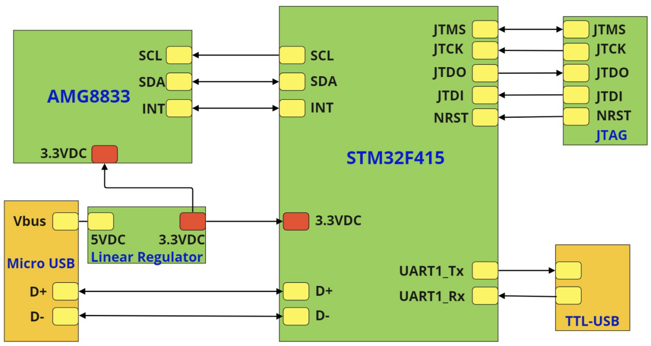


Fig. 2. High-level hardware interface architecture showing the interconnections between the AMG8833 (Grid-Eye) thermopile sensor, STM32F415 microcontroller, power regulation, USB communication, and JTAG debugging interfaces.

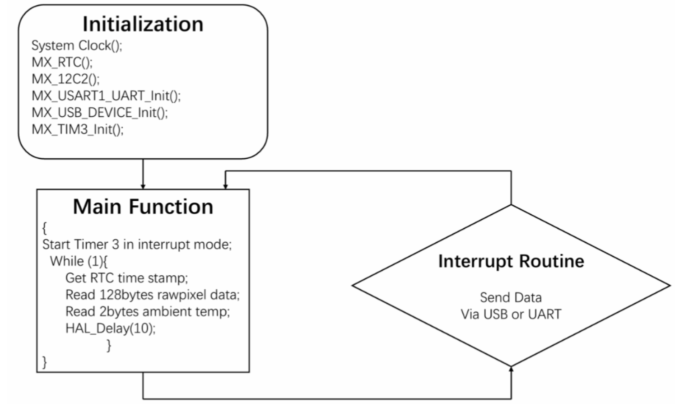


Fig. 4. Flow diagram of the onboard data acquisition and communication process comprising three main stages.

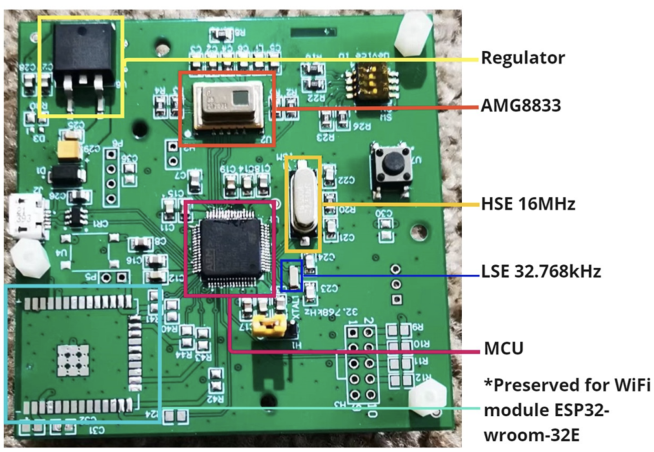


Fig. 3. Annotated photograph of the custom-designed sensing board, showing the integrated AMG8833 (Grid-Eye) thermopile sensor, STM32 microcontroller unit, onboard voltage regulation, 16 MHz HSE and 32.768 kHz LSE crystal oscillators, and reserved footprint for future ESP32-WROOM-32E Wi-Fi module integration.

to the MCU pins to minimize signal delay and noise coupling [40]. Except for through-hole components, 0.8 mm vias were employed for interlayer routing to simplify fabrication. The Gerber files for each layer are available online[1]. Following fabrication and component assembly, the completed sensor module (Figure 3) was ready for firmware and software development.

D. Firmware and Software

Figure 4 illustrates the embedded firmware data acquisition and communication flow for the sensor module. It consists of three major stages.

1) 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.

2) 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.

3) Interrupt Routine: It handles non-blocking data transmission and timing synchronization for the sensing node. Timer-driven interrupts are used to maintain deterministic communication intervals between the sensing module and the host PC. Although the AMG8833 sensor internally updates at 10 Hz (100 ms interval), the host PC polls the sensing node at 25 Hz (40 ms interval). This higher communication rate reduces latency between sensor updates and host-side visualization and minimizes timing-window mismatches between data acquisition and processing. As a result, the system achieves more stable synchronization and lower communication latency during continuous real-time sensing.

A python GUI was created to acquire data from the sensor boards at 25 Hz over serial, display it graphically at 10 Hz, as shown in Figure 5, and log it to a csv file as shown in Figure 6.

After completing the application development, the next step involved configuring the network. Prior to deployment, each sensor was individually tested. Since all sensing data are processed on a central PC, a star-topology network was adopted to provide reliable centralized communication and simplified synchronization between sensing nodes. A USB

[1]https://github.com/shengjunma/thermopile-sensor-pcb

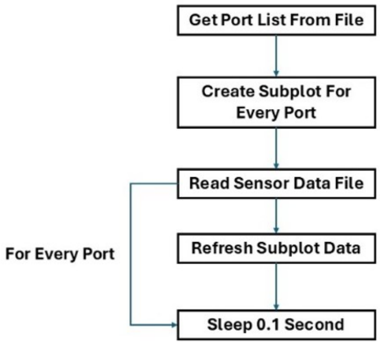


Fig. 5. Flow diagram of the GUI display process. The program iterates through each sensor port, reads the latest data file, and refreshes the corresponding subplot at 10 Hz.

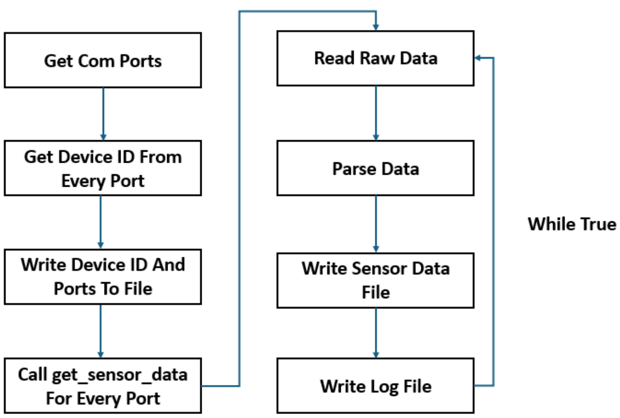


Fig. 6. Program flow for data collection. It is used to create several independent sessions to collect data from multiple devices, write the data into separated log files, and also update the latest data frame from all the sensor devices to data files for display purposes.

3.0 seven-port hub with an external power adapter was used to interface multiple sensor boards with the PC. The USB-based architecture also enabled simultaneous power delivery and high-speed data communication using a single cable per sensing node, simplifying deployment and debugging during system development. To extend the communication range and enable wider sensing coverage, a 10 m active USB cable was employed between the PC and the hub. Each sensor device was then connected to the hub using 5 m USB cables.

IV. System Setup

A. 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 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 $2.4\,\mathrm{m} \times 2.4\,\mathrm{m}$, which, although modest, is consistent with prior studies. In addition, we conducted a large-scale deployment spanning $28.2\,\mathrm{m}^2$, 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 m. At such ceiling height, each Grid-Eye sensor provides an effective coverage footprint of approximately 2.88 m $\times$ 2.88 m. To ensure coverage and eliminate blind spots, adjacent sensor footprints were intentionally designed to overlap by approximately 0.45 m. The resulting layout formed a uniform 2×2 grid, creating a square sensing array covering approximately $28.2\,\mathrm{m}^2$ (see Figure 8). 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.

B. 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 9 and 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 11 and 12. To ensure sufficient variability for training and evaluation of the Machine Learning (ML) models, each participant completed five repetitions of both paths.

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 13.

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2026.3727948

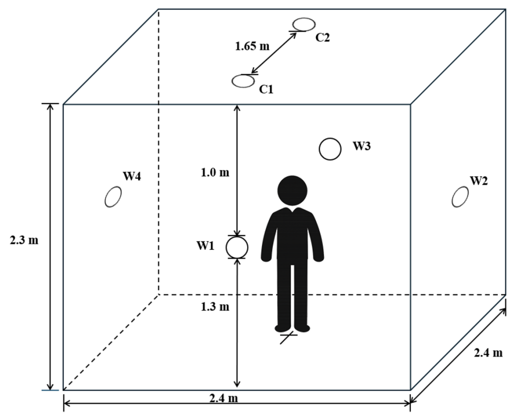


Fig. 7. The sensor layout for the lab setup. Two ceiling-mounted sensors (C_1, C_2) are installed at 2.3 m height with 1.65 m spacing. Four wall-mounted sensors (W_1–W_4) are placed at 1.3 m height around the perimeter. The coverage area is $2.4\,\mathrm{m} \times 2.4\,\mathrm{m}$.

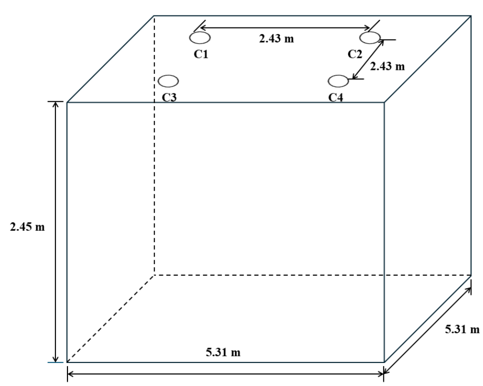


Fig. 8. Layout of the 4 ceiling mounted sensors (C_1–C_4) for the foyer setup at a height of 2.45 m. The coverage area is approximately 28 m^2.

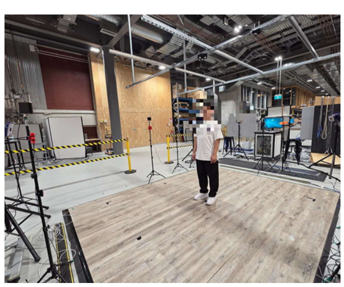


Fig. 9. Photograph of the lab setup showing a participant standing within the sensing area. The thermopile sensors are mounted on tripod stands (wall-mounted) and on the overhead truss (ceiling-mounted). The participant's face is anonymized for privacy.

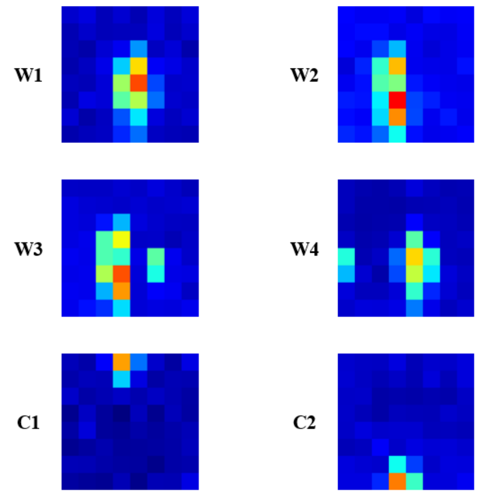


Fig. 10. Thermal data captured by the six thermopile sensors (W_1–W_4: wall-mounted; C_1, C_2: ceiling-mounted) when a participant is standing within the sensing area (as shown in Figure 9).

To provide ground truth reference for evaluating the system's positioning accuracy, a high-precision ultrasonic sensor system, the Marvelmind Indoor Navigation System [41], was deployed. According to the manufacturer specifications, the system provides approximately ± 20 mm differential precision under suitable operating conditions [41]. Independent benchmarking studies have reported similar levels of positioning accuracy [42]. Such ground-truth uncertainty is approximately an order of magnitude smaller than the localization errors reported in this work. Therefore, while the uncertainty of the ground-truth system contributes to the overall reported localization error, its overall influence is likely to be relatively small. The ground-truth system used four ultrasonic beacons at the corners of the test area, sampled at 10 Hz. Strict temporal synchronization between the thermopile array and the ground-truth system was achieved with a bespoke Python control application that simultaneously initialized and triggered acquisition across both systems. The ground-truth data were then resampled to a common 10 Hz base by linear interpolation, as for the thermal data (Section IV-C), but interpolating the (x, y) coordinates rather than pixel values.

C. 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.

1) 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 [18].

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

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

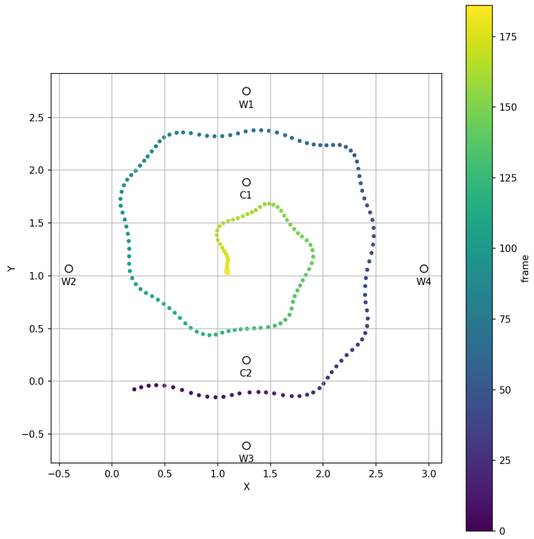


Fig. 11. Example spiral trajectory of a participant in the lab setup, recorded by the ground-truth system. C_1, C_2 are ceiling-mounted sensors, and W_1–W_4 are wall-mounted sensors.

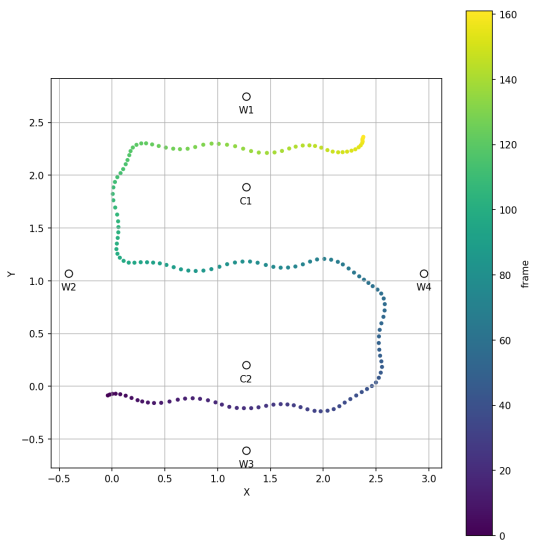


Fig. 12. Example zigzag trajectory of a participant in the lab setup, recorded by the ground-truth system. C_1, C_2 are ceiling-mounted sensors, and W_1–W_4 are wall-mounted sensors.

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

$$y_i^{p,q}(t_{i,k}) = z_i^{p,q}(t_{i,k}) - \bar{b}_i^{p,q} \tag{2}$$

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.

2) 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

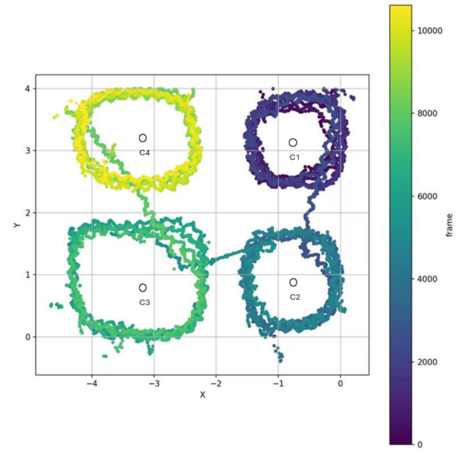


Fig. 13. Trajectory for the foyer setup as captured by the ground-truth system with four ceiling-mounted sensors (C_1–C_4).

of the thermopile sensors (each producing 8×8 frames at a nominal 10 Hz; an empirical analysis of the per-frame timestamps across all recordings showed a mean effective rate of 9.05 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

$$\hat{y}_i^{p,q}(t_n) = [1 - w_{i,k}^{p,q}(t_n)]y_i^{p,q}(t_{i,k}) + w_{i,k}^{p,q}(t_n)y_i^{p,q}(t_{i,k+1}) \tag{3}$$

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

$$w_{i,k}^{p,q}(t_n) = (t_n - t_{i,k})/(t_{i,k+1} - t_{i,k}). \tag{4}$$

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.

Additional experiments were conducted to assess the sensitivity of the localization framework to the choice of temporal resampling strategy. While nearest-neighbour resampling exhibited a modest advantage in certain tail-error metrics, the median localization accuracy obtained using nearest-neighbour and linear interpolation was statistically indistinguishable. These results suggest that the proposed framework is generally robust to the choice of interpolation method under the observed timing variability of the sensing system. A more comprehensive investigation of synchronization and resampling strategies remains an avenue for future research. Researchers adopting the proposed sensing platform should consider this design choice in applications where worst-case localization performance is particularly important.

D. 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 because the localization problem contains both temporal and spatial structure. CNN–LSTM architectures have previously demonstrated strong performance in thermopile-based localization studies [27]. Preliminary experimentation showed that purely spatial architectures such as CNN-only models produced substantially poorer localization performance than recurrent architectures, indicating that temporal evolution of the thermal patterns contains highly informative motion and trajectory cues for localization. This is particularly important for low-resolution thermopile sensing, where instantaneous spatial information is limited. The LSTM component therefore plays a central role by modeling temporal dependencies arising from target motion across consecutive thermal frames. The convolutional layers provide complementary spatial feature extraction by learning localized thermal intensity distributions associated with different target positions. Although temporal modeling contributed most strongly to overall performance, combining CNN-based spatial processing with LSTM-based temporal modeling consistently improved localization robustness and reduced large localization errors relative to purely recurrent architectures. A full architectural ablation study presented in Section V-D further validates the proposed CNN–LSTM family. The architecture of the proposed 2D CNN–LSTM network used for the lab setup is illustrated in Figure 14.

The foyer setup uses a variant of the proposed CNN–LSTM adapted to its single-image input geometry. In the foyer deployment, the four ceiling-mounted sensors form a 2×2 spatial grid and are processed as a single 16×16 thermal image rather than four separate 8×8 channels. This permits a deeper convolutional stack (three 3×3 blocks of 48, 96, and 192 filters with one intermediate 2×2 max-pooling) with a two-stage 1×1 spatial-attention block (1×1 conv of 24 filters $\rightarrow$ ReLU $\rightarrow$ 1×1 conv of 1 filter $\rightarrow$ sigmoid $\rightarrow$ element-wise multiplication), global average pooling, and a stacked two-layer LSTM head (256 units returning a sequence then 128 units returning the last step, with intermediate dropout of 0.3), followed by a 64-unit fully connected layer with ReLU and dropout of 0.2 before the two-unit regression output. The lab architecture in Table III is retained for the laboratory configurations because those sensors are mounted on different surfaces and do not form a contiguous image that benefits from the deeper cross-sensor convolutions. Both variants share the same CNN $\rightarrow$ LSTM $\rightarrow$ regression backbone and differ only in conv-stack depth, spatial attention, and LSTM stacking, tuned to the input geometry of each deployment.

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. Within every fold, the train, validation, and test partitions are subject-disjoint: no participant contributes data to more than one of the three partitions of a given fold. Each participant therefore appears in test for exactly one fold and never simultaneously in training and test. The reported results are computed on per-sample Euclidean errors pooled across the three fold-specific held-out test sets ($n = 836$ samples per configuration); 95% bootstrap confidence intervals (CI) on each cell are reported.

For the foyer setup, a frozen 60%-20%-20% train/validation/test split of the time-series data was used; the split indices are reused across all foyer experiments to allow direct comparison of network variants on the same held-out sequences. 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 III. 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, each minimizing the mean squared 2D Euclidean localization error on its respective training set.

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 III. 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 same architecture was subsequently re-evaluated in the laboratory setup under the two-sensor sub-configurations to assess the impact of a sparser, and more realistic, deployment. All reported results use the hyperparameter values used in Table III. 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.

V. Results and Discussions

In this work, localization performance is evaluated using multiple statistical metrics. The median localization error serves as the indicator of typical accuracy. Localization error distributions are often right-skewed and contain occasional large outliers making the median a robust measure of central tendency. The Root Mean Squared Error (RMSE) is reported as a complementary measure that is more sensitive to large

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2026.3727948

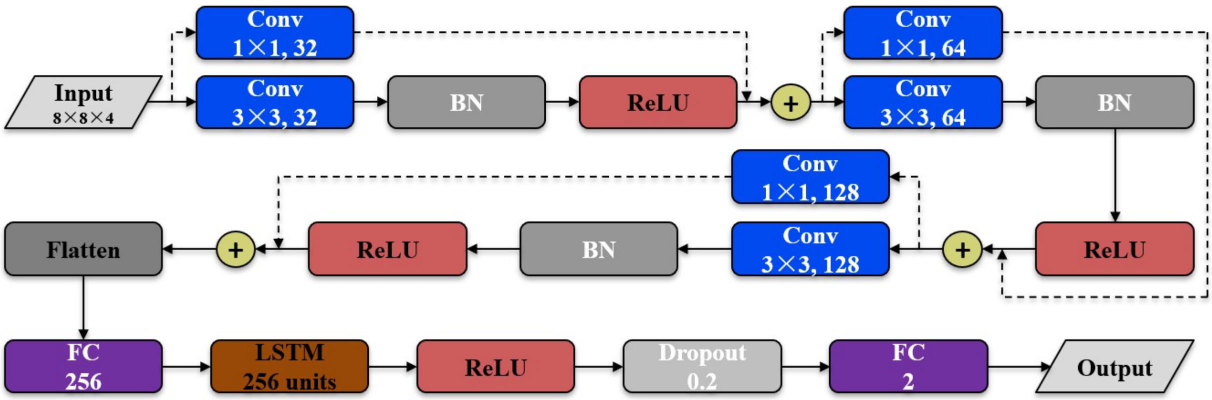


Fig. 14. Architecture of the proposed 2D CNN–LSTM network. The network consists of three convolutional blocks with skip connections (showed using dotted lines) followed by a fully connected layer, an LSTM layer with 256 units, and a final output layer for position estimation. Each convolutional block includes batch normalization (BN) and ReLU activation.

TABLE III
HYPERPARAMETER SEARCH SPACE AND FINAL VALUES FOR THE 2D CNN–LSTM (LAB VARIANT).

Hyperparameter (search range)	Final value
CNN section depth: $[1, 4]$	3
Filter size: $[2, 4]$	3
Number of filters (block 1): $2^{[3,8]}$	32
Number of filters (block 2): $2^{[3,8]}$	64
Number of filters (block 3): $2^{[3,8]}$	128
Spatial attention: {yes, no}	no
Channel attention: {yes, no}	no
Skip connections: {yes, no}	yes
Pooling: {yes, no}	no
Pre-LSTM FC neurons: $2^{[3,10]}$	256
LSTM section depth: $[1, 2]$	1
LSTM neurons: $2^{[5,10]}$	256
Dropout: $[0, 0.5]$	0.2

deviations, so two systems with similar medians can differ substantially in RMSE. The 90th and 95th percentile errors (P_{90} and P_{95}) characterize the tail of the distribution and the frequency of large errors, which is particularly important for practical deployments where occasional large errors degrade reliability.

The localization results for both setups are shown in Table IV. The results indicate that increasing the sensor count from two to four does not yield a uniform improvement across all localization metrics. While the four-sensor configurations (Comb1 and Comb2) achieve some of the lowest mean, RMSE, and P_{90} errors, several two-sensor configurations exhibit comparable or even superior median localization accuracy. For example, Comb7 achieves the lowest median error (0.168 m), slightly outperforming both four-sensor configurations. However, the four-sensor configurations generally demonstrate stronger robustness against large localization errors, particularly Comb2, which achieves the lowest P_{90} (0.434 m) and P_{95} (0.580 m) values. These observations suggest that sensor placement and coverage geometry may be as important as sensor count in determining localization performance.

For the four-sensor Comb1 configuration, the median error is 0.177 m [95% CI 0.168–0.186], for the two-sensor ceiling-only Comb3 it is 0.171 m [0.159–0.182], and for the two-sensor wall-only Comb4 it is 0.198 m [0.186–0.211]. The confidence intervals for Comb1 and Comb3 overlap substantially, indicating broadly comparable median localization accuracy. In contrast, the Comb4 interval is shifted upward and exhibits only limited overlap with the ceiling-based configurations, suggesting poorer localization performance for the wall-mounted arrangement. While Comb3 achieves a slightly lower median error than Comb1, the four-sensor configurations generally achieve lower tail-error statistics, indicating improved robustness against large localization failures.

As a sensitivity check, paired permutation tests were additionally performed on the per-sample errors of a single representative fold (fold 1) for the proposed CNN–LSTM. These tests yielded median differences of -41 mm between Comb1 and Comb3 ($p < 0.001$), -18 mm between Comb1 and Comb4 ($p = 0.051$), and $+23$ mm between Comb3 and Comb4 ($p = 0.029$). The single-fold significance pattern is broadly consistent with the bootstrap confidence intervals. Although some pairwise differences are statistically detectable, the effect sizes between the four-sensor and ceiling-only configurations remain modest. In contrast, the wall-mounted two-sensor configuration consistently exhibits poorer localization performance. Taken together, these analyses suggest that the principal performance trends are unlikely to be attributable solely to statistical noise, while also indicating that several of the observed differences should be interpreted cautiously given the limited participant cohort.

For the foyer setup, which covered a substantially larger sensing area the proposed approach achieved a median localization error of 0.302 m and an RMSE of 0.436 m. Although the localization errors are higher than those observed in the laboratory setup, the degradation is relatively modest given that the monitored area is almost five times larger. These results demonstrate that the proposed sensing framework remains effective when scaled beyond small-room environments and provide evidence that thermopile-based localization can be extended to larger deployment areas through networked sensing.

The results presented in Tables I and IV provide a broad comparison with earlier thermopile-based localization studies. However, direct numerical comparison between studies should be interpreted with caution because the reported results are obtained under substantially different experimental conditions, including differences in sensor layouts, sensing coverage, participant counts, evaluation metrics, and train–test protocols. In particular, many earlier studies do not report participant-independent evaluation using unseen users, making it difficult to assess their ability to generalize beyond the training subjects.

Nevertheless, Table I still provides useful qualitative insight into the progression of thermopile-based localization systems in terms of sensing scale, deployment complexity, and evaluation methodology. The proposed approach demonstrates competitive localization performance while employing participant-independent testing and substantially larger sensing coverage than most prior works.

Considering the foyer setup, we observe a clear advancement in the scalability of low-resolution thermopile localization systems. Previous works mostly relied on a single ceiling-mounted sensor (e.g., Kuki et al. [23], Tariq et al. [24]) or small multi-sensor configurations with overlapping fields of view limited to less than 12 m^2 (Faulkner et al. [18]). In contrast, our reported foyer setup employs four sensors covering more than twice the sensing area of the next largest reported deployment while maintaining high localization accuracy.

Going forward, three of the laboratory configurations – Comb1, Comb3, and Comb4 – are adopted as the running benchmark set for all subsequent analyses (Faulkner-baseline benchmarking, ablation, jacket sensitivity, and per-subject within-subject tests). Together they span the three mounting-geometry archetypes evaluated in this work: the best-performing four-sensor mixed-mount arrangement (Comb1, $\{C_1, C_2, W_2, W_4\}$), the canonical two-sensor ceiling-only arrangement (Comb3, $\{C_1, C_2\}$), and the canonical two-sensor opposing-wall arrangement (Comb4, $\{W_2, W_4\}$). Restricting downstream analyses to this triplet probes the contributions of sensor count, mounting height, and viewpoint diversity without proliferating combinatorial sub-experiments.

A. Experimental Benchmarking

To provide a fairer comparison with prior thermopile localization approaches, we implemented representative algorithms from earlier studies and evaluated them using our datasets and experimental setups. All benchmarked models were retrained using identical training, validation, and test splits as the proposed approach.

1) Two sensor combinations: Table V benchmarks the proposed approach against the CNN–LSTM method reported by Faulkner et al. [27], 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.

Under the three-fold subject-disjoint cross-validation methodology adopted in this work, the proposed approach outperforms the CNN–LSTM approach of Faulkner et al. [27] on all reported metrics for both the Comb3 and Comb4 configurations. For the ceiling-mounted Comb3 configuration, the proposed approach achieves a median localization error of 0.171 m compared with 0.188 m for Faulkner et al., and lower mean (0.253 m vs. 0.295 m), P_{90} (0.479 m vs. 0.650 m), and P_{95} (0.768 m vs. 0.861 m) errors as well. For the wall-mounted Comb4 configuration, the gap is wider: median 0.198 m vs. 0.266 m, with consistent improvements on every other metric.

2) Four sensor combinations: Table VI presents the corresponding benchmark for four-sensor configurations. As shown in the table, the proposed approach consistently achieves the best performance across all reported statistical metrics in the laboratory setup, achieving median errors of 0.177 m and 0.190 m for Comb1 and Comb2 respectively, compared with 0.250 m and 0.226 m for Faulkner et al. The proposed method also outperforms the Faulkner baseline in the large-scale foyer setup, with median error 0.302 m versus 0.341 m, and consistently lower errors across all other reported statistics. The foyer evaluation uses a deeper variant of the proposed CNN–LSTM architecture (see Section IV-D) to take advantage of the spatial adjacency of the four ceiling-mounted sensors as a single 16×16 thermal image; the lab evaluation uses the per-sensor $8 \times 8 \times N_c$ channel architecture described in Table III.

It is worth noting that the CNN–LSTM approach of Faulkner et al. [27] experiences a noticeable degradation in performance in our two- and four-sensor configurations compared to the single-sensor setup reported in their original study, where a median error of 0.036 m was achieved. This observation highlights the challenges associated with scaling thermopile localization systems to larger sensing areas and more complex multi-sensor configurations. In contrast, the proposed approach is specifically designed and evaluated for distributed multi-sensor deployments operating over substantially larger coverage areas.

We also implemented the neural network proposed by Tariq et al. [24]. However, the model performed poorly in our experimental setting, with median localization errors exceeding one meter in both setups. A likely reason is that the network was originally designed for single-sensor 4×4 thermal images and therefore may lack the representational capacity required for the larger and more complex inputs produced by our sensing network.

The models proposed in [18] require substantially larger amounts of training data and therefore could not be fairly trained or evaluated using our dataset. Unfortunately, several other prior studies do not provide sufficient architectural or training details to enable reliable re-implementation and retraining under our experimental conditions.

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2026.3727948

TABLE IV

POOLED-PER-SAMPLE LOCALIZATION ERROR STATISTICS AND 95% BOOTSTRAP CONFIDENCE INTERVALS FOR THE PROPOSED CNN–LSTM IN THE LABORATORY AND FOYER SETUPS. LABORATORY RESULTS ARE COMPUTED FROM THE COMBINED PREDICTIONS OF THREE SUBJECT-DISJOINT CROSS-VALIDATION FOLDS. SEE SECTION IV-A FOR DETAILS OF THE SENSOR ARRANGEMENT COMBINATIONS.

	Median (m)		Mean (m)		RMSE (m)		P_{90} (m)		P_{95} (m)	
Config	**Point**	**[95% CI]**	**Point**	**[95% CI]**	**Point**	**[95% CI]**	**Point**	**[95% CI]**	**Point**	**[95% CI]**
Comb1	0.177	[0.168, 0.186]	0.242	[0.223, 0.260]	0.369	[0.324, 0.414]	0.444	[0.397, 0.487]	0.727	[0.551, 0.862]
Comb2	0.190	[0.183, 0.200]	0.237	[0.223, 0.253]	0.327	[0.295, 0.360]	0.434	[0.397, 0.464]	0.580	[0.491, 0.644]
Comb3	0.171	[0.159, 0.182]	0.253	[0.233, 0.273]	0.388	[0.343, 0.433]	0.479	[0.427, 0.514]	0.768	[0.611, 0.882]
Comb4	0.198	[0.186, 0.211]	0.272	[0.254, 0.293]	0.399	[0.356, 0.441]	0.509	[0.463, 0.568]	0.732	[0.622, 0.996]
Comb5	0.192	[0.182, 0.203]	0.254	[0.237, 0.271]	0.357	[0.321, 0.394]	0.478	[0.438, 0.527]	0.678	[0.598, 0.744]
Comb6	0.196	[0.189, 0.201]	0.251	[0.235, 0.267]	0.349	[0.314, 0.383]	0.450	[0.418, 0.502]	0.676	[0.559, 0.810]
Comb7	0.168	[0.158, 0.179]	0.258	[0.238, 0.278]	0.395	[0.353, 0.437]	0.495	[0.434, 0.546]	0.855	[0.656, 1.098]
Comb8	0.207	[0.189, 0.217]	0.290	[0.272, 0.309]	0.402	[0.365, 0.439]	0.574	[0.529, 0.621]	0.762	[0.685, 0.848]
Comb9	0.195	[0.184, 0.209]	0.282	[0.265, 0.299]	0.380	[0.350, 0.410]	0.549	[0.503, 0.644]	0.787	[0.710, 0.899]
Comb10	0.221	[0.214, 0.231]	0.305	[0.288, 0.324]	0.407	[0.375, 0.440]	0.621	[0.555, 0.670]	0.833	[0.727, 0.925]
Foyer	0.302	[0.284, 0.318]	0.346	[0.336, 0.356]	0.436	[0.421, 0.451]	0.655	[0.627, 0.676]	0.824	[0.802, 0.837]

Note: Confidence intervals are obtained using 10,000 bootstrap resamples of the pooled per-sample errors. As a separate fold-variance diagnostic, the standard deviation of the per-fold median localization error ranges from 16 to 62 mm across configurations.

TABLE V

BENCHMARKING OF TWO-SENSOR LABORATORY CONFIGURATIONS USING POOLED-PER-SAMPLE LOCALIZATION ERROR STATISTICS AND 95% BOOTSTRAP CONFIDENCE INTERVALS. RESULTS ARE COMPUTED FROM THE COMBINED PREDICTIONS OF THREE SUBJECT-DISJOINT CROSS-VALIDATION FOLDS.

Metric	Proposed	Faulkner et al. [27]
Comb3 (two ceiling-mounted sensors)		
RMSE	**0.388** [0.343, 0.433]	0.434 [0.394, 0.476]
Mean	**0.253** [0.233, 0.273]	0.295 [0.273, 0.316]
Median	**0.171** [0.159, 0.182]	0.188 [0.176, 0.200]
P_{90}	**0.479** [0.427, 0.514]	0.650 [0.609, 0.703]
P_{95}	**0.768** [0.611, 0.882]	0.861 [0.749, 1.013]
Comb4 (two wall-mounted sensors)		
RMSE	**0.399** [0.356, 0.441]	0.462 [0.431, 0.494]
Mean	**0.272** [0.254, 0.293]	0.357 [0.337, 0.377]
Median	**0.198** [0.186, 0.211]	0.266 [0.246, 0.287]
P_{90}	**0.509** [0.463, 0.568]	0.739 [0.711, 0.760]
P_{95}	**0.732** [0.622, 0.996]	0.825 [0.791, 0.899]

Note: For each metric and sensor arrangement, the best (lowest) error between the proposed method and Faulkner et al. [27] is highlighted in bold.

B. Benchmarking against other modalities

Localization performance reported across different sensing modalities varies widely in terms of accuracy and coverage (see Table VII). 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.

C. 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 IV-D, two participants in the lab setup contributed data both while wearing jackets and without jackets. Figure 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.

To investigate the impact of clothing on localization performance, a dedicated evaluation protocol was adopted. The clothing-effect results summarized in Table X were obtained by retraining the network on data from the six participants who contributed only no-jacket data (with a 10% randomly drawn internal validation split used for early stopping). The two participants who provided data under both jacket and no-jacket conditions (Subject A and Subject B) were excluded entirely from training and used exclusively for testing. This evaluation was conducted using the Comb1, Comb3, and Comb4 sensor configurations. The same subject-disjoint protocol also underpins the complementary per-subject within-subject and leave-one-jacket-subject-out analyses summarized below. By ensuring that Subject A and Subject B appeared only in the test set, the clothing-effect estimates are free from any leakage of participant-specific thermal signatures through the training data, thereby providing a cleaner assessment of the impact of clothing on localization performance.

As shown, all three configurations exhibit measurable degradation in localization performance when the test subject is wearing a jacket. The median and mean increases are modest in absolute terms (between -5 and $+91$ mm depending on configuration and subject). However, a per-subject within-subject permutation analysis (10,000-permutation two-sample tests), corroborated by a leave-one-jacket-subject-out sensitiv-

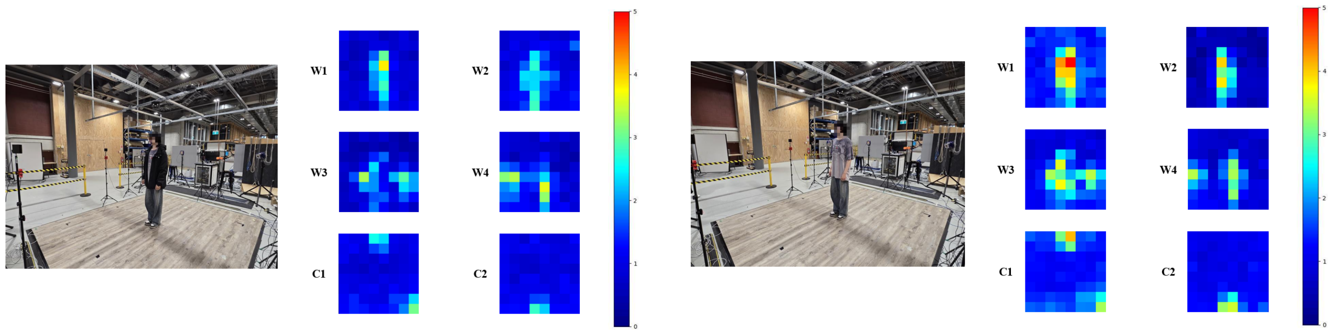
This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2026.3727948

TABLE VI

BENCHMARKING OF FOUR SENSOR ARRANGEMENTS. ALL LOCALIZATION ERROR STATISTICS ARE REPORTED IN METERS AND ARE POOLED-PER-SAMPLE POINT ESTIMATES AGGREGATED ACROSS THREE PARTICIPANT-INDEPENDENT CROSS-VALIDATION FOLDS (LAB) OR THE SINGLE TRAINING RUN (FOYER), WITH 95% BOOTSTRAP CIS IN BRACKETS.

	Comb1 (2 ceiling + left & right wall-mounted sensors)		Comb2 (2 ceiling + front & back wall-mounted sensors)		Foyer	
Metric	Proposed	Faulkner et al. [27]	Proposed	Faulkner et al. [27]	Proposed	Faulkner et al. [27]
RMSE	**0.369** [0.324, 0.414]	0.485 [0.447, 0.527]	**0.327** [0.295, 0.360]	0.410 [0.374, 0.449]	**0.436** [0.421, 0.451]	0.649 [0.619, 0.679]
Mean	**0.242** [0.223, 0.260]	0.348 [0.325, 0.371]	**0.237** [0.223, 0.253]	0.299 [0.280, 0.318]	**0.346** [0.336, 0.356]	0.472 [0.455, 0.489]
Median	**0.177** [0.168, 0.186]	0.250 [0.231, 0.270]	**0.190** [0.183, 0.200]	0.226 [0.219, 0.236]	**0.302** [0.284, 0.318]	0.341 [0.323, 0.360]
P_{90}	**0.444** [0.397, 0.487]	0.680 [0.612, 0.770]	**0.434** [0.397, 0.464]	0.618 [0.547, 0.698]	**0.655** [0.627, 0.676]	0.924 [0.874, 0.978]
P_{95}	**0.727** [0.551, 0.862]	0.999 [0.866, 1.171]	**0.580** [0.491, 0.644]	0.820 [0.744, 0.946]	**0.824** [0.802, 0.837]	1.388 [1.263, 1.521]

Note: For each metric and sensor arrangement, the best (lowest) error between the proposed method and Faulkner et al. [27] is highlighted in bold.



(a) Participant wearing a jacket (b) Participant without a jacket

Fig. 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.

TABLE VII

COMPARISON OF LOCALIZATION SYSTEMS ACROSS COMMON SENSING MODALITIES

Author	Mode	ROI (sqm)	Accuracy (m)
WoW [43]	Visible Light	6.8	Median: 0.12
Fieldlight [16]	Visible Light	28–56	Median: 0.68–1.2
DeepPIRATES [44]	PIR	25–65	Mean 0.37–0.54
Ngamakeur et al. [45]	PIR	9	Mean: 0.2379
SpringLoc [46]	RF (RSSI)	25–44	Median: 0.6–1.2
Hi-loc [10]	RF (CSI)	120	Mean: 0.65–1.03
Mirshekari et al. [12]	Vibration	20	Median 0.38
Alajlouni et al. [47]	Vibration	26	80 pct: 0.7
CapLoc [14]	Capacitive	2.9	Median: 0.026
Tariq et al. [15]	Capacitive	3.24	Mean: 0.307, RMSE: 0.251
Proposed	Thermopile	28.2	Median: 0.30, Mean: 0.35, RMSE: 0.44

ity check, confirms that the median effect is statistically significant ($p \leq 0.040$; effect sizes of roughly +40 to +100 mm) for both held-out participants under the ceiling-mounted (Comb3) and four-sensor (Comb1) configurations. For the wall-mounted (Comb4) configuration, the median effect on individual subjects does not consistently reach significance. Heavy clothing such as a jacket attenuates IR radiation from the human body, reducing the thermal contrast captured by the thermopile sensors particularly as the sensor–target distance increases. The resulting degradation is concentrated in the tail of the error distribution: the P_{90}–P_{95} metrics increase substantially across all three sensor configurations, indicating that wearing a jacket reliably increases the likelihood of large localization errors.

D. Ablation Study

As discussed in Section IV-D, combining convolutional feature extraction with temporal modelling consistently improved robustness relative to purely recurrent architectures. Two complementary ablations probe what the CNN layers contribute. In a pixel-shuffle ablation (randomly permuting the 64 pixels within each 8×8 frame before training and inference), the median localization error changed by less than 10 mm across configurations (between -9 and $+5$ mm), i.e. essentially identical performance. The convolutional layers, therefore, do not exploit local 2D spatial adjacency as conventional natural-image CNNs do. Rather, they act as a complementary nonlinear feature extractor over structured thermal distributions while the LSTM captures the dominant temporal dynamics. A complementary channel-shuffle ablation (permuting the sensor-to-channel ordering while preserving each 8×8 frame) produced only a modest degradation for the four-sensor configuration (+8.3% median) and negligible change for the two-sensor configurations. This indicates that

TABLE VIII

ARCHITECTURAL ABLATION STUDY ON THREE REPRESENTATIVE LABORATORY CONFIGURATIONS. THE RESULTS ARE COMPUTED FROM THE COMBINED PREDICTIONS OF THREE SUBJECT-DISJOINT CROSS-VALIDATION FOLDS.

Architecture	Median (m)	P_{95} (m)
Comb1 (4-sensors : 2 ceiling + left & right wall-mounted)		
Ridge regression	0.338	0.994
MLP	0.490	1.160
2D CNN only	0.410	1.242
LSTM only	0.256	0.731
1D CNN over time	0.384	1.080
Transformer encoder	0.831	1.424
Proposed CNN–LSTM	**0.177**	**0.727**
Comb3 (2 ceiling-mounted sensors)		
Ridge regression	0.520	1.207
MLP	0.486	1.358
2D CNN only	0.397	1.144
LSTM only	0.245	0.832
1D CNN over time	0.394	1.186
Transformer encoder	0.827	1.432
Proposed CNN–LSTM	**0.171**	**0.768**
Comb4 (2 wall-mounted sensors)		
Ridge regression	0.448	1.255
MLP	0.481	1.083
2D CNN only	0.354	0.945
LSTM only	0.260	0.800
1D CNN over time	0.457	1.167
Transformer encoder	0.827	1.432
Proposed CNN–LSTM	**0.198**	**0.732**

Note: Seven architectures are trained under identical splits. Lowest median and P_{95} per configuration in **bold**.

TABLE IX

FOYER-ENVIRONMENT ARCHITECTURAL ABLATION RESULTS.

Architecture	Median (m)	P_{95} (m)
Ridge regression	$1.784^{\dagger}$	$2.796^{\dagger}$
MLP	0.737	1.615
2D CNN only	0.375	1.255
LSTM only	0.381	1.339
1D CNN over time	0.391	1.680
Transformer encoder	$1.849^{\dagger}$	$2.633^{\dagger}$
Proposed CNN–LSTM	**0.302**	**0.824**

Note: † indicates collapse to near-mean prediction. The bottom row shows the proposed CNN–LSTM reported in the main text.

TABLE X

IMPACT OF CLOTHING ON LOCALIZATION PERFORMANCE. ALL LOCALIZATION ERROR STATISTICS ARE REPORTED IN METERS.

Metric	Jacket On	Jacket Off
Comb1 (2 ceiling + left & right wall-mounted sensors)		
RMSE	0.33	0.26
Mean	0.27	0.21
Median	0.21	0.18
P_{90}	0.52	0.40
P_{95}	0.62	0.49
Comb3 (2 ceiling-mounted sensors)		
RMSE	0.42	0.29
Mean	0.33	0.21
Median	0.26	0.18
P_{90}	0.62	0.36
P_{95}	0.73	0.45
Comb4 (2 wall-mounted sensors)		
RMSE	0.35	0.27
Mean	0.29	0.22
Median	0.25	0.19
P_{90}	0.57	0.42
P_{95}	0.70	0.46

fixed sensor identity contributes only weakly and that temporal information remains the dominant discriminative cue.

To validate the choice of the CNN–LSTM family, seven representative architectures were trained on the running benchmark set of three laboratory sensor configurations (Comb1, Comb3, and Comb4) introduced at the start of Section V, under identical training, validation, and test splits. The results are reported in Table VIII. The proposed CNN–LSTM achieves the lowest median errors, with the LSTM-only model a clear second and the purely spatial architectures (2D-CNN-only, MLP, and Ridge regression) together with the Transformer encoder producing substantially larger median and tail errors. This confirms that temporal modelling is the dominant contributor to localization accuracy.

The analogous foyer ablation (single $16 \times 16 \times 1$ image, saved $60/20/20$ split) is reported in Table IX. The same architectures were trained from scratch on the foyer setup under identical training budgets to the lab ablation in Table VIII. On the foyer input geometry, the higher-capacity lab-variant networks collapse to near-mean prediction (median error close to the geometric centroid distance of 1.85 m) and are therefore omitted from Table IX, whereas the hybrid foyer variant introduced in Section IV-D attains a median of 0.302 m, providing empirical justification for the dual-variant design.

E. Computational Scalability and Real-time Performance

The proposed 2D CNN–LSTM contains approximately 2.7×10^6 trainable parameters in the four-sensor configuration, of which about 96% lie in the fixed-size pre-LSTM fully connected layer and the 256-unit LSTM; both are independent of the number of sensors N_c, so doubling N_c from four to eight increases the parameter count by less than 0.1%. Consistent with this, the median per-sequence inference latency on the workstation GPU is essentially unchanged (1.62 ms at N_c=4 versus 1.64 ms at N_c=8), and across workstation, desktop, and embedded (NVIDIA Jetson Xavier NX) targets the network meets the 111 ms real-time budget with between $30\times$ and $68\times$ headroom. On commodity CPUs the latency approaches the sampling period and would benefit from model compression.

Communication overhead from each Grid-Eye sensor is approximately 2.3 kB/s in 32-bit float representation (0.6 kB/s with 8-bit quantization), so even deployments of order 30 sensors require well under 100 kB/s of aggregate bandwidth. The architecture and I/O pipeline therefore both scale efficiently with sensor count, and the real-time margin established here extends to moderately larger deployments without changes to the model or communication layer; beyond approximately 30 sensors, sensor-selection or attention mechanisms and a transition from a star USB topology to Ethernet or wireless connectivity would be beneficial (see Future Work).

VI. 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.

A. Summary of Findings

For the laboratory setup covering $2.4\,\mathrm{m} \times 2.4\,\mathrm{m}$, no single sensor arrangement was uniformly superior across all localization metrics. While the four-sensor configurations (Comb1 and Comb2) achieved some of the lowest RMSE, P_{90}, and P_{95} errors, indicating improved robustness against large localization failures, several two-sensor configurations exhibited comparable or even superior median localization accuracy. In particular, the ceiling-mounted Comb3 configuration achieved a median error comparable to the four-sensor arrangements, whereas the wall-mounted Comb4 configuration exhibited a somewhat higher median error. The proposed method consistently outperformed algorithms reported in earlier studies, particularly in reducing large localization errors as reflected by the improved P_{90} and P_{95} statistics.

Several practical deployment observations follow from these results. Configurations with overlapping fields of view consistently exhibited lower P_{90} and P_{95} errors, as overlap reduces blind spots and large-error events. Ceiling-mounted configurations performed consistently well as overhead placement maintains uniform target visibility. Also from a real-world perspective, such arrangements are more robust to occlusion by furniture or other occupants, and are preferable in larger rooms where wall-mounted sensors can exceed their effective range. Nevertheless, appropriately positioned wall-mounted configurations also remained competitive, and combining complementary ceiling and wall viewpoints reduced geometric ambiguity in the thermal patterns. Overall, practical thermopile deployments should prioritize sufficient spatial coverage, overlapping sensor footprints, and complementary viewing perspectives to minimize large localization errors.

The impact of clothing on localization performance was investigated. A per-subject within-subject analysis on the two held-out participants showed that wearing a jacket produces a statistically significant degradation of the median localization error under the ceiling-mounted and four-sensor configurations, and a substantial degradation of the P_{90}–P_{95} tail metrics across all three sensor configurations. This is consistent with jacket-induced attenuation of infrared radiation emitted by the human body. This finding is important for practical deployments in environments where clothing variations are expected.

The foyer deployment further demonstrated the scalability of the proposed approach. Despite covering a substantially larger sensing area of approximately $28.2\,\mathrm{m}^2$, the system maintained a median localization error of 0.302 m, indicating that useful localization accuracy can be preserved when the sensing network is expanded beyond small laboratory environments. 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.

B. 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.

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.

It should be noted that the reported between-configuration significance analysis was conducted on a representative fold and is intended as a sensitivity study. A comprehensive statistical framework that aggregates uncertainty across all cross-validation folds remains a topic for future investigation.

The experiments reported in this work were conducted under relatively controlled indoor environmental conditions. In long-term real-world deployments, factors such as ambient temperature drift, direct sunlight exposure through windows, HVAC-induced thermal fluctuations, and other environmental heat sources may alter the background thermal distribution and reduce the thermal contrast between occupants and the environment. Such effects could potentially increase localization uncertainty, particularly during prolonged operation. Future work may therefore benefit from adaptive background modelling, periodic recalibration, and environmental compensation strategies to improve robustness under dynamically changing thermal conditions.

Connectivity improvements are possible. While the current USB-based architecture provides reliable high-speed communication and simplified deployment for centralized sensing, it has inherent scalability limitations in terms of cable length and hub size. Standard USB-based communication typically supports cable lengths of approximately 5 m [48]. Ethernet-based connectivity could extend this to 100 m [49], enabling larger-scale distributed sensing deployments. Alternatively, wireless networking using the ESP32 Wi-Fi module, which is already provisioned on the PCB, could support more flexible and easily reconfigurable installations.

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.

Acknowledgment

The authors acknowledge Yan (Jasmine) Wang for her assistance with the hardware and firmware development. Research supported by the NVIDIA Academic Grant Program using the RTX PRO 6000 Blackwell Max-Q GPU. AI tools ChatGPT and Copilot were used for editing and grammar enhancement of the manuscript.

References

[1] 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.

[2] 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.

[3] 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.

[4] 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.

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

[6] 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.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] 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.

[16] 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.

[17] 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.

[18] 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.

[19] N. Neumann and V. Banta, "P12-comparison of pyroelectric and thermopile detectors," in *Proceedings IRS^2 2013*, 2013, pp. 139–143.

[20] 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.

[21] 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.

[22] 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.

[23] M. Kuki, H. Nakajima, N. Tsuchiya, and Y. Hata, "Human movement trajectory recording for home alone by thermopile array sensor," in *2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2012, pp. 2042–2047.

[24] 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.

[25] 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)*. IEEE, 2015, pp. 428–434.

[26] 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.

[27] 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.

[28] 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.

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

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

[31] 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.

[32] 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.

[33] 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.

[34] 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.

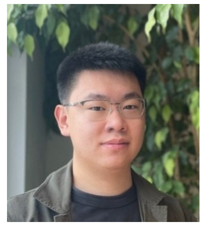
[35] J. L. Honorato, I. Spiniak, and M. Torres-Torriti, "Human detection using thermopiles," in *2008 IEEE Latin American Robotic Symposium*. IEEE, October 2008, pp. 151–157.

[36] 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.

[37] 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.

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

[39] Panasonic, "Amg8833 infrared array sensor grid-eye," Datasheet, 2017, specifications subject to change without notice.
[40] STMicroelectronics, "Getting started with stm32 mcu hardware development," Application Note AN4661, 2021, rev 5.
[41] Marvelmind Robotics, "Marvelmind indoor navigation system operating manual," https://marvelmind.com/pics/marvelmind_navigation_system_manual.pdf, 2024, accessed.
[42] R. Amsters and et al., "Evaluation of low-cost/low-weight indoor positioning systems in warehouse conditions," *Sensors*, vol. 19, no. 20, p. 4503, 2019.
[43] 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.
[44] 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.
[45] 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.
[46] 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.
[47] 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.
[48] USB Implementers Forum, "Universal serial bus specification revision 2.0," USB-IF, 2000, section 7.1.19, Cable Length.
[49] IEEE Computer Society, "Ieee standard for ethernet," IEEE Std 802.3-2018, 2018.



Shengjun (Eric) Ma is a Lecturer at the School of Business Technologies, New Zealand Skills & Education College (NZSE), New Zealand. He received his B.Comp.Sc. and M.Comp.Sc. degrees in 2019 and 2022, respectively, and his Ph.D. degree in Electrical & Electronic Engineering in 2026, all from Auckland University of Technology (AUT), New Zealand. His research interests include smart agriculture, indoor localization, digital signal processing, machine learning, the Internet of Things (IoT), and wireless sensor networks.

Daniel Konings (Senior Member, IEEE) is a Senior Lecturer in Mechatronics at the Department of Mechanical Engineering, School of Engineering, Computer and Mathematical Sciences, Auckland University of Technology, New Zealand. He received his BE (Hons.) and PhD degrees in computer systems and electronics engineering from Massey University, Auckland, New Zealand, in 2014 and 2019, respectively. His research interests include machine learning, engineering education, indoor localization, mobile robotics, IoT & sensor networks.

Edmund M-K Lai (Life Senior Member, IEEE) received his BE(Hons) and PhD from The University of Western Australia, both in Electrical Engineering, in 1982 and 1991 respectively. He is currently Professor in the Department of Data Science and AI at the Auckland University of Technology, New Zealand. He has over 35 years of academic experience, having previously held faculty positions at universities in Australia, Hong Kong, and Singapore. He has published over 150 international refereed journal and conference papers in signal processing, intelligent control, computational intelligence, and artificial neural networks. Dr. Lai is also a Fellow of the Institution of Engineering and Technology (FIET) and a Fellow of Engineers Australia (FIEAust).

Fakhrul Alam (M'17-SM'19) is a Professor at the Department of Electrical & Electronic Engineering, School of Engineering, Computer & Mathematical Sciences, Auckland University of Technology (AUT), New Zealand. He received BSc (Hons) in Electrical & Electronic Engineering from BUET, Bangladesh in 1996 and MS and Ph.D. in Electrical Engineering from Virginia Tech, USA in 1999 and 2002 respectively. His research interests includes indoor localization, IoT & wireless sensor networks, sensor applications.