



Article

A Photovoltaic Light Sensor-Based Self-Powered Real-Time Hover Gesture Recognition System for Smart Home Control

Nora Almania ^{1,2} , Sarah Alhouli ^{1,3}  and Deepak Sahoo ^{1,*} 

- ¹ Department of Computer Science, Swansea University, Swansea SA1 8EN, UK; 2033611@swansea.ac.uk or nalmanee@su.edu.sa (N.A.); 2028812@swansea.ac.uk or sholi@kisir.edu.kw (S.A.)
² Department of Computer Science, Shaqra University, Shaqra 11961, Saudi Arabia
³ Systems and Software Development Department, Kuwait Institute for Scientific Research (KISR), Shuwaikh 13017, Kuwait
* Correspondence: d.r.sahoo@swansea.ac.uk

Abstract

Many gesture recognition systems with innovative interfaces have emerged for smart home control. However, these systems tend to be energy-intensive, bulky, and expensive. There is also a lack of real-time demonstrations of gesture recognition and subsequent evaluation of the user experience. Photovoltaic light sensors are self-powered, battery-free, flexible, portable, and easily deployable on various surfaces throughout the home. They enable natural, intuitive, hover-based interaction, which could create a positive user experience. In this paper, we present the development and evaluation of a real-time, hover gesture recognition system that can control multiple smart home devices via a self-powered photovoltaic interface. Five popular supervised machine learning algorithms were evaluated using gesture data from 48 participants. The random forest classifier achieved high accuracies. However, a one-size-fits-all model performed poorly in real-time testing. User-specific random forest models performed well with 10 participants, showing no significant difference in offline and real-time performance and under normal indoor lighting conditions. This paper demonstrates the technical feasibility of using photovoltaic surfaces as self-powered interfaces for gestural interaction systems that are perceived to be useful and easy to use. It establishes a foundation for future work in hover-based interaction and sustainable sensing, enabling human–computer interaction researchers to explore further applications.

Keywords: visible light sensing; photovoltaic light sensor; hovering hand gesture recognition; machine learning pipeline; real-time model evaluation; smart home control; user experience evaluation



Academic Editor: Mohammad Jafari

Received: 29 July 2025

Revised: 1 September 2025

Accepted: 3 September 2025

Published: 9 September 2025

Citation: Almania, N.; Alhouli, S.; Sahoo, D. A Photovoltaic Light Sensor-Based Self-Powered Real-Time Hover Gesture Recognition System for Smart Home Control. *Electronics* **2025**, *14*, 3576. <https://doi.org/10.3390/electronics14183576>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The integration of gesture recognition systems in smart homes is seen as an innovative and futuristic way to interact with smart home devices. Such systems have emerged recently to control devices in smart home environments through touchless gestures without physical contact, offering effortless and hands-free interaction. Touchless gestural interactions are becoming popular, as users prefer them for reasons of personal hygiene [1–3] and a reduction in the risk of spreading infection [4]. This type of interaction also makes it easier for people to perform their daily tasks [1,5], especially when hands are occupied, such as in the kitchen if hands are dirty or holding objects [6], or for people with mobility impairments or temporary injuries that cause them to struggle to control household appliances [7].

Recently, touchless gestural interaction approaches, such as mid-air interaction sensed by camera-based systems, have been widely investigated for controlling appliances in smart home environments because they can facilitate interaction over a large area [8–11]. However, such sensing technologies can be high in installation costs, bulky, difficult to setup and use, and high in energy consumption, and they cannot be embedded in or installed on many surfaces in a house, leading to a lack of user adoption. Poor and negative user experiences with gesture-sensing technologies can also impact their adoption in smart homes due to unclear system behavior and feedback and confusing device interfaces [12–14].

Recently, photovoltaic (PV) modules have been used in devices such as solar panel-powered sensor systems [15–17] and as self-powered light sensors [18–22]. While harvesting energy from ambient light, PV modules can be used simultaneously to detect changes in light intensities due to shadows made by hovering hand movements above them. Different hand gestures can create unique temporal signatures as modulations in the electrical voltage or current, which can be captured as an alternating voltage signal in parallel with the charging direct current [6,23,24]. PV light sensors are becoming a promising solution for gesture recognition systems due to their battery-free and self-powered technology, which harvests energy from ambient light and can prolong the lifespan of sensor systems and sustain continuous energy consumption [18–21]. These sensors can be flexible, rollable, and portable and can be embedded in various surfaces throughout the home, making them ideal for indoor and outdoor applications due to their ability to function under natural and artificial light conditions [25]. PV sensors have been integrated with everyday objects like wearable devices [26] and interactive surfaces [6], leading to user-friendly and self-sustained electronic systems. Gesture recognition over a large area can be enabled by embedding PV light sensors into various surfaces, making them an attractive alternative to conventional sensor technologies [23]. This integration provides interactive interfaces that facilitate a seamless and aesthetic interaction, enhancing user experience (UX).

Hand gesture recognition using self-powered sensors is becoming increasingly common in human–computer interaction (HCI) research [18,22,27–29]. Supervised machine learning (ML) models are typically used in gesture recognition systems to classify different gesture patterns [21,25]. Recently, suitable semi-supervised and deep learning methods have been explored [22,30]. Previous studies have primarily focused on offline evaluation of ML models for gesture recognition. A detailed examination of the various steps of a suitable ML pipeline, including data collection and preprocessing, gesture engineering, and model selection, evaluation, and deployment, for real-time prototype development is lacking. UX evaluation of gestural interfaces often does not use real-time working prototypes which could provide a more realistic experience. Smart home interaction using a real-time PV sensor-based prototype has not been demonstrated before. In this paper, our work addresses these gaps. Specifically, we contribute the following:

- A development of a novel supervised machine learning pipeline for a smart home hover gesture recognition system.
- A feasibility demonstration of a PV light sensor-based gesture recognition system comparing one-size-fits-all and user-specific models with offline and real-time testing studies.
- An end-user evaluation of user experiences with the real-time gesture recognition system for smart home device control.

2. Related Work

2.1. Self-Powered Sensor Technologies for Gesture Recognition

Self-powered sensor technologies for gesture recognition can use energy harvesting modules to power the gesture recognition system. Alternatively, the energy harvesting

module can serve as the gesture sensor itself, or if the sensor is self-powered, it does not require external power. For example, Shahmiri et al. (Serpentine) used triboelectric nanogenerator cord sensors that are self-powered and can be deformed reversibly to measure various human inputs [29]. Wen et al. used self-powered conductive superhydrophobic triboelectric textile sensors to track finger and hand motions for gesture recognition, which are cheap and lightweight for applications such as virtual or augmented reality (VR/AR) applications [31]. However, these sensors do not generate adequate energy to power the entire gesture recognition system. Energy harvesting sensors that can be used for gesture recognition can use kinetic energy from human activities or actions [32–34] and ambient radio frequency energy [35]. In this work, we focus on an ambient light harvesting sensor, i.e., a PV light sensor.

2.2. Light-Sensor-Based Gesture Recognition

Initial studies focusing on light sensors for gesture recognition used different photodiode (PD) array configurations and specialized located or modulated light, while later studies used ambient light [36]. Li et al. (LiSense, StarLight) and Venkatnarayan et al. [27] used multiple PDs and light-emitting diode (LED) panels in a room to reconstruct body pose or gesture [37,38]. Kaholokula et al. (GestureLite) used a 3×3 array of PDs and detected 10 hand gestures using ambient light [28]. Hu et al. and Mahmood et al. used 10 and 16 PDs, respectively, and detected 8 and 17 hand gestures, respectively, in ambient light using Arduino boards [39,40]. Li et al. estimated the hand pose in real time using an LED panel in a lampshade and a 5×5 array of PDs [41]. These works used signals from each PD in the array to recognize the gestures, which were computationally complex. Zhu et al. used the signal from a single PD but used a modulated LED light source emitting light of four different frequencies [30].

An initial study focusing on PV modules as light sensors used six PV modules to detect finger touches and swipes via their partial shadowing [42]. The researchers used the signal from the series-connected PV modules and also demonstrated self-powered, battery-free operation. Li et al. also used PDs in the PV mode to harvest ambient light energy for self-powered, battery-free devices [26]. Recently, Li et al. (SolarSense) used a 5×5 array of small PV modules in a hand-size device and detected 12 hand gestures, using signals from each sensor. Meena et al. used signals from 12 semi-transparent PV cells to detect four hand gestures (swipe left, right, up, and down) [6]. Varshney et al. used the signal from a small PV module and detected three hand gestures (swipe, two taps, and four taps) [43]. Ma et al. used the signal from a small transparent PV cell and detected five hand gestures (close–open palm, left–right swipe with horizontal palm, left–right swipe with vertical palm (wave), up, and up–down) [44]. Sorescu et al. used the signal from large PV film and detected six hand gestures (swipe clockwise, swipe counterclockwise, left–right swipe once and continuously, move up–down, and close palm or make a fist) [23]. These systems either use signals from multiple PV light sensors, which is computationally complex, or offer a limited number of gestures, which may not be suitable for applications such as controlling a smart home device. We address these issues in this work.

2.3. Machine Learning Pipeline for Gesture Recognition Using PV Light Sensors

The machine learning pipeline, e.g., data collection and preprocessing, gesture engineering, and model selection, evaluation, and deployment, in the recognition system is often influenced by the sensor [45], such as the PV light sensor. In a gesture recognition system, the voltage from the PV modules is often sampled directly to capture the gesture data. Alternatively, a resistor across the PV module ($50 \text{ k}\Omega$ [21,44]) or a shunt resistor in series [18] can be used. Sorescu et al. used a capacitor ($1 \text{ }\mu\text{F}$) to capture the alternating voltage

signal without affecting the direct current charging, which we used in our work. Gesture recognition studies often recruit one [23,44] or a few participants to perform the gestures and capture a gesture dataset. PV sensor signals are often sampled at lower frequencies for a few seconds, for example, at 500 Hz for 1 s [21,44]. Some studies sampled the data at higher frequencies and then downsampled it, e.g., Li et al. downsampled 5000 data points to 128 data points [19]. The gesture execution timing is often controlled to capture the entire gesture within a data window. An overlapping windowing approach can be used to capture gestures in a data window. Data preprocessing techniques include filtering/denoising (constant false alarm rate (CFAR) and global light-change filtering [26]), scaling, (Z-score) normalization/standardization, and segmentation [19]. Popular feature engineering techniques include statistical features [21,44], entropy, dynamic time warping (DTW), wavelet transformation, rasterization, and principal component analysis (PCA). The key results of the model selection studies are as follows: Ma et al. achieved 92–95% accuracy using KNN, SVM, DT, and RF, with RF achieving the highest accuracy. Venkatnarayan et al. achieved 96.36% accuracy with SVM. Li et al. (PowerGest) investigated KNN, SVM, and RF and achieved a gesture recognition accuracy from 96.2% to 98.9%, with SVM achieving the highest accuracy [19]. Thomas et al. achieved 97.71% using SVM [18]. Sorescu et al. investigated various ML models and reported 97.2% accuracy using RF. There is no ML model in the literature that significantly outperforms the others. Therefore, for our experimental prototype, we investigated various ML models and techniques mentioned above.

Although previous research has demonstrated the feasibility of ML-based gesture recognition systems using PV sensors, employing various data sampling, preprocessing, feature extraction, and model selection techniques, there is a lack of research on developing real-time gesture recognition systems through a comprehensive ML pipeline. In addition to this, we used user-elicited gestures for the PV sensor and created a gesture dataset with multiple users; demonstrated the limitations of a one-size-fits-all model; developed a user-specific model development and real-time evaluation approach; and demonstrated the usability of the real-time prototype in a user experience experiment.

3. A Self-Powered Real-Time Gesture Recognition System

We envision a real-time gesture recognition system that uses a large-area PV film or panel that acts as a light sensor to capture gestures and to power a microcontroller board for self-powered operation. In this work, we used InfinityPV solar tape (InfinityPV, <https://infinitypv.com>, accessed on 2 September 2025), which harvests 6 mA (30 mW) continuously in indoor diffused light at a level of 300 lux [23], sufficient to power low-power microcontrollers, such as the MSP430 (400 μ A), nRF54 (3.2 mA), ESP32 (4 mA), and STM32 (21 mA) in active mode, as well as low-power analog-to-digital converters (ADCs) (1.4 mA) and Bluetooth low-energy (BLE) communication (9.8 mA) for smart home control. A battery or super-capacitor may be used to store energy to provide extra power during computation and communication. The maximum ratings of the PV sensor are 57.6–62.4 V, 40–50 mA, and 1440 mW. In this work, we focus on the feasibility of real-time gesture recognition using machine learning for future deployment of such a system. We explore using a digital oscilloscope to collect gesture data and a laptop personal computer (PC) for real-time computation.

Next, we present a hover gesture set that we selected from the literature for the proposed system. Then we present a hardware prototype with a graphical user interface (GUI) developed for user experiments to investigate the deployment of the proposed gesture recognition system. This includes evaluation of five different supervised machine learning models for gesture classification, including associated data collection and preprocessing

issues and offline and real-time testing using one-size-fits-all and user-specific models through two user studies.

3.1. Hover Gestures for PV Light Sensors

We selected 11 dynamic hand gestures that frequently appeared in mid-air gesture elicitation studies. These gestures are considered to be fairly easy to learn and remember as well as comfortable and enjoyable for smart home control [10,46–50]. These gestures have been proposed for hover-based input using PV sensors [18,19,23,24]. The 11 hand gestures which are shown in Figure 1 are as follows: (G1) move down; (G2) move up; (G3) move right; (G4) move left; (G5) close fist; (G6) move forward; (G7) move clockwise; (G8) move counterclockwise; (G9) move down quickly; (G10) split hands; and (G11) combine hands. These gestures are effective for the proposed gesture recognition system because they can be performed with a variety of natural hand motions that do not require excessive arm movement, thus preventing hand fatigue. The space above the PV sensor allows for performing these gestures, and the sensor's light and shadow sensing capabilities generate gesture signals with unique patterns specific to the gestures that can be efficiently detected by ML algorithms.

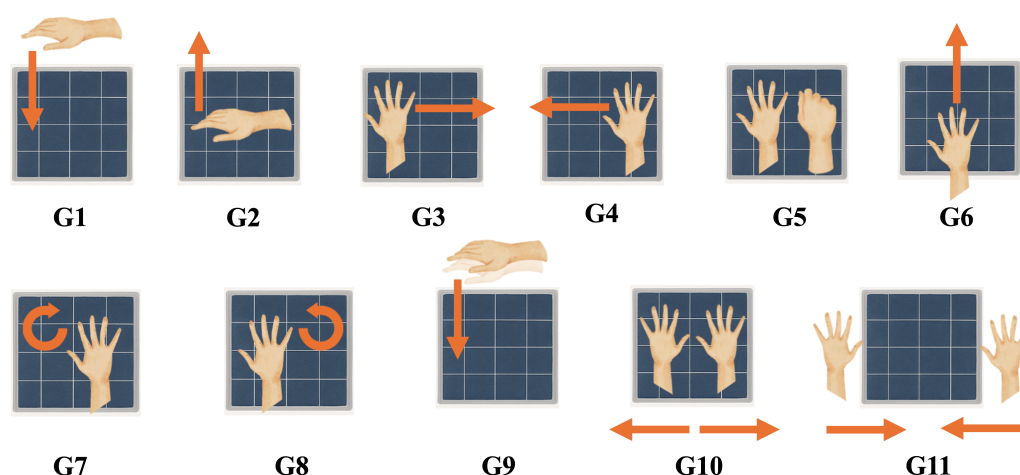


Figure 1. The selected 11 hover gestures—(G1) move down; (G2) move up; (G3) move right; (G4) move left; (G5) close fist; (G6) move forward; (G7) move clockwise; (G8) move counterclockwise; (G9) move down quickly; (G10) split hands; and (G11) combine hands—are shown.

3.2. Experimental Prototype

The experimental prototype is shown in Figure 2. In this setup, InfinityPV solar tape is used as the PV light sensor, and a digital oscilloscope (PicoScope, PicoScope 5442D, <https://www.picotech.com/oscilloscope/5000/flexible-resolution-oscilloscope>, accessed on 2 September 2025) is used to log the data. A laptop PC is used for model training and for real-time model deployment and testing and gesture recognition. It also runs a GUI application that is used to perform these tasks interactively. The GUI includes a music player demo that users can interact with in real time to experience controlling a smart home device with hover hand gestures above the PV sensor. A PC monitor is used to show the GUI, and the music player output is played on a speaker. We also used a light meter to measure the intensity of light in the room.

The GUI was developed using the Python built-in library Tkinter for easy integration with the software development kit of PicoScope, PicoSDK (PicoSDK-Python, <https://github.com/picotech/picosdk-python-wrappers>, accessed on 2 September 2025) and real-time testing of machine learning models. It was used for data recording and model training as well as to display the gesture signal and the model-recognized gesture in real time. It

also includes a music player demo that users can interact with to control various music player features (see Table A1) in real time with hover gestures above the PV sensor.

The data was transmitted over a USB and handled via the PicoSDK C API using Python's ctypes interface. The data logging system was set up to collect 15,000 samples for each gesture, with a buffer size of 500 samples and 30 buffers in total. The sampling rate was 3000 samples per second, and data was captured for 5 s. The end-to-end latency includes 5 s of data capture and less than 2 s of processing, visualization, and prediction, resulting in a total system response time of approximately 6–7 s.

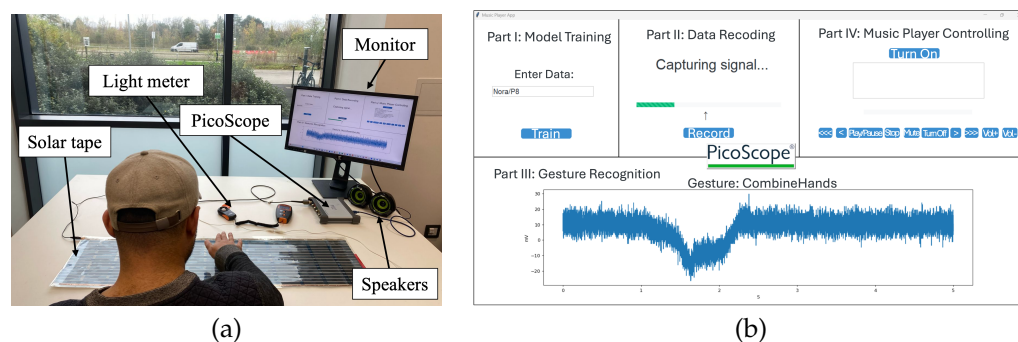


Figure 2. The experimental setup with (a) the hardware prototype and (b) graphical user interface (GUI) developed for the user studies is shown. The main hardware components are InfinityPV solar tape, a PicoScope digital oscilloscope, a PC, a PC monitor, speakers, and a light meter. The main sections of the GUI are data recording, model training, real-time gesture recognition, and music player control.

4. One-Size-Fits-All Model Testing

We conducted a first user study to explore the performance of a “one-size-fits-all model” for real-time deployment for the proposed gesture recognition system [24]. We recruited 48 participants (27 female, 21 male), aged 26–35 (mean 28.4, SD 4.2) years from the university campus and collected a dataset of the 11 hovering gestures. We conducted the experiment in a laboratory environment near a glass window with open blinds and with ceiling lights switched on. As a result, the indoor light level varied in the experiment. We collected 7 samples per gesture from each participant, i.e., $7 \times 11 \times 48 = 3696$ gesture samples.

In the experiment, the raw signal data was collected at a sampling rate of 3000 Hz over 5 s, giving 15,000 data points per gesture sample. Spectrum analysis on the oscilloscope showed that the gesture signals were contained below 5 Hz. Considering the corresponding Nyquist frequency of 10 Hz and a 5-second gesture capture window, the minimum number of samples required is 50. We collected the data at a relatively high sampling rate of 3000 Hz to determine an appropriate number of data points after downsampling required for high recognition accuracy and with low computational load. First, we smoothed the 3 kHz data using a low-pass averaging filter with a rectangular window size of 100 data points. Then, we resampled the data at 600 Hz, reducing 15,000 data points in five seconds to 3000 data points in five seconds, which yielded good recognition accuracy in preliminary attempts. Therefore, the gesture signals were oversampled by a factor of at least 60. In total, 3000 data points per gesture sample were used for ML modeling and testing. We used the Python `scipy.signal.resample()` function, which has a built-in anti-aliasing filter and preserves the signal characteristics. Decimating the smoothed signal would be a simpler approach, and further downsampling may be explored to reduce the computational load for real-time recognition.

In the preprocessing phase, we applied the data augmentation techniques of noise injection, time shifting, and scaling [51,52], which tripled our dataset from 3696 to 11,088 samples. We also employed magnitude normalization using MinMax scaling to generalize the model. For feature extraction, we used statistical features, crossing features, entropy [53],

and multilevel 1D DWT [54], resulting in a total of 144 features. We split the data into 70% for training the models and 30% for testing.

We then trained five ML classifiers—K-nearest neighbor (KNN), decision tree (DT), random forest (RF), gradient boosting (GB), and logistic regression (LR)—and compared them via offline testing. LR had the poorest accuracy (38%), followed by KNN (56%), DT (77%), and GB (79%). RF achieved the highest accuracy of 97%. Figure 3 shows the confusion matrices for the five models. We concluded that we should explore the RF classifier for developing the proposed real-time gesture recognition system.

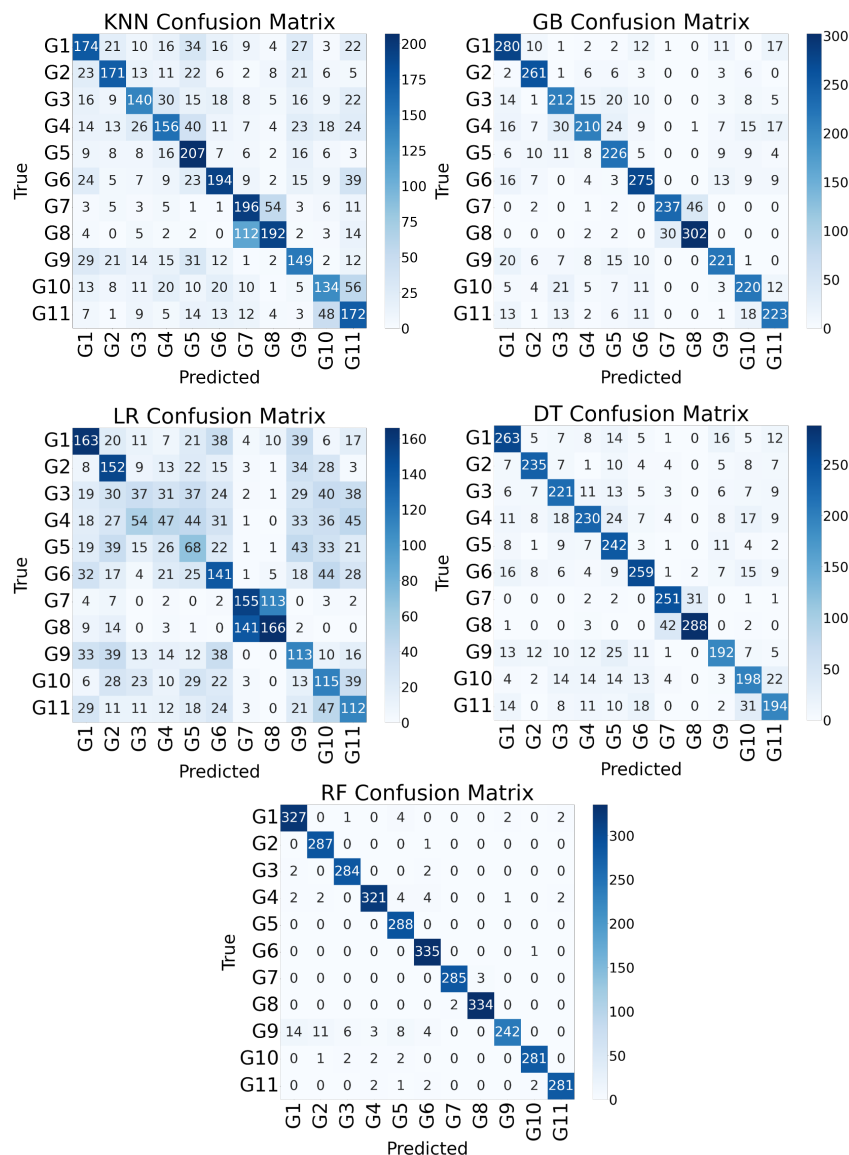


Figure 3. The confusion matrices for the five models (KNN, DT, LR, GB, and RF) to recognize the 11 hover gestures are shown.

We then tested the RF model as a one-size-fits-all model for real-time deployment. For the real-time testing study, we recruited participants different from the original 48 participants as the original testing data was generated by these 48 participants. The new participants performed the 11 gestures in a random order using the same experimental setup, and the gesture signal was collected and processed as in the original experiment. However, we encountered low accuracies of 45%, 54%, and 63% with the first three participants, which was much lower than the original accuracy of 97%. We concluded that a one-size-fits-all model may not be suitable

for real-time gesture recognition. A gesture dataset created by a larger number of participants may be required to record more user-specific variations in gesture signals.

We therefore analyzed the variability in the gesture signal in the data. Gesture G10 (split hands) is the second most misclassified gesture after gesture G9 (move down quickly). Figure 4 shows the normalized signals for G10. It shows the noticeable variations in signal across different participants (see Figure 4a). We consistently observed for all gestures that the signals could vary noticeably across users and would look similar for each user (see Figure 4b). In this preliminary study, we concluded that despite gesture signal variability across the users, the one-size-fits-all model could have a high accuracy, yet the real-time test could show relatively lower accuracy. This led us to investigate user-specific RF classifiers that are trained using a smaller number of gesture samples in real time for each user for their individual use. This approach could be used for first-time setup or to recalibrate the gesture recognition system in little time.

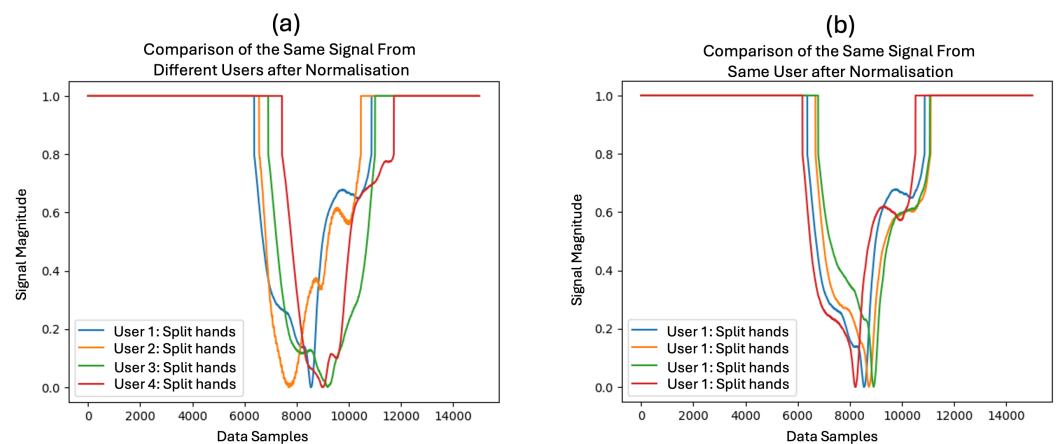


Figure 4. Gesture signals of G10, “split hands”, showing the general trend of (a) variability across users and (b) similarity for a user.

5. Real-Time Testing User Study

We conducted a second user study to test the real-time recognition of gestures using a self-powered PV light sensor with user-specific RF models, as well as to evaluate the usability and acceptability of the proposed gesture recognition system for controlling a smart home device. The experimental setup used with the hardware prototype and the real-time GUI is shown in Figure 2.

We recruited 10 participants (6 female, 4 male), aged 26–35 (mean 31.5, SD 3.8) years, from the university campus through an email invitation. All participants were right-handed with no prior experience of hover hand gestures or touchless interaction. The second user study comprised two sessions: a real-time gesture testing session and a smart home device user experience evaluation session. The entire study took approximately one hour. First, we welcomed the participants and asked them to read the information sheet, sign the consent form, and provide their demographic information, which took about 10 min.

5.1. Real-Time Testing Session

This session took approximately 35 min. We began by showing a training video on how to perform the 11 hover hand gestures, which took around 5 min. We then asked the participants to perform the 11 gestures five times in a random order, which took around 15 min. Thus, we collected 55 gesture samples as training data from each participant. The gesture signals were collected using the PicoScope for 5 s at 3000 Hz, yielding 15,000 data points. To reduce the number of data points for real-time implementation, we smoothed the data using a rectangular sliding window size of 500 data points and resampled at 50 Hz to

decrease the number of data points from 15,000 to 250. As the gesture signals lay below 5 Hz, with the Nyquist frequency of 10 Hz, the gesture signals were oversampled by a factor of 10.

In the preprocessing phase, we used the time-shifting data augmentation technique. We did not use the noise injection and scaling techniques, which were used in the previous study. We did not use the noise injection technique as the data was smoothed. We did not use the scaling technique as the data was user-specific. These techniques may not be required for real-time model development and testing used in this study; however, they may be required for long-term deployment. The time-shifting approach produces a variety of temporal representations that more accurately simulate real-world variations in gesture signals for a user. This prevents overfitting to specific timing patterns and creates a more robust model that can handle the time variability while capturing naturally performed human gestures in real time. Therefore, we used random time shifting of 0–250 data points four times to augment the data from 55 to 220 samples.

For the feature extraction phase, we selected the wavelets feature from the DWT, which increased the number of features from 144 in the previous study to 17,408. This approach could address situations where a small number of samples are collected for quick modeling and the greater variability introduced by real-world deployment conditions. We split the data into 80% for training and 20% for testing and used 5-fold cross-validation to create an RF model. It took less than a minute on a laptop PC with an i7 (2.8 GHz) processor and 8 GB of RAM for each participant. In a commercial product, this short training/calibration step could be delegated to a remote server via a smartphone app to augment the data and achieve quicker results. Model reduction techniques may be required to deploy user-specific models on a microcontroller which has limited RAM and flash memory, and to reduce the computation time.

For real-time testing, the participants performed the 11 gestures in a random order using the same experimental setup, and the gesture signal was collected and processed as mentioned above in this user study. It took approximately 9 min. The GUI was used to collect, display, and process the data, as well as to display and save the results automatically. We then asked the participants to control the music player demo on the GUI using real-time gestures on the PV sensor for a further 5 min.

Real-Time Testing Results

The offline and real-time testing accuracies of the 10 user-specific models for the 10 participants are shown in Table 1, along with the light intensity values recorded during the (offline) data collection and real-time testing. Light intensities differed between the data collection and real-time testing for each participant, with a maximum change of 300 lux. The offline testing accuracy of each user-specific model for each participant is the average value following k(5)-fold cross-validation. The average offline testing accuracy of 10 user-specific models is 82%. P9 achieved the highest offline and real-time testing accuracies of 89.09% and 100% at 1513 lux and 1400 lux, respectively. P3 achieved the lowest offline and real-time testing accuracies of 67.36% and 64% at 1788 lux and 1495 lux, respectively. P6 and P8 achieved higher real-time accuracy of 100% with slightly higher light intensities, while P2 achieved a higher real-time accuracy with a slightly lower light intensity. The confusion matrices of the 10 user-specific models for the 10 participants during offline and real-time testing are shown in Figure 5.

In the real-time testing, 4 out of 10 participants (P2, P6, P8, and P9) achieved 100% accuracy, while P3 attained the lowest at 64% (see Table 1). This participant had misclassification with four gestures, which were G1: move down; G2: move up; G9: move down quickly; and G10: split hands. The remaining five participants had misclassification with G2: move up, with (freq. = 2); G3: move right, with (freq. = 3); G4: move left, with

(freq. = 1); G7: clockwise, with (freq. = 2); and G9: move down quickly, with (freq. = 2). The average real-time test accuracy for all 10 participants was 87.4%. G7 (clockwise) and G8 (counterclockwise) gestures were misclassified the most in offline testing for 7 out of 10 participants. G7 and G8 gestures were misclassified in real-time testing for 3 out of 10 participants and they were also misclassified in the offline test.

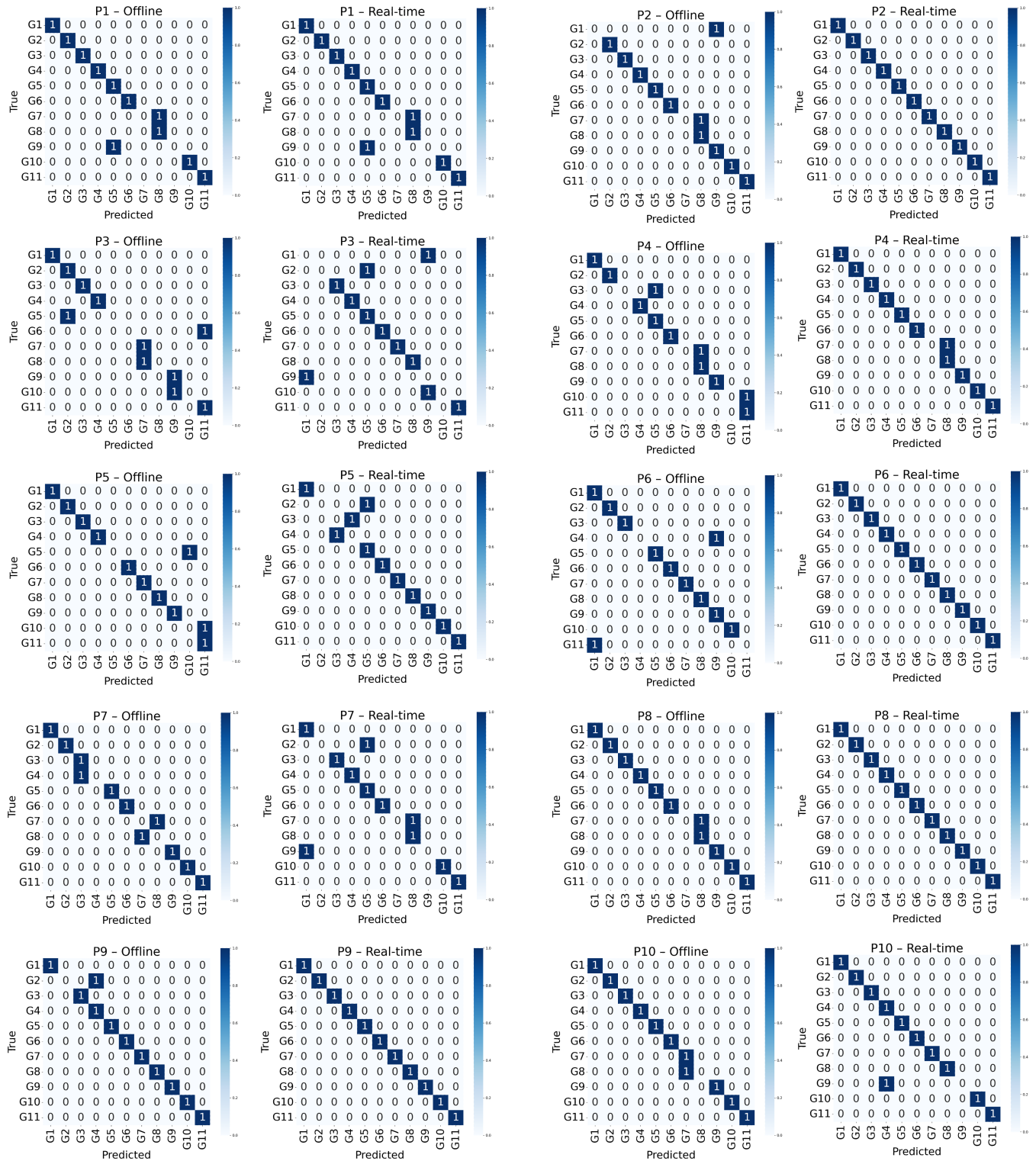


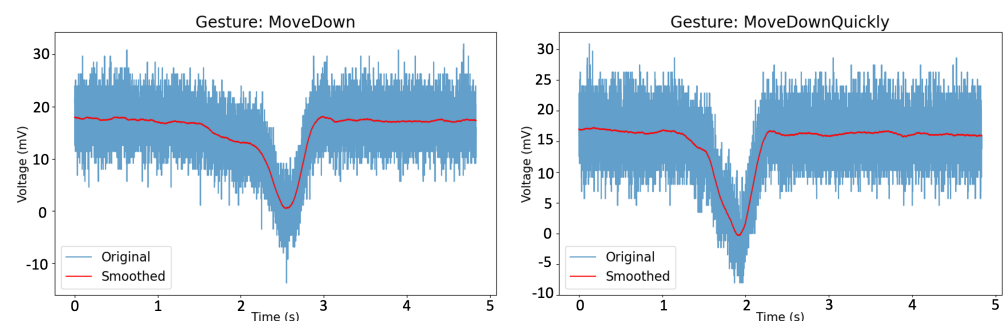
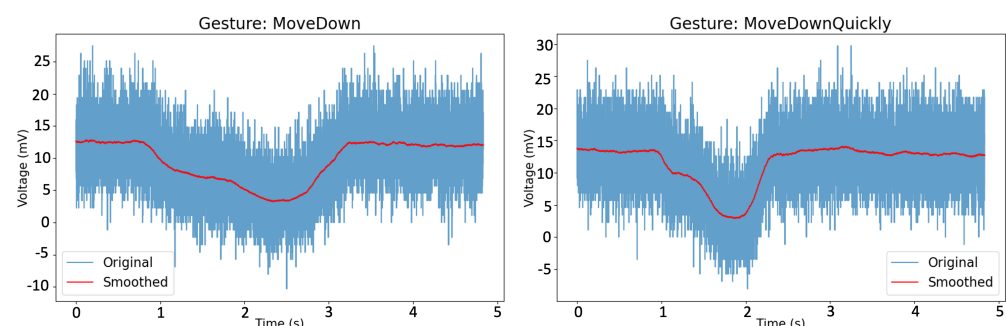
Figure 5. Confusion matrices of the 10 user-specific models for the 10 participants in offline and real-time testing are shown.

Table 1. Offline and real-time testing results of user-specific models.

Participant	Light Intensity (Lux) Data Collection	Light Intensity (Lux) Real-Time Testing	Offline Accuracy (%)	Real-Time Accuracy (%)
P1	1270	1075	83.64	82
P2	1966	1854	87.27	100
P3	1788	1495	67.36	64
P4	1430	1240	74.55	91
P5	1566	1780	85.45	73
P6	1725	1836	81.82	100
P7	1522	1231	69.09	73
P8	1224	1390	87.27	100
P9	1513	1400	89.09	100
P10	1304	1152	85.45	91

We found that the gesture signals of the four participants who achieved 100% real-time accuracy during testing were similar to their training data. In contrast, the gesture signals of the participants with lower real-time accuracy during testing varied noticeably from their training data. Because we collected only five training samples, this issue may arise for participants who can execute gestures with more than five variations or who execute the gestures differently after the training session. We observed that some participants performed the gestures carefully during data collection but were more relaxed afterward in terms of speed and the position and angle of their hand above the sensor. More training with the sensor could have resulted in a closer matching of gesture execution during data collection and real-time testing. Next, we present a similar mismatch in gesture execution.

G1 (move down) and G9 (move down quickly) are semantically similar gestures. However, they were misclassified only once for one participant (P2). Figure 6 shows that G1 and G9 signals look similar for P2, that is, they did not perform G9 significantly quicker than G1. Figure 7 shows the G1 and G9 signals for P9, which look different from each other and consequently are correctly classified.

**Figure 6.** G1 (move down) and G9 (move down quickly) raw and smoothed signals for P2, which look similar to each other.**Figure 7.** G1 (move down) and G9 (move down quickly) raw and smoothed signals for P9, which look different from each other.

We tested whether there was a significant difference in the distribution of offline and real-time misclassifications for each participant using the user-specific model by applying a Chi-square test [55]. We found that there was no statistically significant difference across all 10 participants, with p -values > 0.01 . We tested whether there was a significant difference between the offline and real-time testing accuracies of the user-specific models across the 10 participants. We performed a paired-samples t -test [56] for normally distributed and within-group samples and found no significant differences, with $t = -2.036$ and p -value > 0.01 . These results suggest that the real-time testing accuracy of user-specific models closely reflects their offline testing accuracy. We also tested whether the light intensity affects recognition accuracy in offline and real-time conditions. We applied a linear mixed-effects model [57] to examine the fixed effects of light intensity and condition (offline versus real-time), as well as their interaction. We found that there is no significant effect of light intensity on accuracy in either offline or real-time testing, with all p -values > 0.01 , no significant main effect of condition ($z = -0.11, p = 0.91$), no main effect of light intensity ($z = -0.14, p = 0.89$), and no significance in their interaction ($z = 0.41, p = 0.68$). Therefore, we conclude that the accuracy remained stable across lighting intensities for both offline and real-time conditions and that differences were due to individual variability in performing the gesture rather than illumination.

We had tested a user-specific model in different locations with relatively low to normal indoor light intensities—a kitchen (400 lux), a lab with closed blinds (750 lux), a living room (1000 lux), and a lab with open blinds (1500 lux)—before the user study. The gesture data was captured and model-tested at the respective locations. The respective real-time testing accuracies were 69%, 78%, 83%, and 85%. The confusion matrices are shown in Figure 8. The gesture recognition accuracy decreased below a light intensity of 1000 lux. We explored software-based approaches, including dynamic signal normalization, smoothing, and frequency filtering, to improve performance in low-light conditions but saw limited success due to the low signal-to-noise ratio. Therefore, the real-time testing user study was conducted in a lab with open blinds with variations in natural daylight.

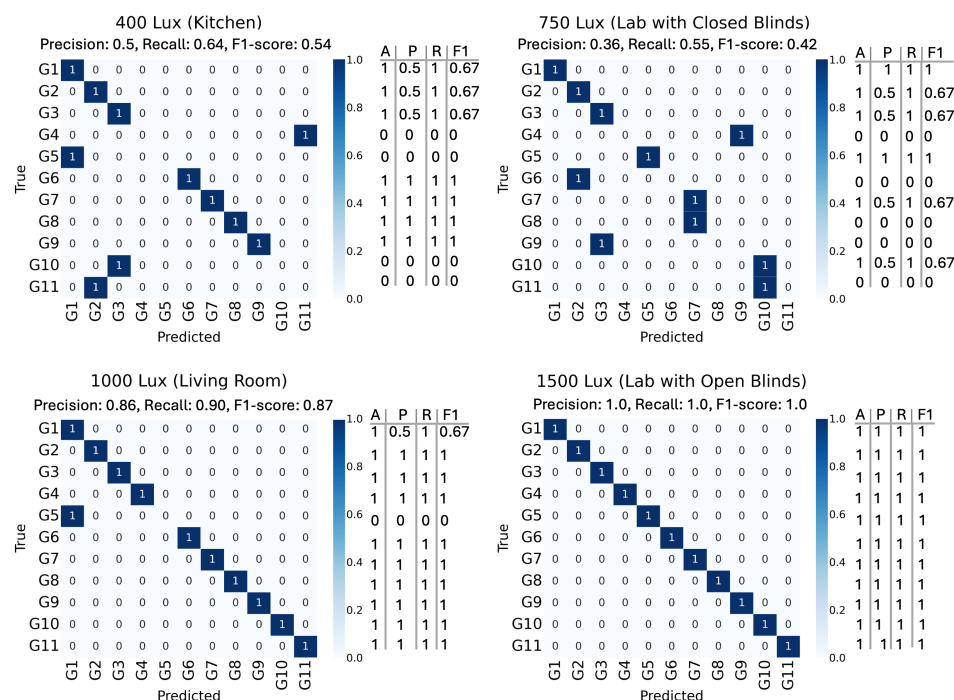


Figure 8. The confusion matrices of different locations with different light intensities, along with four performance matrices for each gesture (A: accuracy, P: precision, R: recall, and F1: F1-score), are shown in four different light light conditions: kitchen (400 lux), lab with closed blinds (750 lux), living room (1000 lux), and lab with open blinds (1500 lux).

5.2. User Experience Evaluation

In this session, which took around 15 min, the participants completed two questionnaires and attended a semi-structured interview. We used a questionnaire by Villanueva et al. as the first questionnaire [58], which focused on three aspects: acceptability, usability, and summary/debriefing (see Table A2). The second questionnaire was the user experience questionnaire (UEQ) [59,60]. For the semi-structured interview, we used the five questions from Shahmiri et al. [29] (see Table A3).

5.2.1. Quantitative Survey Results

The results of the first questionnaire [58] in Table A2 are shown in Figure 9. Using a 5-point Likert scale, the 10 participants indicated mean acceptability, usability, and summary/debriefing scores of 4.63 (SD 0.2), 3.38 (SD 0.4), and 4.63 (SD 0.2), respectively. Regarding acceptability, participants indicated a higher perceived usefulness (A1), ease of use (A2), and intention of use (A4) than an attitude towards using the technology (A3). Regarding the summary/debriefing, participants favored the interaction mode (S1) over the perceived investment (S2 and S3). Regarding usability, participants rated memorability (U1), learnability (U2), and satisfaction (U5) considerably higher than efficiency (U3), errors (U4), and comfort (U6). The A1, U2, and S1 ratings were very high at 4.8 (SD 0.4), 4.6 (SD 0.7), and 4.9 (SD 0.3), respectively, suggesting high perceived usefulness, learnability, and openness to using gestures compared to other available modes of interaction, such as voice commands and remote controls.

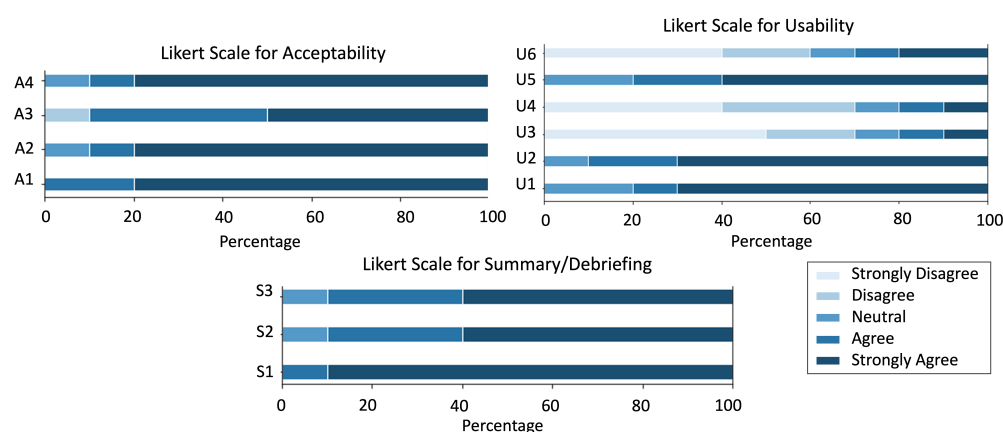


Figure 9. The percentages of responses from the acceptance, usability, and summary/debriefing questionnaire are shown. A 5-point Likert scale ranging from strong disagreement (1) to strong agreement (5) was used.

The results of the second questionnaire (UEQ) are shown in Figure 10. The participant responses from the UEQ were analyzed using the UEQ data analysis tool (UEQ data analysis tool, <https://www.ueq-online.org>, accessed on 2 September 2025). The mean scores for all six scales were interpreted as excellent, ranging from 2.30 to 2.68 on a scale from −3 to 3. This demonstrates that the participants were satisfied with and had a positive experience with the system. The attractiveness and perspicuity categories scored the highest, with means of 2.68 and 2.63, respectively. Both scores are in the excellent range, which highlights how appealing and easy to use the system is.

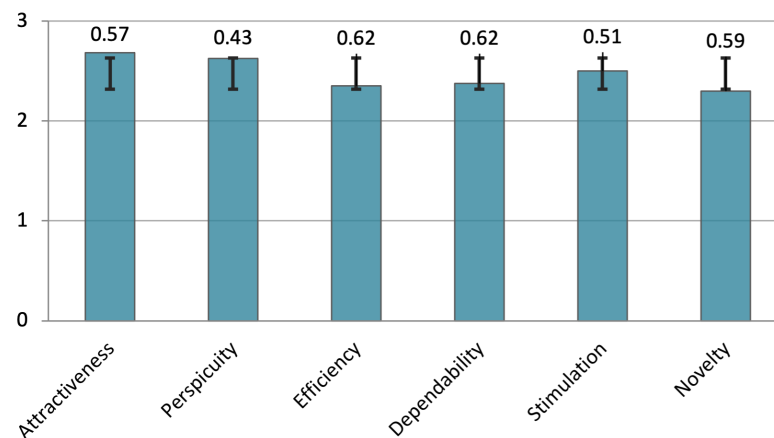


Figure 10. The user experience questionnaire (UEQ) scores for the six dimensions of system usability and user perception—attractiveness, perspicuity, efficiency, dependability, stimulation, and novelty—are shown.

In the stimulation category, users indicated that the system is effective in interaction and engagement by rating it with a high mean of 2.50. The final three categories—efficiency, dependability, and novelty—had slightly lower means of 2.35, 2.38, and 2.30, respectively. Overall, the system was rated as delivering a very positive user experience, as reflected by the ratings in all six categories in the UEQ.

5.2.2. Qualitative Semi-Structured Interview Results

We captured the interview data of the 10 participants (see Table A3 [29]) through notes and voice recordings. We then transcribed the data and conducted a qualitative data analysis on this data using NVivo software. This analysis produced five themes: Natural Interaction, Social Appropriateness, Comfort, Usability, and Sources of Frustration.

- (1) **Natural Interaction:** This theme shows that the participants felt the hovering hand gestures were natural and intuitive. For example, P1 said they are everyday actions (“I really thought it was like a daily expression of my body language. So, when I was interacting or doing the gestures, it was like I was performing natural everyday actions, like moving my hand left or right.”). P6 said they are normal (“I felt they were easy, quick to perform, and very natural, close to human movements in daily life, as when we moved forward, right or left, it felt like normal, everyday human motions.”). P4 said they are commonsense (“I do believe that they are intuitive and natural. Personally, I think most of the gestures come from common sense for different animations, like increasing or decreasing the TV volume, which I would use in my daily life.”).
- (2) **Social Appropriateness:** This theme shows that the participants viewed the hovering gestures as non-intrusive and socially appropriate in various private and public settings. For example, P10 said they are expressive: “They are appropriate and expressive. I do not think there is anything inappropriate about them, as they can be used anywhere and anytime. You can use them at home or in your car.” P6 said they are relatable: “I felt the gestures are relatable to all societies and to people in general. I did not feel they were limited to a specific group of people or a specific context, like just in homes or only in shopping malls. Especially as a mum with kids, I think it would even be easy for children to use at home. I feel it would simplify household tasks and definitely improve our quality of life at home by keeping all surfaces clean.” P1 said they are hygienic: “I think in public places it is nice, maybe because of some diseases or viruses, people do not want to touch things, so this will increase hygiene. But I think it will be more enjoyable at home because you are in your own comfort zone, so it would be convenient in both places.”

- (3) **Comfort:** This theme shows that the participants found the gestures comfortable to perform without much physical effort and causing no strain on the body. For example, P4 said they all are comfortable (*"No, they were all comfortable. I was comfortable applying them. I do not recall any gesture that I felt uncomfortable with."*). P7 said they cause no strain (*"No, not at all. On the contrary, they were simple, easy, and comfortable. There was no issue or strain on the body."*). P10 said they do not need much physical effort (*"No, I do not think so. They are all comfortable and easy, and they do not need any effort to perform."*). P5 said they are easy to perform (*"For me, I thought it was really comfortable, even the gestures were not that hard. It was an easy opening. Nothing very tricky, you just need to remember it."*). P9 said they are simple (easy) to perform (*"I felt completely comfortable, no effort at all. Compared to physically handling a device to perform actions like playing or skipping music, the gestures would be much simpler and not complicated at all."*).
- (4) **Usability:** This theme shows that the participants found the gestures easy to learn and perform after a brief description and demonstration. For example, P8 needed a short demonstration to learn (*"After practicing the gestures a few times, around the third attempt, I learned how to execute them consistently without any mistakes. They were easy to learn and remember after a short period of use."*). P2 said the descriptions were self-explanatory (*"They are all very easy to learn because just a couple of words description was pretty self-explanatory and you could tell straight away what the action needed to be."*). P7 said the gestures were easy to execute (*"The gestures are completely fine, there was not any issue, it was clear and easy to execute."*). P10 said the gestures need no effort to perform (*"They were all easy gestures, easy to perform. It needed no effort, actually."*). Overall, this theme shows that the participants found the gestures easy to learn, remember, recall, and perform.
- (5) **Source of Frustration:** This theme shows that frustration was largely absent in most participants. For example, P3 found no source of annoyance (*"No, I think nothing annoyed me while making the gestures as it was easy."*). P8 found no inconvenience (*"No, everything was clear, and there was no inconvenience during the experience. Even in the future, if these gestures are implemented, I will use them without any issues."*). However, some participants did express a degree of frustration. P1 was scared of making mistakes (*"It was fun, and I liked it, and it was also exciting to do within the application, but I was a bit scared of making some mistakes with the sensor, which frustrated me a little."*). P5 was also worried about making a mistake (*"Never, but I thought if I made a mistake in the move down and the move down quickly gestures, maybe the sensor would not understand it. This is frustrating, but other than that, everything was good for me."*). However, P1 and P5 also said they were excited and happy with the interaction.

6. Discussion

We investigated the development of a novel real-time hovering gesture recognition system using a self-powering PV light sensor. We presented a machine learning pipeline for a PV light sensor to deploy such a system. Working with users, we demonstrated a user-specific modeling approach for the feasibility of real-time hovering gesture recognition with high accuracy. We also deployed a prototype system to evaluate the user experience of using the system when interacting with a smart home device.

6.1. Machine Learning Pipeline for PV-Light-Sensor-Based Hand Gesture Recognition

Hand gesture signals tend to be slow-varying with frequency components below 5 Hz. This allows for capturing data at low sampling rates, such as 50 samples per second. However, the PV sensor signal can be noisy at lower ambient light intensity levels, which may require sampling the gesture signals at a higher rate and then smoothing the data. The

gesture data can be noisy, and the signal magnitude depends on the ambient light intensity and the hand position above the sensor. It can also be affected by user-dependent temporal factors such as the timing and speed of hand movement. Therefore, data augmentation techniques such as noise injection, scaling, and time shifting can be used to create a larger dataset for model training and to avoid overfitting. While these techniques can increase the sample size, feature extraction techniques such as DWT can be used to significantly increase the number of features to help train a model to achieve high accuracy with limited data. We found that the random forest classifier produced the highest accuracy, which aligns with the prior research [19–21,23]. The gesture signal from a PV sensor can vary noticeably from user to user. We found a user-specific modeling approach to be more reliable to reproduce offline testing accuracy in real-time testing. The main limitation of the developed pipeline is that the resulting model is not readily deployable on a microcontroller, though it can be deployed on a microcomputer such as a Raspberry Pi Zero, due to the large size of the exported model (2–3 MB). Development of a low-resource-requiring PV-sensor-based real-time gesture recognition system is a future work.

6.2. User Experience of Using a Real-Time Hovering Gesture Recognition System

Our experimental prototype allowed participants to control a music player by performing hovering gestures above a PV sensor and listening to its output through a speaker with real-time interaction. This real-time prototype offered seamless interaction and a more realistic experience, and it thus allowed for a more realistic evaluation of a hovering gestural interface. We used it to evaluate its user experience, usability, and acceptance in smart home control. Participants perceived the hovering gestures as natural, intuitive, hygienic, comfortable, easy to learn, and socially appropriate with no sources of frustration, showing their willingness to accept such an interactive system. These results align with previous research on touchless gestural interactions [2,49,61,62]. This prototype could be used further to evaluate the application of hovering gestures in other use cases such as interaction with public displays, virtual reality, and gaming.

6.3. Limitations of the User Studies

In the one-size-fits-all model testing study, the real-time performance was unexpectedly low, even though the new testing users were from the same demographic as the users who generated the data. Our work does not indicate how many participants would be required to successfully implement a one-size-fits-all model or whether such an approach would work. We decided against pursuing the one-size-fits-all model approach given the initial results. We did not complete the real-time user testing study after achieving poor results with the first three participants. A comprehensive comparison of the offline and real-time performances of the one-size-fits-all model approach can be obtained by completing the real-time testing study in future work. All of the participants in the real-time testing user study were right-handed young adults. Including a broader range of participants, such as left-handed individuals, children, elderly adults, and people with limited mobility, would have introduced different levels of gesture variability due to factors such as hand size or movement and would have influenced the features of the gesture signals differently. The proposed user-specific models could compensate for these differences. In the future, we would consider including a more diverse range of users from different gender and age groups, cultures, and physical abilities. Due to time constraints, an extensive real-time test could not be performed in the second user study. Therefore, the results, including the accuracy and statistical significance numbers, are only indicative, not conclusive. This issue could be resolved by recruiting a larger number of participants to explore the user-specific modeling approach in the future.

6.4. Limitations of PV Light Sensors

The main limitation of the PV light sensor is low performance in low light intensities. It requires bright and steady lighting to generate gesture signals with good signal-to-noise ratio. The PV sensor will be ineffective in low-light conditions such as watching TV in a dark room where the light intensity would be less than 1000 lux. This was expected, as it is for other light-based sensors such as RGB cameras. Aesthetics could be another limitation of a PV light sensor. Though they can be made available in various colors and patterns, they can be expensive, and most PV sensors look like a solar panel with typical dark-colored patterns.

6.5. Ethical and Privacy Concerns

Using an always-on gesture recognition system at home raises ethical and privacy concerns [63]. For example, users may be concerned about or be unaware of being continuously monitored. While the PV surface harvests energy continuously, the power management and data handling strategies of the gesture recognition system can be designed to address these concerns. The PV signal could serve as an external trigger to put the system into light or deep sleep. The system would wake up when there is a sudden, large change in the signal and would enter a sleep state when there is no change in the signal for an extended period. The system would not store data and would only process it locally. It would also communicate with a smart home device only without an internet connection.

6.6. A Portable Prototype

The current experimental prototype used an appropriate PV light sensor and also used a PC, an oscilloscope, and USB communication. In a commercial product, a small, low-power microcontroller board such as nRF54 or ESP32 with low-power ADCs and low-power communication such as BLE will be used. Inspired by historical ancient scrolls, we imagine a cylindrical shell with a spring mechanism to pull a PV sensor (solar tape) and securely roll and store it around the cylinder [64]. The sensor could be then pulled and unfurled on any surface at home during usage. The core of the shell could house the microcontroller board and other electronic components. There could be vents on either side of the cylinder to dissipate heat. The PV sensor would harvest energy when it is unfurled during normal operation. When it is rolled around the cylinder, it would harvest a limited amount of energy, sufficient to power the electronics in a deep sleep state. There would be a limited amount of heat generation during storage or transport. PV sensors such as flexible organic solar films are designed to be durable with repeated rolling and unrolling, if not frequent rolling and unrolling daily. These sensors are not developed for the proposed application. The PV sensor is electrically insulated by a transparent material such as a plastic film and does not pose a risk of electrical shock. To minimize potential hazards, all electrical components should be properly enclosed away from end-user contact.

7. Conclusions

In this paper, we investigated the development, deployment, and evaluation of a self-powered, real-time hover gesture recognition system for controlling multiple smart home devices using a PV light sensor. We aimed to provide a comprehensive workflow of hover interactions, from gesture classification to real-time gesture recognition and user testing, demonstrating the viability and user acceptance of hands-free interaction via PV interfaces. We presented development of a real-time gesture recognition prototype for smart home control with 11 distinct hover gestures, using user-specific random forest models that worked effectively in real-time testing with an average accuracy of 87.4%. The real-time performance was robust despite variations in gesture timing and speed and in

light conditions. Participants could easily control smart home devices such as a music player through a PV light sensor. Our results from the user testing study emphasized the intuitiveness, comfort, and social acceptability of this technology, as well as the learnability, memorability, and enjoyment of gestural interaction for use in shared environments, such as homes and public places. Overall, we demonstrated the potential of the self-powered PV light sensor to serve as a useful interface in a smart home to control multiple devices through hover gestures. This work shows that the systematic integration of sustainable sensing and machine learning is possible to build a user-friendly interaction experience that is technically feasible. Our work sets the foundation for and opens new research areas to continue the investigation and development of self-powered gesture classification systems for sustainable HCI, low-power ubiquitous computing, and surface-based gestural interaction.

Author Contributions: Conceptualization, N.A. and D.S.; methodology, N.A. and D.S.; software, N.A. and S.A.; validation, N.A., S.A., and D.S.; formal analysis, N.A. and S.A.; investigation, N.A.; resources, D.S.; data curation, N.A.; writing—original draft preparation, N.A., S.A., and D.S.; writing—review and editing, N.A., S.A., and D.S.; visualization, N.A. and S.A.; supervision, D.S.; project administration, N.A. and S.A.; funding acquisition, D.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Swansea University by the Engineering and Physical Sciences Research Council grant (EPSRC) EP/W025396/1.

Institutional Review Board Statement: The study was conducted according to the guidelines of UKRI and the Concordat to support Research Integrity and approved by the Institutional Review Board of the Science and Engineering Ethics Committee, Swansea University (Approval Date: 28 May 2024), Research Ethics Approval Number: 1 2024 9480 8867.

Informed Consent Statement: Informed consent was obtained from all participants involved in the study.

Data Availability Statement: The researchers collected all the data, which is contained and analyzed within the manuscript and is not publicly available.

Acknowledgments: The authors sincerely thank all participants who took part in the study, appreciate the effort of the editors and reviewers, and acknowledge Swansea University and Shaqra University for their provided support.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ADCs	Analogue-to-Digital Converters
BLE	Bluetooth Low Energy
CFAR	Constant False Alarm Rate
DT	Decision Tree
DWT	Discrete Wavelet Transform
DTW	Dynamic Time Warping
EMG	Electromyography
GB	Gradient Boosting
GUI	Graphical User Interface
HCI	Human–Computer Interaction
KNN	K-Nearest Neighbor
LR	Logistic Regression
LED	Light-Emitting Diode

ML	Machine Learning
PV	Photovoltaic
PC	Personal Computer
PD	Photodiode
RF	Random Forest
PCA	Principal Component Analysis
SVM	Support Vector Machine
UEQ	User Experience Questionnaire
UX	User Experience
VR	Virtual Reality
VR/AR	Virtual/Augmented Reality

Appendix A

Table A1. The hovering gestures with their corresponding music player functionalities.

Gestures	Music Player Functions
G1: Move down	Decrease the volume of the song
G2: Move up	Increase the volume of the song
G3: Move right	Play the next song
G4: Move left	Play the previous song
G5: Closing fist	Toggle mute
G6: Move forward	Stop the song
G7: Clockwise	Skip the song forward 5 s
G8: Counterclockwise	Rewind the song 5 s
G9: Move down quickly	Toggle play and pause the song
G10: Split hands	Select the music folder
G11: Combine hands	Clear the song list

Table A2. Likert items on acceptance, usability, and summary/debriefing aspects [58].

No.	Statement	Heuristics	Aspect
A1	Gesture control will improve my overall experience with smart homes.	Perceived usefulness	Acceptability
A2	Gesture control will make interacting with smart homes easier.	Perceived ease of use	Acceptability
A3	I will easily get used to smart home interactions with the help of gestures.	Attitude towards using the technology	Acceptability
A4	Gestures will be a typical way of interacting with technology in the future.	Behavior towards intention of use	Acceptability
U1	The gestures are generally easy to remember.	Memorability	Usability
U2	Most gestures are easy to learn because they correspond well with the commands.	Learnability	Usability
U3	The gestures are generally very complex and complicated to perform.	Efficiency	Usability
U4	It is easy to make errors or mistakes with the current set of gestures.	Error	Usability
U5	I am generally satisfied with the gestures used for smart home interaction.	Satisfaction	Usability
U6	Most gestures are straining to the arms and hands.	Comfort	Usability
S1	Compared to other media (voice command and remote controls), I am open to using gestures to interact with smart homes.	Comparison to other available mode of interactions	Summary/Debriefing
S2	I would buy (or invest in) smart home devices to control my home.	Perceived investment	Summary/Debriefing
S3	I would buy (or invest in) gesture technologies to interact with my smart home.	Perceived investment	Summary/Debriefing

Table A3. Semi-structured interview questions [29].

#	Question
1	How intuitive and natural do you think the gestures are while performing them?
2	How expressive do you think these gestures are and socially appropriate to be used in different environments (e.g., at home or in public)?
3	How comfortable were you while performing the gestures? Are there any gestures that cause discomfort?
4	How easy was it to learn and perform the gestures? Did you encounter any difficulties?
5	Did you encounter any frustration while performing these gestures? If so, what were the causes of frustration?

References

- Iqbal, M.Z.; Campbell, A.G. From luxury to necessity: Progress of touchless interaction technology. *Technol. Soc.* **2021**, *67*, 101796. [CrossRef]
- Pearson, J.; Bailey, G.; Robinson, S.; Jones, M.; Owen, T.; Zhang, C.; Reitmaier, T.; Steer, C.; Carter, A.; Sahoo, D.R.; et al. Can't Touch This: Rethinking Public Technology in a COVID-19 Era. In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, New Orleans, LA, USA, 29 April–5 May 2022. [CrossRef]
- Almania, N.; Alhouli, S.; Sahoo, D. User Preferences for Smart Applications for a Self-Powered Tabletop Gestural Interface. In *Human Interaction and Emerging Technologies (IHET-AI 2024): Artificial Intelligence and Future Applications*; AHFE International: St Honolulu, HI, USA, 2024. [CrossRef]
- Gerba, C.P.; Wuollet, A.L.; Raisanen, P.; Lopez, G.U. Bacterial contamination of computer touch screens. *Am. J. Infect. Control* **2016**, *44*, 358–360. [CrossRef]
- Iqbal, M.Z.; Campbell, A. The emerging need for touchless interaction technologies. *Interactions* **2020**, *27*, 51–52. [CrossRef]
- Meena, Y.K.; Seunarine, K.; Sahoo, D.R.; Robinson, S.; Pearson, J.; Zhang, C.; Carnie, M.; Pockett, A.; Prescott, A.; Thomas, S.K.; et al. PV-Tiles: Towards Closely-Coupled Photovoltaic and Digital Materials for Useful, Beautiful and Sustainable Interactive Surfaces. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 25–30 April 2020; pp. 1–12. [CrossRef]
- Caon, M.; Carrino, S.; Ruffieux, S.; Khaled, O.A.; Mugellini, E. Augmenting interaction possibilities between people with mobility impairments and their surrounding environment. In *Advanced Machine Learning Technologies and Applications, Proceedings of the First International Conference, AMLTA 2012, Cairo, Egypt, 8–10 December 2012*; Proceedings 1; Springer: Berlin/Heidelberg, Germany, 2012; pp. 172–181. [CrossRef]
- Hincapié-Ramos, J.D.; Guo, X.; Moghadasian, P.; Irani, P. Consumed Endurance: A Metric to Quantify Arm Fatigue of Mid-Air Interactions. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Toronto, ON, Canada, 26 April–1 May 2014; pp. 1063–1072. [CrossRef]
- Ruiz, J.; Vogel, D. Soft-Constraints to Reduce Legacy and Performance Bias to Elicit Whole-body Gestures with Low Arm Fatigue. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, Seoul, Republic of Korea, 18–23 April 2015; pp. 3347–3350. [CrossRef]
- Vogiatzidakis, P.; Koutsabasis, P. Frame-Based Elicitation of Mid-Air Gestures for a Smart Home Device Ecosystem. *Informatics* **2019**, *6*, 23. [CrossRef]
- Panger, G. Kinect in the Kitchen: Testing Depth Camera Interactions in Practical Home Environments. In Proceedings of the CHI '12 Extended Abstracts on Human Factors in Computing Systems, Austin, TX, USA, 5–10 May 2012; pp. 1985–1990. [CrossRef]
- He, W.; Martinez, J.; Padhi, R.; Zhang, L.; Ur, B. When Smart Devices Are Stupid: Negative Experiences Using Home Smart Devices. In Proceedings of the 2019 IEEE Security and Privacy Workshops (SPW), San Francisco, CA, USA, 19–23 May 2019; pp. 150–155. [CrossRef]
- Lu, Y.; Wang, X.; Gong, J.; Zhou, L.; Ge, S. Classification, application, challenge, and future of midair gestures in augmented reality. *J. Sens.* **2022**, *2022*, 3208047. [CrossRef]
- Modaberi, M. The Role of Gesture-Based Interaction in Improving User Satisfaction for Touchless Interfaces. *Int. J. Adv. Hum. Comput. Interact.* **2024**, *2*, 20–32. Available online: <https://www.ijahci.com/index.php/ijahci/article/view/17> (accessed on 2 September 2025).
- Amaravati, A.; Xu, S.; Cao, N.; Romberg, J.; Raychowdhury, A. A Light-Powered Smart Camera With Compressed Domain Gesture Detection. *IEEE Trans. Circuits Syst. Video Technol.* **2018**, *28*, 3077–3085. [CrossRef]
- Kartsch, V.; Benatti, S.; Mancini, M.; Magno, M.; Benini, L. Smart Wearable Wristband for EMG based Gesture Recognition Powered by Solar Energy Harvester. In Proceedings of the 2018 IEEE International Symposium on Circuits and Systems (ISCAS), Florence, Italy, 27–30 May 2018; pp. 1–5. [CrossRef]
- Bono, F.M.; Polinelli, A.; Radicioni, L.; Benedetti, L.; Castelli-Dezza, F.; Cinquemani, S.; Belloli, M. Wireless Accelerometer Architecture for Bridge SHM: From Sensor Design to System Deployment. *Future Internet* **2025**, *17*, 29. [CrossRef]

18. Thomas, G.; Pockett, A.; Seunarine, K.; Carnie, M. Simultaneous Energy Harvesting and Hand Gesture Recognition in Large Area Monolithic Dye-Sensitized Solar Cells. *arXiv* **2023**, arXiv:2311.16284. [[CrossRef](#)]
19. Li, J.; Xu, Q.; Xu, Z.; Ge, C.; Ruan, L.; Liang, X.; Ding, W.; Gui, W.; Zhang, X.P. PowerGest: Self-Powered Gesture Recognition for Command Input and Robotic Manipulation. In Proceedings of the 2024 IEEE 30th International Conference on Parallel and Distributed Systems (ICPADS), Belgrade, Serbia, 10–14 October 2024; pp. 528–535. [[CrossRef](#)]
20. Li, J.; Xu, Q.; Zhu, Q.; Xie, Z.; Xu, Z.; Ge, C.; Ruan, L.; Fu, H.Y.; Liang, X.; Ding, W.; et al. Demo: SolarSense: A Self-powered Ubiquitous Gesture Recognition System for Industrial Human-Computer Interaction. In Proceedings of the 22nd Annual International Conference on Mobile Systems, Applications and Services, Tokyo, Japan, 3–7 June 2024; pp. 600–601. [[CrossRef](#)]
21. Ma, D.; Lan, G.; Hassan, M.; Hu, W.; Upama, M.B.; Uddin, A.; Youssef, M. SolarGest: Ubiquitous and Battery-free Gesture Recognition using Solar Cells. In Proceedings of the 25th Annual International Conference on Mobile Computing and Networking, Los Cabos, Mexico, 21–25 October 2019. [[CrossRef](#)]
22. Kyung-Chan, A.; Jun-Ying, L.; Chu-Feng, Y.; Timothy, N.S.E.; Varku, S.; Qinjie, W.; Kajal, P.; Mathews, N.; Basu, A.; Kim, T.T.H. A Dynamic Gesture Recognition Algorithm Using Single Halide Perovskite Photovoltaic Cell for Human-Machine Interaction. In Proceedings of the 2024 International Conference on Electronics, Information, and Communication (ICEIC), Taipei, Taiwan, 28–31 January 2024; pp. 1–4. [[CrossRef](#)]
23. Sorescu, C.; Meena, Y.K.; Sahoo, D.R. PViMat: A Self-Powered Portable and Rollable Large Area Gestural Interface Using Indoor Light. In Proceedings of the Adjunct Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology, Virtual Event, 20–23 October 2020; pp. 80–83. [[CrossRef](#)]
24. Almania, N.A.; Alhouli, S.Y.; Sahoo, D.R. Dynamic Hover Gesture Classification using Photovoltaic Sensor and Machine Learning. In Proceedings of the 2024 8th International Conference on Advances in Artificial Intelligence, London, UK, 17–19 October 2024; pp. 268–275. [[CrossRef](#)]
25. Sandhu, M.M.; Khalifa, S.; Geissdoerfer, K.; Jurdak, R.; Portmann, M. SolAR: Energy Positive Human Activity Recognition using Solar Cells. In Proceedings of the 2021 IEEE International Conference on Pervasive Computing and Communications (PerCom), Kassel, Germany, 22–26 March 2021; pp. 1–10. [[CrossRef](#)]
26. Li, Y.; Li, T.; Patel, R.A.; Yang, X.D.; Zhou, X. Self-Powered Gesture Recognition with Ambient Light. In Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology, Berlin, Germany, 14–17 October 2018; pp. 595–608. [[CrossRef](#)]
27. Venkatnarayan, R.H.; Shahzad, M. Gesture Recognition Using Ambient Light. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **2018**, *2*, 1–28. [[CrossRef](#)]
28. Kaholokula, M. Reusing Ambient Light to Recognize Hand Gestures. Technical Report. 2016. Available online: https://digitalcommons.dartmouth.edu/senior_theses/105/ (accessed on 2 September 2025).
29. Shahmiri, F.; Chen, C.; Waghmare, A.; Zhang, D.; Mittal, S.; Zhang, S.L.; Wang, Y.C.; Wang, Z.L.; Starner, T.E.; Abowd, G.D. Serpentine: A Self-Powered Reversibly Deformable Cord Sensor for Human Input. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, Glasgow, UK, 4–9 May 2019; pp. 1–14. [[CrossRef](#)]
30. Zhu, J.; Liu, Z.; Sun, Y.; Yang, Y. SemiGest: Recognizing Hand Gestures via Visible Light Sensing with Fewer Labels. In Proceedings of the 2023 19th International Conference on Mobility, Sensing and Networking (MSN), Nanjing, China, 14–16 December 2023; pp. 837–842. [[CrossRef](#)]
31. Wen, F.; Sun, Z.; He, T.; Shi, Q.; Zhu, M.; Zhang, Z.; Li, L.; Zhang, T.; Lee, C. Machine learning glove using self-powered conductive superhydrophobic triboelectric textile for gesture recognition in VR/AR applications. *Adv. Sci.* **2020**, *7*, 2000261. [[CrossRef](#)]
32. Khalid, S.; Raouf, I.; Khan, A.; Kim, N.; Kim, H.S. A review of human-powered energy harvesting for smart electronics: Recent progress and challenges. *Int. J. Precis. Eng.-Manuf.-Green Technol.* **2019**, *6*, 821–851. [[CrossRef](#)]
33. Chen, Z.; Gao, F.; Liang, J. Kinetic energy harvesting based sensing and IoT systems: A review. *Front. Electron.* **2022**, *3*. [[CrossRef](#)]
34. Ding, Q.; Rasheed, A.; Zhang, H.; Ajmal, S.; Dastgeer, G.; Saidov, K.; Ruzimuradov, O.; Mamatkulov, S.; He, W.; Wang, P. A Coaxial Triboelectric Fiber Sensor for Human Motion Recognition and Rehabilitation via Machine Learning. *Nanoenergy Adv.* **2024**, *4*, 355–366. [[CrossRef](#)]
35. Ni, T.; Sun, Z.; Han, M.; Xie, Y.; Lan, G.; Li, Z.; Gu, T.; Xu, W. REHSense: Towards Battery-Free Wireless Sensing via Radio Frequency Energy Harvesting. In Proceedings of the Twenty-Fifth International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing, Athens, Greece, 14–17 October 2024; pp. 211–220. [[CrossRef](#)]
36. Alijani, M.; Cock, C.D.; Joseph, W.; Plets, D. Device-Free Visible Light Sensing: A Survey. *IEEE Commun. Surv. Tutor.* **2025**, Early Access. [[CrossRef](#)]
37. Li, T.; An, C.; Tian, Z.; Campbell, A.T.; Zhou, X. Human Sensing Using Visible Light Communication. In Proceedings of the 21st Annual International Conference on Mobile Computing and Networking, Paris, France, 7–11 September 2015; pp. 331–344. [[CrossRef](#)]

38. Li, T.; Liu, Q.; Zhou, X. Practical Human Sensing in the Light. In Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services, Singapore, 26–30 June 2016; pp. 71–84. [\[CrossRef\]](#)
39. Hu, Q.; Yu, Z.; Wang, Z.; Guo, B.; Chen, C. ViHand: Gesture Recognition with Ambient Light. In Proceedings of the 2019 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCOM/IOP/SCI), Leicester, UK, 19–23 August 2019; pp. 468–474. [\[CrossRef\]](#)
40. Mahmood, S.; Venkatnarayan, R.H.; Shahzad, M. Recognizing Human Gestures Using Ambient Light. In Proceedings of the 2020 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCOM/CyberSciTech), Calgary, AB, Canada, 17–22 August 2020; pp. 420–428. [\[CrossRef\]](#)
41. Li, T.; Xiong, X.; Xie, Y.; Hito, G.; Yang, X.D.; Zhou, X. Reconstructing Hand Poses Using Visible Light. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **2017**, *1*, 1–20. [\[CrossRef\]](#)
42. Manabe, H.; Fukumoto, M. Touch sensing by partial shadowing of PV module. In Proceedings of the Adjunct 25th Annual ACM Symposium on User Interface Software and Technology, Cambridge, MA, USA, 7–10 October 2012; pp. 7–8. [\[CrossRef\]](#)
43. Varshney, A.; Soleiman, A.; Mottola, L.; Voigt, T. Battery-free Visible Light Sensing. In Proceedings of the 4th ACM Workshop on Visible Light Communication Systems (VLCS'17), Snowbird, UT, USA, 16 October 2017; pp. 3–8. [\[CrossRef\]](#)
44. Ma, D.; Lan, G.; Hassan, M.; Hu, W.; Upama, M.B.; Uddin, A.; Youssef, M. Gesture Recognition with Transparent Solar Cells: A Feasibility Study. In Proceedings of the 12th International Workshop on Wireless Network Testbeds, Experimental Evaluation & Characterization, New Delhi, India, 2 November 2018; pp. 79–88. [\[CrossRef\]](#)
45. Berman, S.; Stern, H. Sensors for Gesture Recognition Systems. *IEEE Trans. Syst. Man Cybern. Part C (Appl. Rev.)* **2012**, *42*, 277–290. [\[CrossRef\]](#)
46. Vogiatzidakis, P.; Koutsabasis, P. Address and command: Two-handed mid-air interactions with multiple home devices. *Int. J. Hum.-Comput. Stud.* **2022**, *159*, 102755. [\[CrossRef\]](#)
47. Andrei, A.T.; Bilius, L.B.; Vatavu, R.D. Take a Seat, Make a Gesture: Charting User Preferences for On-Chair and From-Chair Gesture Input. In Proceedings of the CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 11–16 May 2024. [\[CrossRef\]](#)
48. Hosseini, M.; Ihmels, T.; Chen, Z.; Koelle, M.; Müller, H.; Boll, S. Towards a Consensus Gesture Set: A Survey of Mid-Air Gestures in HCI for Maximized Agreement Across Domains. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, Hamburg, Germany, 23–28 April 2023. [\[CrossRef\]](#)
49. Hosseini, M.; Mueller, H.; Boll, S. Controlling the Rooms: How People Prefer Using Gestures to Control Their Smart Homes. In Proceedings of the CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 11–16 May 2024. [\[CrossRef\]](#)
50. Vogiatzidakis, P.; Koutsabasis, P. Mid-air gesture control of multiple home devices in spatial augmented reality prototype. *Multimodal Technol. Interact.* **2020**, *4*, 61. [\[CrossRef\]](#)
51. Le Guennec, A.; Malinowski, S.; Tavenard, R. Data Augmentation for Time Series Classification using Convolutional Neural Networks. In Proceedings of the ECML/PKDD Workshop on Advanced Analytics and Learning on Temporal Data, Riva del Garda, Italy, 19–23 September 2016. Available online: <https://shs.hal.science/halshs-01357973> (accessed on 2 September 2025).
52. Islam, M.Z.; Hossain, M.S.; ul Islam, R.; Andersson, K. Static Hand Gesture Recognition using Convolutional Neural Network with Data Augmentation. In Proceedings of the 2019 Joint 8th International Conference on Informatics, Electronics & Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision & Pattern Recognition (icIVPR), Spokane, WA, USA, 30 May–2 June 2019; pp. 324–329. [\[CrossRef\]](#)
53. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [\[CrossRef\]](#)
54. Chaovalit, P.; Gangopadhyay, A.; Karabatis, G.; Chen, Z. Discrete wavelet transform-based time series analysis and mining. *ACM Comput. Surv.* **2011**, *43*, 1–37. [\[CrossRef\]](#)
55. Tallarida, R.J.; Murray, R.B. Chi-square test. In *Manual of Pharmacologic Calculations: With Computer Programs*; Springer: New York, NY, USA, 1987; pp. 140–142. [\[CrossRef\]](#)
56. Ross, A.; Willson, V.L.; Ross, A.; Willson, V.L. Paired samples T-test. In *Basic and Advanced Statistical Tests: Writing Results Sections and Creating Tables and Figures*; SensePublishers: Rotterdam, The Netherlands, 2017; pp. 17–19. [\[CrossRef\]](#)
57. Gałecki, A.; Burzykowski, T. Linear mixed-effects model. In *Linear Mixed-Effects Models Using R: A Step-by-Step Approach*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 245–273. [\[CrossRef\]](#)
58. Villanueva, M.; Drögehorn, O. Using Gestures to Interact with Home Automation Systems: A Socio-Technical Study on Motion Capture Technologies for Smart Homes. In Proceedings of the SEEDS International Conference-Sustainable, Ecological, Engineering Design for Society, 2018. Available online: <https://hal.science/hal-02179898/> (accessed on 2 September 2025).
59. Schrepp, M.; Hinderks, A.; Thomaschewski, J. Construction of a benchmark for the user experience questionnaire (UEQ). *Int. J. Interact. Multimed. Artif. Intell.* **2017**, *4*, 40–44. [\[CrossRef\]](#)

60. Schrepp, M. User experience questionnaire handbook. In *All You Need to Know to Apply the UEQ Successfully in Your Project*; 2015; Volume 10. Available online: <https://www.ueq-online.org/Material/Handbook.pdf> (accessed on 2 September 2025).
61. Gentile, V.; Adjorlu, A.; Serafin, S.; Rocchesso, D.; Sorce, S. Touch or touchless? evaluating usability of interactive displays for persons with autistic spectrum disorders. In *Proceedings of the 8th ACM International Symposium on Pervasive Displays*, Palermo, Italy, 12–14 June 2019. [CrossRef]
62. Khamis, M.; Trotter, L.; Mäkelä, V.; Zezschwitz, E.v.; Le, J.; Bulling, A.; Alt, F. CueAuth: Comparing Touch, Mid-Air Gestures, and Gaze for Cue-based Authentication on Situated Displays. In *ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*; Association for Computing Machinery: New York, NY, USA, 2018; Volume 2. [CrossRef]
63. Guhr, N.; Werth, O.; Blacha, P.P.H.; Breitner, M.H. Privacy concerns in the smart home context. *Appl. Sci.* **2020**, *2*, 1–12. [CrossRef]
64. Gomes, A.; Priyadarshana, L.L.; Visser, A.; Carrascal, J.P.; Vertegaal, R. Magicscroll: A rollable display device with flexible screen real estate and gestural input. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services*, Barcelona, Spain, 3–6 September 2018. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.